

ACCEPTED VERSION – PLEASE CITE PUBLISHED VERSION AS:
Curran, F. C., & Kellogg, A. T. (2016). Understanding science achievement gaps by race/ethnicity and gender in kindergarten and first grade. *Educational Researcher*, 45(5), 273-282.

Understanding Science Achievement Gaps by Race/Ethnicity and Gender in
Kindergarten and First Grade

F. Chris Curran, PhD
Assistant Professor of Public Policy
UMBC School of Public Policy
1000 Hilltop Circle
Public Policy Building 411
Baltimore, MD 21225
615 337-6854
curranfc@umbc.edu

Ann T. Kellogg
Doctoral Student
UMBC School of Public Policy
1000 Hilltop Circle
Baltimore, MD 21225

Abstract

Disparities in science achievement across race and gender have been well documented in secondary and post-secondary school; however, the science achievement gap in the early years of elementary school remains understudied. We present findings from the ECLS-K:2011 which demonstrate significant gaps in science achievement in kindergarten and first grade by race/ethnicity. We estimate the Black-White science gap in kindergarten at $-0.82SD$ but find only a small gender gap by first grade. Large disparities between Asian student performance in science as compared to mathematics and reading are documented. Student background characteristics and school fixed effects explain nearly 60% of the Black-White and Hispanic-White science achievement gaps in kindergarten. Implications for policy and practice are discussed.

Acknowledgements: The authors are thankful to the editors and anonymous reviewers for useful feedback on earlier versions of the manuscript.

Introduction

Improving science achievement is an important initiative for policymakers and educators. Recent job growth in STEM fields has outpaced that in non-STEM fields and is expected to continue to grow (Langdon, McKittrick, Beede, Khan, & Doms, 2011). Unfortunately, not all groups are poised to thrive in this environment equally. It is well documented that minorities and women are underrepresented in college degree attainment and labor force participation in the sciences (Beede et al., 2011; Hrabowski et al., 2011; U.S. News, 2015). These racial and gender gaps represent a pressing policy problem.

Despite the attention given to the importance of science achievement and the related gaps by race and gender, the majority of research on the topic has focused on secondary or post-secondary students. Though large scale assessments such as the National Assessment of Educational Progress (NAEP) and the Trends in International Mathematics and Science Study (TIMSS) document racial and gender achievement gaps as early as fourth grade, such studies do not assess students in the earliest years of elementary school. Recently, several studies have utilized the original Early Childhood Longitudinal Study (ECLS-K) to extend analysis a grade earlier by examining science achievement gaps in the third grade (Kohlhaas, Lin, & Chu, 2010; Morgan, Farkas, Hillemeier, & MacZuga, 2016; Quinn & Cooc, 2015). These studies find that Black students perform approximately one standard deviation lower than White students in science at the end of third grade and that this gap is fairly stable as students progress to eighth grade. The Hispanic-White gap and Asian-White gap are both smaller than the Black-White gap and decrease as students age. The female-male gap is slightly less than one quarter of a standard deviation (Kohlhaas et al., 2010; Morgan et al., 2016; Quinn & Cooc, 2015).

While evidence from these studies and assessments such as the NAEP and TIMSS provide important evidence on science achievement gaps, such work is limited in several ways. First, the data utilized cannot speak to gaps in the earliest years of elementary school. Fryer and Levitt (2004) have demonstrated that the Black-White gap in mathematics and reading is present at the start of school and expands by over 0.1 standard deviations by the end of first grade. This suggests that measuring the science achievement gap in kindergarten and how it changes through first grade may reveal important information on disparities in science achievement and the development of the gap prior to third grade.

A second limitation of the prior studies utilizing the original ECLS-K is the age of the data. The science achievement gap measured in third grade was measured during the 2001-02 school year, just as the No Child Left Behind legislation was passed. The data for eighth grade, collected in 2006-07 is itself nearing a decade old. Importantly, the original ECLS-K data was collected before or concurrent to the implementation of important provisions of NCLB. For

instance, science assessments were not required until school year 2007-08, after the completion of the original ECLS-K data collection (U.S. Department of Education, 2004). Consequently, there is a need to examine science achievement gaps utilizing data that more accurately reflects the current policy environment.

We extend the literature by focusing on science achievement gaps in the first two years of schooling and by utilizing more current data. Specifically, we ask the following research questions:

- 1) What are the science achievement gaps by race/ethnicity and gender in kindergarten and first grade?
- 2) To what extent do individual factors and school fixed-effects explain the kindergarten and first grade science achievement gaps?

Using data from the new Early Childhood Longitudinal Study (ECLS-K:2011), our analysis provides first evidence on science achievement gaps in the first two grades of school. Additionally, by replicating the methodological approach utilized by Quinn and Cooc (2015), we respond to increased calls for replication in the educational research literature (Makel & Plucker, 2014). The results of this study have the potential to extend our understanding of science achievement gaps and the factors that explain these gaps. Doing so may provide insight that can inform policy and practice aimed at increasing equity in science achievement.

We begin by framing the paper in the context of developmental theory and ecological systems theory and reviewing the extant literature on early science achievement. We then discuss the data and methodology followed by a discussion of results and implications for policy and practice.

Theoretical Framework

A substantial body of literature demonstrates the importance of academic achievement in the early years of schooling for later academic success. Across numerous datasets, early academic achievement, particularly math achievement, has been shown to be one of the strongest predictors of later achievement (Duncan et al., 2007). For instance, mathematics achievement has been shown to be predictive of not only later achievement in mathematics but also reading, science, and grade retention (Claessens & Engel, 2013). Similarly, recent evidence on science finds that early science achievement is highly predictive of later science achievement through middle school (Morgan et al., 2016).

This observed relationship between early and later academic achievement exists because early skills provide the foundation for later skills, serving as a building block from which future gains can be made (Cunha, Heckman, Lochner, & Masterov, 2006). Indeed, theoretical perspectives from developmental psychology and human development support the notion that skills beget skills. Though there is some debate over whether skills develop in a continuous process or through a series of stages, both views of development support the notion that

early skills are foundational for later skill development (Hayslip, Neumann, Loudon, & Chapman, 2006; Shaffer & Kipp, 2013). Cunha and colleagues (2006) take a stage approach describing development as a life cycle process in which early skills and early investments in children increase the attainment of later skills and the productivity of later investments in the child.

