The Black-White Gap in Non-Cognitive Skills Among Elementary School Children*

Todd Elder
Michigan State University

Yuqing Zhou
UCLA

December 2017

Abstract

A vast literature has examined black-white gaps in cognitive skills, but racial differences in non-cognitive skills have attracted relatively little attention. Using data from two cohorts of the Early Childhood Longitudinal Study, we find large black-white gaps in teacher-reported measures of non-cognitive skills, even after controlling for detailed student and family characteristics. We show that these measures likely understate true racial disparities in non-cognitive skills because of systematic differences across schools in what teacher reports actually represent. Correcting for the resulting bias nearly doubles the size of the estimated gaps, to roughly the same magnitude as analogous gaps in test scores. Our estimates are remarkably stable across cohorts, suggesting that black children have neither made nor lost ground in recent decades. Finally, supplemental analyses based on the British Cohort Study of 1970 suggest that non-cognitive skills may account for sizeable portions of black-white disparities in adult outcomes.

Keywords: Non-cognitive skills, racial disparities, measurement error

JEL Codes: I21, J15

---

* We thank seminar participants at Cornell University, Michigan State University, and the University of Michigan for helpful comments and suggestions. Jeff Biddle, Steven Haider, and Scott Imberman provided helpful comments on an earlier draft. E-mails: telder@msu.edu; yuqing.zhou.1@anderson.ucla.edu

I. Introduction

Racial disparities in achievement and educational attainment are stubbornly persistent features of the U.S. educational system. An extensive literature spanning several disciplines has studied racial differences in outcomes such as test scores, graduation rates, and college attendance, finding that white students outperform their black classmates in every subject and at every age, including as young as age two (Scott and Sinclair, 1997; Fryer and Levitt, 2013). Moreover, black-white achievement gaps widen sharply in the early years of schooling beyond what would be predicted based on differences in SES and other observable characteristics (Jencks and Phillips, 1998; Fryer and Levitt, 2004, 2006).

As the literature on black-white differences in educational outcomes has evolved, a separate but related literature has highlighted the importance of non-cognitive skills in shaping educational attainment and adult outcomes. For example, Heckman and Rubinstein (2001) present evidence that GED recipients earn less than high school dropouts with similar cognitive skills because GED recipients have systematically lower levels of non-cognitive skills. Similarly, numerous recent studies have documented the effects of non-cognitive skills on both labor market outcomes and a host of measures of social performance (Heckman et al., 2006; Flossmann et al., 2006; Agan, 2011; Segal, 2013). For example, Heckman et al. (2006) establish that both cognitive and non-cognitive skills are important predictors of teenage pregnancy, tobacco and marijuana use, and participation in criminal activities.[1]

Despite the surge in interest in non-cognitive skills within economics, little attention has been paid to documenting how these skills might vary across demographic groups. In one recent exception, Bertrand and Pan (2013) identify an important role for gender in the production of non-cognitive skills, finding that boys' propensity to engage in disruptive behavior stems partly from gender differences in the non-cognitive returns to parental inputs. To our knowledge, Goldhammer (2012) is the lone study to examine the relationship between non-cognitive skills and race. He estimates a dynamic model of skill formation, finding that non-cognitive skills can explain a portion of the Asian advantage in adult economic outcomes relative to whites, blacks, and Hispanics.[2]

[1] Heckman et al. (2008) and Cunha et al. (2010) emphasize the complementary nature of cognitive and non-cognitive skills by developing and estimating models of the joint evolution of cognitive and non-cognitive skills over the life cycle.

[2] Researchers outside of economics have also suggested that non-cognitive skills may play a role in the development and performance of cognitive skills (Diamond, 2000; Raver et al., 2007; Blair et al., 2007; Magnuson and Waldfogel, 2008).

In this paper, we use data from the 2010-2011 cohort of the Early Childhood Longitudinal Study: Kindergarten Cohort (ECLS-K:2011) to estimate black-white gaps in non-cognitive skills. The ECLS-K:2011 is especially well-suited to the study of non-cognitive skills because it includes detailed teacher assessments of multiple aspects of each child's personality. We also use the original ECLS-K cohort of kindergarteners in the 1998-1999 school year (ECLS-K:1999) in order to track how racial gaps in non-cognitive skills have evolved across birth cohorts.

Our analyses of the ECLS-K cohorts produce four substantive findings. First, we find large, statistically significant black-white gaps in several measures of non-cognitive skills. We focus on three such measures: one meant to capture externalizing problem behaviors like arguing and getting into fights, one meant to capture elements of self-control in interacting with peers, and one meant to capture "approaches to learning", which encompasses attentiveness, task-persistence, and motivation.[3] Much like the widely studied test score gaps described above, racial gaps in these measures are present in Kindergarten and grow steadily in the first three years of school. Second, the gaps shrink considerably after conditioning on detailed controls for home and school environments. Measured in standard deviation units, the conditional gaps at the end of third grade are roughly one-half to three-fourths as large as the conditional test score gaps studied by Fryer and Levitt (2006), who use a similar set of control variables. Third, the estimated gaps are remarkably stable across cohorts. For example, at the end of third grade the unconditional black-white gap in the measure of approaches to learning is 0.524 standard deviations for the ECLS-K:2011 cohort, compared to 0.545 standard deviations for the ECLS-K:1999 cohort. This finding mirrors the temporal patterns found in the literature on test scores, with studies such as Neal (2006) and Magnuson and Waldfogel (2008) concluding that, after declining for decades, the black-white test score gap has remained roughly constant since the early 1990s.

Fourth, and perhaps most troubling, we present evidence that our baseline estimates substantially understate true black-white differences in non-cognitive skills. We argue that this understatement arises due to the nature of teacher-provided student evaluations. For example, the measure intended to capture approaches to learning is based on statements such as "the

[3] The importance of task-persistence, sometimes known as "grit", in determining long-run outcomes has recently received attention in several literatures, most notably in psychology. For example, authors such as Duckworth and Quinn (2009) argue that grit is a more powerful predictor of academic success than is IQ.

student persists in completing tasks”, to which teachers could choose one of four responses: “never”, “sometimes”, “often”, or “very often”. Because there is no natural, objective scale for delimiting these choices, teachers might use relative comparisons across students to guide their responses. Specifically, in a classroom of driven, goal-oriented children, a teacher might respond that a particular student persists in completing tasks “sometimes”, while an identically-behaved student in a classroom with less driven classmates may instead receive a rating of “often” (or “very often”).[4]

In order to assess how relative comparisons across students influence teacher reports of non-cognitive skills, we compare the ratio of the between- to within-school variances for these measures to analogous ratios for more “objective” metrics such as achievement test scores, parental education, and SES. For example, when normed by their respective within-school variances, the between-school variance in third grade math test scores is nearly four times larger than that of the approaches to learning index. We argue that these patterns suggest that students in relatively underperforming schools (and classrooms) receive systematically higher non-cognitive skill ratings from their teachers than identically-behaved students who attend high-performing schools. Because black students are overrepresented at underperforming schools, this phenomenon compresses the unconditional black-white gap in non-cognitive skills.

Our approach to addressing the compression of the distribution of observed non-cognitive skills across schools involves treating the underlying skills as latent variables. Specifically, in order to use the observed measures to recover the distribution of latent skills, we construct counterfactuals that answer the question, “What would be the distribution of the non-cognitive skill measures in the ECLS-K if the ratio of between- to within-school variances were the same as they are for less obviously subjective ones, such as test scores or measures of parental SES?” Across all grade levels and all three measures of non-cognitive skills, applying this correction increases the implied black-white gaps substantially, by roughly 0.25 to 0.35 standard deviations.

[4] The ECLS-K also includes parental ratings of measures of non-cognitive skills, but both DiPrete and Jennings (2012) and Elder (2010) argue that these measures are of much lower quality than the teacher ratings. First, the parental ratings are relatively unstable across survey wave, with first-order autocorrelations of only roughly 0.2, suggesting that much of the variation in the parental ratings is due to measurement error. More generally, parental ratings are much more weakly correlated with observable determinants of outcomes than are teacher ratings, across a broad set of outcomes. For example, compared to teacher ratings, parental ratings of a child’s math ability are much more weakly associated with variables that predict math achievement test scores, such as parental education and income. As Elder (2010) argues, some degree of subjectivity in ratings appears to be unavoidable, regardless of the rater, but teacher ratings are at least based on a well-defined frame of reference (peers within the same classrooms).

These corrected gaps are nearly as large as the corresponding gaps in cognitive skills, as measured by reading and math achievement test scores.

Finally, in order to assess how non-cognitive skill gaps contribute to black-white disparities in adulthood, we analyze the longitudinal British Cohort Study of 1970, which includes information on childhood skills and a host of adult outcomes. In conjunction with the growing literature on the importance of non-cognitive skills, our results suggest that non-cognitive skills may account for sizeable portions of black-white gaps in economic outcomes. For example, skills measured by the approaches to learning index alone could account for an 8 to 10 percentage-point gap in the probability of graduating high school, even conditional on several other measures of skills. These findings highlight the potential effectiveness of policies aimed at building non-cognitive skills for reducing black-white inequalities in adult outcomes.

II. Data and Descriptive Findings

Our analysis is based on data from two cohorts of the ECLS-K, administered by the National Center for Education Statistics (NCES). The original ECLS-K cohort is a longitudinal survey that followed a nationally representative sample of roughly 18,600 children who entered kindergarten in the 1998-1999 school year. NCES re-sampled children in the spring of 1999, the fall and spring of the 1999-2000 school year (when most students were in first grade), and again in the spring of 2002, 2004, and 2007 (when most students were in third, fifth, and eighth grade, respectively). NCES also interviewed parents and teachers in each survey wave. Following NCES's convention, we refer to this survey as ECLS-K:1999.

The second cohort of the ECLS-K, denoted by NCES as ECLS-K:2011, includes 18,200 children who are representative of the national population of kindergarten-age children in the 2010-2011 school year. NCES designed the structure of the ECLS-K:2011 to be nearly identical to the ECLS-K:1999 in order to facilitate comparisons across the two cohorts. NCES first surveyed children in the fall of 2010, with follow-up samples in the spring of 2011, the fall and spring of the 2011-2012 school year, the fall and spring of the 2012-2013 school year, and the spring of the 2013-2014 school year (when most students were in third grade). The NCES also conducted surveys in the spring of 2015 and 2016, but those data are not yet available to researchers.

