The Race of a Teacher and Differential Student Achievement: Evidence from Project STAR

Issa Abdulcadir

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Arts

University of Washington

2015

Committee:

Charles Hirschman
Aimée Dechter

Program Authorized to Offer Degree:
Sociology

University of Washington

Abstract

The Race of a Teacher and Differential Student Achievement: Evidence from Project STAR

Issa Abdulcadir

Chair of Supervisory Committee:
Professor Charles Hirschman
Department of Sociology

This study explores the impact of black teachers on the achievement of black and white students in first, second, and third grade. Using data from Project STAR, I exploit the random assignment of students and teachers within schools to classes in order to estimate the effect of a teacher's race on student reading and mathematics performance. Estimates from linear regressions using maximum likelihood estimation indicate that black students perform better when taught by black teachers relative to white teachers, net of family, neighborhood, teacher, and school characteristics.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Number Page

LIST OF TABLES

Table Number Page

I. Introduction

A consequence of the Brown v. Board ruling was the start of a pervasive dismantling of black schools as black students began filtering into a newly "integrated" school system. The integration process was slow, as 98% of black children in the south still remained in all-black schools a full decade after the ruling (Jaynes & Williams 1989). With the passage of the Civil Rights Act of 1964 and the Elementary and Secondary Education Act of 1965, the integration process accelerated as these acts enforced the Brown decision under threat of economic and legal sanctions. As integration gathered speed, a consequence was the massive reduction in the numbers of black educators. Since black schools were now under the control of all-white school boards, black teachers and principals were dismissed, demoted, reassigned, denied new contracts, or denied pay equal to that of their white counterparts (Tillman 2009; Fultz 2004; Spruill 1960). Of the approximately 82,000 black teachers teaching in all-black schools across the nation (Rosenthal 1957), there were at least 39,386 documented cases of black teachers losing their jobs between 1954 and 1972 (Ethridge 1979). Since this secular shift in the nature of schooling for black students has largely been remembered as an unerring improvement in the educational and social lives of black people, few scholars consider the potentially harmful implications of the decline in the number of black students that have the opportunity to be taught by black teachers, as white students continue to be taught by white teachers.

During this period of social change in which desegregation severely reduced the numbers of black educators, academic interest in educational inequality gathered steam. Of the early investigations of educational inequality in the United States, none have proved more substantively and methodologically influential than James Coleman and his colleagues' 1966 report, *Equality of Educational Opportunity* commonly referred to as the Coleman Report.

Although this report and the many investigations that followed have offered clear evidence of educational and consequent social inequalities, the conversation has largely focused on what resources and opportunities black students lack, and less on those resources and opportunities that schools and teachers do not provide. The focus of this study lies within an organizational function under the control of schools: teacher hiring; and interrogates if the current paradigm in which schools are dominated by white teachers contributes to educational inequality. Specifically, this study advances a relatively understudied stratifying mechanism, the race of a teacher, by questioning whether student achievement gains are associated with having a teacher of the same race. If there is a benefit to having a teacher of the same race, then white students, a population overwhelmingly taught by white teachers, are advantaged within schools as black students are largely disadvantaged by the absence of black educators.

Previous work on the impact of the race of a teacher on student educational outcomes suggests that black teachers are less likely to harbor racial biases in their perceptions and evaluations of student behavior, effort, and ability, and hold higher expectations for black students than white teachers. In addition, black teachers are less likely to be racially biased in their expressions of praise and criticism for students as compared to white teachers. Although previous work suggests that there exist differences between black and white teachers that likely disadvantage black students, empirical evidence on the direct impact of the race of a teacher on student performance is subject to a number of issues.

The lack of consensus and empirical support regarding what teacher characteristics improve student performance, due, in part, to the lack of comprehensive measures of teacher characteristics as well as the selection of students and teachers to certain kinds of schools based on race, makes investigations of the impact of a teacher's race on student achievement a difficult

enterprise. In particular, black teachers often select into schools where minority students are concentrated. Often these schools have fewer resources and are more likely to promote lower levels of achievement than schools where white students are disproportionately taught. Overall, students and teachers in these schools with a high minority student composition, are both subject to the same difficult circumstances and the same absence of resources. As a result, in the non-experimental data that is used to study the impact of the race of a teacher, these confounding influences bias estimates of the impact of teachers' race on student performance. Beyond the differential sorting of teachers and students to these schools, within these schools, teachers and students are subject to organizational features and practices that disrupt teaching and learning, among these are crowded classrooms, classrooms composed of high proportions of low-income students or students with learning disabilities, and student tracking and ability grouping, all of which have been shown to disadvantage black students. These within-school conditions further confound estimates of the effects of teacher characteristics on student performance. Given the selection of teachers and students to schools, I estimate within school effects, and given within school practices that disadvantage students and teachers, I use a unique experimental data set that eliminates the estimation issues that result from the unequal treatment of students and teachers within schools by race. These data accomplish this through the randomization of students and teachers to classes.

This study employs data from Tennessee's Project STAR (Student-Teacher Achievement Ratio) one of the largest experiments of class size, wherein teachers and students were randomly assigned within schools to different class sizes. The randomization of both students and teachers to classes within schools eliminates an important source of student and teacher heterogeneity resulting from an unequal allocation of resources and differential treatment by race. This within

school level of heterogeneity cannot be measured completely, and would otherwise confound estimates of the effects of teacher race on student performance.

I estimate linear regressions of student reading and math test scores on the race of a teacher and student, and include covariates for other student and teacher characteristics, classroom size, and a metropolitan-rural classification of school location that distinguishes metropolitan areas according to their concentration of poverty. This model does not capture other school level characteristics. I compare estimates from Maximum Likelihood Estimation (MLE) models to within school estimates from school level fixed effects models. Within school estimates eliminate the influence of unmeasured school characteristics that vary across students, and may be correlated with teacher and student race because of residential segregation and other sources of inequality and discrimination

This study is an improvement over previous studies given the data, which allow for improved estimates of the influence of teacher characteristics on student performance within schools, and methods, which include model specifications that control for important school level differences that drive differential racial achievement. Together, these data and methods allow for an investigation into the proximate learning environments of students, the classrooms where students learn skills necessary for current and future achievement, and the institutional agents most responsible for learning, teachers. In addition, where most studies of differential achievement investigate individual characteristics that lead to human and social capital deficits that are largely embedded in a society where upward mobility is constrained by race, my focus is on a feature of schools, the teachers therein, that is far more amenable to alteration than the systemic dismantling required to address the root causes of achievement gaps that other scholars advance. Further, my focus on the earliest and most important years of schooling, years that

influence later educational and social outcomes, is an improvement over other studies that look at latter educational outcomes without addressing early schooling outcomes that are likely to act as antecedents. Finally, this investigation contests popular notions of the infallible benefits of education and integration by focusing on the institutionalized practice of these often disembodied concepts, practices that are as likely to reduce racial inequality as they are to widen it.

In Section II of this paper, I demonstrate that there are black-white achievement gaps as measured by test scores. In Section III, I demonstrate that prevalent explanations of the racial achievement gap that focus on individual, family, and neighborhood characteristics which only account for a portion of racial achievement differences, are embedded in a racialized social structure that is difficult to change. In Section IV, I argue that not enough is known about why the race of a teacher may affect student performance. Since the racial composition of teachers in schools is more easily changed than the upheaval of the current social structure, knowing more about the influence of a teacher's race offers a path toward reducing educational inequality.

The results of this study suggest that black teachers boost the performance of black students relative to white teachers, net of student and family background characteristics, and teacher quality and school characteristics. The findings of this study also reveal that educational inequality is largest in classrooms under the instruction of white teachers, but that a policy recommendation that increases the recruitment of black teachers offers only a partial solution to the reduction of achievement gaps given the nature of school segregation. The clear racial segregation present in the STAR sample, makes it problematic to offer a definitive account of the influence of black teachers on white students, but an attempt is made in this study to address the apparent negative effects of black teachers on white student performance as well.

