Classroom-based Inequalities and Achievement Gaps in First Grade: The Role of Classroom Context and Access to Qualified and Effective Teachers

GREGORY PALARDY

University of California, Riverside

Background: *An enduring question about achievement gaps is, which aspects of schools contribute most? At the early grade levels, when children spend the vast majority of their school day in a single classroom with a single teacher, school inequities that correlate with achievement gaps likely originate within the classroom. This study examined the degree to which three potential sources of classroom-based inequality contribute to reading and math achievement gaps that develop during first grade, including classroom context, access to a qualified teachers, and access to an effective teacher. The study also estimated the degree to which these effects are manifested among classrooms within the same school and between classrooms at different schools, which has implications for policy and practice.*

Population: *A nationally representative sample of first graders from the Early Childhood Longitudinal Study was used. An important feature of ECLS data is that students were administered reading and math achievement tests near the beginning and near the end of first grade. This sampling design allows for the estimation of student achievement gains during first grade that does not include the summer period, which has confounded past efforts to study achievement gaps.*

Research Design: *Multilevel models were used to estimate classroom-specific and school-specific random effects (i.e., residuals), which are conceptualized as within- and between-school classroom effects. These random effects were then used as outcomes to estimate the degree to which within- and between-school classroom effects contribute to Black–White and Hispanic–White achievement gaps that develop during first grade. Covariates for classroom context, access to a qualified teacher, and access to an effective teacher were entered into the model hierarchically to isolate their effects on the gaps.*

Conclusions: *Classroom inequality within and between schools contributed substantially to achievement gaps that developed during first grade. Inequality in contextual aspects of classrooms was the most prominent school-based factor, the majority of which originated from classrooms in different schools rather than classrooms in the same school. However, compared to White children attending the same school, Black children tended to be members of a classroom with more negative contextual characteristics and a less effective teacher. This within-school inequity likely stems from non-random assignment of students to teachers. Finally, Black and Hispanic children were slightly less likely to be taught by a highly qualified teacher. However, that inequity did not significantly contribute to achievement gaps.*

Teachers College Record Volume 117, 020302, February 2015, 48 pages

0161-4681

An important objective of public education is to provide all children with an opportunity to learn and excel in school. Yet substantial school-based inequalities exist that undermine learning for underrepresented minority (URM[1]) children, contributing to what is often referred to as achievement gaps. The literature has documented a number of school-based factors that contribute to these gaps (Betts, Zau, & Rice, 2003; Coleman et al., 1966; Greenwald, Hedges, & Laine, 1996; Murnane, Willett, Bub, & McCartney, 2006). However, at the early grade levels where children spend the vast majority of their school day in a single classroom with a single teacher, school-based inequity is likely be predominantly transmitted in the classroom. The present study examines the degree to which inequalities in three general types of classroom effects contribute to achievement gaps that develop during first grade, including contextual aspects of the classroom, access to qualified teachers, and access to effective teachers.

A shortcoming in the research literature on how teachers contribute to inequity is that it largely ignores the fundamental issue of access to effective teachers in terms of their contributions to student learning, and instead focuses primarily on access to qualified teachers. Similarly, the most prominent federal policy intervention for addressing achievement gaps (i.e., NCLB) does not address the issue of access to effective teachers, instead focusing on equitable access to qualified teachers in terms of certifications and degrees. This research and policy focus is problematic because, while it is easier to regulate teachers in terms of their qualifications than their effectiveness, the literature on the relationship between teacher qualifications and student learning is conflicting, with many studies concluding that there is little or no relationship (Palardy & Rumberger, 2008). Therefore, it is unclear whether assuring equitable access to qualified teachers will address achievement gaps. Furthermore, that focus distracts from the more essential issue of assuring equitable access to effective teachers.

Another limitation of the research literature on school-based factors that contribute to inequality in the early grades is the lack of differentiation between classroom and school effects. In particular, there is a dearth of research on how contextual characteristics of classrooms, as opposed to contextual characteristics of schools, contribute to achievement gaps (Abt Associates, 1997). This classroom-versus-school distinction is relevant because effective interventions targeting classroom factors will likely differ from those targeting school factors. Variation in the contextual characteristics of classrooms, teacher characteristics, and the effectiveness of teachers in the same school are sources of classroom effects, which are referred to as "within-school" classroom effects. Variation in mean classroom characteristics, mean teacher characteristics, and mean teacher effectiveness among schools are sources of school effects, which is referred to as the "between-school" classroom effect.

While it is obvious that teachers within the same school tend to vary in their effectiveness, it is less evident that the contextual characteristics of classrooms in the same school may vary. Contextual differences, including student composition, may occur arbitrarily during student assignment. Contextual difference may also be the result of strategic planning. For example, some research suggests that principals occasionally assign children with certain characteristics to certain classrooms in an effort to counterbalance other perceived issues (Burns & Mason, 2002). Principals may also assign students who have certain characteristics to teachers who are perceived to have a willingness, capacity, or specialized training to constructively work with those students. These student assignment strategies can create contextual imbalances across classrooms in the same school that potentially contribute to achievement gaps. Ignoring these potential within-school effects when modeling achievement gaps can result in incorrect conclusions about the origin of achievement gaps. This can lead to misguided solutions because the appropriateness of interventions for addressing the gaps depends in part on their origin. For this reason, it is critical to consider both within- and between-school components when modeling achievement gaps.

BACKGROUND AND LITERATURE

Achievement gaps, particularly between Black and White children, have received considerable attention from educational researchers since the beginning of the civil rights movement (Conant, 1961; Jencks & Phillips, 1998). The Coleman Report was the first large-scale study that examined racial differences in achievement, finding significant Black–White gaps at each grade level that tended to increase with grade levels (Coleman et al., 1966). Trend data reveals that much progress was made during the '60s, '70s, and '80s toward reducing achievement gaps, and research has linked this progress to federal and state programs designed to improve educational opportunities for Black and Hispanic children (Grissmer, Flanagan, & Williamson, 1998; Hedges & Nowell, 1998). For example, Title I funding, which targets schools with high numbers or percentages of children from low-income families (these schools also tend to serve high proportions of URM children), increased substantially during these decades (Grissmer et al., 1998). In addition, for an approximately 15-year period beginning in the mid-1960s, progress was made toward desegregating schools through busing, redistricting, and other methods (Orfield, 2005). These and other programs altered the context of minority-serving schools either by increasing the level of available resources or by changing the composition of the student body. Due in part to those programs, the rate of progress toward achievement parity was such that by the mid-1980s some scholars projected that ethnic achievement gaps would close by the turn of the century. However, since the 1990s the rate of progress in closing achievement gaps has fluctuated, at times abating or even reversing (Lee, 2002; Magnuson & Waldfogel, 2008). Although trend data indicates that achievement gaps are smaller today than reported by Coleman et al. (1966), they remain substantial (Gamoran & Long, 2006).

CLASSROOM-BASED FACTORS CONTRIBUTING TO ACHIEVEMENT GAPS

Many factors are known to be associated with URM achievement gaps in the early grades. Of those factors the most prominent are students' family and academic backgrounds (Berliner, 2006, 2014; Lee & Burkam, 2002; Phillips, Brooks-Gunn, Duncan, Klebanov, & Crane, 1998), particularly socioeconomic status (SES) (Brooks-Gunn, Duncan, & Klebanov, 1994; Duncan & Magnuson, 2005; Fryer & Levitt, 2004; Reardon, 2011). However, this should not be taken as evidence that schools do not matter. On the contrary, the research literature suggests that school-based inequities play a major role in achievement gap development because URM children and children from low SES backgrounds tend to receive less academic support at home, and therefore are more dependent on schools for educational progress (Alexander, Entwisle, & Olsen, 2001; Greenwald et al., 1996; Johnson, Kraft, & Papay, 2012; Palardy, 2008). This section outlines the literature on classroom factors that contribute to achievement gaps at the early grades with subsections on the roles of classroom context, access to qualified teachers, and access to effective teachers.

Classroom Context

Classroom context can be defined as characteristics or features of classrooms that do not include the teachers or their teaching. This includes the composition of the student body, classroom structures, and resources.

Student composition. Student composition pertains to aspects of the student body such as mean SES, mean achievement, and rate of student mobility. To the degree that these factors are associated with achievement and differ in the classrooms that minority children and White children are members of, they are sources of inequity that potentially contribute to achievement gaps. The compositional effect that has received the most attention in the literature is the mean SES of the student body. Its robust association with student achievement above and beyond students' own family and academic backgrounds has long been recognized (Borman & Kimball, 2005; Coleman et al., 1966; Palardy, 2008; Rumberger & Palardy, 2005). Research suggests that mean SES is a proxy measure for SES-related peer influences that tend to depress educational performance in low SES settings where peers have lower levels of the educational and cultural attributes that enhance educational performance (Hanushek, Kain, Markman, & Rivkin, 2003; Jencks & Mayer, 1990). Mean SES may also be a proxy measure for school factors that are associated with achievement, such as the quality of the teachers and the rigor of the curriculum (Betts, Rueben, & Danenberg, 2000).

Another compositional factor that has been linked to learning outcomes is student mobility. The vast majority of the research in this area focuses on the impact that changing schools has on the mobile student's academic and behavioral outcomes (Rumberger, Larson, Ream, & Palardy, 1999; Swanson & Schneider, 1999). However, some research suggests that student mobility also negatively impacts other non-mobile members of the classroom (Conniff, 1998). This is because when mobile students enter classrooms mid-year, teachers need to (1) divert their attention from the rest of the class in order to assimilate the new student into the class routine, (2) provide individual instruction on material that was not covered in the mobile student's former school, (3) address behavioral problems that are more common among mobile students, and (4) update records (Lash & Kirkpatrick, 1990; Sanderson, 2003). Although no research has directly linked the rate of student mobility to achievement gaps, mobility rate tends to be higher in high-minority schools (Rumberger et al., 1999; Swanson & Schneider, 1999).

Research has also found that the academic composition of the student body is associated with classroom practices. For example, classrooms composed of mostly high-achieving children typically receive accelerated instruction whereas classrooms with mostly low-achieving children tend to receive a less rigorous version of the curriculum (Gamoran, 1992). While academic composition has not been directly linked to achievement gaps, to the degree that Black and Hispanic children are underrepresented in high-achieving classrooms, it may be a contributing factor.

Classroom Heterogeneity

Variation within classrooms in the characteristics of students, such as their levels of achievement, giftedness, special needs, or English proficiency, may contribute to achievement gaps. That is because such heterogeneity tends to necessitate individualized instruction, which diverts the teacher's instructional focus away from the class at large. However, no research has examined the effects of classroom heterogeneity on student outcomes or achievement gaps, and it is not known whether URM children tend to be members of more heterogeneous classrooms. However, Black and Hispanic youths are overrepresented among children classified as learning disabled and are less likely to be classified as gifted, and Hispanic students are far more likely to be non-native English speakers (Donovan & Cross, 2002; U.S. Department of Education, 2006), suggesting that the classrooms they are members of may tend to be more heterogeneous on those factors.

Structures and Resources

Some structural features of classrooms and the availability of resources have also been linked with achievement gaps. For example, large class size is negatively associated with student learning and tends to have a stronger negative impact on Black children than on White children (Krueger & Whitmore, 2001; Ladd, 2008). Class size is also positively associated with rate of disruptive student behavior (Betts & Shkolnik, 1999; Stasz & Stecher, 2000). Moreover, this research suggests that teachers tend to adapt to student misbehavior by altering their instructional practices to better manage behavior at the expense of the academic rigor of the instruction. Finally, the availability of classroom resources such as books, supplies, and equipment also affect student learning and tend to be in greater scarcity in high minority schools (Betts, Zau, & Rice, 2003; Greenwald et al., 1996), potentially creating inequity that can contribute to achievement gaps.

Teacher Qualifications

Research on the importance of teacher qualifications to student learning is inconclusive. Even reviews of the literature are conflicting. Reviews by Hanushek (1986, 1989, 1997) found that qualifications are generally not associated with student learning. Conversely, a review by Greenwald et al. (1996) found that teachers' education, experience, and academic ability are related to student achievement. A more recent review of 21 studies, all of which controlled for students' prior achievement and SES, concluded that teachers' own academic ability was predictive of student achievement gains (Wayne & Youngs, 2003). However, the same study found inconsistent evidence for the effects of certifications, degrees, and coursework on student learning.

Teacher qualifications have the potential to impact achievement gaps because URM and low SES children tend to be served by less-qualified teachers (Betts, Rueben, & Danenberg, 2000; Clotfelter, Ladd, Vigdor, & Wheeler, 2002). However, there is a dearth of research examining the association between inequitable access to qualified teachers and achievement gaps. Furthermore, because the body of research on the effects of teacher qualifications on achievement is inconclusive, it may be the case that inequitable access to qualified teachers does not contribute to achievement gaps.