Both ECLS-K cohorts include detailed information on family background and home environments, as well as information about children's cognitive skills, including reading, math, and science IRT scores. In addition, teachers provided assessments of students' mastery of specific skills in reading, math, and science. These assessments are measured on a five-point integer scale (from zero to four) known as the Academic Rating Scale (ARS), although the interpretation of the scales differs slightly across the two cohorts. In the ECLS-K:2011, a rating of zero indicates "far below grade level" and four indicates "far above grade level", while in ECLS-K:1999, zero indicates "not yet demonstrated the skill, knowledge, or behavior" and four indicates "consistent and competent demonstration of the skill, knowledge, or behavior".

The ECLS-K includes a comprehensive set of weights designed to make analysis samples nationally representative. NCES produced cross-sectional weights for each survey wave, as well as weights to be used for panel analyses. We conduct all empirical analyses both with and without the appropriate sample weights to assess the sensitivity of our results, but we report weighted estimates below. In each survey wave, we restrict our samples to children who have valid teacher reports of non-cognitive skills. For other covariates, we create missing-variable indicators and set missing values equal to the variable's respective sample mean.

Non-Cognitive Skills Based on Teacher and Parent Reports

Teachers in both ECLS-K cohorts rate individual students on scales from 1 ("never") to 4 ("very often") on 24 different dimensions intended to measure social, emotional, and cognitive development. NCES does not release data on each of these 24 items individually, instead aggregating them to five composite scales known as Social Rating Scales (SRS).[5] For example, the "externalizing problem behaviors" scale is based on information about the frequency with which a child acts impulsively, interrupts ongoing activities, fights with other children, gets angry, and argues. The approaches to learning scale is based on information about a child's attentiveness, task persistence, eagerness to learn, learning independence, flexibility, and organization. The third scale, "self-control", includes four items that measure a child's ability to control his or her behavior. The fourth scale, "interpersonal skills", measures a child's ability to interact with others on the basis of five items, and the last scale, "internalizing problem

[5] The Social Rating Scales used by NCES are adaptations of the scales designed by Gresham and Elliot (1990). Because the scales are copyright-protected, we cannot reproduce their precise wording here; we refer interested readers to Gresham and Elliott (1990).

behaviors", includes four items that rate the presence of anxiety, sadness, loneliness, and low self-esteem. Appendix A provides detailed information about the creation of these variables.

The Social Rating Scales are widely-used survey instruments for detecting social and behavioral problems (Gresham and Elliott, 1990). As Bertrand and Pan (2013) and Neidell and Waldfogel (2011) argue, these scales are highly reliable measures of non-cognitive skills, and are arguably the most comprehensive assessments that are usable in large surveys such as the ECLS-K.[6] Although both ECLS-K cohorts include all five scales, we focus on the approaches to learning, externalizing behavior, and self-control scales in this paper because previous studies have shown that these skills influence future educational and labor market outcomes. Bertrand and Pan (2013) show that externalizing problem behaviors are closely linked to school suspension and illegal activities. Cornwell et al. (2013) find that the approaches to learning scale strongly predicts students' current and future performance in school, even conditional on achievement test scores. Elder (2010) shows that among young children, the self-control scale is closely related to ADHD diagnoses and treatment, which are both associated with later educational outcomes.

Descriptive Statistics

Table 1a presents descriptive statistics for the ECLS-K:2011 sample. The column labeled "Full Sample" lists sample means for all children whose race is listed as either "white, non-Hispanic" or "black / African-American, non-Hispanic" from parent interviews. We standardize the approaches to learning, externalizing problem behaviors, and self-control scales to have zero mean and unit standard deviation in this sample. The first column also includes means and standard deviations of the family background and demographic characteristics that we include in the empirical analyses below, all measured in the first survey wave. Mother's and father's education is measured in years, "Parents Married" is a binary variable equaling one if the child's parents were married, and "Two-Parent Household" is defined similarly as a binary variable equal to one if the child lived with both biological parents. The remaining variables include the number of books the child has (again, as of the first survey wave), a binary measure

[6] Specifically, Neidell and Waldfogel (2010) write that the ECLS-K non-cognitive measures appear to have relatively high "validity as assessed by test-retest reliability, internal consistency, interrater reliability, and correlations with other, more advanced behavioral constructs (Elliott et al., 1988) and are considered the most comprehensive assessment that can be widely administered in large surveys such as the ECLS-K (Demaray et al., 1995)."

of gender, an SES composite index, and the child's birth weight, in ounces. We use the relevant cross-sectional weights provided by NCES throughout; for example, we the first grade cross-sectional weight for first grade non-cognitive skills, and so on.

We follow Fryer and Levitt (2004, 2006) in using this relatively parsimonious set of covariates. As in Fryer and Levitt's analyses, our results below using this set of controls change little if we instead add a much more exhaustive set of controls, primarily due to the fact that the SES composite index is such a powerful predictor of child outcomes. NCES created this index for the ECLS-K:1999 cohort based on parental education, parental occupation, and household income. The SES composite for the ECLS-K:2011 cohort is constructed identically. For both cohorts, we standardize the index to have zero mean and unit standard deviation in our estimation samples.

As columns 2 and 3 show, the averages of many of these variables differ dramatically by race. White children outperform black children on all three measures of non-cognitive skills; we invert the scale of the externalizing problem behavior measure so that higher scores represent "better" behavior. The differences are statistically significant, with (unreported) *t*-statistics greater than 10 in all cases. On average, white children perform better than black children on the approaches to learning scale by 0.331 (= 0.063 + 0.268) standard deviations in Kindergarten, with the gap widening gradually and growing to 0.524 standard deviations in third grade. The patterns for the externalizing problem behaviors and self-control scales are similar, with initial gaps of 0.398 and 0.399 standard deviations that grow to 0.521 and 0.524 standard deviations, respectively.

As the remaining rows of the table show, black children in the ECLS-K:2011 sample grow up in more disadvantaged households than do white children, along many dimensions. For example, only 35.5 percent of black children grow up in families with both biological parents present in the household, compared to 76.6 percent among white families. Likewise, the black-white difference in the SES composite index is 0.735 standard deviations.

Table 1b presents descriptive statistics for the ECLS-K:1999 sample. The central patterns are similar to those shown in Table 1a: average non-cognitive skills and home environments differ markedly by race. As in the 2011 cohort, the non-cognitive skill gaps grow as children progress in school, although they appear to stabilize between third and fifth grade (the 1999 cohort includes a fifth-grade wave but no second-grade wave). The non-cognitive skill gaps are generally larger in the 1999 cohort than in the 2011 cohort, although the cross-cohort differences are small and statistically insignificant in all cases; for example, the third grade approaches to learning gap is 0.545 standard deviations in 1999 compared to 0.524 standard deviations in 2011. The home-environment gaps are also similar across the two cohorts, with a gap in the SES index of 0.764 (= 0.149 + 0.615) standard deviations in 1999, compared to 0.735 in 2011. Because these home-environment variables differ so dramatically by race, we turn next to assessing their roles in explaining the racial differences in non-cognitive skills.

III. Baseline Estimates of Black-White Gaps in Non-Cognitive Skills

In order to measure black-white gaps in non-cognitive skills, we begin by estimating equations of the form

(1) $$y_{ij} = \gamma Black_i + X_{ij}^{'}\Theta + \varepsilon_{ij},$$

where *i* indexes children and *j* indexes schools. The vector X_{ij} denotes the control variables included in the regressions, and ε_{ij} denotes unobserved determinants of skills. We limit our sample to non-Hispanic black and white children, so the coefficient on the indicator variable $Black_i$ measures average black-white differences in outcomes y_{ij}.

Table 2 presents the estimates of γ from equation (1). Columns (1) and (4) include estimates from models that include no covariates (corresponding to the black-white differences in means shown in Tables 1a and 1b), columns (2) and (5) add the home-environment controls, and columns (3) and (6) add school fixed effects. All standard errors in the table are robust to heteroscedasticity and serial correlation within schools.

The first three columns of Panel A show estimates for the approaches to learning index in the ECLS-K:2011 cohort. The raw mean difference of -0.331 in Kindergarten declines to -0.109 in column (2), implying that the large disparities in home environment shown in Tables 1a and 1b account for two-thirds of the gap in non-cognitive skills (all estimates in the table are statistically different from zero at conventional significance levels).[7] In column (3), we also include school fixed effects as controls. In contrast to the home-environment controls, inclusion

[7] We have also estimated specifications that include variables meant to directly capture parental time inputs, with little effect on the estimates. In these specifications, controls for parental time inputs include time spent reading to the child, telling stories, singing songs, helping the child create art, helping with chores, playing games, teaching nature or science, building something with the child, engaging in sports, visiting the library, going to a concert, visiting a museum, visiting a zoo, attending a sporting event, helping with homework, and helping children practice numbers.

of the school indicators increases the estimated black-white gap; taken on its face, this pattern suggests that black students systematically attend better schools (in terms of producing non-cognitive skills) than do white students. This is counterintuitive, given our priors and the findings from the literature on cognitive skills. For example, Fryer and Levitt (2004) show that including school fixed effects eliminates two-thirds of the black-white difference in test score trajectories between Kindergarten and first grade. Below we argue that the relative magnitudes of the estimates in columns (2) and (3) are not due to differences in school quality that favor black students, but instead capture systematic differences across schools in the interpretation of the teacher-reported measures of non-cognitive skills.[8]

The estimates in Panel A for grades 1, 2, and 3 in ECLS-K:2011 are similar to those from Kindergarten, in that the inclusion of the home-environment controls substantially reduces the estimated gaps in each grade, while the inclusion of school fixed effects increases those gaps. Although the raw gap widens monotonically between Kindergarten and third grade, there are no clear temporal patterns in the estimates in column (3); the third-grade estimate is only slightly larger than the corresponding estimate for Kindergarten, and smaller than the estimate in first grade.

Turning next to the estimates in columns (4)-(6) from the ECLS-K:1999 cohort, the same broad patterns emerge. Again, the estimated gaps shrink with the inclusion of the home-environment controls but enlarge with the inclusion of school fixed effects, and there is no clear temporal pattern for the estimates in columns (5) and (6). The estimates in columns (3) and (6) are similar in magnitude when comparing the same grades (recall that the ECLS-K:1999 cohort includes a follow-up in grade 5 but not in grade 2).