That early skills are important for later development, then, is supported both theoretically and empirically; however, the opportunity to develop early skills is not uniform across students. Ecological systems theory suggests that individuals develop in context such that their development is not solely a product of their own characteristics but is also a function of environmental factors such as the family, neighborhood, and schooling contexts (Bronfenbrenner, 1979; Bronfenbrenner & Ceci, 1994; Bronfenbrenner & Morris, 1998). Indeed, the empirical literature demonstrates the importance of early experiences both in and out of school on later educational outcomes (Entwisle, Alexander, & Olson, 2005; Heckman, Moon, Pinto, Savelyev, & Yavitz, 2010). The result of systematic differences in both within and out of school factors across groups of students results in what we term “achievement gaps” between such groups.

In short, then, ecological theory points to the importance of contextual factors for differential early student development across groups while developmental theory points to the potentially lasting and compounding impacts of these early gaps. Despite the attention given to these theories and early achievement in general, the preponderance of empirical work has tended to focus on mathematics and reading, with less attention focused on science. In the next section, we review the literature on early science achievement gaps and point to the need for the current study.

Background

The science workforce has been and continues to be characterized by an over-representation of White males (Hrabowski, 2011; Riegle-Crumb & King, 2010; Shen, 2013). As reviewed by Quinn & Cooc (2015), such disparities in the workforce can be linked to differential likelihoods to study and persist in science majors in college which, in turn, can be linked to differential preparation, achievement, and attitudes towards science in high school (Chen & Weko, 2009; Griffith, 2010; Riegle-Crumb, King, Grodsky, & Muller, 2012; Price, 2010; Tai, Liu, Maltese, & Fan, 2006; Wang, 2013). The literature on science achievement and science gaps is far sparser, however, with regard to the earlier years of schooling.

Some of the best evidence to date on early science achievement gaps comes from the National Assessment of Educational Progress (NAEP) and the first iteration of the Early Childhood Longitudinal Study (ECLS-K) as well as international assessments such as the Trends in International Mathematics and Science Study (TIMSS) and the Program for International Student Assessment

(PISA). Results from the NAEP, a nationally representative study, demonstrate that racial/ethnic and gender achievement gaps in science exist in fourth and eighth grade (National Center for Education Statistics, 2012). For instance, in 2011, White eighth graders scored approximately a standard deviation higher than their Black peers while males scored approximately 0.15 standard deviations higher than females (National Center for Education Statistics, 2012; U.S. Department of Education, 2012). Results for fourth graders were similar for the Black-White gap though the gender gap was approximately half that observed in eighth grade (National Center for Education Statistics, 2011). Similarly, data from the TIMSS study has confirmed racial achievement gaps in science at both fourth and eighth grade while results from PISA demonstrate a small gender gap at age 15 (Lee & Buxton, 2010).

Like these other assessments, the first iteration of the ECLS-K allowed for an examination of science achievement in upper elementary and middle school. Kohlhaas, Lin, and Chu (2010) examined racial achievement gaps at third grade finding the Black-White gap to be slightly larger than a standard deviation and the Hispanic-White gap to be approximately 0.9 standard deviations. They also found males to score approximately a quarter of a standard deviation higher than females (Kohlhaas et al., 2010).

Subsequent work by Quinn and Cooc (2015) replicated these findings for third grade while also examining fifth and eighth grade as well factors that explain such gaps. They found the Black-White gap to be fairly consistent across the grades examined while both the Hispanic-White and Asian-White gaps decrease. In addition to documenting such gaps, the authors also explored the explanatory power of student socioeconomic status, prior mathematics and reading achievement, and classroom fixed effects to explain the gaps. They found that controlling for prior mathematics achievement explained the entire eighth grade science gender gap while the inclusion of all factors reduced the Black-White and Hispanic-White gaps to non-significance in eighth grade (Quinn & Cooc, 2015). Other recent work has examined general knowledge tests from the original ELCS-K in earlier grades as a proxy for science achievement; however, such tests do not exclusively measure science achievement (Morgan et al., 2016).

Despite the insight gleaned from analysis of NAEP and the original ECLS-K data, there are significant limitations in the aforementioned studies. In particular, the NAEP does not conduct science assessments below fourth grade and the original ECLS-K data only administered science examinations one year earlier, in third grade. This limitation is a particularly important shortcoming as a number of studies demonstrate the presence of racial achievement gaps in mathematics and reading at the start of kindergarten (Fryer & Levitt, 2004; Quinn, 2015; Reardon & Galindo, 2009). Furthermore, some evidence suggests that achievement gaps can change over time as cohorts of students' progress through

school. For instance, Fryer and Levitt (2004) demonstrate that the reading and mathematics achievement gaps by race change during the first few years of formal schooling. Such early changes in achievement gaps may also be expected in science given evidence from the Explore, PLAN, and ACT assessments that demonstrate a widening of racial achievement gaps in science as students progress through secondary schooling (ACT, 2012)

Developmental stages theory suggests the importance of early science skills for the development of future skills and for the perpetuation of achievement gaps (Cunha et al., 2006). To date, little research has systematically examined science achievement in the earliest grades of school, instead focusing on third grade and above (Kohlhaas et al., 2010; Morgan et al., 2016; National Center for Education Statistics, 2012; Quinn & Cooc, 2015). Given the potential importance of early achievement for later outcomes and the evidence that achievement gaps can change in the earliest years (Fryer & Levitt, 2004), there is a need to document and better understand science achievement gaps beginning in kindergarten and throughout the first years of formal schooling.

The present study fills this void by utilizing data from the newest iteration of the Early Childhood Longitudinal Study, making use of the newly added science assessments in the earliest grades. This study seeks to document science achievement gaps by race/ethnicity and gender in kindergarten and first grade while also exploring both the in and out of school factors that explain such gaps. In the next sections, we turn to an explanation of the methodological approach and an overview of the data.