Teacher Effectiveness

Teacher effectiveness pertains to the impact teachers have on student outcomes, particularly achievement. *Teaching*, namely the practices teachers employ and the attitudes and beliefs they convey once in the classroom, is a more direct measure of teacher effectiveness than are teacher qualifications. Despite its greater relevance, the effect of teaching on student outcomes has received far less attention in the research literature than teacher qualifications. That is likely because teaching involves a large set of complex social interactions that take place over an extended period, and therefore is far more difficult to measure than qualifications. Nonetheless, researchers have examined teaching by breaking it down into specific practices and attitudes. These studies found significant associations between a number of instructional practices and student learning (Guarino, Hamilton, Lockwood, & Rathbun, 2006; Palardy & Rumberger, 2008; Xue & Meisels, 2004). Similarly, specific teacher attitudes and beliefs that are predictive of their students' achievement gains have been identified (Goodard, Hoy, & Hoy, 2000; Palardy & Rumberger, 2008).

Research has also linked teaching with achievement gaps, finding that teachers of URM children tend to use practices and have attitudes that may be less conducive to cognitive development (Stipek, 2004). For instance, Stipek (2004) found that teachers at schools with higher concentrations of URM students tend to use instructional practices that focus on developing discrete skills and factual and procedural knowledge. In contrast, teachers at schools with greater concentrations of White children tend to use constructivist teaching methods that have been advocated by such prominent theorists as Piaget and Vygotsky, and that emphasize the development of higher-order thinking and social skills. Similarly, teachers tend to have lower expectations of Black students (Baron, Tom, & Cooper, 1985; Jussim, Eccles, & Madon, 1996), which may negatively impact their learning via a self-fulfilling prophecy (Rosenthal, 1994).

While research on teaching has identified specific practices and attitudes that contribute to learning, and to a much lesser extent, achievement gaps, it does not address the more fundamental issue of the overall effectiveness of teachers in terms of the degree to which they contribute to student learning. Because of the complexity of teaching, individual practices and attitudes, or even sets of practices and attitudes that are predictive of overall effectiveness, may yield unreliable estimates of overall teaching effectiveness. This problem has created the need for methods that more directly measure the overall effectiveness of individual teachers. Value-added modeling (VAM) has been designed for that purpose (Harris, 2011; McCaffrey, Lockwood, Koretz, Louis, & Hamilton, 2004).

In addition to estimating the performance of individual teachers, VAM has been used to study the degree to which teachers vary in their effectiveness and how that can impact students' achievement. For instance, research that employs VAM has shown that being taught by an effective teacher versus an ineffective teacher can result in a full grade-level difference in achievement in a single school year (Hanushek, 1992). Further, research using VAM revealed that having a succession of ineffective teachers can have an even larger cumulative impact (Konstantopoulos, 2009; Rivkin, Hanushek, & Kain, 2005). These findings are particularly concerning because the impact of a good or bad teacher may persist for several subsequent years (Konstantopoulos & Chung, 2011; Sanders & Horn, 1995).

METHODOLOGICAL ISSUES

Researchers have noted a few issues related to the measurement of achievement and the statistical modeling of achievement gaps that can impact gap estimates. One issue is the scale or metric of the achievement test. Seltzer, Frank, and Bryk (1994) show that estimates of change in an outcome over time depend in part on the scale of the outcome. This conclusion also applies to estimates of change in achievement gaps (Murnane et al., 2006; Reardon, 2008). When estimating achievement gaps, it is important to ascertain that the test instrument is on an interval scale. However, in practice that is rarely verified, likely because doing so is complicated.

Achievement gap estimates are also sensitive to model specification. In order to effectively address achievement gaps, it is important to differentiate the part of the gap that originates from differences in the quality of schools that URM and White children attend (i.e., between-school effects) from within-school differences in classroom quality. Some recent research suggests that estimates of the within- and between-school effects are particularly sensitive to model specification. This point is illustrated by two recent studies that employed Early Childhood Longitudinal Study (ECLS) data, but used different model specifications and drew essentially opposite conclusions on the degree that within- and between-school effects contribute to Black–White gaps. Fryer and Levitt (2004, 2006) used a model specification known as "school fixed effects" to conclude that between-school effects have little bearing on Black–White achievement gaps during the early grades, whereas within-school effects contribute substantially. Hanushek and Rivkin (2006) use the same data source and a standard model specification to draw the opposite conclusion and show that the school fixed effects specification will underestimate the between-school component to the degree that schools differ in the proportion of students who are Black. It is widely known that there is a considerable degree of racial/ethnic segregation in American schools and therefore, the school fixed effects specification is generally inappropriate for estimating the within- and between-school components. Reardon (2008) concurs that Fryer and Levitt's (2004, 2006) method has a strong tendency to underestimate the degree to which between-school effects contribute to gaps, but also argues that Hanushek and Rivkin's (2006) specification tends to overestimate the between-school effects. He shows that without a clear theoretical or conceptual framework guiding the model specification, a substantial proportion of the achievement gap cannot clearly be attributed to within- or between-school effects. Currently there is no consensus on the best model specification or most appropriate conceptual framework. More research is needed addressing this important issue.

ACHIEVEMENT GAP RESEARCH BASED ON ECLS DATA

Several studies have examined achievement gaps using ECLS data (Fryer & Levitt, 2004, 2006; Lee & Burkam, 2002; Murnane et al., 2006; Reardon, 2008; Reardon & Galindo, 2009; Reardon & Robinson, 2008; Rumberger, 2007; Rumberger & Arellano, 2009). Some of these studies are descriptive, focusing on the magnitude of the gaps and the degree to which they change during elementary school. Lee and Burkam (2002), who were perhaps the first to use ECLS data to examine achievement gaps, estimate the magnitudes of the gaps at the start of kindergarten and factors associated with those gaps. They also show that, on average, URM and low SES children attend kindergarten at schools that are lower quality in terms of mean student achievement, school resources, the qualifications and attitudes of the teachers, and neighborhood deprivation. Rumberger (2007) found that while Black–White gaps increase markedly during elementary school (K–fifth grade), Hispanic–White gaps close substantially. He also found that both Black and Hispanic children are far more likely to attend high-poverty schools, which presumably are inferior on a number of dimensions of school quality, and that Black children have lower academic learning behaviors compared with Hispanic children. He hypothesizes that these factors contribute to achievement gaps and their patterns of change over time. A series of papers by Reardon and his colleagues document the magnitude of the ethnic and socioeconomic achievement gaps during the first six years of schooling (Reardon, 2008; Reardon & Galindo, 2009; Reardon & Robinson, 2008). Besides estimating the magnitude of the socioeconomic and ethnic achievement gaps, this work shows that the ethnic gaps differ by subgroups. Hispanic–White gaps vary depending on the child's country of origin, immigration and English language status, and above all, SES (Reardon & Galindo, 2009). Furthermore, they found that most disadvantaged Hispanic children start school with the largest gaps, but also make the most progress in closing them during the early grades. Reardon (2008) also shows that the rate at which the Black–White reading gap increases between kindergarten and fifth grade depends on children's initial levels of achievement, with the gaps for initially high-achieving children growing approximately twice as fast as the gaps for initially low-achieving children. While these descriptive studies are essential for documenting the magnitude of the gaps and their change over time, they do not attempt to estimate the degree to which school-based mechanisms contribute to achievement gaps.

Fryer and Levitt (2004) tested a number of hypotheses about the causes of Black–White achievement gaps. As described above, they reject the hypothesis that differences in the quality of the schools that Black and White children attend substantially contribute to the gaps. Instead, they found that within-school differences contribute to the growing gaps in kindergarten and first grade. This means that Black children tend to progress at slower rates than White children attending the same school, even after controlling for differences in students' backgrounds. While the authors were able to rule out discrimination by White teachers toward Black students as a factor contributing to the within-school effect, they did not find any specific factors that did contribute. Fryer and Levitt (2006) extended their 2004 study to include an additional wave of data collected near the end of third grade. They note the Black–White gaps continued to grow at a rate of approximately 0.10 standard deviations per year, and otherwise draw conclusions similar to their previous study. Murnane et al. (2006) designed their study to extend upon Fryer and Levitt's 2004 study to test a more extensive set of school effects on achievement growth during the first four years of schooling (K–third grade). They found that while teacher qualifications were not associated with achievement growth, school socioeconomic composition was. However, they did not directly examine the degree to which classroom or school factors contribute to achievement gaps. Rumberger and Arellano (2009) studied Latino–White reading and math achievement gap trajectories during kindergarten and first grade in California using a subsample of ECLS data. They focused on the role of student background and school factors, finding that student SES and

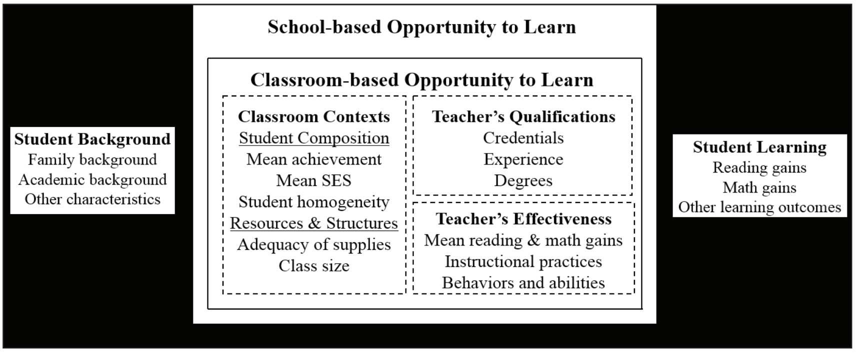
English language status account for over 80% of the reading gap and over two-thirds of the math gap that exist when children enter kindergarten. After controlling for those two student background variables, the reading gap was not significant and the math gap was marginally significant ($p < 0.10$). They also found that Latino–White achievement gaps did not grow during the first two years of schooling.

As described above, estimates of change in achievement gaps are sensitive to the metric of the test scores. It is important that the test scores are on an interval scale. ECLS data achievement test scores are provided in the t-score metric, which is standardized to a mean of 50 and standard deviation of 10 at each data collection wave. The test scores are also provided in the IRT scale score metric, which uses an ability-specific subset of the test items to estimate each child's score on the full set of items and is not standardized. Reardon (2008) examined the scale of these metrics for the ECLS data, concluding that the IRT scale score metric is sensitive to the set of test items that were administered and that the size of a unit is larger at the high end of the test scale. Hence, the IRT scale scores are not interval scale and are not inappropriate for estimating change in achievement gaps. The wave-standardized scores are approximately interval scale and are recommended for modeling change in the gaps between two waves (Reardon, 2008). However, the wave-standardized scores are inappropriate for growth modeling, which utilizes more than two waves of data. For growth modeling it is recommended that the scores be transformed to the IRT theta score metric and vertically linked across waves.

CONCEPTUAL FRAMEWORK

This study is guided by a conceptual framework for the role that opportunity to learn (OTL) in early elementary school classrooms plays in the formation of achievement gaps (see Figure 1). The framework is informed by the research literature (described above), which has identified aspects of classrooms that impact achievement and where there may be inequity that contributes to achievement gaps. Children enter school with a set of family, academic, and other background characteristics that directly impact their learning. To the degree that these background characteristics differ by ethnic group, they may contribute to achievement gaps without regard to school-based inequity. Student learning is also a function of school-based factors. At the early grades where children spend the majority of their school day in a single classroom with a single teacher, OTL at school may be largely a classroom effect. The framework identifies three aspects of classrooms for which ethnic inequality may impact OTL and lead to the formation of achievement gaps, including the contextual characteristics of the classroom, access to qualified teachers, and access to effective teachers in terms of their contribution to students' learning. While teachers' effectiveness is likely a function of a number of factors including instructional practices, classroom behaviors and attitudes, teaching ability, and a large number of more subtle factors, it is operationalized as the value teachers add to student achievement. Although the framework recognizes that non-classroom-based school effects (e.g., per pupil expenditures or principal leadership style) may impact OTL even at the early grades, the role of those factors is not elaborated upon because the present study focuses on classroom effects.

Figure 1. Conceptual Framework for Studying the Role of Classroom-based Inequality in Opportunity to Learn on the Formation of Achievement Gaps



Note that each subheading is followed by a list of examples, which are not intended to be comprehensive.

RESEARCH QUESTIONS

The present study uses data from the Early Childhood Longitudinal Study Kindergarten Class of 1998–99 (ECLS-K) to examine the degree to which inequity in school-based OTL contributes to achievement gaps that accumulate during first grade. The study is designed to partition the effects of classrooms within the same school (within-school classroom effects) from classrooms in different schools (between-school classroom effects). As described by the conceptual framework in Figure 1, school-based OTL is gauged based on whether URM children have equitable access to classrooms with contextual factors that facilitate learning, to qualified teachers, and to effective teachers. The following research questions are addressed.

1. What are the magnitudes of the URM achievement gaps in reading and math at the beginning of first grade, and to what degree do they change during first grade?
2. Do student inputs vary across classroom and schools? If so, to what degree do those differences account for school-based achievement gaps that accrue during first grade?
3. What are the magnitudes of school-based achievement gaps that develop during first grade?
4. Do the URM children have equitable access to: (a) classrooms with contextual characteristics conducive to learning, (b) highly qualified teachers, and (c) effective teachers? If not, to what degree does each contribute to the achievement gaps that accumulate during first grade?