II. The Achievement Gap

Trends in the Black-White Test Score Gap

In the large body of literature on racial test score gaps, researchers have largely been concerned with understanding the trends and major determinants of black-white achievement differences. According to data from the National Assessment of Educational Progress Long Term Trend Assessment (NAEP-LTT), a national testing series regularly administered to a nationally representative sample of nine, thirteen, and seventeen year olds since 1971 a persistent black-white test score gap has always existed in the United States, but the size of the black-white test score gap in reading and mathematics has fluctuated in the last 40 years. Using these data, I display these patterns in Figures 1a and b.

Figure 1a. National Trends in the Black-White Reading Test Score Gap for Students of Different Ages.

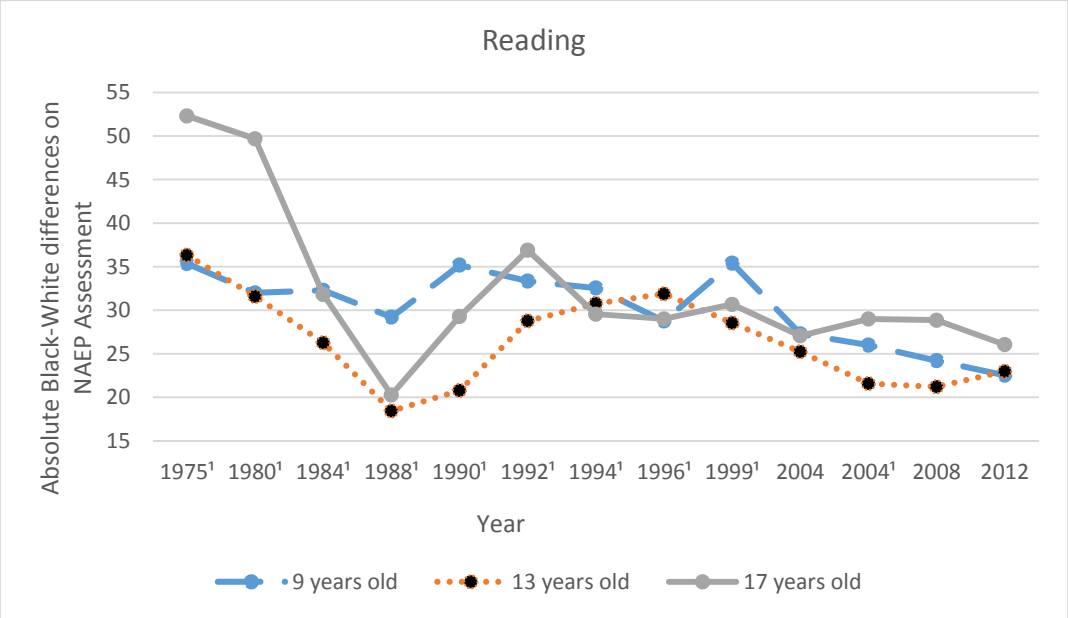
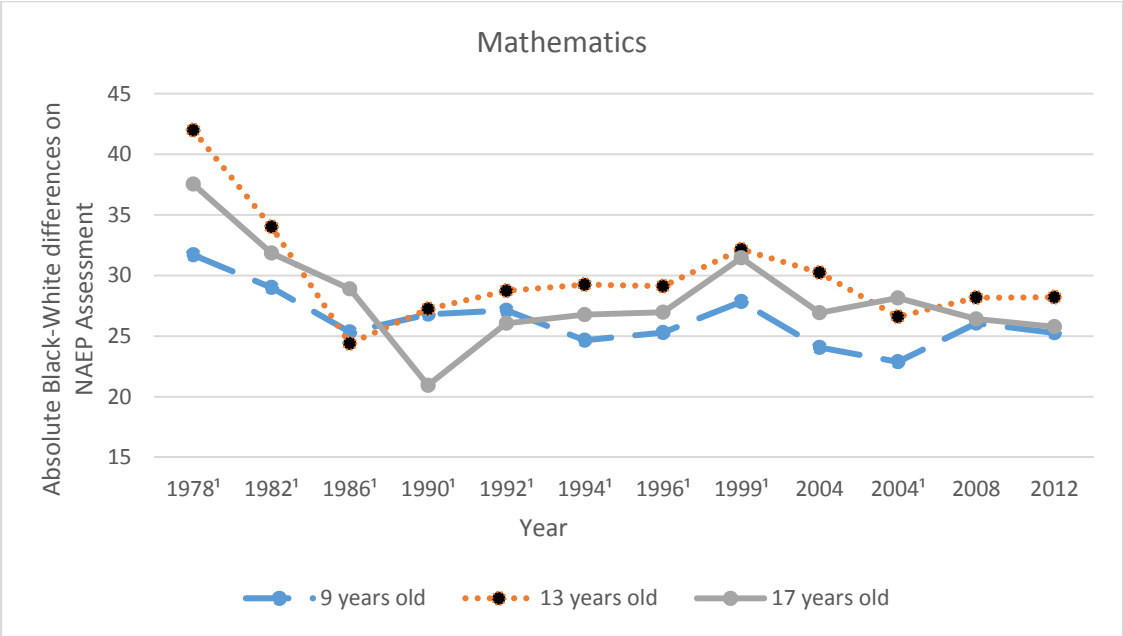


Figure 1b. National Trends in the Black-White Mathematics Test Score Gap for Students of Different Ages.



Notes for Figures 1a and 1b: ¹ Original assessment format.

a) Race categories exclude Hispanic origin.
b) The NAEP Long-Term Trend Reading and Mathematics scales range from 0 to 500.
c) Some apparent differences between estimates may not be statistically significant.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1978-2012 Long-Term Trend Reading and Mathematics Assessments.

Between 1971 and 1988, the reading gap narrowed for each age group, and most substantially for thirteen and seventeen year olds, and between 1978 and 1986 the math gap narrowed among all age groups, with a continued reduction in the gap until 1990 for seventeen year olds. Through the 1990s the test score gap in mathematics widened, depressed in the early 2000s and flattened since 2004

Given these trends, researchers have investigated the reasons behind the reduction of the gap during the 1970s and 1980s, its widening in the 1990s, and its relative stability in the last decade. Although conventional wisdom suggests that the conditions that served to create gaps among younger students would serve to exacerbate gaps as students progress through grade

levels, these NAEP show inconsistent fluctuations in the gap across age groups, but overall they reflect a persistent gap between black and white students.

Test Score Gaps in the Earliest Years of Schooling

Studies of the black-white test score gap have revealed two important findings that are still debated among educational researchers: first, that black students begin school with test scores that are, on average, lower than that of their white peers, and second, that these learning deficits expand during the school year. These findings indicate that black students begin school with less of the requisite skills needed to successfully navigate elementary school and beyond, and implicate schools in the growth of these differences.

Previous research has indicated a number ways that that schools expand learning gaps. School segregation by race and socioeconomic status significantly disadvantages students as schools where minorities and low-income students are overrepresented are often without important resources that white students and wealthier students are more likely to enjoy (Condron & Roscigno 2003). Low quality schools are likely to expand gaps over time as students do not develop important basic skills that are required as they progress through subsequent grade levels, and these students are more likely to be taught by less experienced teachers than those in higher quality schools (Clotfelter et al. 2005). In addition to the schools that students attend, within-class ability grouping and curricular track placements are likely to disadvantage students as well (Lucas 2001). Also, the learning environment at home affects how much of the lessons from the school year are retained during the summer time, as a larger share of black students and low income students have been shown to accrue larger learning setbacks over the summer than their white and wealthier peers (Entwistle & Alexander 1992). As these deficits accumulate over the

course of a students' educational tenure, racial and socioeconomic achievement gaps are thought to expand, although the picture is unclear in the earliest years of schooling.