Finally, Panels B and C show estimates from specification (1) based on the externalizing problem behaviors and self-control indices as dependent variables. Most of the estimates are similar in magnitude to the analogous estimates in Panel A, even though these scales are intended to capture much different aspects of personality than the approaches to learning index. Unlike in Panel A, including school fixed effects does not uniformly inflate the estimated gaps in the bottom two panels. Across all three measures, the magnitudes of the point estimates are dishearteningly similar for the two ECLS-K cohorts, especially in kindergarten and first grade,

[8] In alternative specifications, we include indicators for teachers, rather than for schools. The resulting estimates are quite similar to those in column (3) in all cases. We return to this issue below in the context of a discussion about the mechanisms that drive differences across schools in reported non-cognitive skills.

implying that black students made little measurable progress in catching up to their white counterparts between 1999 and 2011. Given this similarly, we focus primarily on the ECLS-K:2011 cohort hereafter.

In sum, the baseline estimates from Tables 1 and 2 show that black elementary school students lag behind their white counterparts on several dimensions of non-cognitive skills. By the end of third grade, the raw gaps are roughly 0.5 to 0.6 of a standard deviation, but conditioning on controls for home and school environments reduces these gaps by 45 to 66 percent. We turn next to assessing whether these estimates accurately capture the differences in non-cognitive skills between black and white children.

IV. Teachers' Subjective Ratings of Cognitive and Non-Cognitive Skills

Objective and Subjective Measures of Cognitive Skills

The estimates in Table 2 are consistent with large black-white differences in several dimensions of non-cognitive skills. An important limitation of these findings, however, is that the measures of non-cognitive skills available in the ECLS-K cohorts (and, to our knowledge, in any existing data set) are arguably subjective, in that they are based on teacher-provided ratings. In contrast, achievement test-based measures of cognitive skills available in the ECLS-K have not only been found to have high levels of reliability and content validity, but are regularly de-identified before grading, eliminating one potential source of subjectivity.[9] Due to the nature of teacher ratings, de-identification is impossible, which introduces the possibility that subjectivity or implicit biases may directly influence the ratings; for example, one might suspect that at least some fraction of the black-white gaps stems from teachers' biases in favor of white students.

As noted in Section III, one puzzle that emerges from Table 2 is why the estimated gaps in the three indices of non-cognitive skills do not systematically decline after the inclusion of school fixed effects. Because of the close links between residential location and school attendance, including school indicators is roughly similar to including neighborhood indicators (Fryer and Levitt, 2006), and previous research has found strong associations between skills and neighborhood characteristics. Moreover, in our own (unreported) estimates based on the ECLS-

[9] Recent research by Bond and Lang (2013, 2017), among others, has highlighted the debate about what achievement tests actually measure. Nonetheless, the achievement tests available in ECLS-K are among the most reliable of any survey-based test; NCES psychometric reports for the original ECLS-K:1999 suggest that both math and reading IRT tests have test-retest reliabilities of over 0.9 in all grades (NCES, 2002).

K:2011, we find that the inclusion of school indicators substantially reduces black-white gaps in achievement test scores, conditional on home environment. For example, in the spring of third grade, inclusion of school fixed effects reduces the black-white gap in math test scores by 0.19 standard deviations (from 0.81 to 0.62), conditional on the same home-environment variables included above. Similarly, the reading gaps decline by 0.16 standard deviations (from 0.43 to 0.27). It is natural to ask why the inclusion of school fixed effects does not produce similarly large declines in any of the specifications in Table 2.

In order to better understand how teachers form ratings of non-cognitive skills, we first look to teachers' analogous ratings of cognitive skills, with a particular focus on how these ratings are related to achievement test scores. Figure 1 displays averages of four metrics of skills among third grade black students in ECLS-K:2011, separately by categories of school racial composition. For example, the left-most bar in the figure shows that black students who attended schools with fewer than 25 percent minorities had average math achievement test scores approximately 0.44 standard deviations below the full-sample mean. The average standardized score falls to -0.63 among black students in schools with 25-50 percent minority populations, to -0.72 in schools with 50-75 percent minority populations, and to -1.01 in schools with more than 75 percent minority students. In sum, black students who attend heavily-minority schools have lower achievement test scores than black students who attend relatively "white" schools.

The second bar from the left in the figure, labeled "Math Rating", measures students' math skills as rated by teachers' responses to the Academic Rating Scale (ARS) surveys. The ARS are meant to reflect students' proficiency in a subject, and in principle should capture much of the variation that IRT scores do. However, in contrast to the IRT scores, there is essentially no relationship between school racial composition and black students' average ARS scores. Even though black students' IRT scores in the "over 75 percent minority" schools are more than 0.5 standard deviations lower than in the "under 25 percent minority" schools, the average ARS ratings of the two groups differ by only 0.07 standard deviations. Results for reading proficiency, not shown in the figure, are nearly identical: IRT test scores are 0.53 standard deviations lower in the "over 75 percent minority" group than in the "under 25 percent minority" groups, but the average ARS scores differ by only 0.06 standard deviations.

One interpretation of the patterns in Figure 1 is that ARS ratings and IRT test scores do not measure the same dimensions of skills. Table 3 presents evidence counter to this interpretation. Each entry in the table corresponds to an estimate from a separate regression of a measure of third-grade cognitive skills (either standardized ARS ratings or standardized IRT test scores) on a measure of home environment, using the sample of white students only. All estimated models control for school fixed effects. For example, the entry in column (1) in the "Mother's Education" row is the estimated coefficient on mother's education in a regression of the math ARS rating on mother's education and school fixed effects. The entry in column (2) is from an identical regression, except the dependent variable is the child's math IRT test score. The two estimates are remarkably similar: an additional year of mother's education is associated with a 0.129 standard deviation increase in ARS math ratings and a 0.121 standard deviation increase in IRT math test scores, on average. Column (3) presents the *p*-value under the null that the estimates in columns (1) and (2) are equal. Columns (4)-(6) are identical to columns (1)-(3) except they are based on reading skills.

The table shows that ARS ratings and IRT scores have nearly identical relationships with the home-environment variables among white students in ECLS-K:2011. The point estimates are similar in all cases, and we reject the equality of the two coefficients at the 10 percent level in only one case (for math skills' relationship with the number of books that the child has in the home).[10] Judging by the point estimates in the table, ARS ratings have, if anything, slightly stronger associations with home environment variables than do IRT test scores – the estimates in the ARS ratings columns are larger in 12 of the 14 cases. Thus, it is even more surprising that Figure 1 shows the opposite pattern – ARS ratings are more weakly associated with school racial distributions than are IRT scores. Finally, we note that the ARS ratings and IRT scores are closely related to each other within schools, with within-school correlations of greater than 0.6 for both math and reading for all grades.

If ARS ratings and IRT test scores capture the same dimensions of skills, what explains the patterns in Figure 1? One potential mechanism is that the ARS ratings of cognitive skills are driven primarily by comparisons across children in the same school. In ECLS-K:2011, ARS ratings are based on an integer scale from 0 to 4, where the ratings of 0, 1, 2, 3, and 4 represent "far below grade level", "below grade level", "at grade level", "above grade level", and "far above grade level", respectively. The definition of "at grade level", however, is potentially a

[10] Because the distribution of *p*-values is uniform over the [0,1] interval under a true null hypothesis, the probability of finding at least one *p*-value below 0.1 when running 14 hypothesis tests is roughly 65.1 percent.

function of the classroom's achievement distribution. In particular, teachers might implicitly define "at grade level" as the mean of the within-classroom distribution of proficiency. The insensitivity of mean ARS ratings to the racial distributions of schools shown in Figure 1 – despite the strong association between IRT test scores and those racial distributions – is consistent with this interpretation.

Implications for Non-Cognitive Skills

The preceding discussion compared two measures of cognitive skills, one arguably more subjective than the other, but it likely has implications for our measures of non-cognitive skills, for which we only have relatively subjective measures. To see why, we return to Figure 1. The third and fourth bars in the figure show the average values of the approaches to learning and externalizing problem behaviors scales among black students, labeled "LEARN rating" and "EXTERN rating", respectively. As was the case for the math ARS ratings, both measures of non-cognitive skills are relatively insensitive to school racial distributions. For example, the average of the approaches to learning scale is slightly *lower* in the "<25 percent minority" schools than in the most heavily-minority schools (-0.44 versus -0.42). The externalizing problem behaviors scale is relatively flat as well. The similarity of both series to the math ARS ratings is striking, given that all three series are based on qualitative ratings without obvious objective scales. If ARS ratings understate differences across schools in cognitive skill levels, then it is likely that teacher reports of non-cognitive skills understate differences across schools in non-cognitive skill levels.

Figure 2 presents additional evidence about the comparability of non-cognitive skill ratings across schools. Unlike Figure 1, this figure focuses on home environment characteristics that are important for non-cognitive skills. The first bar shows the SES composite index, and the second and third bars show regression-based predictions of the approaches to learning and externalizing problem behavior indices, respectively. These predicted indices are based on regressions of each non-cognitive skill rating on the vector of home environment variables and school fixed effects. In order to focus on the home environments themselves, we exclude the school fixed effects when we form the predicted indices.

The figure shows that black students in predominantly minority schools are substantially more disadvantaged than black students in predominantly white schools; specifically, they are disadvantaged along dimensions that predict non-cognitive skill ratings within schools. For example, the average predicted approaches to learning scale is 0.57 standard deviations lower (-0.08 versus -0.65) in the ">75% minority" schools than in the "<25% minority" schools. The corresponding gradient for the externalizing problem behaviors scale is 0.54 standard deviations (-0.27 versus -0.81). Both of these differences are similar to the 0.57 standard deviation difference in average math IRT scores between the ">75% minority" schools and the "<25% minority" schools shown in Figure 1.[11]

The patterns shown in Figures 1 and 2 strongly suggest that teachers' opinions for what constitutes "normal" levels of achievement and behavior are systematically different in schools with disadvantaged student populations as compared to more advantaged schools. As a result, the unconditional gaps in teacher-reported skills (both cognitive and non-cognitive) might understate true black-white disparities in these skills. We turn next to our approach to correct for this source of potential bias.

V. Estimates of the Latent Racial Gaps in Non-Cognitive Skills

In order to describe our procedure for refining estimates of non-cognitive skill gaps, we first introduce some notation. Let $f(y_j)$ denote the probability density of a skill measure y_j, where $j \in \{1,...,J\}$ indexes dimensions of cognitive and non-cognitive skills (such as mathematics, externalizing problem behaviors, and so on). We may express this density as

$$(2) \qquad f(y_j) = \int_s f(y_j \mid s) dF(s),$$

where $f(y_j \mid s)$ is the density in a school s and $dF(s)$ represents the distribution of schools across students (so that $dF(s)$ captures the number of students enrolled in each school). Expression (2) shows that the overall distribution of skills can be written as the within-school distributions integrated over all schools in the population. Because school-level averages of skills are constant for all students within a school, we can write (2) equivalently as

$$(3) \qquad f(y_j) = \int_{\mu_{js}} f(y_j \mid \mu_{js}) dF(\mu_{js}),$$

[11] The central implications shown in Figures 1 and 2 are unchanged if we instead define the *X*-axis based on other school-level variables, such as quantiles of the SES composite variable.

where μ_{js} is the average level of skill *j* within school *s*.