Because of the implications for policy and practice, the degree to which achievement gaps originate from factors among classrooms in the same schools, and from factors between classrooms in different schools, is also addressed for each research question.

SIGNIFICANCE OF THIS STUDY

This study was designed to extend the literature on how classroom-based inequality in OTL contributes to achievement gaps. First, the study was designed to provide evidence on the degree to which inequality on three major aspects of classrooms (context, teacher qualifications, and value-added teacher effectiveness) contribute to achievement gaps. There is surprisingly little research on this important issue. As outlined above, the vast majority of the research on achievement gaps either estimates the size of gaps and their change over time, or estimates the effects of classroom factors on achievement rather than on achievement gaps. No studies that estimate the combined effects of any of the three major aspects of classrooms considered here could be found in the literature. While there is a body of research examining the effects of individual measures of teacher qualifications (e.g., full certification) on achievement and achievement gains, no research examines the effect of inequity on a set of measures of combined teacher qualifications (i.e., certifications, experience, and degrees) on *achievement gaps*. Similarly, some research has investigated the effects of specific measures of classroom context (e.g., class size) on achievement, but no studies address the combined effect of inequity on a set of contextual characteristics of classrooms on achievement gaps. Moreover, although access to an effective teacher is perhaps the most fundamental aspect of school-based OTL during the early grades, there is surprisingly little research examining this and the degree to which it contributes to achievement gaps. The research on this issue addresses it indirectly by focusing on the effectiveness of specific instructional practices and teacher attitudes. While informative, this approach does not capture the teachers' overall effectiveness and the degree to which that contributes to achievement gaps. As a result, despite the many studies on achievement gaps, the somewhat fragmented literature does not provide a clear understanding of the relative importance of key aspects of classrooms to achievement gaps.

Second, this study distinguishes between classroom effects that stem from differences among classrooms in the same school (i.e., within-school classroom effects) and differences among classrooms in separate schools (i.e., between-school classroom effects). Previous research on achievement gaps has focused primarily on between-school effects and has largely ignored within-school effects (Abt Associates, 1997).[2] The vast majority of the studies that do examine within-school effects treat classroom effects as if they were measures of individual differences among students, which raises methodological concerns. The current study partitions the variance in student achievement gains into student, classroom, and school components to produce accurate estimates of within-school classroom effects. This is an important shortcoming in the literature because understanding the origin of the effects is critical for developing successful policies and practices for reducing achievement gaps.

Third, this study examines the achievement gains that accrue during first grade with summer excluded. The inclusion of summer when studying school-based factors that contribute to achievement gaps raises methodological concerns because research indicates that URM children tend to learn less than others over summer (Alexander et al., 2001; Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996), which generally cannot be attributed to teacher, classroom, or school effects. Hence, including summer in achievement gains calculations may bias estimates of school-based achievement gaps and the effectiveness of teachers and schools that serve URM children. ECLS is the only national database for which school-year achievement gains can be calculated excluding summer. However, the extant research on achievement gaps, including studies that use ECLS data, calculate achievement gains or growth with summer included.

METHODOLOGY

This section describes the data sample, variables, and statistical models, as well as how the sequence of models that are estimated address the research questions.

DATA SOURCE

This study uses first-grade data from ECLS. When the appropriate sampling weights are applied, ECLS-K is nationally representative sample of 1998 kindergarteners (NCES, 2002).[3] A vast array of survey items and test scores were collected from the children, their parents, teachers, and school principals. The full first-grade longitudinal sample has 5,034 students. Students without teacher or school IDs were omitted, as were a small number of students who had missing achievement test scores or changed teachers during first grade. The final sample used in this study included 3,496 students, 887 classrooms, and 253 schools.[4]

VARIABLES

Outcome Variables

Achievement gains in reading and math during first grade are the outcomes in this study. The achievement tests were administered near the beginning and end of first grade so the gains do not include summer. Achievement gains were computed by subtracting the fall T-distribution test score from the spring T-distribution test scores. The achievement gains were then standardized to a mean of zero and standard deviation of one. This allows for estimates of the achievement gaps to be in units of standard deviations or effect size. It is worth noting, however, that the expected mean achievement gain in math and reading during first grade is zero.

Independent Variables

Several measures of student characteristics, classroom context, and teacher qualifications were used. The full list of variables with means and standard deviations broken down by ethnicity are shown in the Appendix Table. Note that no variables measuring aspects of teaching effectiveness are used because teacher effectiveness is estimated using VAM. Below are additional details on the independent variables.

Student control variables. While the current study focuses on classroom effects, many student characteristics that are associated with learning also vary across classrooms and schools. Ignoring these differences in student inputs across classrooms can bias estimates of teacher, classroom, and school effects. For this reason, it is recommended that differences in student inputs across classrooms be statistically controlled when modeling classroom and school effects (Ballou, Sanders, & Wright, 2004). The research literature has identified a variety of student characteristics that are associated with student achievement gains and may vary across classrooms. Those variables are classified into four types: demographics, academics, behavioral, and parental. Several measures of each type are used to control for student inputs in this study (see the Appendix Table). Note that student ethnicity is also controlled in an effort to produce estimates of classroom effects that are not the result of ethnic differences in student inputs.

Classroom context. Variables measuring four aspects of classroom context were used: student composition, student heterogeneity, structures, and resources. Measures of student composition include mean SES, mean reading and math achievement, and the proportion of students who enter the classroom after the start of the school year. Student heterogeneity includes variance in reading and math achievement, and the proportion of students with each of the following characteristics: gifted, English language learner, and disabled. One measure of classroom structure, class enrollment (often referred to as "class size"), was used. In addition, four measures of classroom resources were considered, including the adequacy of the instructional supplies, materials, technology, and physical facilities. Each of the resource measures is a factor score developed from related survey items.

Teacher qualification. Three measures of teacher qualifications are examined, including whether the teacher is fully certified, has earned a graduate degree, and has five or more years of teaching experience. These measures frequently appear in the research literature and are consistent with qualifications commonly used in licensing, hiring, and tenure decisions, as well as criteria used for educational policy and legislation designed to assure classrooms are led by highly qualified and effective teachers.

STATISTICAL MODELING

Because students in the ECLS database are nested in classrooms and classrooms are nested in schools, multilevel modeling (MLM) was used. MLM was developed for analyzing nested data and is optimally suited for isolating the variance in student achievement gains that are due to classroom and teacher effects. Four multilevel models were fit sequentially to address the research questions. The process began with the Achievement Gains model, which only included two covariates: prior achievement and the duration between the fall and spring achievement test measurements.[5] This model is considered the unconditional model for change in student achievement during first grade. Second, the Student model was run, which adjusts classroom and school mean achievement gains for differences in student inputs. Third, the Classroom Context model was fit to the data, which, in addition to student inputs, controls for classroom differences in the student body characteristics, structural features, and resources that impact student learning but are generally not attributable to the effectiveness of the teacher. Fourth, the Teacher Qualifications model was run, which includes all Classroom Composition model variables plus three common measures of teacher qualifications (described above).

The multilevel equations for the Teacher Qualifications model are shown below. Note that the other three models are reduced forms of the Teacher Qualifications model for which sets of variables are omitted.

Level 1: Student Level

$$\text{Achievement}_{ics} = \pi_{0cs} + \pi_1 \text{Fall Achievement} + \pi_2 \text{Time Adjustment} + \sum_{q=3}^{9} \pi_q \text{Demographics}_q + \sum_{q=10}^{11} \pi_q \text{Academics}_q + \sum_{q=12}^{16} \pi_q \text{Behavioral}_q + \sum_{q=17}^{22} \pi_q \text{Home}_q + e_{ics} \qquad e_{ics} \sim N(0, \sigma^2)$$

Reading and math achievement are the outcomes and are modeled separately. The subscript, *ics*, denotes the nested structure of the data; individuals, *i*, are nested in classrooms, *c*, which are nested in schools, *s*. The model controls for fall achievement, the duration between fall and spring achievement tests (see note 4), seven demographic, two academic, five behavioral, and six parental home involvement variables that research has indicated are associated with achievement gains and teacher effects. A list of the student control variables appears in Table 1 and also with additional detail in the Appendix Table. Continuous measures were grand mean centered and dummy variables were left uncentered to preserve the dummy variable interpretation. Centering the variables this way adjusts classroom achievement gains for differences in student inputs, which is important for the purpose of reducing the potential for biases in the estimates of classroom and school effects that may result from differences in student inputs. The coefficients for the Level 1 student variables (π_1 - π_{22}), which are all fixed, represent the linear associations between the respective variable and the outcome.

While the Level 1 model is critical for separating the variance in student achievement gains due to individual differences among children from classroom and school effects, the substantive interest of the current study is in Level 2 and 3 models. Level 2 is the within-school classroom model. The outcomes at Level 2 (π_{0cs}) are the adjusted mean classroom achievement gains in reading or math.

Level 2: Classroom Model

$$\pi_{0cs} = \beta_{00s} + \sum_{k=1}^{9} \beta_{0k}\text{Classroom Context}_k + \sum_{k=10}^{12} \beta_{0k}\text{Teacher Qualifications}_k + r_{0cs} \qquad r_{0cs} \sim N(0,\tau_\beta)$$

$$\pi_1 \text{ through } \pi_{24} = \beta_1 \text{ through } \beta_{24}$$

The intercept, β_{00s}, is the adjusted mean gain on the outcome for each school. Nine covariates measuring aspects of classroom context and three measuring teacher qualifications were included. The beta coefficients (β) represent the linear relationship between the classroom and teacher measures and mean classroom achievement gains. As was the case at Level 1, continuous variables were grand mean centered and dummy variables were uncentered. r0cs is the classroom residuals, which is used in the residual analysis described below. The variance in the classroom residuals is represented by τ_β.

At Level 3, the school level, school mean achievement gains (B0cs) in reading or math are the outcomes. These means are adjusted for classroom differences in the student and classroom covariates in the model. The school model is unconditional (no covariates) because the focus of this study is on classroom effects rather than on school effects per se. Including an unconditional school level, as opposed to using a two-level model with no school level, partitions the variance in mean classroom achievement gains into within- and between-school components. This provides an opportunity to examine the degree to which classroom inequity that impacts achievement gaps is manifested within schools and between schools. The model-building sequence provides estimates of the degree to which each successive class of variables entered into the model accounts for within- and between-school variance in mean student achievement gains.

Level 3: Unconditional School Model

$$\beta_{0cs} = \gamma_{000} + u_{00s} \qquad u_{00s} \sim n(0,\tau_\gamma)$$

$$\beta_1 \text{ through } \beta_{24} = \gamma_1 \text{ through } \gamma_{24}$$

$$\beta_{01} \text{ through } \beta_{012} = \gamma_{01} \text{ through } \gamma_{012}$$

The intercept, γ_{000} , is the adjusted grand mean achievement gain. The Level 3 model also provides coefficient estimates for each student covariate (γ_1 through γ_{22}) and classroom covariate (γ_{01} through γ_{012}). The school residuals (u_{00s}) are also used in the residual analysis. τ_γ represents the variance in those residuals.

RESIDUAL ANALYSES

This section describes the model residuals and provides an explanation for how they are used to address the research questions.[6]

Defining the Residuals

As described above, there are three residuals in the MLM including the student, classroom, and school residuals, which capture sources of unexplained variance in the achievement gains outcome. These residuals can be considered latent variables and are assumed to be approximately normal in distribution with a mean of zero. The Level 1 model residual describes each student's achievement gain relative to other children in the same classroom, controlling for the student covariates in the model. Similarly, the Level 2 residual describes the mean student achievement gain for each classroom compared with other classrooms in the same school, controlling for the student and classroom covariates in the model. A Level 2 residual of zero indicates the given classroom's achievement gain is average for the school it is in, whereas a large positive residual indicates the classroom's achievement gain is much larger than average for classrooms in the schools, again controlling for the covariates in the model. The Level 3 residuals are interpreted similarly, except that they measure the deviation of the mean achievement gain for the classrooms at a given school compared with the mean achievement gain for all schools.

Addressing the Research Questions

The residuals from the four sequential models were used to address most of the research questions in this study. The details of the residual analyses and rationale for the interpretations of the results are provided next.

Research question 1. The Achievement Gains model residuals were used to estimate the magnitude of achievement gaps that developed during first grade. The total gap that developed during first grade was estimated by summing the student, classroom, and school residuals and conducting an ANOVA with post-hoc tests to determine the magnitude and significance of mean differences in Black–White and Hispanic–White achievement gains.[7] The gaps that can be attributed to classroom factors within- and between-schools were calculated using the Achievement Gains model classroom and school residuals, respectively. ANOVA with post-hoc tests were conducted on each residual.

Research question 2. To address the degree to which differences in student inputs contribute to school-based achievement gaps that accrue during first grade, the change in the gaps that result from adding the student control variables was computed. This entailed estimating the Student model gaps for the classroom and school residuals, which was accomplished using the same method described above for the Achievement Gains model in research question 1, and then computing the changes in the gaps by subtracting Student model gap estimates from Achievement Gains model gap estimates.