Using data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), a nationally representative sample of 1000 public and private schools from which a random sample of 25 kindergarteners in each school were drawn, Lee and Burkam (2002) found that black students began kindergarten .62 standard deviations behind white students in mathematics, and .4 standard deviations behind their white peers in reading.

These findings are mirrored by Fryer and Levitt (2004) in their analysis of the ECLS-K in which they found a .638 standard deviation black-white math score gap that increases to .728 standard deviations by the spring of their first grade year. Similarly, the black-white reading score gap begins at .401 standard deviations in kindergarten and increases to .529 standard deviations by the end of first grade. Even when a host of covariates[1] are included, black students appear to fall behind whites by .107 standard deviations in math during the kindergarten year, and by the spring of their first grade year, they fall behind an additional .49 standard deviations. Similarly, black students fall behind their white peers in reading scores by .126 standard deviations during the kindergarten year, and lose an additional .062 standard deviations by the end of the first grade year.

In a follow up study that analyzes the black-white test score gap from kindergarten through students' third grade year, Fryer and Levitt (2006) found that the black white test score gap continues to grow at a rate of .10 standard deviations during the school year for the first four

[1] Covariates included in Fryer and Levitt 2004: baseline child characteristics including gender and age of entrance into kindergarten; geography and urbanicity, family composition, parental education and income, child's early home environment including measures of birth weight, books at home, language spoken at home; type of childcare, neighborhood characteristics, parental involvement in education, and the duration of their kindergarten program. In addition they found that a parsimonious set of family background covariates were able to explain as much as this nearly exhaustive set, and further, this parsimonious set of covariates was able to fully explain the black-white gap in reading, and more than 80% of the gap in mathematics, with a noted caveat that these same covariates are only able to explain a small portion of differences on a test of general knowledge.

years of schooling, net of school, family, neighborhood characteristics. If the trends in these losses for black students continue, by their ninth grade year, black students will be a full standard deviation behind their white peers in math and nearly two-thirds of a standard deviation behind in reading, net of standard socioeconomic, home, family, and school characteristics. These trends are consistent with other studies of black-white achievement gaps (Phillips, 2000; Phillips, Crouse, and Ralph, 1998). These findings indicate that learning gaps are present at the outset of schooling, with black students beginning school at a greater learning deficit than their white peers, and schools swell this achievement gulf.

In their reanalysis of the work of Fryer and Levitt (2004, 2006), Murnane et al. (2006) use data from ECLS-K and longitudinal data from the National Institute of Child Health and Human Development (NICHD) and find, in contrast to Fryer and Levitt's study, that a substantial black-white test score gap in mathematics and English Language Arts (ELA) skills is present at the beginning of kindergarten even after controlling for the same set of background characteristics as Fryer and Levitt. Murnane and colleagues indicate that differences between the tests administered in the ECLS-K, which narrowly focus on reading and math skills learned in school, and in the NICHD, which measure a broader range of skills, are the most plausible explanation for the difference in findings. This persistent gap at the beginning of kindergarten is also supported in other studies using other data as well (Rouse et al. 2005). In contrast to Fryer and Levitt, Murname and colleagues found that the black-white math score gap declines during the first four years of school, while the ELA score gap does grow, as Fryer and Levitt found, but at a slower rate than ECLS-K data predicts. The authors suggest that the observed reduction of black-white math score gap in the NICHD sample as compared to the ECLS-K sample given similar covariates is an artifact of the difference in assessment tests of the two data sets. In the

NICHD sample, the Woodcock-Johnson test measures skills that are somewhat different than those in the ECLS-K sample. In addition, the measures of the gaps used by Fryer and Levitt, standard deviations from the from the test score distribution at a particular age, produce different patterns than the vertically equitable IRT test scale metric used by Murname et al.

While there is disagreement among these studies about the growth of test score gaps during the school year, there is evidence to support the claim that gaps grow as students progress through early years of schooling, as well as evidence to support that black-white achievement gaps are present at the beginning of schooling. This study departs from most studies of the within school determinants of black-white achievement gaps by advancing an altogether understudied mechanism that facilitates differential racial achievement, the race of a teacher.

III. What are some proposed causes of the Black-White test score gaps?

Most explanations of gaps in achievement focus on individual, family, or neighborhood characteristics in their influence on the black-white achievement gap. All of these characteristics are a part of a social structure that is difficult to disrupt, much less dismantle. This is most apparent through the failure of interventions that attempt to supplement household incomes for the disadvantaged (Morris et al. 2001), those aimed at preventing teen childbearing and single parenthood (Kane & Sawhill 2003), those that promote parental education (McGroder et al. 2000), and those that attempted to address the negative effects of neighborhood poverty like the Moving to Opportunity program (Orr et al. 2003). Many of these interventions proved ineffective or too costly to implement on a large scale. Further, this myopic focus on student, family, and neighborhood characteristics has rarely demonstrated that these characteristics are able to fully

explain racial achievement differences; often they explain only a portion of the achievement gap. In the following subsection, I briefly review studies from this mold.

The Influence of Student, Family, and Neighborhoods on Student Achievement

The negative influence of poverty, and differences in family, neighborhood, and school characteristics have long been cited as causes of achievement differences between white and black children. Given the strong correlation between poverty and race, and the many and varied ways that poverty affects the family, neighborhood, and school resources linked with academic success, studies of black-white test score gaps have revealed some important ways that persistent family and neighborhood poverty forecast lower test scores for black students.

Previous studies indicate that children living in poor households are more likely to be raised by single parents, have parents with lower educational attainment and lower occupational status, and have parents that are more likely to be unemployed (McLoyd & Wilson, 1990; Wilson, 1987; Garfinkel & McLanahan, 1986; Featherman & Hauser, 1978). Each of these family characteristics have been found to lower the test scores of children from poor families relative to their more privileged peers (Duncan et al., 1994; McLoyd, 1990), and given the much larger proportion of black children living under the poverty line (38% of black compared to only 10.1% of white children[2]) these conditions are more likely to impose themselves on the educational careers of black students than whites.

Researchers have demonstrated that children raised in single-parent families are at greater risk of academic and behavioral problems than children raised in two-parent households (Carlson & Corcoran, 2001; McLanahan, 1997), that these families are at greater risk of residential

[2] U.S. Census Bureau, Current Population Survey, 2014 Annual Social and Economic Supplement

mobility which is associated with poorer academic performance (Pribesh & Downey, 1999; Rumberger & Larson, 1998; McLanahan & Sandefur, 1994). In addition, dual-parent households are associated with higher family incomes, which are important for educational investments, and studies have revealed that single-parenthood has a persistent negative effect on adolescent outcomes, net of family income levels (Sandefur et al., 1992; Hauser & Featherman, 1977). Among the many family characteristics that disadvantage students, the negative effects of living in single-parent households are more likely to be felt by black students than their white counterparts, as only 34 percent of black children were living with two parents, compared with 75 percent of white children in 2014.[3]

The consensus among educational researchers who study the impact of family income and parental education on student achievement is that lower levels of each negatively impact the test scores of children (Duncan & Brooks-Gunn, 1997; Danziger & Danziger, 1995). This negative impact on achievement is facilitated by a number of mechanisms, including class-based differences in extracurricular and educational investments of parents in children (Lareau 2002; Ainsworth 2002) and class-based differences in children's orientations toward schooling (Lareau, 2002; Alexander et al., 1997; Rumberger et al., 1990). In addition, parental human capital coupled with the inflexibility of low-status jobs, prevents low-income parents from performing the roles that schools expect of them i.e. attending school meetings, tutoring children, checking their homework, and otherwise intervening in school matters and advocating for their children (Lareau 2002).

Beyond these resources and practices that families provide, neighborhood poverty has been shown to lower student achievement levels (Turley, 2003; Menacker, 1990), although very

[3] Ibid.

young children are likely less susceptible to negative impact of a poor neighborhood than adolescents (Brooks-Gunn et al., 1997). Neighborhoods in poverty are more likely to augur poor student achievement as a result of neighborhood stressors like violence, drugs, or by poor institutional resources like low-quality schools and negative police presence, or by negative peer influences and a lack of positive role models, yet empirical studies of the influence of neighborhood characteristics on student achievement claim that neighborhoods explain only a small portion of achievement variation (Leventhal & Brooks-Gunn, 2000).