For our purposes, the central hurdle to estimating the distributions of non-cognitive skills is that they are potentially scaled differently in different schools. As such, we think of y_j and μ_{js} as being latent measures: y_j represents a latent measure of a skill *j* for a particular individual, implying that $\mu_{js} \equiv E(y_j \mid s)$ is also latent (we suppress individual subscripts throughout). In contrast, certain measures of skills, such as test scores, arguably capture scales that do not vary by school; a student who correctly answered all of the questions on a standardized math test performed identically to all other students who correctly answered all of the questions on that test, regardless of school. Thus, for test scores, y_{ij} and μ_{js} are both observed, not latent.[12]

Our goal is to recover the distribution of latent non-cognitive skills, and in particular, to recover the latent *average* skills across schools, from the corresponding observed distributions. In order to do so, we assume that the following condition holds:

Condition 1: The distributions of latent mean skills across schools, $dF(\mu_{js})$, are common for all skills *j* up to a proportionality factor equal to the within-school standard deviation of skills σ_j, i.e., $dF(\mu_{js}/\sigma_j)$ is constant for all skills *j*.

Condition 1 implies that the ratio of the within-school variance to the between-school variance in a given skill *j* is equal across all skills. Although different measures of skills are measured on different scales, Condition 1 is useful because it norms school-level averages for all of those skills to identical metrics.

We further assume a variance-components structure such that, in each school, each student's observed measure of skills $\tilde{y}_{js}$ differs from the latent measure y_{js} by a factor, κ_{js}, that varies across skills and schools but is constant for all students within the school:

(4) $\tilde{y}_{js} = y_{js} + \kappa_{js}$.

[12] We note that test *scores* are likely to be noisy measures of latent cognitive skills, as argued by Heckman and co-authors. We return to this point below.

Even though latent skills are unobservable for all measures of non-cognitive skills (and for all skill measures that have a subjective component), the within-school standard deviation σ_j is identified based on observed skills in a school *s*,

$$\sigma_j^2 \equiv E(y_{js} - \mu_{js})^2 = E(\tilde{y}_{js} - \tilde{\mu}_{js})^2, \tag{5}$$

where the equality in (5) holds because κ_{js} is constant within schools.

To provide some intuition for how Condition 1 can identify the latent distribution of non-cognitive skills, Table 4 presents the ratio of between-school variance to total variance in several observed skill measures in ECLS-K:2011. Panel A shows results for our three central measures of non-cognitive skills and the two ARS ratings. Specifically, the first entry in column (1) shows that between-school variation accounts for only 8.5 percent of the total sample variation in the Kindergarten approaches to learning index, so that the implied estimate of $\sigma_j^2 / Var(\tilde{y}_j)$ is 0.915 (= 1 – 0.085). Columns (2)-(4) show that the between-school variance is less than 10 percent of the total variance of this index in the other three grades as well. The remaining four rows in Panel A show analogous estimates for the other four teacher-reported skill assessments, with similar results: between-school variation accounts for little of the overall variation in any of the indices, as Figure 1 implied.

Panels B and C show that, in comparison to the teacher-reported assessments, between-school variation plays a much larger role in relatively objective measures of skills. Panel B shows results for IRT test scores. In second and third grade, the estimated contribution of between-school variance to IRT test scores is more than three times as large as that of the analogous teacher ARS ratings. Panel C shows estimates for the predicted teacher assessments described in Section IV above. These entries are estimates of the degree of sorting across schools on home-environment characteristics, with those characteristics weighted by a regression index to reflect how they influence teachers' perceptions within a given school. As a comparison of the predicted to actual teacher assessments shows, there is much more sorting on the characteristics that predict non-cognitive skills than on the observed non-cognitive skills themselves. For example, while between-school variation accounts for only 8.5 percent of the total variation in the Kindergarten approaches to learning index, it accounts for 32.2 percent of the total variation in the predicted version of that index.

The intuition underlying Condition 1 is that the degree of sorting on the characteristics that predict non-cognitive skills is informative about the degree of sorting on the *latent* non-cognitive skills themselves. As an illustration of how to implement Condition 1, consider a situation in which we had access to only one objective measure of skills, such as the predicted approaches to learning index. Based on that measure, we would estimate that 30.9 percent of the overall variation in cognitive skills is due to between-school sorting. Condition 1 then implies that in third grade, 30.9 percent the overall variation in *any* skill is due to between-school sorting. To recover the distribution of the latent third grade approaches to learning index, we first use expression (5) to estimate the within-school variance of this index. Then, we re-scale the distribution of school-level means $\tilde{\mu}_{js}$ so that the estimate of $\sigma_j^2 / Var(\tilde{y}_j)$ = 1 – 0.309 = 0.691. In practice, this procedure involves using the between-school distribution of one measure of skills to anchor the between-school distribution of the other measure.

When there are multiple objective measures of skills, the between-school distribution of skills is overidentified, as there is no obvious best choice of an anchoring variable. In practice, we will use the simple average across all seven of the objective measures shown in Table 4: math and reading IRT test scores, as well as the five predicted teacher assessments. We repeat this process separately for each grade level.

Table 5 presents estimates of the black-white gaps in latent non-cognitive skills in ECLS-K:2011 based on Condition 1. Columns (1) and (2) replicate the first two columns of Table 2, showing the gaps in observed skills both with and without home-environment controls. Columns (3) and (4) present analogous gaps in the estimated *latent* non-cognitive skills, as estimated using Condition 1. Column (5) adds school fixed effects, which results in identical estimates from those in column (3) of Table 2 by construction – the latent and observed skill measures have identical distributions within schools.

For each of the three skills, the estimated latent gaps are substantially larger than the corresponding observed gaps. For example, in Panel A, a comparison of columns (1) and (3) of the top row implies that the raw gap in the latent approaches to learning index is 0.330 standard deviations larger (-0.661 versus -0.331) than the gap in the observed index. The relative sizes of the unconditional latent and observed gaps are relatively stable over time, with differences of 0.304, 0.305, and 0.306 standard deviations in first, second, and third grade, respectively.

In all cases, the inclusion of the home-environment controls reduces the estimated latent gaps more than it reduces the corresponding observed gaps. For example, a comparison of columns (1) and (2) in the top row of Panel A shows that the inclusion of the controls reduces the estimated gap by 0.222 standard deviations, from -0.331 to -0.109, while the corresponding reduction based on a comparison of columns (3) and (4) is 0.367 standard deviations, from -0.661 to -0.294. We suspect that this phenomenon arises because a portion of the observed gaps reflects variation across schools in what constitutes "normal" behavior, and this variation has the opposite-signed relationship to the control variables as the within-school variation. For example, we argued above that students in high-SES schools will tend to have lower approaches to learning scores than identically-behaved students in low-SES schools. The estimated latent scales are purged of this source of variation, with the result that the home-environment controls have more explanatory power in the latent scales than in the observed scales. We note that the (unreported) r^2 values from the regressions in Table 5 support this conjecture; for example, in Panel A the partial r^2 of the home-environment controls is 0.128 in column (2), compared to 0.174 in column (4) .

In Figure 3, we return to the question of how the four metrics of skills shown in Figure 1 vary by school racial composition, but we replace the observed teacher-reported skills (math, approaches to learning, and externalizing problem behaviors) with the corresponding estimated latent skills. Unlike Figure 1, Figure 3 shows a strong relationship between school racial composition and the skill levels of black children. The estimated latent approaches to learning scores among black children in the "over 75 percent minority" and "under 25 percent minority" schools differ by 0.53 standard deviations (-0.91 versus -0.38), on average, which is similar in magnitude to the corresponding differences in math and reading IRT test scores (0.57 and 0.53 standard deviations, respectively).

In sum, Figure 3 and Table 5 show that black-white gaps in estimated latent non-cognitive skills are much larger than the corresponding gaps in teacher-reported skills. By the spring of third grade, the black-white unconditional gaps are larger than 0.8 standard deviations for all three latent measures of non-cognitive skills, with the inclusion of home environment controls reducing those gaps by roughly 50 percent. Additionally, our approach to estimating latent non-cognitive skills resolves two puzzling aspects of the baseline results. First, the observed gaps implied that black students in high- and low-minority schools had the same average levels of non-cognitive skills, in spite of large differences in both cognitive skills and observable determinants of non-cognitive skills. Second, the comparison of models with and without school fixed effects in Table 2 implied that black students systematically attend "better" schools than do white students, in terms of producing non-cognitive skills. The results in Table 5 reverse this finding: the inclusion of school fixed effects uniformly reduces the estimated black-white gaps. Taken together, these estimates suggest that the puzzles in Table 2 stemmed from systematic differences across schools in the interpretation of the teacher-reported measures of non-cognitive skills.

Finally, we note that Condition 1 is a strong assumption that is unlikely to hold exactly. Nonetheless, we view it as a better approximation to reality to the implicit assumption underlying the baseline results in Table 2: that the standards teachers use to evaluate achievement and behavior are invariant to the composition of the school's student body. We view our estimates based on Condition 1 as a complement to the baseline analysis, rather than as a replacement for it.

VI. Non-cognitive Skills and Adult Outcomes in BCS70

Although the ECLS-K datasets are valuable for studying non-cognitive skills among children, our ability to draw inferences about adult outcomes is limited because both cohorts only track children through the primary school years. In order to gauge how non-cognitive skills influence later outcomes, we follow Goldhammer (2012) in analyzing the longitudinal British Cohort Study of 1970 (BCS70), which includes teacher-reported behavior ratings of students.

The BCS70 is a longitudinal survey of all infants born in the United Kingdom during the week of April 5-11, 1970. The original survey included 17,198 children, and the BCS70 resampled children when they were 5, 10, 16, 26, 30, 34, and 38 years old. The age-10 survey included teacher-reported non-cognitive skill ratings that are similar (but not identical) to those found in the ECLS-K cohorts. The age-10 survey also included an administration of the Friendly Math Test and the Edinburgh Reading Test, both widely-used achievement tests in the U.K., and we standardize these scores to have zero mean and unit standard deviation in the estimation samples.