Methods

We followed the methodology utilized by Quinn and Cooc (2015), applying the approach to the examination of kindergarten and first grade science achievement. In short, we fit an ordinary least squares regression model in which we predicted standardized science achievement from gender and race/ethnicity indicators. Our primary model was as follows:

$$1) \text{ScienceAchievement}_i = \beta_0 + \beta_1 \text{Female}_i + \beta_2 \text{Black}_i + \beta_3 \text{Hispanic}_i + \beta_4 \text{Asian}_i + \beta_5 \text{OtherRace}_i + e_i$$

Where *ScienceAchievement* is the standardized science test score in either the spring of kindergarten or the spring of first grade for student *i*, and the covariates are binary indicators of gender or race/ethnicity. In this model, β_1 can be interpreted as the standardized science gender gap controlling for race while coefficients β_2 through β_5 can be interpreted as the standardized science race/ethnicity gap between the named race and non-Hispanic Whites controlling for gender. The estimates are weighted to account for the complex sampling design of the ECLS-K and standard errors are estimated using Taylor series

linearization. This primary model allows for an estimation of the raw race/ethnicity and gender gaps in science at both kindergarten and first grade.

In addition to the primary model, we also sought to understand the factors that explain the gaps. To do so, we fit versions of model 1 with controls for fall of kindergarten mathematics and reading, student socioeconomic status (SES), and school fixed effects. Models were fit including each of these factors individually as well as a fully specified model including all of these controls. Collectively, these models allow for an exploration of the degree to which these factors explain the racial and gender gaps in early science achievement.

Finally, in an effort to put early science achievement gaps in the context of the more frequently examined mathematics and reading gaps, we estimated versions of model 1 predicting standardized spring of kindergarten mathematics and reading achievement scores. These models allow for a comparison of the magnitude of race/ethnicity and gender gaps in science to those in mathematics and reading.

Data

We utilized data from the newly released Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011 (ECLS-K:2011). The survey includes a nationally representative set of 18,174 kindergartners from the 2010-2011 school year who will be followed through fifth grade. The ECLS-K:2011 employed a complex multi-stage sampling strategy to produce nationally representative estimates of public and private school students (Tourangeau et al., 2015a; 2015b). The ECLS-K:2011 provides the most current evidence on the experience of elementary school students, updating the ECLS-K data utilized by earlier studies which began collection in 1998-99.

Unlike the original ECLS-K, the ECLS-K:2011 includes science achievement tests for the spring of kindergarten and the spring of first grade. Though the original ECLS-K contained general knowledge tests that assessed science along with social studies concepts, the original ECLS-K did not include true science achievement tests until students were in third grade complicating examination of science achievement in the earliest grades of school (Tourangeau et al., 2009). The ECLS-K:2011 science assessment included questions pertaining to the physical sciences, life sciences, environmental sciences, and scientific inquiry (Tourangeau et al., 2015a; 2015b). In the kindergarten year, the science assessment included a consistent set of items for all students while in the first grade year the assessment utilized a two-stage approach in which routing questions ensured that the questions could precisely measure the child's skills (Tourangeau et al., 2015a; 2015b). Given the age of students, the assessments were administered verbally and did not require written responses. We utilized standardized versions of the science theta score which were derived from item response theory (IRT) measures. In particular, we standardized the measures

within each year using sampling weights and survey settings to account for the complex sampling design of the ECLS-K:2011.

Our primary independent variables of interest were indicators of student race and gender. We created mutually exclusive race and ethnicity variables using the ELCS-K:2011 variable (X_RACETH_R) constructed from the variables for race and ethnicity collected in the parent interview. We coded mutually exclusive dummy variables (0/1) to represent each of the following racial/ethnic categories: White, Black, Hispanic, Asian, and other. For our analyses, we utilized White as the omitted race/ethnicity category. We operationalized gender as a single binary indicator for female. Consequently, the omitted category for gender was males.

In order to explore the factors that explain race/ethnicity and gender gaps in early science achievement, we also utilized measures of prior math and reading achievement, a composite measure of socioeconomic status (SES), and school fixed effects. Measures of prior achievement in mathematics and reading came from tests administered during the fall of the kindergarten year. The reading examination included items measuring basic skills (such as letter and word recognition), vocabulary knowledge, and reading comprehension. The mathematics examination included items measuring number sense, properties, operations, measurement, geometry, data analysis, statistics, probability, patterns, algebra, and functions (Tourangeau et al., 2015a; 2015b). Given the age of students, the assessments were administered verbally and did not require written responses. Further details on the assessments can be found in the ECLS-K users' manual (Tourangeau et al., 2015a; 2015b). For our analyses, we utilized standardized versions of the mathematics and reading theta scores which were derived from item response theory (IRT) measures. In particular, we standardized the measures using sampling weights and survey settings to account for the complex sampling design of the ECLS-K:2011.

The composite measure of SES was derived from five items administered on the parental survey and compiled by the administrators of the ECLS-K:2011. In particular, the SES measure consists of indicators of the two primary parents' education, the two primary parents' occupational prestige score, and household income. This composite measure is an established metric utilized in prior studies using data from the ECLS-K (i.e. Quinn & Cooc, 2015). Further details on the construction of the SES composite measure can be found in the ECLS-K users' manual (Tourangeau et al., 2015a; 2015b)

Finally, we used school identifiers that uniquely identify schools, allowing for the inclusion of school-fixed effects in the analysis. The entire ECLS-K sample consisted of 1,319 schools with approximately 13 kindergartners sampled per school on average (Tourangeau et al., 2015a; 2015b). These school identifiers are arbitrary numbers that uniquely identify schools. They are used in fixed effect

models to restrict estimates to variation within schools, thereby implicitly controlling for any fixed aspect of that school.

We focused on available data from the kindergarten and first grade years. We utilized appropriate weights for analysis spanning the kindergarten and first grade years utilizing student assessment data and parent survey data. Standard errors were adjusted using Taylor series linearization. The full weighted sample size was 10,353.

Our analytic sample ($n=10,050$) included students with science test scores in kindergarten and first grade as well as available data on race/ethnicity and gender and dropped observations missing prior math and reading achievement scores, information on SES, or school indicators. We handled missing data through case wise deletion and report weighted sample sizes that account for the complex sampling structure of the ECLS-K:2011. This approach reduced our sample size from a weighted sample size of 10,353 by 3%.

Though notably smaller than the entire ECLS-K:2011 sample ($n=18,174$), this weighted analytic sample size reflects non-response on science assessments in kindergarten or first grade as well as missing demographic data. The weights provided with the ECLS-K and used in this analysis, however, are designed to account for such missing data. Indeed, in robustness checks using just the data from the kindergarten year with kindergarten weights ($n=14,895$) or with no weights ($n=16,863$), the results are qualitatively no different.