Studies of the achievement gap, like that of Fryer and Levitt mentioned above have demonstrated that after kindergarten, socioeconomic factors and other observable characteristics fail to fully explain racial achievement differences. In addition, other studies of the achievement gap have found that racial differences in family characteristics like parental education and household income only explain around a third of the black-white test score gap (Hedges & Nowell 1999; Phillips et al. 1998).

IV. The Race of a Teacher and Student Achievement

While it is clear that black students face more barriers to academic preparedness at the outset of schooling and are more likely to contend with environments outside of school that are less supportive of learning that happens in the classroom, what is less clear is how practices within schools, particularly those associated with teacher characteristics contribute to racial achievement inequalities. Commonly known within-school stratifying processes like school tracking and classroom ability grouping are, at least in part, a consequence of teacher recommendations and assessment practices, and these practices are a consequence of the perceptions and expectations that teachers have for their students. Previous studies of teacher

race have demonstrated that the race of a teacher is likely to affect these perceptions and expectations. In particular, for black students, having a black teacher may benefit the academic performance of black students as black teachers are likely to act as role models for students, to create and maintain a better rapport with students by virtue of shared experiences, to be less likely to discriminate on the basis of stereotypes, to hold higher expectations and perceptions of students and their academic promise, and to offer more positive attention to black students that benefits learning perhaps in the form of praise or even extra-curricular investment in student learning.

Role modeling

In her article outlining the impact of the declining numbers of black teachers, Cole (1986) asserts that black students and students from poor families are less likely to have positive, black role models, a sentiment echoed in studies outside of education (Anderson 1999; Duneier 1999). As role models, black teachers are thought to inspire student effort, self-belief, and enthusiasm (Clewell & Villegas, 1998; King, 1993; Waters, 1989; Cole, 1986) and reduce feelings of alienation (Graham 1987). The proposed effects of black role models make for a particularly interesting solution to black-white achievement differences: if role models boost student achievement, and if black students are more sensitive to the impact of role models given the lack of role models in their neighborhoods, then the introduction of black role models in schools will likely boost achievement gains for black students more than for whites, and thereby reduce the black-white achievement gap. In addition, black teachers will likely benefit white students as well since exposure to a diverse teaching force will likely prepare them to live in a multiracial society, and dispel negative beliefs about the abilities and capacities of members of other racial groups (Waters 1989; Irvine 1988). On the whole, it seems that the re-introduction of black

teachers to schools where they have been disproportionately absent, benefits both black and white students.

Student-Teacher Race Matching and Test Scores

Within the broader literature of the impact of minority teachers on minority educational outcomes, findings suggest that black students benefit from the presence of black teachers on a number of outcomes and on a number different levels of analysis. For example, Farkas et al. (1990) found that black students had lower instances of absenteeism when taught by black teachers, compared to white teachers. At the school level, Ehrenberg et al. (2005) found that increases in the percentage of black teachers in schools produced test score gains for black students even if black students were not paired with these teachers in classrooms. This finding is likely representative of a feature of school or school district practice or culture that is associated with the hiring of black teachers that additionally benefits black student performance. In addition, at the district level, Hess and Leal (1997) found that higher concentrations of non-white teachers in large urban school districts significantly increased overall college matriculation rates among students. Together these findings suggest that black teachers offer important educational benefits for black students.

In his investigation of the impact of the race of a teacher on student test scores, Dee (2004) employs data from the Tennessee STAR experiment, as does this study, to investigate the impact of racial pairings of students and teachers on reading and math test scores. He finds that a one year assignment to an own-race teacher significantly increases math and reading scores of both black and white children net of teacher, school, and family background characteristics, and he notes that this finding is robust to different model specifications that account for school and classroom differences. He also notes that subsequent years of exposure offer modest, but clearly

additive increases in student performance. By pooling observations of students from kindergarten through third grade, a departure from my own use of these data, he increases the precision of estimates of the impact of student-teacher racial pairings, but accomplishes this at the cost of an analysis that reveals the impact of student-teacher race matching in each of the three compulsory grades of schooling in the experiment: first, second, and third grade. In addition, the comparisons he makes between black and white students with white teachers are not as direct as he asserts given the racial segregation of black and white students in different neighborhoods and different schools with varying levels of poverty. The experimental data used in this analysis and that of Dee offer a major benefit over those of previous studies, benefits which will be discussed in the data section of this paper. The key finding of Dee's investigation is that race matching increases student test scores in the earliest years of schooling, something that may, if the benefits are greater for blacks than whites, reduce black-white test score gaps.

In another study of students at the beginning of their educational careers, Hanushek (1992) examined the impact of the race of a teacher on the performance of students in second through six grade using data from the Gary Income Maintenance Experiment, one of four federally funded negative income tax experiments from the first half of the 1970s aimed at testing the impact of an alternative to the welfare system for poor black families in Indiana. Approximately 59% of the over 1700 black families in the sample were headed by single mothers. He found that black teachers significantly increased the vocabulary and reading scores of black students, while white teachers were ineffective in this aim. This study reinforces the idea that for black students, and particularly poor black students from single-parent families, black teachers improve achievement outcomes. Although the other does not 'unpack' the finding

of less effective white teachers, it may be possible that their ineffectiveness is the result of lower-quality white teachers being attracted to schools that are primarily attended by black students.

Overall, these studies suggest that there are important test-score gains made by black students as a result of their instruction by black teachers.

Racial biases in perception and expectations of teachers for students

An alternative approach to studying the direct, positive impact of student–teacher race matching, is to address how having a teacher of the same race mitigates the negative impact of having a teacher of a different race. Specifically, I review studies that advance different mechanisms through which student-teacher race matching improves student achievement, particularly for black students. The two main mechanisms whereby the race of a teacher impacts educational outcomes in these studies are perceptions and expectations.

The perceptions and expectations that teachers have for students have been shown to influence year-end achievement (McKown & Weinstein, 2007; Kuklinski & Weinstein, 2001). Not only are teacher expectations important to student achievement, but previous studies have found that black students and students with low socioeconomic statuses are more sensitive to the perceptions and expectations of their teachers (McKown & Weinstein, 2002; Jussim et al., 1996; Alexander et al., 1987). Unfortunately, studies of the intersection of race and teachers' perceptions and expectations indicate that overall, white teachers have lower perceptions of black students' academic performance than their white peers even after accounting for prior student performance (Oates 2009).

In his review of empirical evidence of the influence of teachers' perception and expectations on the black-white test score gap, Ferguson (1998) found evidence that teachers held lower expectations for black students than white students, and posited these differences in

expectation as determinants of racial achievement gaps. These findings of racial differences in student expectations are also counterbalanced by opposing findings, with naturalistic studies finding that once students' previous achievement and other student characteristics are considered, teachers are unbiased in their expectations of black and white students, as their perceptions of student effort were often accurate when compared with student self-reports (Madon et al. 1998). While many studies have been unable to convincingly link teacher expectations directly to student achievement, others have noted that expectations are particularly consequential to within classroom tracking in the form of ability grouping, with black children more likely to be assigned to lower ability groups (Condron, 2007; Hallinan, 1996).

In his observational study of a kindergarten classroom, Rist (1970) found that teachers organized their classrooms according to their expectations of student success, expectations that were formed as early as the eighth day of school and dictated the differential treatment of students for the remainder of the year. Teachers labelled students as fast learners and offered them preferential treatment and preferential workgroups, based on their own "normative reference groups" which are informed by their own ascribed characteristics and childhood experiences, of which race and socioeconomic status are particularly salient. For those students that they considered as lower status, primarily black students and students from poor families, teachers offered less interaction time, less favorable seating arrangements, and were more often hostile to these children, and these apparent hostilities were mirrored by students toward their low-status peers.