The non-cognitive skill measures in BCS include teacher ratings of students from 1 to 100 on a variety of scales, including items measuring externalizing problem behaviors (whether a child fights with other children, destroys belongings, teases other children, and bullies other children), self-control (whether the child displays outbursts of temper, and whether the child is excitable or impulsive), and approaches to learning (whether the child is attentive in class, completes tasks on time, or is forgetful or irresponsible). We use the average of all items to create three scales that are analogous to the ECLS-K scales; for example, to create the externalizing problem behaviors scale, we take the simple average of each of the "child fights with other children", "destroys belongings", "teases other children", and "bullies other children" scales. Finally, as with the standardized tests, we standardize the non-cognitive skill measures to have zero mean and unit standard deviation in the estimation samples.

To measure outcomes, we use the 2008 (age-38) follow-up, which includes a variety of measures of economics and social success. We focus on the following variables:

- A binary measure of whether the respondent obtained a high school degree, as measured by passing the British O-level exams ("HS Graduate")
- A binary measure of whether the respondent graduated from university ("College Degree")
- A binary measure of whether the respondent was arrested at any point in the previous 10 years ("Arrest")
- A binary measure of whether the respondent was unemployed in the previous year ("Unemployment")
- The logarithm of labor earnings in the previous year ("Log(Wages)")

We then estimate linear regression models of each of these five outcomes as a function of the five non-cognitive and cognitive skill measures.[13]

Table 6 presents the resulting estimates. We report estimates separately by gender because the functions linking skills and outcomes are likely to differ for males and females. Column (1) presents estimates for high school graduation. For both males and females, the largest point estimates are those associated with the approaches to learning index, which imply that a one-standard-deviation increase in the index increases the probability of graduating high school by at least 10 percentage points, even conditional on the other non-cognitive and cognitive skill measures. This finding is consistent with the results from the psychology

[13] Estimates of marginal effects from probit models for the binary outcomes are nearly identical to the linear probability estimates in all cases.

literature (see, e.g., Duckworth and Quinn (2009)) that task-persistence and attentiveness are even more powerful predictors of educational attainment than are cognitive skills. The externalizing problem behavior index (which is again inverted so that higher scores represent "better" behaviors) and both math and reading test scores are significant predictors of high school graduation for men, but only reading scores are significant for women.

The relative importance of the skill measures varies somewhat across outcomes, but in all cases, non-cognitive skills significantly affect outcomes. The approaches to learning scale loads most heavily on education and log(wages), and the externalizing problem behaviors scale loads most heavily on the probability of arrest. For men, a one standard deviation increase in the externalizing problem behaviors index is associated with a 0.078 percentage point decrease in the probability of arrest, which is 26 percent of the baseline sample mean of 0.300. For women, the analogous point estimate is -0.024, a 42 percent reduction compared to the sample mean of 0.057. We find it remarkable (yet intuitively appealing) that teachers' assessments of a child's behavior at age 10 are powerful predictors of arrests twenty years later. In contrast, the cognitive measures do not significantly impact arrest rates for either men or women.

We caution against using the results of Table 6 to draw firm conclusions about racial disparities in the U.S., as there are several reasons to believe that the mapping between skills and outcomes in the U.S. population and the BCS70 are likely to differ considerably. Keeping these concerns about external validity in mind, the estimates are still potentially useful to provide a rough sense of the magnitude of the role that skill gaps may play in black-white gaps in adult outcomes. For example, Table 5 implies that the raw black-white gap in (latent) externalizing problem behaviors is roughly 0.8 standard deviations in third grade. Table 6 implies that a gap of this magnitude, even conditional on other skill gaps, could produce a 3.12 ($= 100 \times 0.039 \times 0.8$) percentage-point difference in high school graduation rates of black and white men.

Based on the estimates in Table 6, racial differences in the approaches to learning index appear to play the most significant role in producing black-white differences in adult outcomes among the non-cognitive and cognitive skills that we consider. A 0.8 standard-deviation difference in approaches to learning potentially translates into roughly 8 and 10 percentage-point gaps in the probability of graduating high school for men and women, respectively, and earnings differences of roughly 6 and 12 log points, respectively. We again stress the important caveat that the BCS70 is likely not representative of the U.S. population, but our results suggest that black-white differences in non-cognitive skills could potentially be major factors in producing differences in adult outcomes in the U.S.

VII. Discussion and Conclusions

Using two nationally representative datasets from the ECLS-K, we find evidence of significant differences in observed measures of non-cognitive skills between white and black students, even after controlling for a large set of background variables. The raw gaps in these skills are roughly 0.5 standard deviation units by the end of third grade. Controlling for family background and home environment reduces the estimated gaps but does not eliminate them entirely.

Given that observed non-cognitive skills are based on subjective judgments by teachers, it is natural to ask whether the large estimated black-white gaps stem from teachers' biases against black students. Although our prior belief was that this mechanism might be empirically relevant, we instead found evidence suggesting that the baseline estimates substantially understate true black-white disparities. This understatement arises because our measures of non-cognitive skills are not strictly comparable across schools. Specifically, teachers appear to base their responses on the skills of "typical" students in their classrooms, so that a student in a high-achieving school will tend to have lower teacher-reported skills than an identical student in a low-achieving school. Because white students are disproportionately likely to attend high-achieving schools, white students have lower average teacher-reported skills than otherwise identical black students.

In order to correct for the school-level measurement error in observed non-cognitive skills, we adopt an approach that treats the underlying skills as latent variables. This approach assumes that the distributions of latent mean skills across schools are common for all skills up to a proportionality factor given by the within-school standard deviation of skills, which allows us to recover the latent distributions from the corresponding observed distributions.

After using our approach to estimate the latent distributions of skills, we then estimate black-white gaps in these measures. For each of the three skills we consider, the estimated latent gaps are substantially larger than the corresponding observed gaps. By the end of third grade, the raw gaps in these skills are roughly 80 percent of the corresponding full-sample standard deviations. Unlike the baseline teacher-reported measures, the estimated latent distributions imply that black students in high-minority schools have much lower levels of skills than black students in schools with low minority populations. Moreover, black students systematically attend worse schools in terms of producing non-cognitive skills than do white students. These results suggest that differential school quality is at least partly responsible for the growth in the non-cognitive skill gaps between Kindergarten and third grade.

Finally, we analyze the longitudinal British Cohort Study of 1970 (BCS70) in order to assess how non-cognitive skill gaps contribute to disparities in adult outcomes in the United States. Although our findings are only suggestive because the BCS70 is not meant to be representative of the U.S. population, they are consistent with the possibility that non-cognitive skills account for sizeable portions of black-white gaps in economic outcomes. Most notably, skills measured by the approaches to learning index could account for a roughly 8 to 10 percentage-point difference in the probability of graduating high school and a 6 to 12 log-point difference in earnings.

The magnitudes of the estimated black-white gaps are worrisome, but they provide a new perspective for understanding and addressing racial inequality in the United States. A nascent literature has established that non-cognitive skills have large impacts on adult outcomes, and our own estimates using the BCS70 are consistent with these findings. Importantly, these findings highlight a potentially powerful tool to reduce black-white inequalities because non-cognitive skills are possibly much more malleable than are cognitive skills, especially after age 5. As a result, interventions aimed at reducing non-cognitive skill gaps among school-age children might be an effective policy tool for ameliorating black-white disparities in adult outcomes.

References

Agan, Amanda. 2011 “Non-Cognitive Skills and Crime.” IZA Conference Paper.

Bertrand, Marianne, and Jessica Pan. 2013. “The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior.” *American Economic Journal: Applied Economics*, 5(1): 32-64.

Blair, C., & Razza, R. P. 2007. “Relating Effortful Control, Executive Function, and False Belief Understanding to Emerging Math and Literacy Ability in Kindergarten.” *Child Development*, 78: 647-663

Bond, T., and K. Lang. 2013. “The Evolution of the Black-White Test Score Gap in Grades K–3: The Fragility of Results.” *The Review of Economics and Statistics* 95(5): 1468-1479.

Bond, T., and K. Lang. 2017. “The Black-White Education Scaled Test-Score Gap in Grades K-7. *Journal of Human Resources*: 0916-8242R.

Brooks-Gunn, J., and G. J. Duncan (Eds.). 1997. *The Consequences of Growing Up Poor* (New York: Russell Sage).

Brooks-Gunn, J., and G. J. Duncan. 2000. “Family Poverty, Welfare Reform and Child Development,” *Child Development* 71(1): 188-196.

Brooks-Gunn, J., G. J. Duncan, and P. Klebanov. 1994. “Economic Deprivation and Early-Childhood Development,” *Child Development* 65(2): 296-318.

Brooks-Gunn, J., G. J. Duncan, and P. Klebanov. 1996. “Ethnic Differences in Children’s Intelligence Test Scores: Role of Economic Deprivation, Home Environment and Maternal Characteristics,” *Child Development* 67 (2), 396-408.

Burkett, John; McMillen, Marilyn; Owings, Jeffery and Pinkerton, Daniel B. 1995. “Statistics in Brief: Making the Cut: Who Meets Highly Selective College Entrance Criteria?” Washington, DC: National Education Longitudinal Study.

Campbell, Ernest Q.; Coleman, James S.; Hobson, Carol J.; McPartland, James; Mood, Alexander M.; Weinfeld, Frederic D. and York, Robert L. 1996. *Equality of Educational Opportunity*. Washington D.C.: U.S. Government Printing Office (GOP).

Cook, M., and W. Evans. 2000. “Families or Schools? Explaining the Convergence in White and Black Academic Performance,” *Journal of Labor Economics* 18(4): 729-754.

Cook, P., and J. Ludwig. 1998. “The Burden of Acting White: Do Black Adolescents Disparage Academic Achievement?” (pp. 375–400), in C. Jencks and M. Phillips (Eds.), *The Black-White Test Score Gap* (Washington, DC: The Brookings Institute).

Cornwell C., D. Mustard D, and J. Van Parys. 2013. “Noncognitive skills and the gender disparities in test scores and teacher assessments: evidence from primary school,” *Journal of Human Resources* 2013(48): 236-64.

Cunha, Flavio, James J. Heckman, and Susanne M. Schennach. 2010. “Estimating the Technology of Cognitive and Noncognitive Skill Formation.” *Econometrica*, 78(3): 883–931.

Delpit, L., 1995. *Other People’s Children: Cultural Conflict in the Classroom* (New York: The New Press,).