Table 1 shows descriptive statistics for our full analytic sample as well as broken out by student race/ethnicity and gender. As shown, the sample is nearly evenly split between White students and minority students. White students and Asian students, on average, come from households with higher socioeconomic status and score higher on the fall of kindergarten math and reading assessments. As a group, Asians have higher incoming math and reading scores; however, Whites have the highest science scores in kindergarten. The descriptive table provides initial evidence for disparities in science achievement across groups. We now turn to results from regression models exploring these gaps in more detail.

Results

We find that significant science achievement gaps are present in the first two years of formal schooling, particularly by race/ethnicity. Furthermore, several of the gaps appear to change in size between kindergarten and first grade. Figure 1 shows the gaps and corresponding confidence intervals estimated from regressions predicting spring of kindergarten and spring of first grade science achievement from race/ethnicity and gender indicators. As shown, the female-male gap is statistically insignificant in kindergarten and small in magnitude ($\beta=-0.062$, $p<0.01$) by the end of first grade. This suggests that there is no gender gap in science achievement during kindergarten but that the gap begins to emerge as students move through first grade.

For race/ethnicity, the largest gap in kindergarten is seen for Hispanic students ($\beta=-0.937, p<0.001$) though the Black-White gap is also considerable in size ($\beta=-0.815, p<0.001$). Interestingly, the Asian-White gap in kindergarten is also considerably large, at over half a standard deviation. As students progress through first grade, notable changes in the Hispanic-White and Asian-White gap are observed. As shown in Figure 1, the Hispanic-White gap reduces by almost 0.2SD while the Asian-White gap reduces by nearly 50%, an approximately 0.3SD decline. In contrast, the Black-White gap appears relatively stable across grades. As a result, by the end of first grade, the Black-White science achievement gap replaces the Hispanic-White achievement gap as the largest of the race/ethnicity gaps.

In addition to estimating the raw gaps, we also explored the extent to which student background characteristics and school fixed effects accounted for the observed gaps. Table 2 shows results of regressions predicting spring of kindergarten science achievement progressively controlling for prior achievement in mathematics and reading, student socioeconomic status, and school fixed effects. We find that the race/ethnicity gaps remain statistically significant even after controlling for student characteristics and school fixed effects, though the magnitude of many of the gaps are significantly reduced. For instance, the Black-White and Hispanic-White gaps are reduced by approximately 30% controlling for either socioeconomic status, prior mathematics achievement, or school fixed effects. In the fully specified model (column 9), the Black-White and Hispanic-White gaps are reduced by nearly 60%.

Interestingly, the Asian-White gap behaves differently than the other race/ethnicity gaps in models controlling for prior achievement and socioeconomic status. Unlike Black and Hispanic students, who, on average, had lower math and reading scores and came from a lower socioeconomic background, Asian students had slightly higher prior scores than Whites and came from more advantaged backgrounds. Consequently, models controlling for prior achievement (columns 4-6) and socioeconomic status (column 3) actually result in a slight increase in the Asian-White gap. While the inclusion of school fixed-effects (column 8) did reduce the Asian-White gap, the fully specified model (column 9) resulted in only a slight decline in the Asian-White gap compared to the nearly 60% decline in the Black-White and Hispanic-White gaps.

Finally, we sought to provide context to the size of the race/ethnicity and gender gaps in science by comparing them to the corresponding gaps in mathematics and reading, subjects that have traditionally been the focus of achievement gap studies. Figure 2 displays coefficients representing the race/ethnicity and gender gaps in science, mathematics, and reading for the kindergarten year. As shown, the kindergarten achievement gaps in science tend to be as large as or larger than those for reading and mathematics. The difference

is particularly pronounced for the Hispanic-White and Asian-White gaps. Notably, though Asians outperform Whites in reading and mathematics in kindergarten, they perform, on average over half a standard deviation lower than their White peers in science. In short, the coupling between mathematics/reading gaps and science gaps appears loose in some cases, with science gaps tending to be larger in magnitude.

Discussion

Our results demonstrate substantial science achievement gaps in kindergarten by race/ethnicity, gaps that are as large as or larger than the corresponding mathematics and reading gaps. Furthermore, we have shown that, for some subgroups, such gaps change significantly between kindergarten and first grade. In this section, we discuss the results in the context of prior literature.

Though we find no gender gap in science achievement in kindergarten and only a small gap emerging by the end of first grade, the absence of a more substantial gender gap is actually notable. Our finding contrasts with the science gender gap observed at the end of third grade, estimated at approximately one quarter of a standard deviation (Kohlhaas et al., 2010; Quinn & Cooc, 2015). Quinn & Cooc (2015) find the gender gap in science to be relatively stable between third and eighth grade; however, our results demonstrate that this stability does not continue back to the beginning of formal schooling. Rather, this gap appears to develop as students progress through the first years of formal schooling. Our finding, then, suggests that the disparity in female participation in STEM related fields later in life (Beede et al., 2011; Hrabowski et al., 2011; U.S. News, 2015) may have its foundation in the experiences of female students in the earliest grades of school and that this early stage of life may be a policy amenable time to affect trajectories that result in later gender disparities in science.

With regard to race/ethnicity gaps, our findings demonstrate particularly large disparities in science achievement for Black and Hispanic students. While achievement gaps in reading and mathematics have dominated the research and policy conversation (Fryer & Levitt, 2004; Quinn, 2015; Reardon & Galindo, 2009), our findings suggest that the corresponding gap for science is actually larger in the earliest grades. This is particularly worrisome given concerns that the increased accountability prompted by the policy context of No Child Left Behind has resulted in decreased time on science in the elementary grades (Bassok, Latham, & Rorem, 2016; Bowdon & Desimone, 2014; Marx & Harris, 2006).

That we see changes in the race/ethnicity gaps between kindergarten and the end of first grade aligns with prior research. The decreases in the Hispanic-White gap between kindergarten and first grade may be partially attributable to lower proficiency with the English language among some Hispanic students. In discussing the Hispanic-White gap in mathematics and reading, Fryer and Levitt

(2004) suggest that such a lack of familiarity with the language may artificially lower scores during the first year of schooling. Indeed, empirical work has shown that Hispanic students from homes where English is not spoken show the greatest gains in academic achievement in the first several years of schooling (Reardon & Galindo, 2009). Our results suggest that a similar mechanism may be at work for Hispanic student performance in science. Our finding that the Hispanic-White gap in science is significantly reduced when controlling for prior reading achievement lends some preliminary support to this hypothesis.