Research question 3. The classroom and school residuals from the Student model were also used to address research question 3 on the magnitudes of the school-based achievement gaps that accumulate during first grade. After controlling for differences in student inputs, the classroom and school residuals describe the degree to which within- and between-school classroom factors contribute to achievement gains. Using the method described above for research question 1, ANOVA with post-hoc tests were used to estimate the magnitude of the within- and between-school classroom gaps and the magnitude of the total school-based gaps.

Research question 4. Corresponding with the method of residual analysis described for research question 2, the effect of classroom context on achievement gaps is addressed by examining the change in the gaps that occurred after controlling for the set of variables measuring classroom context. That was accomplished by subtracting the gaps estimated from the Classroom Context model residuals from the gaps estimated from the Student model residuals. Likewise, the effects of teacher qualifications on achievement gaps was derived by subtracting the gaps estimated using the Teacher Qualifications model residuals from the gaps estimated using the Classroom Context model residuals.

To address the degree to which access to effective teachers contributes to achievement gaps, first the effectiveness of each individual teacher was estimated. Given the extensive student and contextual controls included in the Classroom Context model, the classroom residuals from that model can be considered estimates of each teacher's effectiveness relative to other teachers in the same school. However, to the degree that the average effectiveness of teachers in each school varies among schools, using the classroom residuals alone will underestimate the magnitude of the teacher effectiveness estimates. Therefore, the classroom residuals may be considered the lower boundary of the teacher effectiveness estimates. The upper boundary is more difficult to define; however, the sum of the classroom and school residuals may be a conservative estimate.[8] Using the residual analysis method described for research question 1, ANOVA with post-hoc tests were used to estimate the degree to which access to effective teachers impacts achievement gaps. The lower boundary of the effect was estimated from the classroom residuals alone, whereas the upper boundary was estimated using the sum of the classroom and school residuals.

MODEL ASSUMPTIONS AND MULTICOLINEARITY

The validity of the results from the statistical models used in this study depend in part on whether a set of assumptions are met, including: (1) the model residuals are approximately normal, independent, and identically distributed, and (2) the associations between independent and dependent variables are approximately linear. The residuals and variables from each model were examined to verify that those assumptions were met. Moreover, the Variance Inflation Factor (VIF) for each variable was computed to test for multicolinearity. The VIF did not exceed 3.0 for any variable, which is far below the conventionally suggested level for concern of 10.0.

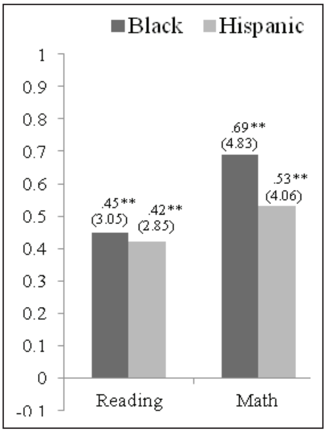
RESULTS

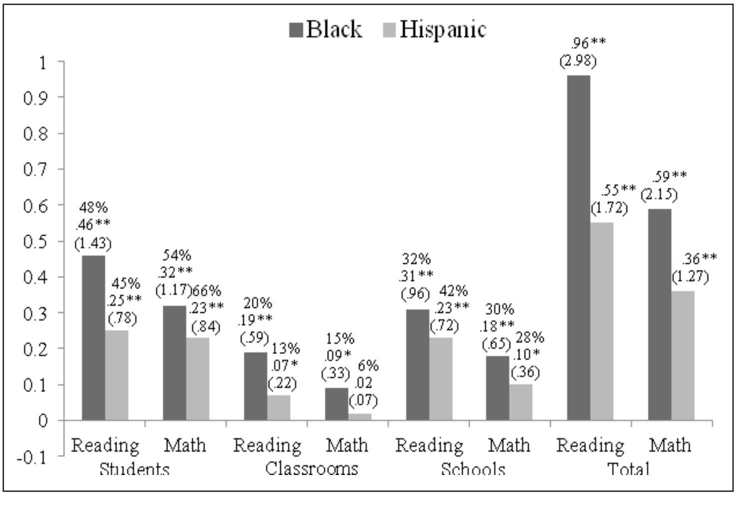
RESEARCH QUESTION 1

Figure 2a shows Black–White and Hispanic–White achievement gaps at the beginning of first grade in units of standard deviations (SD) and months of mean achievement gains. For both ethnic groups, the gaps were larger for math than reading. At the start of first grade, Black children lagged behind White children by 0.45 SD or 3.05 months in reading and 0.69 SD or 4.83 months in math. At the start of first grade the gaps were only slightly smaller for Hispanic children: 0.42 SD or 2.85 months for reading and 0.53 SD or 4.06 months for math. All four gaps are statistically significant ($p < 0.01$). Note that these results are similar to other recent studies that have estimated first-grade achievement gaps (Bali & Alvarez, 2004; Fryer & Levitt, 2004) and are also consistent with previous research indicating URM children begin elementary school lagging behind White children in terms of cognitive skills (Lee & Burkam, 2002).

Figure 2b uses the student, classroom, and school residuals from the Achievement Gains model to illustrate the degree to which each of those sources contributed to achievement gaps that developed during first grade. The Total indicates that all three sources combine to increase the gaps by 1.27–2.98 months as compared to the average learning rate, depending on the outcome and ethnic group. All estimates of the total effects are statistically ($p < 0.01$) and substantively significant. Note that the reading gap increases were larger than the math gap increases for both ethnic groups, which is likely because reading is the primary instructional focus during first grade, and therefore differences in the quality and quantity of instruction may vary more, potentially resulting in a greater impact on school-based inequity.

Figure 2a (left). Achievement Gaps at the Beginning of First Grade.
Figure 2b (right). Changes in Achievement Gaps During First Grade





* = $p < 0.05$, ** = $p < 0.01$. The numbers are for Black–White and Hispanic–White differences in achievement in units of standard deviations and (months) and the percentage of the total gap that is attributable to student, classroom, and school factors. On the left, Figure 2a shows the gaps at the start of first grade. On the right, Figure 2b shows change in the gaps during first grade. The total change is partitioned into student, classroom, and school components. Standard errors, and by extension, the statistical significance levels, were adjusted to account for the design effects that resulted from cluster sampling design used by NCES. For each bar, the bottom number is the gap in months of average achievement gains, and the second number is the gap in standard deviations. For the bars on the left side of Figure 2b, the top number is the percent of the total gap.

The percentage of the total variance in the outcome attributable to classroom and school factors is often cited in school effectiveness research to provide an estimate of the degree to which classrooms and schools matter (see Raudenbush & Bryk, 2002, p. 228).[9] The results in Table 1 (see under "Variance Components" for Achievement Gains model) indicate that 20% of the variance in reading achievement gains and 13% of the variance in math gains are attributable to classroom and school factors combined. These figures can be compared with the degree to which classroom and school factors contribute to achievement gaps that accrue during first grade. On the left side of Figure 2b the total change in the achievement gaps is partitioned into student, classroom, and school components.[10] While individual differences among students accounted for the largest proportion of the total change, Figure 2b also shows that approximately 52% (20% for classroom + 32% for school) of the Black–White reading gap that accumulated during first grade and 45% (15% classroom + 30% school) of the math gap were due to classroom and school factors. Similarly, 55% (13% classroom + 42% school) of the Hispanic–White reading gap and 34% (6% classroom + 28% school) of the math gap were due to classroom and school factors. These figures indicate that an approximately three times greater percentage of the achievement gaps is due to classroom and school factors than to the achievement gains. Note that these findings are indirectly supported by previous research that shows schools have a greater impact on learning for URM and low SES children (Alexander, Entwisle, & Olsen, 2001; Greenwald et al., 1996; Palardy, 2008).

Table 1. Achievement Gains and Student Model Results

	Achievement Gains (1)		Student (2)	
	Reading	Math	Reading	Math
Coefficient Estimates				
Intercept	-0.01	0.01	0.01†	0.16*
Assessment Gap	0.15**	0.23**	0.15**	0.23**
Fall Reading Achievement	0.81**	—	0.80**	0.11**
Fall Math Achievement	—	0.80**	0.22**	0.77**
Demographic Background				
SES	—	—	0.03†	0.04*
Age	—	—	-0.02**	-0.01**
Female	—	—	-0.05†	-0.17**
Asian	—	—	-0.21**	-0.38**

	Achievement Gains (1)		Student (2)	
	Reading	**Math**	**Reading**	**Math**
Black	—	—	-0.18**	-0.22**
Hispanic	—	—	-0.14*	-0.20**
Other	—	—	-0.01	-0.24**
Academic Background				
Disability	—	—	-0.09†	-0.05
Head Start	—	—	-0.18**	0.03
Classroom Behavior				
Approach to Learning	—	—	0.28**	0.18**
Interpersonal Skills	—	—	-0.08**	0.00
Internalizing Problem Behaviors	—	—	-0.01	-0.04*
Externalizing Problem Behaviors	—	—	-0.05†	-0.07**
Self-control	—	—	-0.03	-0.09**
Parental Home Involvement				
Attendance	—	—	0.00	-0.01†
Reading outside of school	—	—	0.05**	-0.01
Practice numbers with child	—	—	0.01	0.00
Belief in reading ability	—	—	0.21**	-0.01
Belief in math ability	—	—	-0.06**	0.12**
Ignore child when annoying	—	—	0.03*	0.00
Variance Components (% of variance explained from prior model)				
Student				
Variance	0.7170	0.7667	0.5745(19.87)	0.6911(9.86)
% of Total	79.44	86.81	77.35	85.50
Classroom				
Variance	0.0833	0.0406	0.0967(-16.09)	0.0681(-67.73)
% of Total	9.23	4.60	13.02	8.37
School				
Variance	0.1022	0.0759	0.0715(30.04)	0.0541(28.72)
% of Total	11.33	8.59	9.63	6.65
Total	0.9025	0.8832	0.7427(17.71)	0.8133(7.91)
Model Summary				

	Achievement Gains (1)		Student (2)	
	Reading	Math	Reading	Math
Deviance Statistic	9268.49	9332.25	8586.43	9053.36
# of parameters	6	6	27	27
Deviance change[1]	—	—	682.06**	278.89**

† $p < .10$, * $p < .05$, ** $p < .01$. [1]Compared with previous model.

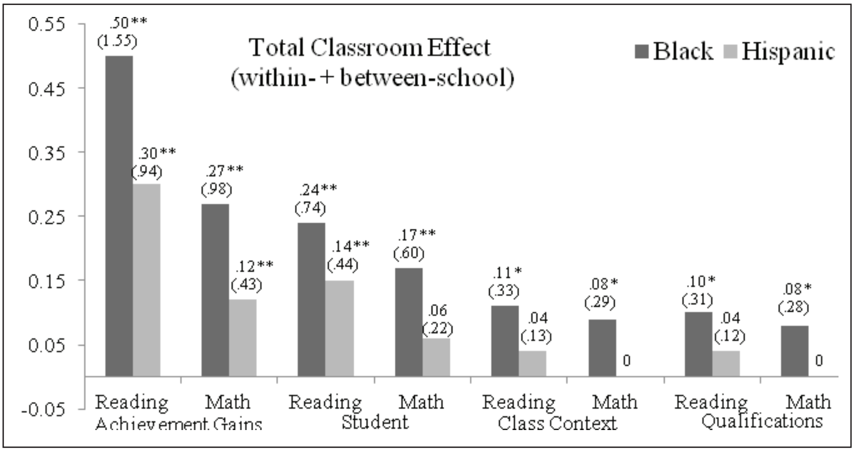
RESEARCH QUESTION 2

The Appendix Table shows that URM and White children differed on several background measures. For example, URM children tended to enter first grade with far less developed reading and math skills, were less likely to have attended preschool, had lower teacher ratings on interpersonal skills and self-control, and spent less time reading outside of school. Each of these differences may contribute to achievement gaps. Moreover, student inputs for each classroom and school may impact school-based achievement gaps, but generally are not attributable to teacher or school effects. For this reason it is critical to control for student inputs to obtain relatively unbiased estimates of teacher and school effects (Ballou et al., 2004).

Figure 3 uses the sum of the classroom and school residuals to illustrate the magnitudes of school-based achievement gaps for each of the four sequential models. Note that the Achievement Gains model gaps in Figure 3 are the sums of the classroom and school gaps in Figure 2b. A comparison of the gaps for the Achievement Gains and Student models shows the degree to which difference in student inputs contributed to the school-based gaps. This comparison indicates that the Black–White reading gap was reduced from 1.55 to 0.74 months (52%) and the Hispanic–White reading gap was reduced from 0.93 to 0.47 months (50%). The math gaps were reduced 37% and 50% for Black and Hispanic children, respectively. These findings indicate that a substantial proportion of school-based achievement gaps that accrue during first grade are due to differences in the background characteristics of the students who enroll, rather than to classrooms and schools effects.