Together, these studies indicate that having a teacher of the same race substantially benefits students, as the practices that result from the racially-biased perceptions and expectations of teachers are likely to increase black-white learning differentials. Beyond biased

expectations and perceptions that inform racially biased practices in schools, I briefly review two other mechanisms through which the race of a teacher may affect student outcomes.

Racial biases in subjective and behavioral evaluations of students

Previous studies of the race of a teacher on subjective ratings of students indicate that black students are likely to receive poorer evaluations from white teachers than black teachers (Bates & Glick, 2013; Downey & Pribesh, 2004; Ehrenberg et al., 1995; Farkas et al., 1990; Alexander et al., 1987; Coates, 1972), while the evaluations of white students by black teachers are no different than those of white teachers (Eaves 1975). These negative evaluations are likely to inform decisions about student behavior and ability which may be consequential in decisions about student progress, tracking, ability grouping, or even how students should be disciplined.

Studies of racial biases in the evaluations of elementary school students by teachers also reveal that black students are more likely to be rated as exhibiting poorer behavior than white students. In their analysis of a sample of kindergartners from the ECLSK, Downey & Pribesh (2004) find that black kindergarteners are more likely to be rated as having poor behavior when evaluated by white rather than black teachers. Similarly, Alexander et al. (1987) found that black first graders were more likely to be considered immature when evaluated by white teachers or teachers with high socioeconomic statuses rather than black teachers, although this study did not control for prior student achievement, nor current student performance.

These evaluations are also likely to reflect teachers' interactions with students as well, with teachers that evaluate students poorly, more likely to treat these students poorly. In addition, these evaluations may reflect important components of student-teacher interaction like racially motivated biases in the praise and criticism offered to students by teachers.

Racial biases in praise and criticism

In a study using qualitative data from 16 first-grade teachers, 8 black and 8 white, in an urban public school system, Simpson and Erikson (1983) find that white teachers were significantly more likely to offer more nonverbal criticism toward black males than black females, white female, and white males. In addition, the study found that black teachers offered more verbal neutral and nonverbal praise to female students than did white female teachers. Generally, studies have found verbal criticism to be negatively associated with student performance (Brophy & Good, 1984; Brophy & Evertson, 1974), so while criticism might not be considered something altogether negative in a schooling context, Simpson and Erikson frame praise and criticism as a binary, that positions criticism as something that negatively impacts students while praise bolsters achievement.

Similarly, in his observational study of 7th grade students in integrated classrooms in southeastern Louisiana, Casteel (1998) found that black students endured more negative interactions with white teachers than did white students, and he found that white students received more praise and positive feedback than did black students.

These studies indicate that in classroom settings dominated by white teachers, it is more likely that black students will endure negative interactions.

V. Research Questions and Hypotheses

Given the literature reviewed, it is clear that the race of a teacher is likely to affect the achievement of students. Specifically, black teachers are likely to reduce black-white achievement gaps relative to white teachers, because they are less likely to exhibit biases that negatively impact black student achievement. Another possibility is that the presence of black

teachers simply inspires increased black student performance. In addition, it seems unlikely that black teachers will substantially impact the achievement of white students, although there are very few studies that examine this particular racial pairing. In this manner, by improving the performance of black students and maintaining the current performance levels of white students, black teachers may reduce black-white test score gaps. The focus of this study is not to uncover the specific mechanisms that facilitate achievement differences, instead, this study lays a foundation for future research into these mechanisms by first addressing whether or not race matching affects racial achievement gaps within schools, net of other important background and school covariates. The focus of this investigation is the impact of black teachers on black and white students.

The research question that guides this analysis is the following:

1) What is the impact of having a black teacher for black and white students?

Given the literature reviewed, I propose the following hypotheses:

H1) Black teachers increase black student test scores relative to white teachers, net of family income, residence, teacher quality, and class size.

H2) The racial composition of teachers (disproportionally white) is a significant factor accounting for the black-white gap in school performance.

VI. Data and Methods

Background on Project STAR

Project STAR (Student-Teacher Achievement Ratio) was a school reform experiment conducted from 1985-1989, designed to test the impact of class size reduction on student

achievement. The initial selection of participants for the study began at the school level with Tennessee's Commissioner of Education inviting all school systems within the state to participate in the project. The guidelines issued to potential schools indicated that the state would cover additional costs for project teachers and teachers' aides, but that schools would receive no other special considerations other than the class size conditions required by the project. Within the criteria for selecting schools, two important conditions reduced the initial number of 180 schools that volunteered to participate. The first condition required that each school had enough kindergarten students for three distinct class-size assignments: a small class (S) with 13-17 students that served as the experimental treatment, and two regular class-size control groups: a regular class (R) with 22-25 students and a part-time teacher's aide, and a regular class 22-25 students and a full-time teacher's aide (RA). Within schools, students and teachers in the STAR experiment were randomly assigned to one of the three aforementioned classroom conditions: a small class size, a regular class size with a part-time teacher's aide, or a regular class size with a full-time teacher's aide. Once assigned to a class type students were to remain in that classroom condition as long as they remained in the project. The STAR experiment followed the cohort of students entering the sample of STAR schools in kindergarten in 1985 for four years through the end of their third grade year in 1989.

This kindergarten grade size condition reduced the sample of schools from 180 to around 100. A second condition imposed by the legislation that funded this experiment stated that schools from inner city, urban, suburban, and rural regions within the state needed to be represented in the school sample, so of the 100 eligible schools, 79 schools were selected from 42 of Tennessee's 141 school systems such that at least one school from each district that volunteered was present in the sample, and so that enough schools from each of the four regions

were included to allow for regional comparisons.

As I will discuss more fully later in this section (see Tables 4a and 4b), the regions included in the STAR sample were subject to significant racial segregation, with about two-thirds of black students concentrated in low-income metropolitan areas while around 3% or fewer white students live in these areas, in addition around two-thirds of white students are located in rural areas while less than 10% of black students are found in these regions. In the sample, black and white students are clustered within schools located in distinctly different locations. Given the likelihood that schools located in poorer areas are more likely to have fewer resources than those schools in less poor areas, it is important to control for school-level differences between schools, as these unobserved school level differences likely impact black student achievement since black students and teachers are clustered within schools in poor areas. As a result of this clustering, I employ models with and without school-level fixed effects, in order to compare how much of the race gap is explained by unobserved school effects.

Given the unique design of the STAR experiment, Table 1 presents selected characteristics of schools in order to assess how schools in the Project STAR sample compare to all public elementary schools in Tennessee and in the United States.

[Table 1 about here]

Students in the state of Tennessee are more likely to be black and from low income families as compared to students in the nation as a whole, but the location of the STAR project does not explain the very high representation of black and low income students in the STAR sample. In fact, the students in the STAR sample are twice as likely to be poor and 50% more

likely to be black than are all students in Tennessee. It appears that many schools from middle and higher income areas declined to participate in the research project. The skewed composition of the STAR sample means that population parameters (averages or proportions) are biased; however, measures of effect (regression coefficients) may still be unbiased if response rates are not correlated the effects of the treatment (race of teacher) on test scores.

Teachers in the sample are much less likely to have completed more than a bachelor's degree, and the STAR sample contains fewer teachers with advanced degrees than current state and national numbers. In addition, average student performance on the ACT is slightly worse than state and national averages. Given the composition of these schools and the unobserved aspects of school quality likely associated with these compositions that affect student performance, between school variation is controlled for in some models in order to produce unbiased estimates of racial groups within schools. As previously stated, I employ school level fixed effects, discussed in subsequent sections, to account for this potential bias, while at the same time introducing models that include school-level regional differences in order to compare models.