In contrast to the Hispanic-White gap, the Asian-White gap in science achievement is not reduced by the inclusion of controls for reading or mathematics achievement. Notably, we find that Asian students significantly underperform their White peers in science, despite performing equally well or better in mathematics and reading. While the inclusion of prior mathematics and reading scores reduces the Hispanic-White science gap in kindergarten, it actually increases the Asian-White gap. While seemingly counterintuitive, our findings do fit with prior literature in this area. For instance, the third grade Asian-White gap in science has been estimated at approximately $-0.3SD$ (Kohlhaas et al., 2010; Quinn & Cooc, 2015) despite the fact that Asians outperform Whites in reading and mathematics at the beginning of kindergarten (Quinn, 2015). Quinn & Cooc (2015) demonstrate that the science gap between Asians and Whites significantly decreases as students progress through upper elementary and middle school such that it is statistically insignificant by the end of eighth grade. Our findings suggest that the decline in the gap they observe is a continuation of a decreasing gap that is even larger at kindergarten.

Further understanding of the fairly loose coupling of the science achievement gap with the mathematics and reading achievement gaps, particularly for the Asian-White gap, may be an important area of further inquiry. Recent work has demonstrated that the Asian advantage in academic achievement appears early (Hsin & Xie, 2014). For mathematics and reading, this advantage is gained after starting on what appears to be relatively equal footing with their White peers. In contrast, however, the case of science achievement appears to be a case of a rapid rise from behind. Understanding the mechanisms that facilitate such gains in science achievement for Asian students may yield insight into ways to improve science achievement for all students.

Our results suggest that science achievement gaps in elementary school may be affected by both in and out of school factors. Our finding that socioeconomic status explains approximately a third of the Black-White and Hispanic-White gaps is consistent with recent evidence on the increasing importance of income disparities for academic achievement (Reardon, 2011). Given the high correlation between these minority groups and lower socioeconomic status, the reduction of the gap when controlling for SES was

expected. It is important to note, however, that a significant gap remains after controlling for SES. The racial gaps observed, then, cannot be solely explained by differences in socioeconomic status across racial groups.

Furthermore, the results of models controlling for prior academic achievement in reading and mathematics demonstrate that the observed science achievement gaps are not merely reflections of differential foundational skills in reading or math. Though controlling for such skills explains a portion of the Black-White and Hispanic White science gap, the majority of it remains, suggesting real differences in science content knowledge and science reasoning above and beyond mathematics and reading ability. These results are similar to those found in previous research examining the eighth grade science gap, which also found that prior achievement left the majority of such racial gaps unexplained (Quinn & Cooc, 2015).

Finally, our results suggest that school factors and, in particular, the differences in such factors across schools likely play an important contributing role to science achievement gaps as students progress through their first years of formal schooling. It is well documented that quality teachers are not uniformly distributed across schools (Clotfelter, Ladd, & Vigdor, 2005; Engel, Jacob, & Curran, 2014; Ingersoll, 2003; Lankford, Loeb, & Wyckoff, 2002; Rockoff, 2004). Coupled with increasing segregation of students across schools (Clotfelter, Ladd, & Vigdor, 2003; Orfield, 2001), systematic differences in educational opportunity arise. The results of our school fixed-effects models demonstrate that restricting variation to that which occurs within schools eliminates approximately a third of the Black-White and Hispanic-White gap. These models implicitly control for fixed aspects within a school thereby controlling for differences that vary across schools. Though the school fixed effect models do not allow for an exploration of the relationships between specific school characteristics and science outcomes, they do suggest that school level variables may matter for explaining the gaps. Consequently, addressing disparities in school resources and segregation across schools may represent one mechanism for alleviating the observed science achievement gap.

Conclusion

Science related fields promise to continue to be an important area for both individual and societal economic growth (Langdon et al., 2011). Our results point to important and policy relevant gaps in science achievement that are present in the earliest grades of formal schooling, a potential foundation for disparities observed in later educational settings and the workforce. To our knowledge, no prior study has examined science achievement in the earliest grades with nationally representative data. We find that gaps by race/ethnicity are particularly pronounced and that these gaps change as students progress through the first years

of formal schooling. Furthermore, we find little gender gap at the start of school, suggesting that gaps by gender arise later in the schooling process.

Our results have important implications for policy, practice, and research. First, the finding that science gaps are as large as or larger than their respective mathematics and reading gaps points to the need for expanding attention to early science gaps as an important policy problem. Accountability policies that have emphasized mathematics and reading at the expense of other subjects may be expanded to include a greater emphasis on science. Furthermore, interventions designed to enhance science learning, such as professional development for teachers, extracurricular experiences for students, or other curriculum interventions may hold promise for enhancing equity in early science achievement. As with achievement gaps in other subjects, the finding that science gaps are partially explained by across school variation suggests a need for attention to segregation of students across schools and the inequitable distribution of resources and opportunity across schools.

With regard to research, our findings point to the need for further study of early science experiences across subgroups of students. In particular, further understanding the instructional mechanisms and opportunity to learn science across school settings may yield insights into the causes of the disparities documented in this study. Further research on particular subgroups may yield further insight into the mechanisms at work. For instance, probing the rapid growth in science achievement among Asians and why their kindergarten science achievement varies so much from their mathematics and reading achievement may yield important insights into ways to improve science achievement for all students. Furthermore, examining variation within race/ethnicity, such as by language status, immigration status, or country of origin is also important.

As science related fields continue to be sources of job growth and technological innovation (Langdon et al., 2011), the importance of ensuring that all groups are in position to contribute to and benefit from such growth is critical. Our findings provide insight into this important area and lay a foundation for further work that can enhance educational equity in science.