To probe their origins, the school-based gaps shown in Figure 3 were partitioned into classroom effects within schools and between schools, which are presented in Figures 4a and 4b. A comparison of the within-school results (Figure 4a) for the Achievement Gains and Student models reveals that controlling for student inputs reduces the reading gap by 37% for Black children (from 0.59 to 0.37 months) and by 82% for Hispanic

Figure 3. Parsing the School-based Learning Gap into Student, Contextual, and Teacher Effects



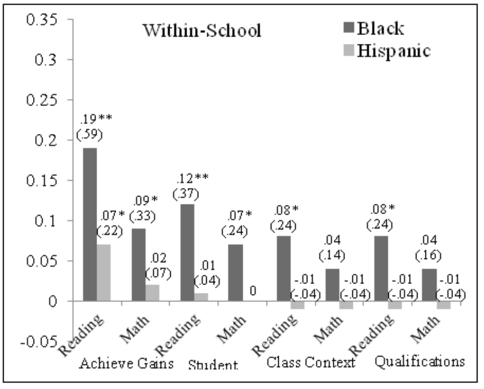
* = $p < 0.05$, ** = $p < 0.01$. The numbers were computed from model residuals and represent Black–White and Hispanic–White achievement gaps that can be attributed to school-based inequity and that accumulated during first grade. The numbers are in units of standard deviations, and (months) of achievement gains. Standard errors, and by extension, the statistical significance levels, were adjusted to account for the cluster sampling design used by NCES.

children, and the math gaps by 27% and by 100% for Black and Hispanic children, respectively. This suggests that students in the same school are not randomly assigned to classrooms, but rather placements are based in part on students' academic and family background including ethnicity. Controlling for student inputs reduced the between-school (Figure 4b) reading gaps by 61% and 44% for Black and Hispanic children, and the between-school math gaps by 45% and 39% for Black and Hispanic children.

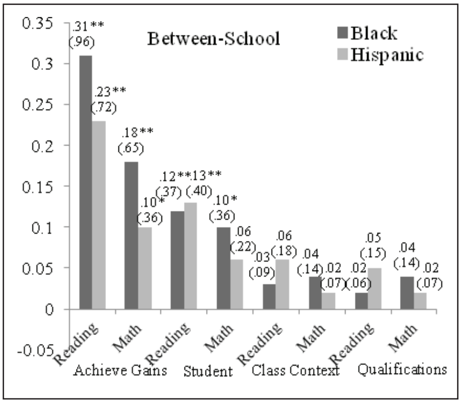
RESEARCH QUESTION 3

The Student model results in Figure 3 show that, even after controlling for an extensive number of student inputs, significant school-based gaps accumulated during first grade. Moreover, the increase is substantially larger for Black children than for Hispanic children. The school-based gaps for Black children increase 0.74 months in reading and 0.60 months in math, compared with 0.47 and 0.22 months for Hispanic children. Each of these increases is statistically significant ($p < 0.01$) with the exception of the math gap for Hispanic children. These increases provide estimates of the

Figure 4a (left). Within-school First Grade Learning Gaps
Figure 4b (right). Between-school First Grade Learning Gaps



Standard errors, and by extension, the statistical significance levels, were adjusted to account for the cluster sampling design used by NCES.



degree to which inequity in school-based opportunity to learn contributes to gaps that accumulate during first grade.

The school-based gaps can be partitioned into classroom effects within and between schools to better understand their origins. A comparison of the length of the bars for the Student model in Figures 4a and 4b indicates that after controlling for differences in student inputs, between-school factors continue to contribute a far greater percentage of the achievement gaps than classroom effects within schools. The one exception is the Black–White reading gap, which is equal parts within- and between-school effect. After controlling for student inputs, the within-school classroom component of the Hispanic–White achievement gaps is not statistically significant ($p > 0.10$), though a substantial within-school effect for the Black–White gaps persists. Between-school effects continue to contribute to reading and math gaps for Black children and the reading gap for Hispanic children, adding 0.22–0.40 months depending on the outcome and ethnic group. However, the between-school Hispanic–White math gap is not statistically significant ($p > 0.10$).

RESEARCH QUESTION 4

Equitable Access to Learning-Conducive Classroom Contexts

The Appendix Table shows pervasive ethnic differences in classroom context that systematically challenges the learning environment of URM children. For example, the mean SES for classrooms that Black children are members of is 0.58 SD lower than for White children (-0.32 vs. 0.26). Black and Hispanic children are also members of classrooms with significantly higher mobility rates, lower mean achievement in reading and math, larger class enrollment, and less adequate instructional supplies. In addition, Hispanic children are members of classrooms with higher percentages of English learners, and Black children are members of classrooms with higher levels of teacher-reported misbehavior.

A comparison of the Student model and Classroom Context model in Figure 3 shows that differences in classroom context account for a large proportion of school-based achievement gaps that accrued during first grade. For example, the figure shows that after controlling for nine contextual factors shown in Table 2, the reading gap for Black children was reduced 55% (from 0.74 to 0.33 months) and the math gap for Hispanic children was erased.

Again, partitioning the school-based gaps into classroom effects within and between schools provides evidence of the origins of the gaps. The change in the gaps from the Student model to the Classroom Context model in Figure 4a indicates classroom context accounted for 35% of the within-school Black–White reading gap (reduced from 0.37 to 0.24 months) and 43% of the math gap, which became non-significant ($p > 0.10$). The within-school Hispanic–White gaps were non-significant after controlling for student inputs, and controlling for classroom context did not change that.

Classroom context accounted for a substantially larger percentage of the between-school gaps (Figure 4b) than the within-school gaps, reducing the Black–White reading gap by 76% (from 0.37 months to 0.09 months), the Black–White math gap by 56%, and the Hispanic–White reading and math gaps by 55% and 73%, respectively. Moreover, after controlling for classroom context, none of the between-school gaps were statistically significant. These findings indicate that inequitable access to learning-conducive classrooms plays a significant role in between-school gaps for both Black and Hispanic children.

Table 2. Teacher Effects Model Results

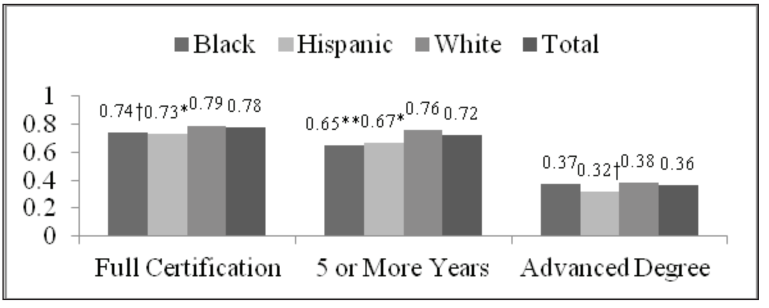
	Classroom Context (3)		Teacher Qualifications (4)	
	Reading	**Math**	**Reading**	**Math**
Intercept	0.06	0.15*	-0.08†	0.12†
Classroom Context				
Student Composition				
Mean SES	0.07*	0.03	0.06*	0.03
Mean reading achievement	0.07	0.09*	0.07	0.09*
Mean math achievement	0.11**	0.08†	0.10**	0.08†
Student heterogeneity				
Reading heterogeneity	-0.08**	-0.09**	-0.08**	-0.09**
Proportion new	-0.15†	-0.13†	-0.16†	-0.13†
Proportion gifted	0.17	-0.57†	0.19	-0.56†
Structures and Resources				
Small	0.05	-0.01	0.05	-0.01
Large	-0.22**	-0.16*	-0.22**	-0.16*
Adequacy of equipment and supplies	0.00	0.03†	0.00	0.03†
Teacher Characteristics				
Full certification	—	—	0.11†	0.03
Five or more years experience	—	—	0.06	0.02
Advanced degree	—	—	0.04	-0.01
Variance Components (% variance explained from prior model)				
	Reading	Math	Reading	Math
Student	0.5745(—)	0.6911(—)	0.5745(—)	0.6909(—)
Classroom	0.0906(6.3)	0.0621(8.8)	0.0891(1.7)	0.0621(0.0)
School	0.0559(21.8)	0.0457(15.5)	0.0542(3.0)	0.0452(1.1)
Total	0.7210(2.9)	0.7989 (1.8)	0.7178(0.4)	0.7982(0.1)
Model Summary				
Deviance Statistic	8512.64	8995.06	8503.37	8994.28
# of parameters	36	36	39	39
Deviance change from prior model	73.78**	58.30**	9.27*	0.78

† $p < .10$; * $p < .05$; ** $p < .01$.

Access to Highly Qualified Teachers

Figure 5 shows the proportions of students from each ethnic group that were served by qualified teachers in terms of full certification, having five or more years of teaching experience, and holding an advanced degree. The disparities are generally minor, 6% or less. One exception is that URM children are approximately 10% less likely than White children to be served by an experienced teacher.

Figure 5. Ethnic Differences in Access to Qualified Teachers



† = p < 0.10, * = p < 0.05, ** = p < 0.01. Numbers are proportions of teachers with the given qualification. Tests are of differences in the proportion for Black or Hispanic students compared with White. Standard errors, and by extension, the statistical significance levels, were adjusted to account for the cluster sampling design used by NCES.

Comparing the results of the Classroom Composition and Teacher Qualifications models in Figure 3 indicates that inequitable access to qualified teachers has little or no impact on the achievement gaps after controlling for student inputs and classroom context. These findings are likely due to the fact that teacher qualifications have weak or null associations with student learning, which the results of the Teacher Qualifications model shows (see Table 2). Note that these findings change very little if the effects of Teacher Qualifications are estimated without the Classroom Context variables in the model. Regardless of the order of entry into the model, Teacher Qualifications have very minor effects on achievement gaps.

Access to Effective Teachers

The Classroom Context model residuals are used to determine the degree to which inequitable access to effective teachers contributes to first-grade achievement gaps. As explained above, ethnic differences in the within-school residuals, shown in Figure 4a, provide an estimate of the lower boundary of the teacher effect, which was estimated at 0.24 months ($p < 0.05$) for the Black–White reading gap, 0.16 months ($p > 0.10$) for the math gap, and approximately zero for the Hispanic–White gap.[11]

The sum of the within- and between-school residuals (Figure 3) arguably provides an estimate of the upper boundary of the teacher effect. Figure 3 indicates a teacher effect of 0.33 months for Black children in both reading and math, but no statistical effect for Hispanic children. These findings suggest that inequitable access to an effective teacher plays a significant role in Black–White achievement gaps that accrue during first grade; however, it has little to no effect on Hispanic–White gaps.

DISCUSSION

This section discusses several aspects of the results, including Black–Hispanic differences in the findings, the effects of classroom heterogeneity and classroom context, the implications of the results to policy and practice, and the limitations of the study.

BLACK–HISPANIC DIFFERENCES

Substantial differences in the magnitudes and origins of the Black–White and Hispanic–White first-grade achievement gaps were noted in this study. At the start of first grade, Black–White gaps were only marginally larger than Hispanic–White gaps (see Figure 2a; 7% larger for reading and 30% larger for math), but the change in the Black–White gaps that accrue during first grade were substantially larger (see Figure 2b, Total; 75% larger for reading and 64% larger for math). Another noteworthy difference is that after controlling for student inputs (see Figure 2b, Student), inequity among classrooms within the same schools continued to contribute to Black–White gaps, but not to Hispanic–White gaps (Figure 4a, Student). This within-school component to the Black–White gap appears to be due to two sources. First, inequity in the contextual characteristics of classrooms significantly contributed to Black–White gaps, but not to Hispanic–White gaps, accounting for 33% of the Black–White within-school reading gap and 43% of the math gap (see Figure 4a; based on the change in the gaps estimated from the Student model to the gaps estimated from the Classroom Context model). Second, Black children tended to be assigned to teachers who were less effective compared with teachers that White children in the same school were assigned to (see Figure 4a, Classroom Context), which increased the reading gap by approximately 0.24 months.

CLASSROOM HETEROGENEITY

The results indicate that classroom heterogeneity tended to undermine student learning, but did not appreciably contribute to achievement gaps. Three measures of classroom heterogeneity—variation in reading achievement, proportion of new students, and proportion of gifted students—were negatively associated with achievement gains (see Table 2, Classroom Context model). Of those, variation in reading achievement had the largest effect. However, as the Appendix Table shows, White children, on average, were members of classrooms with slightly greater heterogeneity in reading achievement (variance = 0.57) compared with Black (0.53) or Hispanic (0.46) children. Hence, controlling for classroom heterogeneity in reading achievement had little effect on the Black–White achievement gaps while narrowing the Hispanic–White gaps slightly. And although the classrooms that Black and Hispanic children were members of had slightly higher proportions of new and gifted students, those factors were only marginally associated with achievement gains and thus also did not appreciably contribute to the achievement gaps.