Strengths of the STAR Data

In randomized experiments, the treatment assignment is unrelated to observed or unobserved characteristics that might otherwise confound the relationship between treatment and the outcome of interest. In this study, students and teachers were randomly assigned to class types. Although this experiment was not designed to assess the effect of the race of a teacher on the learning outcomes of students, the effects of both a teacher's race and classroom size on student performance within schools should be independent of observed and unobserved

classroom characteristics since both students and teachers were randomly assigned to class types. These data provide a unique opportunity to assess the effects of a teacher's race on student learning without potential biases due to unobserved confounding variables that are present in conventional data on student achievement. In addition, because of the overrepresentation of black students and teachers in these data, these STAR data are one of the only datasets with enough statistical power to examine the effect of black teachers, although the selection process of the schools may limit how representative our findings are of the state and national level. Beyond this issue of representativeness, there are other issues that may threaten the validity of these data.

Violations of the Sampling Design

At the end of the first year of the study, due to teacher-identified discipline problems and some parent complaints the STAR Consortium decided to randomly assign approximately one-half of students from the regular classes to regular with aide classes and approximately one-half of regular with aide students to regular classes in the second year of the study (students' first grade year), where they were to remain through the end of the study. Before redistributing students, the STAR Consortium found that there were no significant differences in achievement between students assigned to regular-sized classes (R) and teacher-aided regular-sized classes (RA). According to the data documentation, "no students were purposely reassigned into or out of small (S) classes," and "no further modifications of this sort were made in subsequent years" (Word et al., 1990). Unfortunately, the data reveal that approximately 10 percent of students changed between small and regular class types between every year of the study**.** For a discussion of the rigorous tests used to ensure that reassignment and sample attrition do not affect estimates see Krueger (1999).

Fortunately, in using these data to analyze race matching, it is less likely that reassignment or attrition is a consequence of the race of a teacher and more likely to be a consequence of being stuck in a less favorable class size, a class size that one is expected to remain in for the next three years, whereas teachers change each year.

Sample

Although these data offer the opportunity for longitudinal analysis, I focus on within-grade estimates of the impact of student-teacher race matching; as a result, I constructed three samples, one for each grade level of grades 1-3. Table 2 illustrates how this sample was constructed. I restrict my sample to black and white students who took SAT reading and math subject tests at the end of the school year and were not missing values on other covariates used in this analysis. Students within the sample, and those excluded did not significantly differ on available measures of demographic and treatment characteristics.

[Table 2 about here]

Measures

DEPENDENT VARIABLE: Reading and Mathematics Percentile Rank Scores

The dependent variables used in this analysis are percentile rank scores on year-end mathematics and reading subject tests.[4] To construct this variable, I transformed the original raw scores from the Stanford Achievement Test (SAT) in reading and mathematics into percentiles based on all students' scores in the grade (Krueger 1999). These percentile rank scores reveal

[4] The Reading Comprehension Stanford Achievement Test (SAT) subject-test assesses literary, informational and functional comprehension of short reading passages, and includes questions requiring sentence completion, interpretation and critical analysis, while the Mathematics SAT subject-test tests knowledge of concepts foundational to mathematical problem solving, including number recognition, pattern recognition, basic mathematical concepts, and basic mathematical reasoning.

how students have performed during a given year in relation to the rest of the within-grade sample. In addition, because these subject tests differed in difficulty across grades, percentile ranks were calculated for each grade, which allows for between grade comparisons. In other words, change from year to year can be interpreted as movement up (or down) relative to same age peers.

INDEPENDENT VARIABLE: Race

The main independent variables of interest in this analysis are the race of a student and the race of a teacher. The race of a student was originally measured by six category variable measuring if a student is white, black, Asian, Hispanic, Native American, or some "other race" (the details of the latter category are unspecified).

Since students who are neither black nor white comprise less than 0.8% of the sample, these students were removed from the sample, and the student race variable was recoded as a dichotomous variable indicating whether a student is black or white. Similarly, only one teacher in the sample was neither black nor white, so for the purpose of testing race matching between black and white students, this lone Asian, third-grade teacher was removed from the sample. The race of a teacher was recoded as a dichotomous variable indicating whether a teacher is black or white.

INDEPENDENT VARIABLES: Gender, Family Income, and Residence

A number of individual level background measures were included in this study that have proven significant predictors of educational achievement. These variables include gender, family income, and residence. Gender is coded as a dichotomous variable indicating whether a student is male or female. Family income is measured by whether or not a student participated in a

subsidized lunch program. This measure is the only individual level measure of the socioeconomic status of a child in the sample. Unfortunately, because it indicates participation in a subsidized lunch program and not eligibility, it likely undercounts the actual numbers of students from poor families, as not all eligible low-income students participate in the subsidized lunch program. This participation, an indicator of family poverty status, is still able to act as proxy for family income, as those that participate in the subsidized lunch program are largely from families with lower incomes than those who do not. As a result, this measure is included in this analysis as dichotomous variable indicating whether or not a student participates in a subsidized lunch program or not. As indicated in Table 2, data were missing for this variable in the first (2.6%), second (3.57%), and third grade (4.15%) year.

A measure of residence is included in this analysis that offers some notion of the concentrated poverty around schools and within student neighborhoods. The variable was originally constructed to represents four regional classifications: inner city schools were designated as "those located in metropolitan areas (Nashville, Memphis, Knoxville, or Chattanooga) with more than half of their students on free or reduced price lunch;" suburban schools were "those located in the outlying areas of metropolitan cities;" urban schools were those located in towns of over 2,500 persons in non-metropolitan areas that serve a primarily urban population according to the definition provided by the U.S. Census; and all other schools in non-metropolitan areas were classified as rural (Word et al., 1990). In this analysis, this variable has been reconstructed as a four category variable with metropolitan low-income schools representing those schools in metropolitan areas with more than half of their students participating in a free or reduced price lunch program; metropolitan non low-income schools are those located in metropolitan areas with fewer than half of their students participating in a free or

reduced price lunch program; non-metropolitan urban schools are those located in non-metropolitan areas with populations of over 2,500 persons; all other schools located in non-metropolitan areas were classified as rural. This reconstruction better delineates which metropolitan schools are located in areas of concentrated poverty vs. those outside of these poverty zones, and it more clearly designates which schools are located in non-metropolitan areas that are more populated vs. those that are not. Given the clear impact of neighborhood poverty on student achievement, comparisons between metropolitan low-income schools and the other three categories are of significant interest. Further, because public school enrollment is based on residence in these regions, students that attend schools in metropolitan low-income schools are those living in the most impoverished neighborhoods, and as a result, those whose achievement is most likely disrupted by the many and varied effects of neighborhood poverty.

INDEPENDENT VARIABLES: Within School Covariates-Teacher Quality and Class Size

Two variables were included in this analysis to measure the quality of a teacher: teaching experience and educational attainment. The variable for teaching experience was initially a continuous variable measuring years of teaching experience, but given the sizeable literature on the influence of experience on student achievement, in this analysis the variable was reconstructed as a dichotomous variable indicating if a teacher is within their first two years of teaching, or if a teacher has more than two years of experience. Teachers above this two year threshold of experience have been associated with achievement gains in previous studies (Rivkin et al., 2005; Murnane & Phillips, 1981).

The variable for the educational attainment of a teacher was initially coded as a four-category variable measuring the highest degree a teacher has earned eg., a bachelor's degree

(B.A., B.S.), a master's degree (M.A., M.S.,), a specialist's degree (Ed.S.), or a doctoral degree (Ph.D.). Previous research on the influence of a teacher's educational attainment on student achievement suggests that there is no systematic benefit of having teachers with an advanced degree on student learning outcomes, but this non-finding is likely an artifact of the absence of measures indicating whether teachers are teaching within their degree field (Clotfelter et al., 2007; Hanushek 2003, 1997). As a result, this teacher education variable was reconstructed into a dichotomous variable indicating whether a teacher only attained a bachelor's degree or a master's degree and above.