References

- ACT. (2012). Do race/ethnicity-based student achievement gaps grow over time? Issue Brief for ACT Research and Policy.
- Bassok, D., Latham, S., & Rorem, A. (2016). Is kindergarten the new first grade?. *AERA Open*, 2(1), 2332858415616358.
- Beede, D. N., Julian, T. A., Langdon, D., McKittrick, G., Khan, B., & Doms, M. E. (2011). Women in STEM: A gender gap to innovation. *Economics and Statistics Administration Issue Brief*, (04-11).
- Bowdon, J. & Desimone, L. (2014). More work, less play: Kindergarten, post-NCLB. *Teachers College Record*.
- Bronfenbrenner, U. (1979). *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge, MA: Harvard University Press.
- Bronfenbrenner, U., & Ceci, S. J. (1994). Nature-nuture reconceptualized in developmental perspective: A bioecological model. *Psychological review*, 101(4), 568.
- Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. In R. Lerner (Vol. Ed.) *Handbook of child psychology: Theoretical models of human development* (5th ed., Vol. 1, pp 993-1028). New York, N.Y.: John Wiley.
- Chen, X., & Weko, T. (2009). Students Who Study Science, Technology, Engineering, and Mathematics (STEM) in Postsecondary Education. Stats in Brief. NCES 2009-161. *National Center for Education Statistics*.
- Claessens, A., & Engel, M. (2013). How important is where you start? Early mathematics knowledge and later school success. *Teachers College Record*, 115(6), 1-29.
- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. (2003). Segregation and resegregation in North Carolina's public school classrooms, 81 *North Carolina Law Review* 1463-1512.
- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. (2005). Who teaches whom? Race and the distribution of novice teachers. *Economics of Education review*, 24(4), 377-392.
- Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education*, 1, 697-812.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... & Sexton, H. (2007). School readiness and later achievement. *Developmental psychology*, 43(6), 1428.
- Engel, M., Jacob, B. A., & Curran, F. C. (2013). New evidence on teacher labor supply. *American Educational Research Journal*, 0002831213503031.

- Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2005). First Grade and Educational Attainment by Age 22: A New Story. *American Journal of Sociology*, *110*(5), 1458-1502.
- Fryer Jr, R. G., & Levitt, S. D. (2004). Understanding the black-white test score gap in the first two years of school. *Review of Economics and Statistics*, *86*(2), 447-464.
- Griffith, A. L. (2010). Persistence of women and minorities in STEM field majors: Is it the school that matters?. *Economics of Education Review*, *29*(6), 911-922.
- Hayslip Jr, B., Neumann, C. S., Louden, L., & Chapman, B. (2006). Developmental stage theories. *Comprehensive handbook of personality and psychopathology*, 115.
- Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of public Economics*, *94*(1), 114-128.
- Hrabowski, F., Ammons, J., Begay-Campbell, S., Clewell, B., Grasmick, N., Gutierrez, C., ... Zoback, M. (2011). *Expanding underrepresented minority participation*. Committee on Underrepresented Groups and the Expansion of the Science and Engineering Workforce Pipeline and the Committee on Science, Engineering, and Public Policy and Global Affairs. The National Academies Press: Washington, D.C.
- Hsin, A., & Xie, Y. (2014). Explaining Asian Americans' academic advantage over whites. *Proceedings of the National Academy of Sciences*, *111*(23), 8416-8421.
- Ingersoll, R. (2003). Out-of-field teaching and the limits of teacher policy. *GSE Publications*, 143.
- Kohlhaas, K., Lin, H. H., & Chu, K. L. (2010). Science equity in third grade. *The Elementary School Journal*, *110*(3), 393-408.
- Langdon, D., McKittrick, G., Beede, D., Khan, B., & Doms, M. (2011). STEM: Good jobs now and for the future. ESA Issue Brief# 03-11. *US Department of Commerce*.
- Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational evaluation and policy analysis*, *24*(1), 37-62.
- Lee, O. & Buxton, A. (2010). *Diversity and Equity in Science Education: Research, Policy, and Practice*. *Multicultural Education Series*. Teachers College Press.
- Makel, M. C., & Plucker, J. A. (2014). Facts are more important than novelty replication in the education sciences. *Educational Researcher*, 0013189X14545513.

- Marx, R. W., & Harris, C. J. (2006). No Child Left Behind and science education: Opportunities, challenges, and risks. *The Elementary School Journal*, *106*(5), 467-478.
- Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2016). Science achievement gaps begin very early, persist, and are largely explained by modifiable factors. *Educational Researcher*, *45*(1), 18-35.
- National Center for Education Statistics. (2011). The Nation's Report Card: Science 2009. Institute of Education Sciences, U.S. Department of Education, Washington DC.
- National Center for Education Statistics. (2012) The Nation's Report Card: Science 2011. Institute of Education Sciences, U.S. Department of Education, Washington DC.
- Orfield, G. (2001). Schools more separate: Consequences of a decade of resegregation. Civil Rights Project, Harvard University. Retrieved from: <http://files.eric.ed.gov/fulltext/ED459217.pdf>
- Quinn, D.M. (2015). Kindergarten Black–White Test Score Gaps Re-examining the Roles of Socioeconomic Status and School Quality with New Data. *Sociology of Education*. OnlineFirst.
- Quinn, D. M., & Cooc, N. (2015). Science achievement gaps by gender and race/ethnicity in elementary and middle school: Trends and predictors. *Educational Researcher*, 0013189X15598539.
- Price, J. (2010). The effect of instructor race and gender on student persistence in STEM fields. *Economics of Education Review*, *29*(6), 901-910.
- Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. *Whither opportunity*, 91-116.
- Reardon, S. F., & Galindo, C. (2009). The Hispanic-White achievement gap in math and reading in the elementary grades. *American Educational Research Journal*, *46*(3), 853-891.
- Riegle-Crumb, C., & King, B. (2010). Questioning a white male advantage in STEM examining disparities in college major by gender and race/ethnicity. *Educational Researcher*, *39*(9), 656-664.
- Riegle-Crumb, C., King, B., Grodsky, E., & Muller, C. (2012). The more things change, the more they stay the same? Prior achievement fails to explain gender inequality in entry into STEM college majors over time. *American Educational Research Journal*, *49*(6), 1048-1073.
- Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, 247-252.
- Shaffer, D., & Kipp, K. (2013). *Developmental psychology: Childhood and adolescence*. Cengage Learning.