CLASSROOM COMPOSITION

The student composition of the classroom (i.e., mean SES, mean reading achievement, and mean math achievement) was the type of contextual variable most strongly associated with achievement gaps (see Table 2). The effects of these variables estimate their impact on learning over and above the students' own SES and achievement background. This finding is not surprising because an extensive body of research shows that the composition of the student body affects educational outcomes (Anderson, 2010; Coleman et al., 1964; Jencks & Phillips, 1998; Palardy, 2013; Rumberger & Palardy, 2005). That literature suggests socioeconomic composition impacts educational outcomes primarily through peer effects and instructional practices. That is, during interactions with other children at school, students convey social norms, educational values, and academic skills (e.g., vocabulary) that impact their classmates' behaviors, attitudes, and achievement (Coleman et al., 1966; Hanushek et al., 2003; Jencks & Mayer, 1990; Palardy, 2013). These peer influences tend to undermine student outcomes at schools with high concentrations of low SES students and low-achieving students. Student composition, particularly socioeconomic composition, also tends to be associated with the rigor of instructional practices and curriculum (Betts, Rueben, & Danenberg, 2000; Clotfelter et al., 2002; Stipek, 2004). Some have argued that such differences in instructional practices between high and low SES classrooms can be constructive adaptations that match rigor with students' cognitive development and classroom behaviors (Coleman, 1966; Thrupp, 1999). However, other research indicates that when coupled with appropriate social supports, rigorous instructional practices and curricula can increase achievement in disadvantaged schools (Lee & Smith, 1999).

Within- and Between-School Compositional Effects

Between-school inequity in classroom composition plays a substantially greater role in the development of achievement gaps than within-school inequity in classroom composition (to illustrate, compare Figures 4a and 4b on change from Student model to Classroom Context model). There are two reasons for this. First, classroom composition varies far more between schools than within schools, so there is more potential for between-school ethnic inequity. For example, mean classroom SES varies 17 times more between schools than within schools, and mean classroom reading and math achievement vary 2.1 and 2.6 times more between schools. Second, Black and Hispanic children are significantly more likely than White children to attend schools with compositional characteristics that research has linked to negative peer influences (Greenwald et al., 1996; Hanushek et al., 2003; Jencks & Mayer, 1990), ineffective teaching practices (Baron et al., 1985; Jussim et al., 1996; Stipek, 2004), and academically undemanding curricula (Gamoran, 1992). Table 3 was constructed using ECLS first-grade data to facilitate ethnic comparisons of the distributions of key measures of student composition. The table shows the percent of students from each ethnic group that attended schools in each quintile of four compositional variables, where 1 is the lowest quintile and 5 is the highest. For example, the Black and Hispanic children are more than three times more likely than White children to attend a school in the lowest quintile for mean SES. Conversely, White children are roughly three times more likely than Black and Hispanic children to attend a school in the highest quintile for mean SES. Similar distinct ethnic inequity in the student composition of schools was noted for mean reading and math achievement, and misbehavior rate. The magnitude of these between-school differences in student composition is the primary reason why the between-school effects on achievement gaps are larger than the within-school effects.

Table 3. Ethnic Percentages by Quintile of School Composition Measure

Quintile	Mean SES			Mean Reading			Mean Math			Mean Misbehavior		
	B	H	W	B	H	W	B	H	W	B	H	W
1	36	39	11	33	39	8	43	33	8	12	16	24
2	23	22	17	25	18	19	27	24	16	12	16	23
3	20	14	22	27	19	19	18	20	21	16	24	20
4	12	15	23	7	14	27	8	11	27	25	20	18
5	9	10	27	8	10	27	4	12	28	35	24	15

All estimates are weighted to produce an approximately nationally representative sample of first graders.

IMPLICATIONS FOR POLICY, PRACTICE, AND RESEARCH

Within-School Effects

A major finding of this study is that Black children had less effective teachers than White children in the same schools and were members of classrooms with compositional characteristics less conducive to learning. These within-school sources of inequity contributed to Black–White achievement gaps that accrued during first grade. Moreover, they likely stem from non-random assignment of students to classrooms. Non-random student assignment practices in elementary school have been documented in the research literature (Burns & Mason, 1998, 2002; Praisner, 2003; Rothstein, 2010). This typically occurs when principals purposefully assign students with certain characteristics or special needs to teachers who are believed to be most able or willing to productively work with them (Burns & Mason, 1998). While purposeful assignment may benefit some children, it tends to create imbalances in student composition that can negatively impact average learning in certain classrooms. That these student placement practices appear to systematically disadvantage Black children is troubling. More research is needed to ascertain the prevalence of purposeful student assignment practices, how it impacts classroom composition within schools, and how it potentially creates inequity that contributes to achievement gaps.

Between-School Effects

As described above, Black and Hispanic children are far more likely to attend schools with classroom compositional characteristics that undermine student learning. This inequity was the strongest school-based factor contributing to achievement gaps that accrued during first grade. To the degree that these between-school compositional effects are due to peer influences or to concomitant teaching and school practices, addressing them fully will likely require redistributing children among schools so that schools are more equal in terms of student composition. To that end, a focus on equalizing schools on socioeconomic composition may be the best strategy because previous research has shown that socioeconomic composition is the most robust school predictor of educational outcomes (Coleman et al., 1966; Rumberger & Palardy, 2005), and because it is correlated with other compositional measures. For example, among the 254 schools in the ECLS data, mean SES had a moderate to strong positive correlation with mean reading and math achievement (0.68 and 0.66, respectively) and a moderate negative correlation with the percentage of new students who transfer into the school mid-year (-0.43) and percent URM (-0.45). Successful strategies for redistributing students among schools will have to overcome structural barriers such as socioeconomic segregation in neighborhoods, and district boundaries (Mantil, Perkins, & Aberger, 2012). Many past efforts to redistribute students to reduce compositional differences relied on compulsory transportation plans (i.e., desegregative busing) and were unpopular, costly, and arguably ineffective. More effective strategies may include redrawing district and school boundaries, deliberately locating new schools to maximize compositional diversity, and using school assignment criteria based partially on student SES, the latter of which is already being used in at least 83 school districts in the United States (Kahlenberg, 2012). In the longer term, a concerted effort to increase low-income housing developments in middle class neighborhoods is imperative if neighborhoods and schools are to diversify and resolve this matter.

Teacher Qualifications Versus Teacher Effectiveness

While URM children tend to be served by slightly less qualified teachers in terms of certifications, experience level, and graduate training, qualifications are weak indicators of teacher effectiveness and are only scarcely associated with first-grade achievement gaps. This suggests policies and practices that focus on addressing achievements gaps by assuring equitable access to qualified teachers (e.g., NCLB) will not succeed in closing gaps nor will they have an appreciable impact on average achievement.

In this study, Black children had significantly less effective teachers than White children in terms of their contributions to student learning in both reading and math (see Figure 3, Class Context model). It is worth noting that the magnitudes of those effects were likely underestimated. That is because teachers are not randomly assigned to schools. In fact, research suggests that less effective teachers are more likely to end up at schools with contextual characteristics that are negatively associated with learning (Boyd, Lankford, Loeb, & Wyckoff, 2005; Carroll, Reichardt, Guarino, & Mejia, 2000; Haberman & Rickards, 1990; Johnson et al., 2012). If so, teacher effectiveness and classroom context are expected to be correlated. To the degree that those two factors are correlated, controlling for classroom context may bias downward the magnitude of the effect of access to effective teachers on achievement gaps. This is particularly the case for the between-school effect of access to effective teachers.

Class Size

Another contextual factor that contributed to between-school achievement gaps is class size. Large class size (enrollment > 25) was negatively associated with student achievement gains (Table 2). As the Appendix Table indicates, Black and Hispanic children are 50%–60% more likely to be members of large classes than White children. Moreover, previous research reveals large class size has a stronger negative impact on Black children's achievement than on White children's (Krueger & Whitmore, 2001). Together these findings indicate that the children who are most negatively affected by large classes are also most likely to be placed in them, and this likely contributes to achievement gaps.

Class size reduction reforms began gaining prominence in the 1990s and were federally supported in 2000. By 2010, 35 states had laws limiting the number of students per class (Sparks, 2010). However, several states have recently relaxed or eviscerated class size policies due to budgetary pressures (Kirst, 2011). As a result, over the past two years average class size has rapidly increased in some states. The results of the current study and of previous research suggest that these recent increases in class size will disproportionally affect URM children and lead to increases in achievement gaps. While an argument can be made that class size reduction policies are not cost effective, inequity in class size seems indefensible, particularly if it contributes to achievement gaps. Therefore, state and federal policies that require class size equality are recommended. In addition, new research is needed on the effects that past class size reduction policies had on achievement gaps, as much of the research on class size is based on data collected at least two decades ago, before major reforms were implemented.

LIMITATIONS

This study has limitations that are important to acknowledge. While ECLS-K is an outstanding data source for addressing the research questions, it is observational data, and therefore causal inferences are difficult to ascertain. Although a careful effort was made to minimize selection biases of students and teachers into schools and classrooms when estimating the effects of school-based inequity, it is difficult to verify whether other confounding factors have been omitted. Also, while the results of this study likely generalize to early elementary school grades, they may not be applicable to higher grade levels, particularly middle school and high school where different aspects of school-based inequity may matter. Finally, the current study probes average achievement gaps, but does not test for heterogeneity in the gaps. For example, the Black–White reading gap may differ among low and high SES children, which could have implications for addressing the gaps.

SUMMARY AND CONCLUSIONS

The results of this study provide new evidence on the degree to which inequitable access to learning-conducive classrooms, highly qualified teachers, and effective teachers contribute to achievements gaps that accumulate during first grade. Substantial achievement gaps exist between URM and White children at the start of first grade and they increase significantly during first grade (Bali & Alvarez, 2004; Fryer & Levitt, 2004; Reardon, 2008; Reardon & Galido, 2009). The results show a substantial proportion of the gaps that accumulate during first grade can be attributed to school-based inequalities even after controlling differences in a large number of student background characteristics. In fact, school-based factors account for an approximately three times greater proportion of achievement gaps that develop during first grade than proportion of achievement gains.

While URM children tend to be served by slightly less-qualified teachers in terms of certification, years of experience, and advanced degrees, that inequity contributes only a miniscule amount to achievement gaps. This suggests that current educational policies that focus on equitable access to qualified teachers are misguided, at least for first grade. A far more critical factor is the contextual characteristics of classrooms, primarily the composition of the student body, which systemically disadvantages Black and Hispanic children and is the largest school-based factor that contributes to achievement gaps. Controlling for differences in contextual characteristics of classrooms reduces the school-based achievement gaps by 55%–75%, depending on the subject of the test and the ethnicity group. While these compositional effects mostly operate between schools, there is also a significant within-school effect for Black children, who tend to be assigned to classrooms with more negative contextual factors compared with White children. Moreover, again compared with White children in the same school, Black children tend to be assigned to teachers that are less effective in terms of the average achievement gains of their students. In light of this within-school inequity, a greater awareness of the potentially negative consequences of student assignment practices is needed. However, addressing the between-school inequality, which is primarily due to differences in student composition, will likely require redistributing students among schools so that schools are more equal in terms of student composition.

NOTES

1. In this study *underrepresented minority* refers to Black and Hispanic children. Similarly, *achievement gap* refers to the difference in mean achievement levels of URM and White children.

2. A qualified exception is the research on within-school tracking (for example, see Oakes, 2005). However, this literature does not focus on the classroom unit per se, but rather curricular tracks.

3. The two-stage stratified sampling design used to collect ECLS data involved oversampling certain groups to assure adequate sample sizes for statistical analyses of those groups. NCES developed a number of sampling weights that, when applied to the appropriate subsamples, correct for the oversampling and yield nationally representative samples of 1998 kindergarteners. This study uses the weight for the longitudinal first-grade sample (C3C4cw0).

4. To investigate whether the sample selection criteria introduced sampling bias, the weighted full first-grade longitudinal sample and the weighted sample used in the present study were compared on key variables including the achievement outcomes, SES, and proportion URM. Variable means (or proportions) and standard deviations were highly similar in the two samples, and no statistical differences were found.

5. The duration between fall and spring achievement test administrations varied widely among children in ECLS data. Therefore, accurate estimation of achievement gains necessitated statistically controlling for differences in that duration.

6. Empirical Bayes (EB) classroom and school residuals were used. The EB residuals are less influenced by estimation error than OLS residuals and are "shrunken" or adjusted based on their reliability (see Raudenbush & Bryk, 2002).

7. Note that the standard errors for the ANOVAs were adjusted to account for design effects.

8. While the upper boundary of the teacher effectiveness estimate is more ambiguous to define than the lower boundary, the sum of the classroom and school residuals ($r_{0cs} + u_{00s}$) may be a plausible or even conservative estimate. At the core of the ambiguity is that teachers are not randomly assigned to schools. Indeed, it is reasonable to expect that teachers tend to self-select into the schools based in part on their effectiveness. That is, effective teachers likely have greater opportunity to determine the school they work at. Moreover, recent research indicates that when given a choice

of schools to teach at, teachers tend to select a school with favorable contextual characteristics (Johnson et al., 2012). This results in teacher effectiveness being correlated with the contextual characteristics of the school. Hence, if the model controls for those contextual factors, as is the case in the present study, the magnitude of teacher effectiveness estimates may be biased downward. However, the alternative approach—not controlling for classroom context when estimating teacher effectiveness—would be of even greater concern because it assumes that context does not independently impact student learning. On the contrary, research indicates that the independent effect of context is substantial (Willms, 2010). This is perhaps why the methodology literature on VAM recommends controlling for context when estimating teacher and school effectiveness (Willms & Raudenbush, 1989).