A final school covariate included in this analysis was class size. As noted earlier, the students in the STAR sample were randomly assigned to a small or regular sized classrooms. Class size is measured by a categorical variable indicating whether a student was randomly assigned to a small class of approximately 13-17 students, a regular sized class of approximately 22-25 students with a part-time teacher's aide, or a regular sized class of approximately 22-25 students that included a full-time teacher's aide. As a matter of statewide practice in Tennessee, regular sized classes included a part time teacher's aide for a portion of students' elementary school years, so for grades 1-3 regular sized classes employed the services of a part-time aide that was present "25-33 percent of the time on average" (Krueger 2001), while the regular with full-time aide class type in the STAR sample employed a full-time teacher's aide.

Table 3 offers summary statistics of the STAR sample. The sample is composed of around 33% black students and 66% white students. Around half of the sample is female, and half of the sample receives subsidized school lunches. The majority of students live in rural areas (approximately 46%), followed by a sizeable percentage living in metropolitan low-income areas (approximately 24%), and a comparable percentage living in metropolitan non low-income areas

(approximately 22%). The least populated region is the non-metropolitan urban region where less than 10% of students reside.

Within schools, approximately 20% of students are taught by black teachers, and over 90% of students are taught by teachers who have more than two years of teaching experience. For students in the first and second grade samples, approximately 35% of students are taught by teachers with a master's degree or above, while in the third grade year around 43% of students are taught by teachers with these credentials.

The distribution of students to classroom sizes fluctuates across grade levels, with around 29% of students located in small class sizes in first and second grade, and 32% in the third grade. In the first grade approximately 38% of students were located in normal class sizes with a part-time aide, and in subsequent grades this percentage decreased to around 34% in second grade and 30% in third grade. Meanwhile the opposite trend is true of students taught in normal sized classes with a full-time aide, as around 32% are located in these classes in first grade, 35% in second grade, and 38% in third grade.

While these summary statistics offer a useful general picture of the home and schooling contexts of children, it is not yet clear how these background and school characteristics differ according to the race of students.

[Table 3 about here]

Considering that black and white students are likely to have very different home and schooling experiences, Tables 4a and 4b were constructed to show the differences between black and white students in background and school characteristics. These tables suggest that the home and school contexts of black children are very different from those of white children in the sample.

[Table 4a about here]

While gender distributions are similar across race and grade with a slightly higher percentage of males in the sample than females, there are massive departures when considering family income. Higher percentages of black students are from low-income families than white students with around 80% of black children living beneath the poverty line, while approximately 33% of white students endure similar financial circumstances. The over 45 percentage point difference in the indicator of poverty suggests that black students are far more likely than whites to have parents who are either unemployed or employed in jobs with low occupational statuses, and are more likely to live in single-parent households. These differences in family background mean that black students are less likely to have parents who can effectively advocate for them in schools, and suffer from low investment in educational activities that promote achievement.

As discussed previously, there are large racial differences in where students live. The majority of black students are concentrated in low-income metropolitan areas, while around 3% or fewer white students live in these areas. These low-income metropolitan areas are more likely to expose students to factors that depress achievement like violence, drugs, low-quality schools, and negative peers. Meanwhile around two-thirds of white students are located in rural areas, places that are less likely to disrupt achievement. Overall, given the distribution of students to the residence categories, it is unlikely that the STAR sample is representative of schools across the state.

[Table 4b about here]

Black students are taught by teachers who, in the first and second grade samples, are less educated and less experienced than those of their white peers, which may contribute to differential black-white achievement in those years. In addition, within schools there are slightly

fewer black students in smaller class sizes which have been shown to improve year-end achievement (see Mishel & Rothstein, 2002 for a comprehensive review of the impact of class size), although there do not appear to significant differences in class size distributions by race.

Within schools, white students are overwhelmingly taught by teachers of the same race with an average of 93% of white students taught by white teachers in each grade-level sample, while less than 45% of black students are matched with black teachers in the first and second grade samples, and only 51% in the third grade sample.

In addition, the majority of black teachers who teach black students are primarily within metropolitan low-income schools (around 75%) and secondarily within Metropolitan non low-income schools (around 22%), while the black teachers who teach white students are primarily in rural or metropolitan non low-income neighborhoods. Since the black teachers teaching black students are primarily located in low-income areas, it is likely that these teachers face more challenges to teaching by virtue of low neighborhood and school quality, while the black teachers teaching white students in rural areas are more likely to face challenges associated with being the only black people in schools surrounded by white students and teachers. In these situations, it is possible that black teachers are afforded fewer resources and support from schools that may undermine their effectiveness in elevating student achievement.

White teachers who teach black students are primarily located in metropolitan low-income neighborhoods, where they are joined by nearly equal numbers of black teachers, and in metropolitan non low-income neighborhoods white teachers are the majority. White teachers who teach white students are primarily located in rural areas and metropolitan non low-income areas. This distribution of white teachers suggests that all students are disproportionately exposed to white teachers, who in the aggregate, are more likely to raise student achievement

because they face fewer barriers to teaching as they teach a greater share of students who are not from poor families. In the following section, I look at how these family, neighborhood, and school differences are represented by mean gaps in achievement.

Mean Differences in Achievement

Tables 5a and 5b present black-white test score gaps in reading and math within categories of all the background variables. Across each covariate, gender, family income, residence, and teacher quality, black-white test score gaps exist. A consistent gender gap exists in reading scores but not in math scores for both black and white students, and the influence of our limited measures of teacher quality, teachers' educational attainment and years of experience, are inconsistent across grades and racial categories. Although these gaps certainly offer support for their inclusion as controls in this analysis, what is most interesting are the gaps in family income and residence variables. Relative to the other residence categories, mean test scores for black and white students are lowest in low-income schools, and the race gap is small as well. This suggests that schools in low-income neighborhoods substantially depresses achievement levels for both black and white students. An alternate view is that these schools are simply low-quality schools attended by students who reside in the poor neighborhoods that surround them, as a result, both neighborhood and school deficits act in concert to degrade the achievement of all students in those areas. On the opposite end of the wealth spectrum, students who are not from poor families have the highest mean scores within each racial category, but black students are still outscored by white students. In addition, within racial categories the mean score difference between being poor and non-poor is large, and often larger for blacks than whites. As Table 3 revealed earlier, around only 20% of black students are from families that are not poor as compared to over 65% of white students, and around 65% of black students attend

schools in poor neighborhoods compared to 3% of white students. Since black students are more likely to be from poor families and attend schools in poor neighborhoods, it is likely that class and the low-quality schools extant in these poor neighborhoods drive a considerable amount of the black-white achievement gap.

Figure 2 presents differences in mean test scores according to the race of a teacher with error bars representing 95% confidence intervals. According to these figures white students with white teachers score significantly higher than white students taught by black teachers, while black students with black teachers seem to significantly outscore black students who have white teachers only in the first grade sample. In addition, these figures suggest that black teachers reduce the black-white test score gap, but do so by reducing the achievement of white students. While it is unclear what specific dynamics inform this underachievement, for example black teachers might be less effective than white teachers along unobserved measures of teacher quality or maybe white students in classrooms with black teachers simply undermine black teachers, perhaps some school or parental influence conspires to lower the effectiveness of black teachers teaching white students, it seems clear that white students fair less well under the instruction of black teachers. What remains unanswered is whether or not this dynamic holds when we observe this score gap net of our measures of family income, residence, teacher quality, and class size.

[Figure 2 about here]

Models

I estimate linear regression models and regress a cross-classification variable for the race of the teacher and the race of the student on reading and mathematics percentile rank scores, including covariates for gender, location of school, family income, a teacher's educational

attainment, a teacher's years of experience, and class size. I estimate the models with and without school fixed effects (FE) since Tennessee public schools are spread across the state and within neighborhoods that vary according to a wide variety of characteristics, often resulting from residential segregation by race and socioeconomic status. The process through which pupils are selected into schools is likely selective and difficult to capture. The residuals in these linear regressions using maximum likelihood estimation, e_i, are assumed to be normally distributed with a mean of zero and a constant variance. The covariance between the individual characteristics of students and the residuals are assumed to be zero. However, these assumptions are violated if pupils are nested within schools that vary by measured and unmeasured characteristics. In this context, I examined the robustness of the MLE regression against school fixed effects models, and the substantive findings did not differ. I also examined the implications of the results of the fixed effect models for racial inequality and student performance within schools.