- Shen, H. (2013). Inequality quantified: Mind the gender gap. *Nature*, 495, p. 22–24.
- Tai, R. H., Liu, C. Q., Maltese, A. V., & Fan, X. (2006). Planning early for careers in science. *Science*, 312, 1143–1144.
- Tourangeau, K., Nord, C., Lê, T., Sorongon, A.G., & Najarian, M. (2009). Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K): Combined User's Manual for the ECLS-K Eighth-Grade and K-8 Full Sample Data Files and Electronic Codebooks. NCES 2009-004. *National Center for Education Statistics*.
- Tourangeau, K., Nord, C., Lê, T., Sorongon, A.G., Hagedorn, M.C., Daly, P., Najarian, M., Mulligan, G. M. (2015a). Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K: 2011), User's Manual for the ECLS-K: 2011 Kindergarten Data File and Electronic Codebook (NCES 2013-074). Washington, DC: US Department of Education. *National Center for Education Statistics*.
- Tourangeau, K., Nord, C., Lê, T., Wallner-Allen, K., Hagedorn, M.C., Leggitt, J., Najarian, M., & Mulligan, G. (2015b). Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K: 2011), User's Manual for the ECLS-K: 2011 Kindergarten-First Grade Data File and Electronic Codebook, Public Version (NCES 2015-078). Washington, DC: US Department of Education. *National Center for Education Statistics* (2012).
- U.S. Department of Education. (2004). Testing: Frequently asked questions. Retrieved from: <http://www2.ed.gov/nclb/accountability/ayp/testing-faq.html>
- U.S. Department of Education. (2012). National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 Science Assessment. Retrieved from: https://nces.ed.gov/programs/digest/d12/tables/dt12_168.asp
- U.S. News. (2015). 2015 STEM index shows gender, racial gaps widen. Retrieved from: <http://www.usnews.com/news/stem-index/articles/2015/06/29/gender-racial-gaps-widen-in-stem-fields>
- Wang, X. (2013). Why students choose STEM majors motivation, high school learning, and postsecondary context of support. *American Educational Research Journal*, 50(5), 1081-1121.

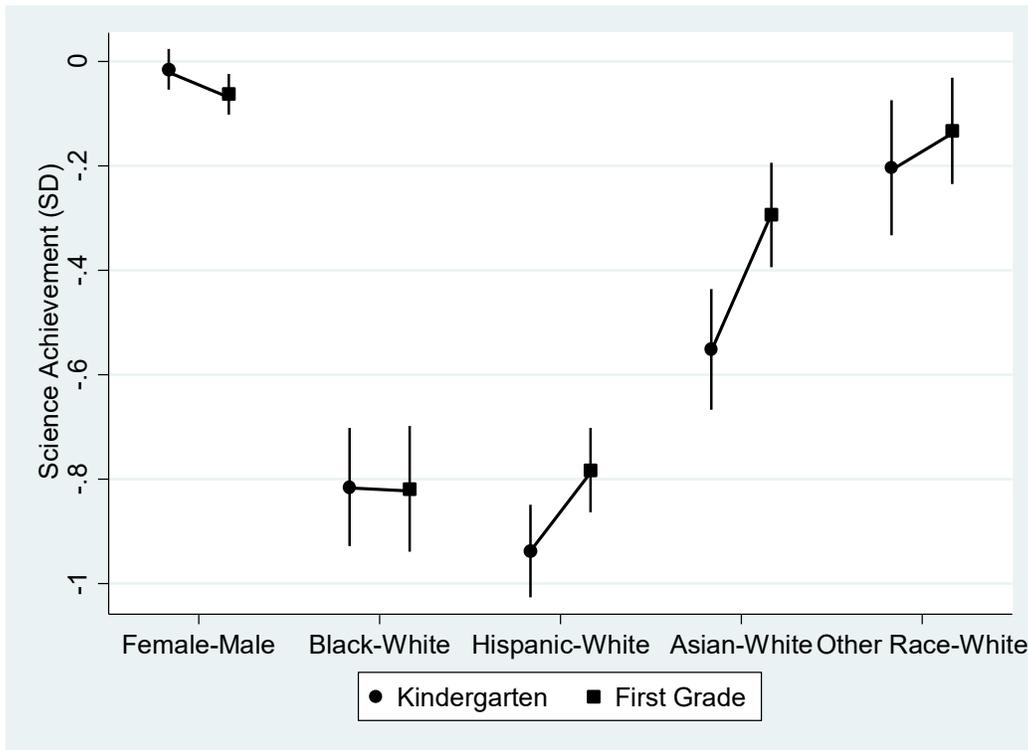


Figure 1. Science Achievement Gaps by Race/Ethnicity and Gender for Kindergarten and First Grade (n=10,050)

Note. Vertical lines show 95% confidence interval. Estimates are adjusted to account for the complex sampling design of the ELCS-K:2011 using sample weights and Taylor linearized estimates of the standard errors.

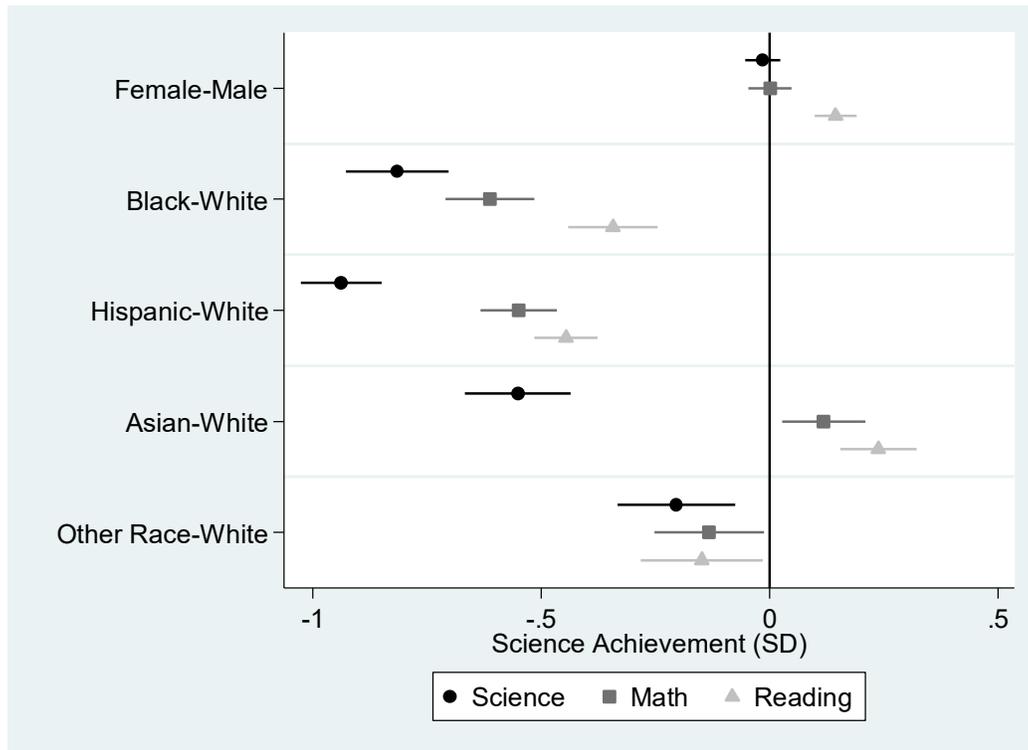


Figure 2. Science, Mathematics, and Reading Achievement Gaps by Race/Ethnicity and Gender for Kindergarten (n=10,050)

Note. Horizontal lines show 95% confidence interval. Estimates are adjusted to account for the complex sampling design of the ELCS-K:2011 using sample weights and Taylor linearized estimates of the standard errors.