Demaray, Michelle, Stacey Ruffalo, John Carlson, R. T. Busse, Amy Olson, Susan McManus, and Amy Leventhal, 1995. “Social Skills Assessment: A Comparative Evaluation of Six Published Rating Scales,” *School Psychology Review* 24: 648-671.

Diamond, A., 2000. “Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex.” *Child Development*, 71: 44-56.

DiPrete, T.A. and Jennings, J.L., 2012. “Social and behavioral skills and the gender gap in early educational achievement.” *Social Science Research*, 41(1): 1-15.

Duckworth, A.L. and Quinn, P.D. 2009. “Development and validation of the Short Grit Scale (GRIT–S).” *Journal of Personality Assessment*, 91(2), 166-174.

Elder, T.E., 2010. “The importance of relative standards in ADHD diagnoses: evidence based on exact birth dates.” *Journal of health economics*, 29(5), 641-656.

Elliott, Stephen, Frank Gresham, Terry Freeman, and George McCloskey. 1988. “Teacher and Observed Rating of Children's Social Skills: Validation of the Social Skills Rating Scales.” *Journal of Psychoeducation Assessment*, 6 (1988), 152-161.

Ferguson, R. F., 1998. “Teachers’ Perceptions and Expectations and the Black White Test Score Gap” (pp. 273–317), in C. Jencks and M. Phillips (Eds.), *The Black-White Test Score Gap* (Washington, DC: The Brookings Institute).

Flossman, Anton L., Remi Piatek and Laura Wichert. 2006. “Going Beyond Returns to Education: The Role of Noncognitive Skills on Wages in Germany.” Working Paper, University of Konstanz, ZEW.

Fordham, S., and J. Ogbu. 1986. “Black Students’ School Successes: Coping with the Burden of Acting White.” *The Urban Review* 18(3): 176–206.

Fryer, R. 2002. “An Economic Approach to Cultural Capital.” University of Chicago working paper.

Fryer, Roland G., Jr. and Steven D. Levitt. 2004. “Understanding the Black-White Test Score Gap in the First Two Years of School.” *Review of Economics and Statistics* 86: 447-64.

Fryer, Roland G., Jr. and Steven D. Levitt. 2006. “The Black-White Test Score Gap Through Third Grade.” *American Law and Economics Review* 8: 249-81.

Fryer, Roland G. Jr., and Steven D. Levitt. 2013. “Testing for Racial Differences in the Mental Ability of Young Children.” *American Economic Review*, 103(2): 981-1005.

Gresham, F. M., & Elliott, S. N., 1990. *Social Skills Rating System Manual*. American Guidance Service, Inc., Circle Pines, MN.

Heckman, J. J. 2008. “Schools, Skills and Synapses.” *Economic Inquiry*, 46: 289–324.

Heckman, James J., Yona Rubinstein. 2001. “The Importance of Noncognitive skills: Lessons from the GED testing program.” *American Economic Review*, 91(2): 145-149.

Heckman, James J., Jora Stixrud and Sergio Urzua. 2006. “The Effects of Cognitive and Non-Cognitive Abilities on Labor Market Outcomes and Social Behavior.” *Journal of Labor Economics*, 24(3): 411-482.

Hernstein, R. J., and C. Murray. 1994: *The Bell Curve: Intelligence and Class Structure in American Life* (The Free Press, 1994).

Jencks, Christopher and Meredith Phillips. 1998. “The Black-White Test Score Gap: An Introduction,” in Christopher Jencks and Meredith Phillips (Eds.), *The Black-White Test Score Gap* (Washington, DC: Brookings Institution Press).

Jensen, A. 1969. “How Much Can We Boost IQ and Scholastic Achievement?” *Harvard Educational Review* 39, 1–123.

Mayer, S. E. 1997. *What Money Can’t Buy: Family Income and Children’s Life Chances* (Harvard University Press).

National Center for Education Statistics 2002. “Early Childhood Longitudinal Study-Kindergarten Class of 1998–99 (ECLS–K), Psychometric Report for Kindergarten Through First Grade,” NCES Working Paper No. 2002–05.

Neidell, M., and J. Waldfogel. 2010. “Cognitive and Noncognitive Peer Effects in Early Education.” *Review of Economics and Statistics,* 92(3): 562-576.

Neal, D., and W. R. Johnson. 1996. “The Role of Pre-market Factors in Black-White Wage Differences.” *Journal of Political Economy* 104, 869–895.

O’Neill, J. 1990. “The Role of Human Capital in Earnings Differences between Black and White Men.” *Journal of Economic Perspectives* 4(4): 25–46.

Raver, C. C., Gershoff, E. T., & Aber, J. L. 2007. “Testing equivalence of mediating models of income, parenting, and school readiness for White, Black, and Hispanic children in a national sample.” *Child Development.* 78: 96–115

Rodgers, W., and W. Spriggs. 1996. “What Does AFQT Really Measure: Race, Wages, Schooling and the AFQT Score.” *The Review of Black Political Economy* 24(4): 13 –46.

Rushton, Philippe J. 1995. “Race and crime: international data for 1989-1990.” *Psychological Reports,* 76(1): 307-12.

Scott, R. and Sinclair, D. 1997. “Ethnic-related cognitive profiles of black and white preschool children.” *Homo- Journal of Comparative Human Biology*, 28: 116-20.

Segal, C. 2013. “Misbehavior, education, and labor market outcomes.” *Journal of the European Economic Association,* 11(4): 743-779

Steele, C., and J. Aronson. 1998. “Stereotype Threat and the Test Performance of Academically Successful African Americans,” in C. Jencks and M. Phillips (Eds.), *The Black-White Test Score Gap* (Washington, DC: The Brookings Institute).

Table 1a: Summary Statistics by Race of Non-Cognitive Skills and Family and Student Characteristics: ECLS-K:2011

Variable	Full Sample (N = 10,885)	White (N = 8489)	Black (N = 2396)
Approaches to Learning			
Kindergarten	…	0.063	-0.268
1st Grade	…	0.058	-0.358
2nd Grade	…	0.049	-0.382
3rd Grade	…	0.062	-0.462
Externalizing Problem Behaviors			
Kindergarten	…	0.034	-0.364
1st Grade	…	0.013	-0.419
2nd Grade	…	0.012	-0.462
3rd Grade	…	0.021	-0.500
Self-Control			
Kindergarten	…	0.073	-0.325
1st Grade	…	0.069	-0.383
2nd Grade	…	0.052	-0.440
3rd Grade	…	0.054	-0.530
Family Background Characteristics			
Mother's Education	14.257	14.507	13.295
	(2.301)	(2.263)	(2.190)
Father's Education	14.235	14.362	13.304
	(2.421)	(2.408)	(2.310)
Parents Married	0.616	0.713	0.267
	(0.486)	(0.452)	(0.442)
Two-Parent Household	0.684	0.766	0.355
	(0.465)	(0.423)	(0.479)
Number of Books Child Has	83.737	94.045	42.509
	(59.535)	(58.947)	(41.226)
SES Composite Index	0.098	0.152	-0.583
	(1.000)	(0.963)	(0.919)
Child's Birth Weight (oz.)	116.481	118.084	109.689
	(21.522)	(20.926)	(22.667)

Table 1b: Summary Statistics by Race of Non-Cognitive Skills and Family and Student Characteristics: ECLS-K:1999

Variable	Full Sample (N = 12293)	White (N = 9824)	Black (N = 2469)
Approaches to Learning			
Kindergarten	…	0.108	-0.346
1st Grade	…	0.084	-0.357
2nd Grade	…	0.088	-0.458
3rd Grade	…	0.087	-0.451
Externalizing Problem Behaviors			
Kindergarten	…	0.066	-0.358
1st Grade	…	0.056	-0.383
2nd Grade	…	0.075	-0.552
3rd Grade	…	0.070	-0.554
Self-Control			
Kindergarten	…	0.107	-0.412
1st Grade	…	0.090	-0.380
2nd Grade	…	0.088	-0.530
3rd Grade	…	0.082	-0.528
Family Background Characteristics			
Mother's Education	13.725	13.929	12.883
	(1.938)	(1.932)	(1.724)
Father's Education	13.919	14.025	13.096
	(2.129)	(2.138)	(1.864)
Parents Married	0.680	0.775	0.302
	(0.466)	(0.417)	(0.459)
Two-Parent Household	0.591	0.671	0.270
	(0.492)	(0.470)	(0.444)
Number of Books Child Has	84.628	95.307	45.521
	(60.091)	(59.173)	(47.365)
SES Composite Index	0.000	0.149	-0.615
	(1.000)	(0.948)	(0.973)
Child's Birth Weight (oz.)	118.589	120.360	110.697
	(21.264)	(20.604)	(22.341)

Table 2: Estimated Racial Gaps in Non-Cognitive Skills, 1999 and 2011 ECLS-K Cohorts

	ECLS-K:2011				ECLS-K:1999		
	(1)	(2)	(3)		(4)	(5)	(6)
				A: Approaches to Learning			
Kindergarten	-0.331	-0.109	-0.242	Kindergarten	-0.454	-0.200	-0.242
	(0.026)	(0.028)	(0.040)		(0.025)	(0.035)	(0.050)
Grade 1	-0.416	-0.158	-0.284	Grade 1	-0.441	-0.178	-0.251
	(0.030)	(0.032)	(0.048)		(0.027)	(0.037)	(0.055)
Grade 2	-0.431	-0.129	-0.175	Grade 3	-0.545	-0.206	-0.350
	(0.032)	(0.034)	(0.053)		(0.034)	(0.047)	(0.073)
Grade 3	-0.524	-0.218	-0.274	Grade 5	-0.538	-0.182	-0.221
	(0.035)	(0.037)	(0.058)		(0.036)	(0.052)	(0.082)
				B: Externalizing Problem Behaviors			
Kindergarten	-0.398	-0.172	-0.260	Kindergarten	-0.424	-0.249	-0.270
	(0.027)	(0.029)	(0.042)		(0.027)	(0.035)	(0.050)
Grade 1	-0.432	-0.197	-0.274	Grade 1	-0.439	-0.222	-0.198
	(0.031)	(0.034)	(0.050)		(0.027)	(0.037)	(0.055)
Grade 2	-0.475	-0.231	-0.205	Grade 3	-0.626	-0.290	-0.354
	(0.033)	(0.036)	(0.056)		(0.034)	(0.047)	(0.072)
Grade 3	-0.521	-0.246	-0.201	Grade 5	-0.623	-0.309	-0.221
	(0.036)	(0.039)	(0.060)		(0.037)	(0.053)	(0.082)
Home Environment Controls		X	X			X	X
School Fixed Effects			X				X

Table 2: Estimated Racial Gaps in Non-Cognitive Skills, 1999 and 2011 ECLS-K Cohorts (cont'd)

	ECLS-K:2011				ECLS-K:1999		
	(1)	(2)	(3)		(4)	(5)	(6)
				C: Self-Control			
Kindergarten	-0.399	-0.197	-0.262	Kindergarten	-0.519	-0.330	-0.272
	(0.027)	(0.030)	(0.042)		(0.026)	(0.037)	(0.051)
Grade 1	-0.452	-0.223	-0.285	Grade 1	-0.470	-0.242	-0.241
	(0.031)	(0.034)	(0.050)		(0.027)	(0.039)	(0.056)
Grade 2	-0.492	-0.242	-0.162	Grade 3	-0.619	-0.320	-0.381
	(0.034)	(0.037)	(0.056)		(0.034)	(0.049)	(0.074)
Grade 3	-0.584	-0.284	-0.195	Grade 5	-0.610	-0.282	-0.239
	(0.036)	(0.039)	(0.061)		(0.038)	(0.054)	(0.084)
Home Environment Controls		X	X			X	X
School Fixed Effects			X				X

Notes:
Each entry in the table is an estimate from a separate regression of a measure of non-cognitive skills on an indicator equaling 1 for black children and 0 for white children, corresponding to equation (1) in the text.