Another source of ambiguity regarding the upper boundary is that some part of the school residuals reflects school effects that are unrelated to teacher effectiveness. For example, the effectiveness of the principal's leadership or per-pupil expenditures may be associated with student achievement gains independent of teacher effectiveness. To the degree that such "true" school factors impact achievement gains in first grade, the sum of the classroom and school residuals may overestimate the teacher effect. However, at the early elementary school level where children spend the vast majority of their school day in a single classroom and are taught by a single teacher, "true" school effects are expected to be small.

9. The multilevel modeling convention is to describe the percentage or proportion of the variance at each level based on the unconditional model. However, differences in student inputs among classrooms and schools will partially account for classroom and school variance, and the classroom and school variance explained by those factors is arguably not a classroom or school effect. Yet controlling for differences in student inputs will also reduce student variation, and therefore the proportion of the variance at each level tends to be somewhat consistent whether or not student inputs are controlled.

10. SD metric in Figure 2a is not comparable with the other figures because the outcome is achievement status, while the outcome in the other figures is achievement gains. However, the month metric is comparable across figures.

11. The square root of the classroom variance component can be interpreted as the standardized regression coefficient or effect size estimate for teacher effectiveness on student *achievement gains* (Nye, Konstantopoulos, & Hedges, 2004). The within-school variance component results for the classroom composition model (see Table 2), which can be considered the lower boundary, indicate effect sizes of 0.30 (i.e.,$\sqrt{0.0906}\sqrt{0.0906}$) and 0.25 for teacher effectiveness on reading and math achievement gains. These estimates are consistent with estimates by Nye et al. (2004) using data from a large-scale random experiment (0.26 and 0.36) and with estimates from their literature review. Combining within- and between-school variance components yields the proposed upper boundary, which suggests effect sizes of 0.38 (i.e., $\sqrt{0.0906+0.0559}$) and 0.33 for reading and math. These estimates are also similar to those of Nye et al. (2004). This consistency of results can be considered validation of the estimation approach used in the current study.

REFERENCES

Abt Associates. (1997). *Prospects: Final report on student outcomes* (Report prepared for U.S. Department of Education). Cambridge, MA: Abt Associates.

Alexander, K. L., Entwisle, D. R., & Olsen, H. R. (2001). Schools, achievement, and inequality: A seasonal perspective. *Educational Evaluation and Policy Analysis, 23*, 171–191.

Anderson, E. (2010). *The imperative of integration.* Princeton, NJ: Princeton University Press.

Bali, V. A., & Alvarez, R. M. (2004). The race gap in student achievement scores: Longitudinal evidence from a racially diverse school district. *The Policy Studies Journal, 32*, 393–415.

Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value added assessment of teachers. *Journal of Educational and Behavioral Statistics, 29*, 37–65.

Baron, R., Tom, D. Y. H., & Cooper, H. M. (1985). Social class, race and teacher expectations. In J. B. Dusek (Ed.), *Teacher expectancies.* Hillsdale, NJ: Lawrence Erlbaum.

Berliner, D. C. (2006). Our impoverished view of educational reform. *Teachers College Record, 108*, 949–995.

Berliner, D. C. (2014). Effects of inequality and poverty vs. teachers and schooling on America's youth**.** *Teachers College Record.* Retrieved on June 26, 2013 from http://www.tcrecord.org

Betts, J. R., Rueben, K. S., & Danenberg, A. (2000). *Equal resources, equal outcomes? The distribution of school resources and student achievement in California.* San Francisco, CA: Public Policy Institute of California.

Betts, J. R., & Shkolnik, J. L. (1999). The behavioral effect of variations in class size: The case of math teachers. *Educational Evaluation and Policy Analysis, 21*, 193–216.

Betts, J. R., Zau, A. C., & Rice, L. A. (2003). *Determinants of student achievement: New evidence from San Diego.* San Francisco, CA: Public Policy Institute of California.

Borman, G. D., & Kimball, S. M. (2005). Teacher quality and educational quality: Do teachers with higher standards-based evaluation ratings close student achievement gaps? *Elementary School Journal, 106*, 3–20.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). Explaining the short careers of high-achieving teachers in schools with low-performing students. *American Economic Review, 95*, 166–171.

Brooks-Gunn, J., Duncan, G. J., & Klebanov, P. (1994). Economic deprivation and early-childhood development. *Child Development, 65*, 296–318.

Burns, R., & Mason, D. (1998). Class formation and composition in elementary schools. *American Educational Research Journal, 35,* 739–772.

Burns, R., & Mason, D. (2002). Class composition and student achievement in elementary schools. *American Educational Research Journal, 39*, 207–233.

Carroll, S. J., Reichardt, R. E., Guarino, C. M., & Mejia, A. (2000). *The distribution of teachers among California's school districts and schools* (MR-1298.0-JIF). Santa Monica, CA: RAND Corporation.

Clotfelter, C. T., Ladd, H. F., Vigdor, J. L., & Wheeler, J. (2002). High poverty schools and the distribution of teachers and principals. *North Carolina Law Review, 85*, 1345–1379.

Coleman, J. S. (1966). Equal schools or equal students? *The Public Interest, 4,* 70–75.

Coleman, J. S., Campbell, E. A., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., & York, R. (1966). *Equality of educational opportunity.* Washington, DC: U.S. Government Printing Office.

Conant, J. B. (1961). *Slums and suburbs: A commentary on schools in metropolitan areas.* New York, NY: McGraw Hill.

Conniff, R. (1998). Bouncing from school to school. *The Progressive, 62*(11), 21–25.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research, 66*, 227–268.

Donovan, M. S., & Cross, C. T. (Eds.). (2002). *Minority students in special and gifted education.* Washington, DC: National Academy Press.

Duncan, G. J., & Magnuson, K. A. (2005). Can family socioeconomic resources account for racial and ethnic test score gaps? *The Future of Children, 15*, 35–54.

Fryer, 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*, 447–464.

Fryer, R. G., & Levitt, S. D. (2006). The black-white test score gap through third grade. *American Law and Economics Review, 8*, 249–281.

Gamoran, A. (1992). Is ability grouping equitable? *Educational Leadership, 50*, 11–17.

Gamoran, A., & Long, D. A. (2006). *Equality of educational opportunity: A 40-year retrospective.* Wisconsin Center for Education Research (Working Paper No. 2006-9). Retrieved on October 18, 2009 from www.wcer.wisc.edu/publications/workingPapers/papers.php

Goodard, R. D., Hoy, W. K., & Hoy, A. W. (2000). Collective teacher efficacy: Its meaning, measure, and impact on student achievement. *American Educational Research Journal, 37,* 479–507.

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effects of school resources on student achievement. *Review of Educational Research, 66*, 361–396.

Grissmer, D., Flanagan, A., & Williamson, S. (1998). Why did Black test scores rise rapidly in the 1970s and 1980s? In C. Jencks & M. Phillips (Eds.), *The Black White test score gap* (pp. 182–228). Washington, DC: The Brookings Institution.

Guarino, C. M., Hamilton, L. S., Lockwood, J. R., & Rathbun, A. H. (2006). *Teacher qualifications, instructional practices, and reading and mathematics gains of kindergartners* (NCES 2006-031). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Haberman, M., & Rickards, W. (1990). Urban teachers who quit: Why they leave and what they do. *Urban Education, 25*, 297–303.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature, 24*, 1141–1177.

Hanushek, E. A. (1989). The impact of differential expenditures on school performance. *Educational Researcher, 18*, 45–62.

Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of Political Economy, 100*, 84–117.

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis, 19*, 141–164.

Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics, 18*, 527–544.

Hanushek, E. A., & Rivkin, S. G. (2006). *School quality and the black-white achievement gap* (NBER Working Paper 12651). Cambridge, MA: National Bureau of Economics Research.

Harris, D. N. (2011). *Value-added measures in education: What every educator needs to know.* Cambridge, MA: Harvard Education Press.

Hedges, L. V., & Nowell, A. (1998). Black-White test score convergence since 1965. In C. Jencks and M. Phillips (Eds.), *The Black White test score gap* (pp. 149–181). Washington, DC: The Brookings Institution.

Jencks, C., & Mayer, S. E. (1990). The social consequences of growing up in a poor neighborhood. In L. Lynn Jr. & M. G. H. McGeary (Eds.), *Inner-city poverty in the United States* (pp. 111–186). Washington, DC: National Academy Press.

Jencks, C., & Phillips, M. (1998). The Black-White test score gap: An introduction. In C.

Jencks and M. Phillips (Eds.), *The Black White test score gap* (pp. 1–51). Washington, DC: The Brookings Institution.

Johnson, S. M., Kraft, M. A., & Papay, J. P. (2012). How context matters in high-need schools: The effects of teachers' working conditions on their professional satisfaction and their students' achievement. *Teachers College Record, 114,* 1–39.

Jussim, L., Eccles, J., & Madon, S. (1996). Social perception, social stereotypes, and teacher expectations: Accuracy and the quest for the powerful self-fulfilling prophecy. *Advances in Experimental Social Psychology, 28,* 281–387.

Kahlenberg, R. D. (2012). Turnaround schools and charter schools that work: Moving beyond separate but equal. In R. D. Kahlenberg (Ed.), *The future of school integration: Socioeconomic diversity as an education reform strategy* (pp. 283–308). Washington, DC: The Century Foundation.

Kirst, M. (2011). California plunges into the unknown in expanding class size. Thoughts on Public Education. Silicon Valley Education Foundation. Retrieved on January 22, 2013 from http://toped.svefoudation.org

Konstantopoulos, S. (2009). Effects of teachers on minority and disadvantaged students' achievement in the early grades. *The Elementary School Journal, 110,* 92–113.

Konstantopoulos, S., & Chung, V. (2011). The persistence of teacher effects in elementary grades. *American Educational Research Journal, 48,* 361–386.

Krueger, A. B., & Whitmore, D. M. (2001). *Would smaller classes help close the Black-White achievement gap?* Princeton, NJ: Princeton University Press.

Ladd, H. (2008). School policies and the test score gap. In K. Magnuson and J. Waldfogel (Eds.), *Steady gains and stalled progress: Inequality and the Black-White test score gap* (pp. 289–319). New York, NY: Russell Sage Foundation.

Lash, A. A., & Kirkpatrick, S. L. (1990). A classroom perspective on student mobility. *The Elementary School Journal, 91*(2), 177–191.

Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity? *Educational Researcher, 31,* 3–12.

Lee, V. E., & Burkam, D. T. (2002). *Inequality at the starting gate.* Washington, DC: Economic Policy Institute.

Lee, V. E., & Smith, J. B. (1999). Social support and achievement for young adolescents in Chicago: The role of school academic press. *American Educational Research Journal, 36,* 907–945.

Magnuson, K., & Waldfogel, J. (2008). *Steady gains and stalled progress: Inequality and the Black-White test score gap.* New York, NY: Russell Sage Foundation.

Mantil, A., Perkins, A. G., & Aberger, S. (2012). The challenge of high-poverty schools: How feasible is socioeconomic integration? In R. D. Kahlenberg (Ed.), *The future of school integration: Socioeconomic diversity as an education reform strategy* (pp. 155–222). Washington, DC: The Century Foundation.

McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics, 29,* 67–101.

Murnane, R. J., Willett, J. B., Bub, K. L., & McCartney, K. (2006). Understanding trends in the Black-White achievement gaps during the first years of school. *Brookings-Wharton Papers on Urban Affairs, 7,* 97–135.

National Center for Educational Statistics. (2002). *User's guide to the longitudinal kindergarten-first grade public-use data file.* Washington, DC: U.S. Department of Education.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26,* 237–257.

Oakes, J. (2005). *Keeping track: How schools structure inequality* (2nd ed.). New Haven, CT: Yale University.

Orfield, G. (2005). The Southern dilemma: Losing Brown, fearing Plessy. In J. C. Boger and G. Orfield (Eds.), *School resegregation: Must the South turn back?* (pp. 1–25). Chapel Hill, NC: University of North Carolina Press.

Palardy, G. J. (2008). Differential school effects among low, middle, and high social class composition schools: A multilevel, multiple group latent growth curve analysis. *School Effectiveness and School Improvement, 19,* 21–49.

Palardy, G. J. (2013). High school socioeconomic segregation and student attainment. *American Educational Research Journal, 50,* 714–754.

Palardy, G. J., & Rumberger, R. W. (2008). Teacher effectiveness in the first grade: The importance of background qualifications, attitudes, and instructional practices for student learning. *Educational Evaluation and Policy Analysis, 30,* 111–140.