Table 1
Descriptive Statistics on Key Independent, Dependent, and Control Variables

	Full Sample	White	Black	Hispanic	Asian	Other Race	Female	Male
Kindergarten Science Score (STD)	0.010 (0.031)	0.384 (0.026)	-0.431 (0.057)	-0.553 (0.040)	-0.168 (0.058)	0.180 (0.061)	-0.003 (0.036)	0.023 (0.030)
First Grade Science Score (STD)	0.022 (0.029)	0.345 (0.023)	-0.475 (0.060)	-0.438 (0.039)	0.047 (0.051)	0.210 (0.050)	-0.014 (0.033)	0.056 (0.029)
Female	0.486 (0.006)	0.477 (0.008)	0.497 (0.019)	0.486 (0.010)	0.552 (0.024)	0.507 (0.023)	1.000 -	0.000 -
White	0.520 (0.020)	1.000 -	0.000 -	0.000 -	0.000 -	0.000 -	0.509 (0.020)	0.530 (0.021)
Black	0.138 (0.015)	0.000 -	1.000 -	0.000 -	0.000 -	0.000 -	0.141 (0.016)	0.135 (0.016)
Hispanic	0.242 (0.014)	0.000 -	0.000 -	1.000 -	0.000 -	0.000 -	0.242 (0.016)	0.242 (0.014)
Asian	0.040 (0.006)	0.000 -	0.000 -	0.000 -	1.000 -	0.000 -	0.045 (0.007)	0.035 (0.006)
Other Race	0.060 (0.008)	0.000 -	0.000 -	0.000 -	0.000 -	1.000 -	0.063 (0.009)	0.058 (0.008)
Composite SES Measure (STD)	-0.060 (0.022)	0.152 (0.024)	-0.340 (0.034)	-0.444 (0.029)	0.343 (0.073)	0.033 (0.069)	-0.059 (0.023)	-0.061 (0.023)
Fall Kindergarten Math Score (STD)	0.010 (0.026)	0.244 (0.031)	-0.288 (0.044)	-0.407 (0.041)	0.418 (0.053)	0.081 (0.057)	-0.002 (0.028)	0.021 (0.029)
Fall Kindergarten Reading Score (STD)	0.013 (0.023)	0.171 (0.031)	-0.099 (0.048)	-0.348 (0.039)	0.484 (0.055)	0.051 (0.063)	0.069 (0.027)	-0.040 (0.025)
n	10,050	5,358	1,108	2,282	705	597	4,946	5,104

Note. Standard errors in parentheses. Estimates are adjusted to account for the complex sampling design of the ELCS-K:2011 using sample weights and Taylor linearized estimates of the standard errors. Standard errors in parentheses are clustered to account for grouping of students in schools.

Table 2
 Estimated Coefficients and Standard Errors from Regressions Predicting Standardized Spring Kindergarten and First Grade Science Achievement

	Kindergarten		1st Grade		Covariate Adjusted Kindergarten Models				
	1	2	3	4	5	6	7	8	9
Female	-0.0147 (0.0196)	-0.0623** (0.0194)	-0.0158 (0.0197)	-0.00343 (0.0172)	-0.0658*** (0.0151)	-0.0268 (0.0162)	-0.0251 (0.0165)	-0.0369* (0.0175)	-0.0444** (0.0145)
Black	-0.815*** (0.0571)	-0.818*** (0.0609)	-0.615*** (0.0479)	-0.535*** (0.0431)	-0.686*** (0.0526)	-0.562*** (0.0448)	-0.498*** (0.0417)	-0.548*** (0.0479)	-0.334*** (0.0370)
Hispanic	-0.937*** (0.0450)	-0.782*** (0.0410)	-0.695*** (0.0409)	-0.595*** (0.0425)	-0.691*** (0.0486)	-0.592*** (0.0448)	-0.521*** (0.0402)	-0.551*** (0.0342)	-0.283*** (0.0274)
Asian	-0.551*** (0.0586)	-0.293*** (0.0507)	-0.628*** (0.0561)	-0.643*** (0.0629)	-0.695*** (0.0568)	-0.674*** (0.0605)	-0.695*** (0.0605)	-0.374*** (0.0441)	-0.472*** (0.0364)
Other Race	-0.203** (0.0654)	-0.132* (0.0515)	-0.155* (0.0754)	-0.118* (0.0591)	-0.145* (0.0651)	-0.119 (0.0612)	-0.106 (0.0675)	-0.113** (0.0423)	-0.113*** (0.0331)
Composite SES Measure (STD)			0.407*** (0.0151)				0.169*** (0.0171)		0.132*** (0.0127)
Fall Kindergarten Math Score (STD)				0.526*** (0.0111)		0.380*** (0.0153)	0.356*** (0.0152)		0.334*** (0.0138)
Fall Kindergarten Reading Score (STD)					0.473*** (0.0116)	0.188*** (0.0144)	0.162*** (0.0147)		0.201*** (0.0133)
Constant	0.391*** (0.0274)	0.374*** (0.0258)	0.330*** (0.0270)	0.257*** (0.0222)	0.334*** (0.0272)	0.272*** (0.0228)	0.256*** (0.0234)	0.290*** (0.0139)	0.171*** (0.0117)
School Fixed Effects								X	X
R^2	0.182	0.147	0.268	0.433	0.393	0.447	0.460	0.048	0.346
Adjusted R^2								0.048	0.345
Observations	10,050	10,050	10,050	10,050	10,050	10,050	10,050	10,050	10,050

Note. Science achievement outcome is standardized. Estimates are adjusted to account for the complex sampling design of the ELCS-K:2011 using sample weights and Taylor linearized estimates of the standard errors. Standard errors in parentheses are clustered to account for grouping of students in schools.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$