Table 3: Associations between Home Environment Controls and Cognitive Skills among White Third Grade Students, ECLS-K:2011

	Math Skills			Reading Skills		
	ARS ratings	IRT test scores	*p*-value for Ho: (1)=(2)	ARS ratings	IRT test scores	*p*-value for Ho: (4)=(5)
	(1)	(2)	(3)	(4)	(5)	(6)
Mother's Education	0.129	0.121	0.433	0.136	0.127	0.388
	(0.008)	(0.008)		(0.008)	(0.008)	
Father's Education	0.124	0.105	0.116	0.127	0.117	0.381
	(0.008)	(0.008)		(0.008)	(0.008)	
Parents Married	0.349	0.296	0.246	0.345	0.327	0.695
	(0.039)	(0.038)		(0.040)	(0.039)	
Two-Parent HH	0.425	0.387	0.467	0.426	0.397	0.557
	(0.045)	(0.043)		(0.046)	(0.044)	
Number of Books Child Has	0.002	0.003	0.035	0.003	0.003	0.952
	(0.000)	(0.000)		(0.000)	(0.000)	
SES Composite Index	0.486	0.432	0.185	0.505	0.471	0.362
	(0.026)	(0.025)		(0.026)	(0.026)	
Child's Birth Weight (oz.)	0.005	0.006	0.119	0.003	0.003	0.920
	(0.001)	(0.001)		(0.001)	(0.001)	

Notes:
1) Each entry in the table corresponds to an estimate from a separate regression of a measure of cognitive skills on a measure of home environment
2) All estimated models also control for school fixed effects.

Table 4: The Ratio of Between-School Variance to Total Variance of Skill Measures in ECLS-K:2011

	Kindergarten (1)	1st Grade (2)	2nd Grade (3)	3rd Grade (4)
	A: Teacher Assessments			
Approaches to Learning	0.085	0.066	0.069	0.079
Externalizing Problem Behaviors	0.116	0.085	0.106	0.144
Self-Control	0.127	0.101	0.088	0.110
Math ARS Rating	0.063	0.093	0.099	0.078
Reading ARS Rating	0.040	0.117	0.070	0.092
	B: IRT Test Scores			
Math	0.249	0.248	0.296	0.310
Reading	0.212	0.277	0.314	0.316
	C: Predicted Teacher Assessments			
Approaches to Learning	0.322	0.311	0.307	0.309
Externalizing Problem Behaviors	0.324	0.281	0.271	0.263
Self-Control	0.323	0.314	0.291	0.262
Math Ability	0.340	0.347	0.330	0.334
Reading Ability	0.336	0.338	0.327	0.329

Notes:
1) Each entry in the table corresponds to an estimate of the ratio of the between-school variance of a particular skill measure to its total variance.
2) Predicted teacher assessments are generated from linear regressions of teacher assessment on the home environment variables and school fixed effects, then forming predicted values based on the estimated coefficients on the home environment variables.

Table 5: Estimates of Racial Gaps in Latent Non-Cognitive Skills, ECLS-K:2011

	(1)	(2)	(3)	(4)	(5)
			A: Approaches to Learning		
Kindergarten	-0.331	-0.109	-0.661	-0.294	-0.242
	(0.026)	(0.028)	(0.025)	(0.025)	(0.040)
Grade 1	-0.416	-0.158	-0.720	-0.339	-0.284
	(0.030)	(0.032)	(0.029)	(0.030)	(0.048)
Grade 2	-0.431	-0.129	-0.736	-0.316	-0.175
	(0.032)	(0.034)	(0.032)	(0.031)	(0.053)
Grade 3	-0.524	-0.218	-0.830	-0.396	-0.274
	(0.035)	(0.037)	(0.034)	(0.034)	(0.058)
			B: Externalizing Problem Behaviors		
Kindergarten	-0.398	-0.172	-0.668	-0.312	-0.260
	(0.027)	(0.029)	(0.032)	(0.034)	(0.042)
Grade 1	-0.432	-0.197	-0.708	-0.353	-0.274
	(0.031)	(0.034)	(0.033)	(0.036)	(0.050)
Grade 2	-0.475	-0.231	-0.765	-0.399	-0.205
	(0.033)	(0.036)	(0.033)	(0.036)	(0.056)
Grade 3	-0.521	-0.246	-0.813	-0.405	-0.201
	(0.036)	(0.039)	(0.034)	(0.036)	(0.060)
Home Environment Controls		X		X	X
Estimated Latent Skills			X	X	X
School Fixed Effects					X

Table 5: Estimates of Racial Gaps in Latent Non-Cognitive Skills, ECLS-K:2011 (cont'd)

			C: Self-Control		
Kindergarten	-0.399	-0.197	-0.665	-0.297	-0.262
	(0.027)	(0.030)	(0.032)	(0.034)	(0.042)
Grade 1	-0.452	-0.223	-0.713	-0.359	-0.285
	(0.031)	(0.034)	(0.034)	(0.037)	(0.050)
Grade 2	-0.492	-0.242	-0.748	-0.389	-0.162
	(0.034)	(0.037)	(0.034)	(0.037)	(0.056)
Grade 3	-0.584	-0.284	-0.823	-0.406	-0.195
	(0.036)	(0.039)	(0.035)	(0.037)	(0.061)
Home Environment Controls		X		X	X
Estimated Latent Skills			X	X	X
School Fixed Effects					X

Notes:
1) Each entry in the table is an estimate from a separate regression of a measure of non-cognitive skills on an indicator equaling 1 for black children and 0 for white children, corresponding to equation (1) in the text.
2) Models in columns (3)-(5) use the estimated latent non-cognitive skills, rather than teacher-reported skills, as a dependent variable. These skills are estimated based on Condition 1 in the text.

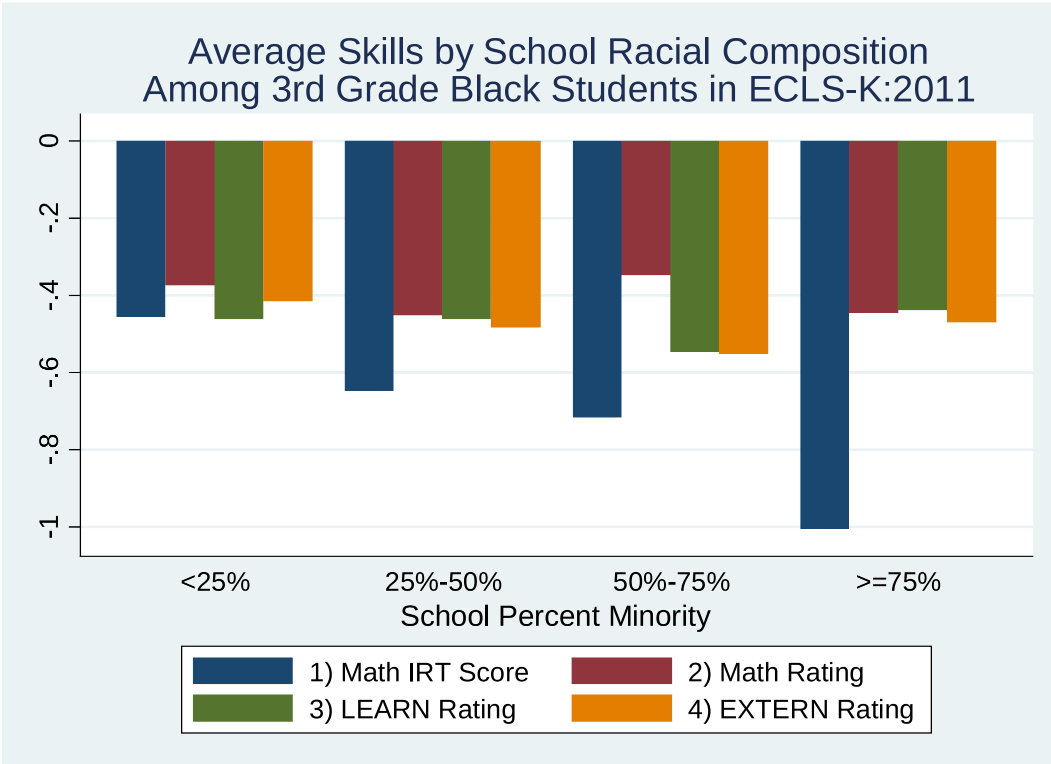
Table 6: Estimates of the Effects of Non-Cognitive Skills on Outcomes in BCS70

	Outcomes				
	HS Graduate	College Degree	Arrest	Unemployment	Log(Wages)
	(1)	(2)	(3)	(4)	(5)
			A: Males		
Externalizing Problem Behaviors	0.039	0.019	-0.078	-0.013	-0.024
	(0.010)	(0.009)	(0.010)	(0.005)	(0.018)
Approaches to Learning	0.101	0.079	-0.026	-0.008	0.077
	(0.008)	(0.007)	(0.008)	(0.004)	(0.014)
Self-Control	-0.022	-0.008	-0.001	0.000	-0.015
	(0.010)	(0.008)	(0.010)	(0.005)	(0.017)
Math Test Scores	0.034	0.064	-0.019	-0.013	0.090
	(0.017)	(0.014)	(0.017)	(0.008)	(0.030)
Reading Test Scores	0.080	0.027	-0.014	-0.003	0.040
	(0.017)	(0.014)	(0.017)	(0.008)	(0.030)
			B: Females		
Externalizing Problem Behaviors	0.007	0.005	-0.024	-0.011	0.001
	(0.011)	(0.010)	(0.006)	(0.005)	(0.030)
Approaches to Learning	0.122	0.084	-0.007	-0.015	0.156
	(0.009)	(0.007)	(0.005)	(0.004)	(0.023)
Self-Control	-0.022	-0.007	-0.018	0.006	-0.050
	(0.011)	(0.009)	(0.006)	(0.004)	(0.027)
Math Test Scores	0.007	0.034	0.001	-0.009	0.116
	(0.017)	(0.014)	(0.009)	(0.007)	(0.042)
Reading Test Scores	0.109	0.137	-0.006	0.002	0.047
	(0.016)	(0.014)	(0.009)	(0.007)	(0.041)

Notes:

1) Each column in each Panel corresponds to a separate regression of an outcome in BCS70 on the measures of cognitive and non-cognitive skills in column (1).
2) Means of dependent variables across columns are 0.626, 0.210, 0.300, 0.047, and 9.903 for males and 0.670, 0.197, 0.057, 0.027, and 9.327 for females. N = 4,166 for males and 4,384 for females.