Phillips, M., Brooks-Gunn, J., Duncan, G. J., Klebanov, P., & Crane, J. (1998). Family background, parenting practices and the Black-White test score gap. In C. Jencks & M. Phillips (Eds.), *The Black-White test score gap* (pp. 103–145). Washington, DC: Brookings Institution Press.

Praisner, C. (2003). Attitudes of elementary school principals toward the inclusion of students with disabilities. *Exceptional Children, 69,* 135–145.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Reardon, S. F. (2008). *Thirteen ways of looking at the Black-White test score gap* (Working Paper No. 2008-08). Stanford, CA: Stanford University, Institute for Research on Educational Policy and Practice.

Reardon, S. F. (2011). The widening of the socioeconomic status achievement gap: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 91–116). New York, NY: Russell Sage Foundation.

Reardon, S. F., & Galindo, C. (2009). The Hispanic-white achievement gap in math and reading in the elementary grades. *American Educational Research Journal, 46,* 853–891.

Reardon, S. F., & Robinson, J. P. (2008). Patterns and trends in racial/ethnic and socioeconomic achievement gaps. In H. A. Ladd & E. B. Fiske (Eds.), *Handbook of research in education finance and policy.* Mahwah, NJ: Lawrence Erlbaum.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73,* 417–458.

Rosenthal, R. (1994). Interpersonal expectancy effects: A 30-year perspective. *Current Directions in Psychological Science, 3,* 176–179.

Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *Quarterly Journal of Economics, 125,* 175–214.

Rumberger, R. W. (2007). Parsing the data on student achievement in high poverty schools. *North Carolina Law Review, 85,* 101–119.

Rumberger, R. W., & Arellano, B. (2009). Understanding and addressing the California Latino achievement gap in early elementary school. In P. Zavella and R. Gutiérrez (Eds.), *Mexicans in California: Transformations and challenges* (pp. 61–76). Champagne, IL: University of Illinois Press, 2009.

Rumberger, R. W., Larson, K. A., Ream, R. K., & Palardy, G. J. (1999). *The Educational Consequences of Mobility for California Students and Schools.* Berkeley, CA: Policy Analysis for California Education, University of California Berkeley/Stanford University.

Rumberger, R. W., & Palardy, G. J. (2005). Does segregation still matter? The impact of social composition on academic achievement in high school. *Teachers College Record, 107,* 1999–2045.

Sanders, W. L., & Horn, S. P. (1995). The Tennessee Value-Added Assessment System

(TVAAS): Mixed model methodology in educational assessment. *Journal of Personnel Evaluation in Education, 8*, 299–311.

Sanderson, D. R. (2003). Veteran teachers' perspectives on student mobility. *Essays in Education, 4* (Winter). Retrieved on June 6, 2011 from http://www.usca.edu/essays/vol42003/sanderson.pdf

Seltzer, M. H., Frank, K. A., & Bryk, A. S. (1994). The metric matters: The sensitivity of conclusions about growth in student achievement to choice of metric. *Educational Evaluation and Policy Analysis, 16*(1), 41–49.

Sparks, S. D. (2010). Class size shows signs of growing. *Education Week*. Retrieved on January 16, 2013 from www.edweek.org/ew/articles/2010/11/24/13size_ep.h30.html

Stasz, C., & Stecher, B. M. (2000). Teaching mathematics and language arts in reduced size and non-reduced size classrooms. *Educational Evaluation and Policy Analysis, 22*, 313–329.

Stipek, D. (2004). Teaching practices in kindergarten and first grade: Different strokes for different folks. *Early Childhood Research Quarterly, 19*, 548–568.

Swanson, C. B., & Schneider, B. (1999). Students on the move: Residential and educational mobility in America's schools. *Sociology of Education, 72*, 54–67.

Thrupp, M. (1999). *Schools making a difference: Let's be realistic!* Buckingham, England: Open University Press.

U.S. Department of Education. (2006). *26th annual report to Congress on the implementation of the Individuals with Disabilities Education Act, 2004.* Jessup, MD: Ed Pubs.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*, 89–122.

Willms, J. D. (2010). School composition and contextual effects on student outcomes. *Teachers College Record, 112*, 1008–1037.

Willms, J. D., & Raudenbush, S. W. (1989). A longitudinal hierarchical linear model for estimating school effects and their stability. *Journal of Educational Measurement, 26*, 209–232.

Xue, Y., & Meisels, S. J. (2004). Early literacy instruction and learning in kindergarten: Evidence from the early childhood longitudinal study—Kindergarten class of 1998–1999. *American Educational Research Journal, 41*, 191–229.

APPENDIX

Table 1. Variable Descriptive by Ethnic Group, Means and Standard Deviations

Variable Name	Total	Black	Hispanic	White	Description (ECLS variable label)
Student Variables (Level-1, N = 3,496)					
Measurement Control					
Time Adjustment**	6.84 (0.67)	6.74 (0.60)	7.04 (0.64)	6.78 (0.67)	Months between fall and spring testing dates (c4asmt-c3asmt)
Achievement Test Score Outcomes					
Reading Achievement**	0.00 (1.00)	-0.47 (0.97)	-0.33 (1.01)	0.12 (0.96)	Spring first-grade reading achievement (c4rrscal)
Math Achievement**	0.00 (1.00)	-0.53 (0.96)	-0.33 (1.02)	0.21 (0.94)	Spring first-grade math achievement (c4rmscal)
Demographic Characteristics					
SES**	0.00 (1.00)	-0.46 (0.89)	-0.41 (0.91)	0.22 (0.97)	Composite of family income, parents' educational and occupational prestige (w1sesl)
Female	0.49	0.47	0.51	0.49	Recoded to 0 = male, 1 = female (gender)
Asian	0.02	—	—	—	(race = 5)
Black	0.14	—	—	—	(race = 2)
Hispanic	0.16	—	—	—	(race = 3 or 4)
Other	0.05	—	—	—	(race = 6, 7, or 8)
Age*	86.97 (4.21)	86.38 (4.20)	86.92 (4.18)	87.19 (4.20)	Age in months at spring of first grade (R4AGE)
ELL**	0.08	0.01	0.39	0.01	Non-English home language (WKLANGST = 1)
Academic Background					
Fall Reading Ach**	0.00 (1.00)	-0.34 (0.87)	-0.31 (0.91)	0.10 (0.97)	Fall first-grade reading IRT scale score (c3rrscal)
Fall Math Ach**	0.00 (1.00)	-0.50 (0.91)	-0.39 (0.95)	0.19 (0.96)	Fall first-grade math IRT scale score (c3rmscal)
Preschool**	0.44	0.37	0.33	0.50	Preschool was primary pre-K care (P1PRIMPK = 6)
Head Start**	0.07	0.18	0.09	0.04	Head Start was primary pre-K care (P1PRIMPK = 5)

Variable Name	Total	Black	Hispanic	White	Description (ECLS variable label)
Disability†	0.13	0.12	0.08	0.14	Parent reported disability (P1DISABL = 1)
Classroom Behavior					
Interpersonal Skills**	0.00 (1.00)	-0.28 (1.03)	-0.08 (0.97)	0.10 (0.99)	Level of interpersonal skills (T4INTERP)
Internalizing	0.00 (1.00)	0.09 (1.09)	0.01 (0.94)	-0.02 (0.99)	Internalizing problem behaviors (T4INTERN)
Externalizing**	0.00 (1.00)	0.33 (1.21)	-0.02 (0.88)	-0.07 (0.96)	Externalizing problem behaviors (T4EXTERN)
Self-control**	0.00 (1.00)	-0.31 (1.09)	-0.09 (0.95)	0.10 (0.97)	Level of self-control (T4CONTROL)
Approaches to Learning**	0.00 (1.00)	-0.33 (1.06)	-0.06 (0.99)	0.09 (0.96)	6-measure scale on learning engagement behaviors (T4LEARN)
Absenteeism*	7.97 (7.63)	7.81 (6.37)	8.81 (7.45)	7.59 (7.48)	Total days absent during first grade (U4ABSN)
Parental Home Involvement and Beliefs					
Outside Reading †	0.00 (1.00)	-0.06 (1.07)	-0.11 (1.06)	0.06 (0.96)	Frequency of reading outside of school (P4HEQ030)
Practice numbers with child**	0.00 (1.00)	0.13 (0.95)	-0.11 (1.09)	0.01 (0.97)	How often practice numbers (P4RDWRNM)
Reading Ability	0.00 (1.00)	0.09 (0.93)	0.06 (0.98)	-0.05 (1.02)	Parental beliefs in reading ability (P4BELRDG)
Math Ability	0.00 (1.00)	0.04 (0.97)	0.06 (1.04)	-0.03 (1.00)	Parental beliefs in math ability (P4BELMTH)
Ignore	0.00 (1.00)	0.01 (1.06)	0.00 (1.05)	-0.01 (0.98)	Don't listen to child when impatient (P4IMPATN)
Teacher/Classroom Variables (Level-2, N = 877)[a]					
Classroom Context Student Composition					
Mean SES**	0.08 (0.95)	-0.32 (0.83)	-0.23 (0.87)	0.26 (0.94)	Classroom Mean of SES (w1sesl)
High Minority**	0.28	0.67	0.56	0.09	Greater than 40% of students are Latino or African American in class (a4pmin)

Variable Name	Total	Black	Hispanic	White	Description (ECLS variable label)
Mean Math**	0.14 (0.86)	-0.28 (0.78)	-0.23 (0.89)	0.36 (0.79)	Mean classroom math achievement (classroom mean on c3mrscal)
Mean Reading**	0.10 (0.86)	-0.15 (0.81)	-0.18 (0.88)	0.26 (0.81)	Mean classroom reading achievement (classroom mean on c3rrscal)
Misbehavior**	0.03 (1.01)	-0.17 (0.99)	0.14 (0.96)	0.08 (1.01)	Teacher ratings of class misbehavior (A4behvr)
Student Heterogeneity					
Math Heterogeneity	0.54 (0.46)	0.58 (0.47)	0.48 (0.45)	0.54 (0.46)	Classroom variance in math achievement (classroom variance on c3mrscal)
Reading Heterogeneity	0.54 (0.58)	0.53 (0.57)	0.46 (0.54)	0.57 (0.60)	Classroom variance in reading achievement (classroom variance on c3rrscal)
Proportion New**	0.15	0.17	0.20	0.13	% new students (a4new/ a4totag)
Proportion Left†	0.09	0.11	0.11	0.08	Proportion of the students who left the classroom during the school year (a4left/a4totag)
Proportion Gifted	0.02	0.03	0.03	0.02	Proportion of students classified as gifted (a4gift/a4totag)
Proportion Disability†	0.08	0.07	0.06	0.08	% of students with disability (a4disab/a4totag)
Structures and Resources					
Large†	0.12	0.16	0.15	0.10	More than 25 students in class (a4totag > 25)
Small†	0.11	0.10	0.07	0.13	Fewer than 17 students in class (a4totag < 17)
Adequate supplies**	0.04 (0.96)	-0.30 (1.23)	-0.01 (1.00)	0.13 (0.86)	Adequate basic equipment and supplies (FS)
Adequate materials†	0.03 (0.97)	0.02 (0.99)	-0.07 (0.96)	0.07 (0.94)	Adequate instructional materials (FS)
Adequate technology†	0.02 (0.96)	-0.04 (1.11)	-0.06 (0.95)	0.04 (0.92)	Adequate instructional technology (FS)
Adequate facilities	0.00 (0.98)	0.03 (1.11)	-0.03 (0.99)	0.02 (0.93)	Adequate physical facilities (FS)

Variable Name	Total	Black	Hispanic	White	Description (ECLS variable label)
Teacher Qualifications					
Advanced Degree	0.36	0.37	0.32	0.38	Teacher has advanced degree (b4hghstd = 4 or 5)
Experienced Teacher*	0.72	0.65	0.67	0.76	GT five years teaching experience at school (b1yrsc > 5)
Full Certification†	0.78	0.74	0.73	0.79	Regular/advanced certification (b4typcer = 4 or 5)

Note: Continuous variables were standardized to a mean of 0 and standard deviation of 1 unless the units of measurement were inherently meaningful (e.g., age in units of months). Student variables were standardized at the student level and classroom variables were standardized at the classroom level.[a] The classroom sample size is 877; however, the means and standard deviations for classroom variables on this table have been disaggregated to the student unit of analysis for the purpose of showing the distributions by ethnic group. All variables on the table were weighted using the student weight, C34CW0. FS = factor score. Standard errors, and by extension statistical significance levels, were adjusted to account for the cluster sampling design used by NCES.

GREGORY PALARDY is Assistant Professor of Education at the University of California, Riverside. His research centers on teacher and school effectiveness, educational equity, and quantitative methods for modeling those topics. His recent publications include: Palardy (2013), High school socioeconomic segregation and student attainment, *American Educational Research Journal*; Palardy & Vermunt (2010), Multilevel growth mixture models for classifying groups, *Journal of Educational and Behavioral Statistics*; and Palardy & Rumberger (2008), Teacher effectiveness in the first grade: The importance of background qualifications, attitudes, and instructional practices for student learning, *Educational Evaluation and Policy Analysis*.