Figure 1



Average Skills by School Racial Composition
Among 3rd Grade Black Students in ECLS-K:2011
0
-.2
-.4
-.6
-.8
-1
<25%
25%-50%
50%-75%
>=75%
School Percent Minority
1) Math IRT Score
2) Math Rating
3) LEARN Rating
4) EXTERN Rating

Figure 2

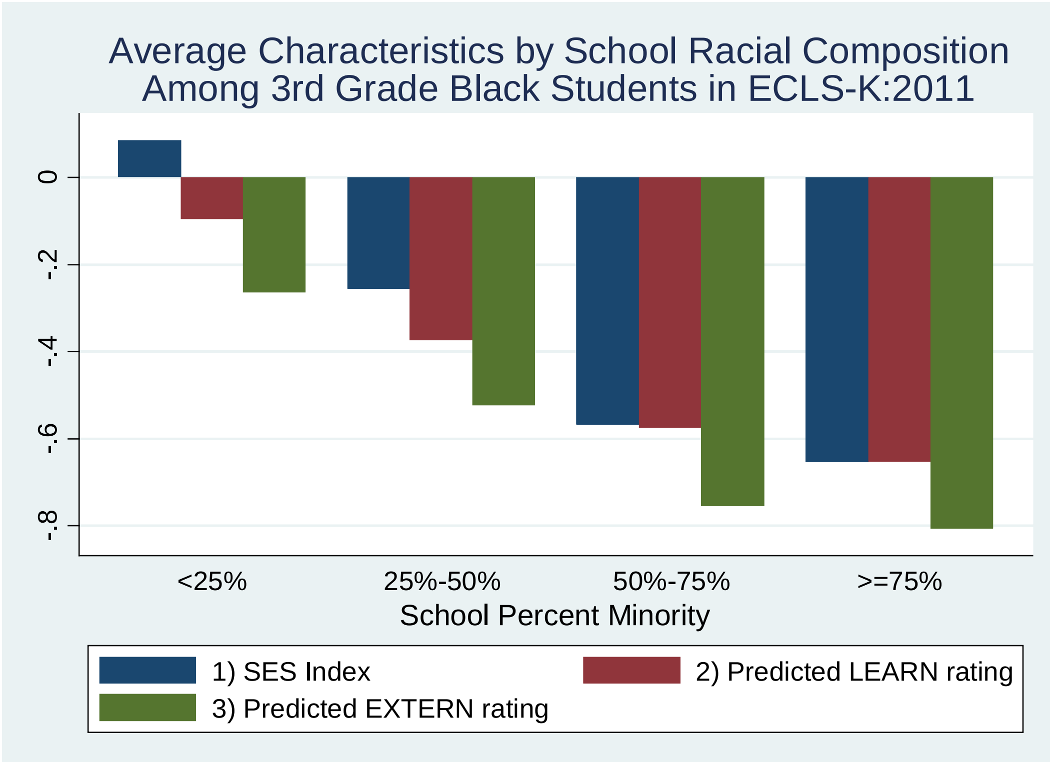
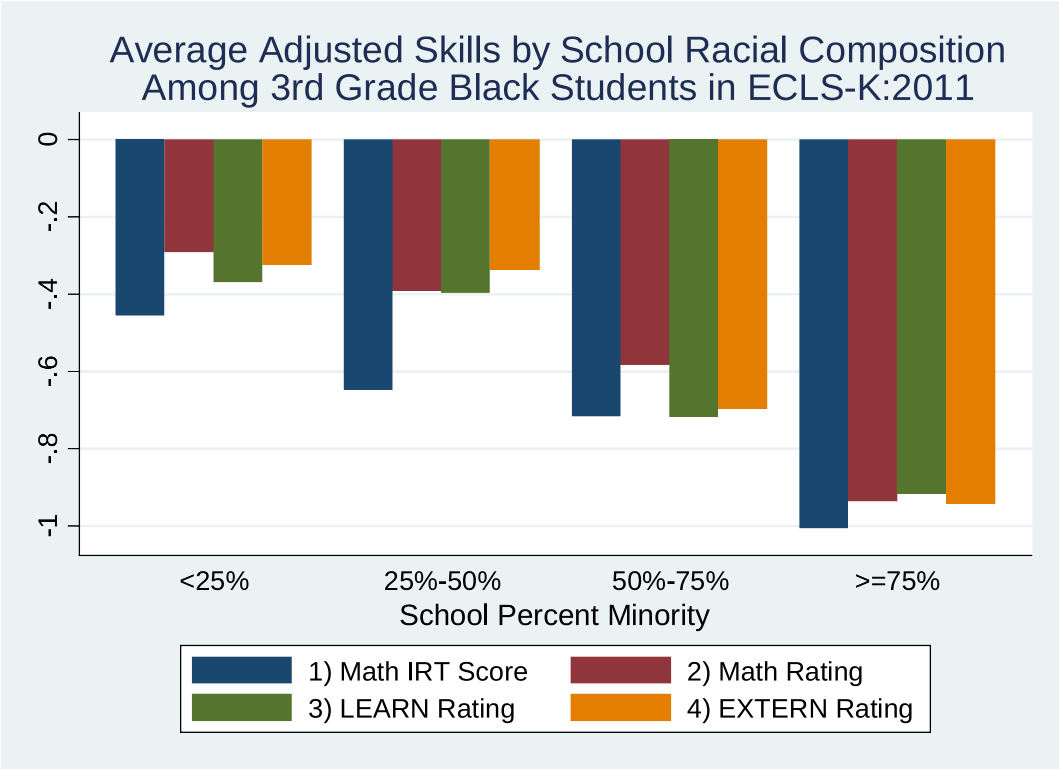

Average Characteristics by School Racial Composition
Among 3rd Grade Black Students in ECLS-K:2011
0
-.2
-.4
-.6
-.8
<25%
25%-50%
50%-75%
>=75%
School Percent Minority
1) SES Index
2) Predicted LEARN rating
3) Predicted EXTERN rating

Figure 3



Average Adjusted Skills by School Racial Composition
Among 3rd Grade Black Students in ECLS-K:2011
0
-.2
-.4
-.6
-.8
-1
<25%
25%-50%
50%-75%
>=75%
School Percent Minority
1) Math IRT Score
2) Math Rating
3) LEARN Rating
4) EXTERN Rating

Appendix A: Data Construction and Description

Teacher and parental Social Rating Scales:

As described in section II of the text, teachers completed Social Rating Scale measures in all survey waves of both ECLS-K:1999 and ECLS-K:2011. Respondents used four-point frequency scales to report how often a student demonstrates a particular behavior (such as getting into fights with peers), with a numerical value of 1 denoting "never", 2 denoting "sometimes", 3 "often", and 4 "very often". NCES aggregates the 24 teacher-reported scales into 5 composites: "approaches to learning" (measured by ECLS-K variable *T1LEARN* in the fall 2010 survey of ECLS-K:2011), "externalizing problem behaviors" (*T1EXTERN*), "self-control" (*T1CONTRO*), "interpersonal skills" (*T1INTERP*), and "internalizing problem behaviors" (*T1INTERN*).

Importantly, NCES does not release the individual scales, even in restricted-use versions of the data – only the composite scales are available. As described in the ECLS-K Base Year User's Guide (NCES, 2012),

- *T1LEARN* measures six items that rate the child's attentiveness, task persistence, eagerness to learn, learning independence, flexibility, and organization.
- *T1CONTRO* has four items that rate whether the child respects the property rights of others, controls his or her temper, accepts peer ideas for group activities, and responds appropriately to peer pressure.
- *T1EXTERN* includes five items that rate the frequency with which a child argues, fights, gets angry, acts impulsively, and disturbs ongoing activities.
- *T1INTERP* has five items that rate the child's skill in forming and maintaining friendships, getting along with people who are different, comforting or helping other children, expressing feelings, ideas and opinions in positive ways, and showing sensitivity to the feelings of others.
- *T1INTERN* includes four items that rate the apparent presence of anxiety, loneliness, low self-esteem, and sadness.

Control Variables used in the Analyses:

Control variables in selected specifications include indicators for gender, race, ethnicity, family structure, the marital status of the child's primary caregiver, Census region, urbanicity, parental education, log family income, and family size.

- The gender, race, and ethnicity variables include indicators for whether a respondent is female, Asian, Hispanic, black, Native American, multiracial, or has missing information on race.

- Family structure variables include indicators for whether the child's mother and father both live with the child, the mother only, the father only, or if some other family member lives with the child. Indicators for the marital status of the child's parents include married, separated, divorced, never married, and not reported.

- There are four indicators for Census region (Northeast, Midwest, South, and West), three for urbanicity of the child's residence (urban, suburban, or rural), and one each for missing Census region and urbanicity, respectively.

- Maternal and paternal education levels are measured as continuous variables ranging from 8 to 18 years. Log family income is created using the midpoints of the ranges of the categorical family income variable provided by NCES. Family size is measured as a continuous variable. Parental education, family income, and family size are set equal to their respective sample means when missing, and new 0-1 indicators for missing values are created for each of the original variables. The SES composite index provided by NCES is based on parental education, parental occupation, and household income.