Equation 1 represents a linear regression model where x_{ij} a vector of observed student and teacher characteristics within schools for individual i, and school j, and z_i a vector of individual level controls. The assumptions of the linear regression model are violated because of the school level characteristics, and shared school characteristics of students within the same school. By introducing school level fixed effects in the form of α_j, we account for school level influences that would otherwise bias our estimates (Agresti & Finlay 2008).

$$y_{ij} = \mu_i + \beta x_{ij} + \gamma z_i + \alpha_j + e_{ij} \quad (1)$$

In a fixed effects model, no assumptions are made about error terms across schools and so the effects of school characteristics on student performance is not relevant in these models. There are

two kinds of fixed effects estimators: one that includes a dummy variable as an explanatory variable for each school to represent the school effects, and another that "differences" out the school effects. I used the former. The equation is as follows:

$$yij = x'ij + \beta + \sum_{j=1}^{J} \gamma_j D_j + e_{ij} \tag{2}$$

where *i* is the individual student, *j* is an individual school, and D_j is the school *j* dummy with coefficient γ_j for $j = 1,\ldots, J$. These school dummies account for school level differences.

Student-Teacher Interaction Variables

Models 1'-5 include an interaction between the race of teacher and race of student. For the purpose of this investigation, this variable was coded according to the indicator formula below:

$$\sum_{kl} \beta_{kl} \; I(S_{RACE} = k, T_{RACE} = l) \tag{3}$$

Specifically, for students of race *k* in classrooms with teachers of race *l* , the associated coefficients displayed below offer estimates of the score differences that result from different student-teacher racial pairings.

Table 6. Indicator Coding for Models 1'-5

Race Interaction Coefficients	Coefficient for
β_{00} I(S_{RACE}=0, T_{RACE}=0)	White Students with White Teachers (Referent)
β_{01} I(S_{RACE}=0, T_{RACE}=1)	White Students with Black Teachers
β_{10} I(S_{RACE}=1, T_{RACE}=0)	Black Students with White Teachers
β_{11} I(S_{RACE}=1, T_{RACE}=1)	Black Students with Black Teachers

According to this indicator coding system, the difference between the performance of white students with white teachers and whites students and black teachers is considered to be the impact of black teachers on the performance white students. The estimate of this impact is the difference between β_{00} and β_{01}. Since β_{00} is the referent category, this impact, β_{01}, and the race gap, β_{10}, are directly observed in in Models 1'- 5. The associated p-value indicates the statistical significance of this estimate.

The difference in the performance of black students when they have black teachers compared to white teachers is interpreted as the impact of black teachers on black students. This impact is equal to the difference between β_{11} and β_{10}. In order to test the significance of this difference, I use a Wald test to test if the impact of black teachers on the performance of black students is equal to the impact of white teachers on the performance of black students (Ho: $\beta_{11} = \beta_{10}$; Ha: $\beta_{11} \neq \beta_{10}$). These estimates are available in Table 8.

I also examine whether there is a net "gain" to overall student performance when black teachers are placed in classrooms. The net effect of having a black teacher instead of a white teacher for black and white students together is equal to the sum of the the impact of a black teacher for white students, β_{01}, and the impact of a black teacher for black students, $\beta_{11} - \beta_{10}$. I calculate the significance of this estimate using a Wald test to test if the impact of black teachers

on the performance of white students is equal to the impact of black teachers on black student performance (Ho: $\beta_{01} = \beta_{11}-\beta_{10}$; Ha: $\beta_{01} \neq \beta_{11}-\beta_{10}$). These results are displayed in Table 8.

Multivariate Analysis: Models 1 – 5

I use the following model building strategy to decompose black-white test score gaps in first, second, and third grade. The following models are estimated by ordinary least squares (OLS) with a disturbance term (e_{ij}) containing individual and school-level components while α_j , introduced in Model 5, represents school-level fixed effects.

Model 1: The baseline model presents an estimate of race and gender gaps in achievement. In addition this model is similar to Tables 5a and 5b which illustrate test score gaps by race and gender without introducing a number of important covariates.

$$Y = \beta o + \beta 1(\text{Student Race}) + \beta 2(\text{Gender}) + e_{ij} \quad (4)$$

Model 1': This model introduces student-teacher race interactions which allow us to observe test score gaps within different racial pairings.

$$Y = \beta o + \beta_{01}(\text{White Students with Black Teachers}) + \beta_{10}(\text{Black Students with White Teachers}) + \beta_{11}(\text{Black Students with Black Teachers}) + \beta 2(\text{Gender}) + e_{ij} \quad (5)$$

Model 2: This model adds two measures of socioeconomic status: family income and residence. The difference between the racial gap in student performance in Model 1' and Model 2 is an estimate of the impact of family SES and neighborhoods where children live on race gaps within student-teacher racial pairings.

$Y = \beta o + \beta_{01}$ (White Students with Black Teachers) + β_{10} (Black Students with White Teachers) + β_{11} (Black Students with Black Teachers) + β2(Gender) + β3(Family Income) + β4(Residence) + e_{ij} (6)

Model 3: This model adds to Model 2 by introducing the only available measures of teacher quality within these data: the educational attainment of a teacher and years of teaching experience. The difference in racial gaps observed in these student-teacher racial pairings between Model 2 and Model 3 offers an estimate of the impact of measureable attributes of teacher quality.

$Y = \beta o + \beta_{01}$ (White Students with Black Teachers) + β_{10} (Black Students with White Teachers) + β_{11} (Black Students with Black Teachers) +β2(Gender) + β3(Family Income) + β4 (Residence) + β5(Teacher Educational Attainment) + + β6(Teaching Experience) + e_{ij} (7)

Model 4: This model introduces our class assignment variable. The difference in racial gaps observed in these student-teacher racial pairings between Model 3 and Model 4 is an estimate of the impact of class size reduction.

$Y = \beta o + \beta_{01}$ (White Students with Black Teachers) + β_{10} (Black Students with White Teachers) + β_{11} (Black Students with Black Teachers) +β2(Gender) + β3(Family Income) + β4 (Residence) + β5(Teacher Educational Attainment) + + β6(Teaching Experience) + β7(Class Size) + e_{ij} (8)

Model 5: This model introduces school-level fixed effects because of the potentially large differences between schools with regard to resources, students, and teachers, and because randomization of students and teachers to class sizes was done within schools. The difference in racial gaps observed in these student-teacher racial pairings between Model 4 and Model 5 is an estimate of the impact of between school differences on student achievement.

$Y = \beta o + \beta_{01}$ (White Students with Black Teachers) + β_{10} (Black Students with White Teachers) + β_{11} (Black Students with Black Teachers) +β2(Gender) + β3(Family Income) + β4 (Residence) + β5(Teacher Educational Attainment) + + β6(Teaching Experience) + β7(Class Size) + $e_{ij} + \alpha_j$ (9)

VII. Results

The Impact of Race of Teacher on Black Student Performance and the Black-White Student Gap

Tables 7a, 7b, and 7c present these regression results

[Tables 7a, 7b, and 7c about here]

According to the baseline model, Model 1, there is an enormous and significant black-white test score gap, with black students scoring from 17 to 21 percentage points less than their white peers on reading and mathematics tests in every grade of the sample, net of gender differences. This indicates that black students score nearly a quintile behind their white peers. In addition, there are persistent gender differences in reading scores at each grade level, with girls scoring around 6 percentage points higher than boys, yet there are no significant differences in mathematics scores between girls and boys across grade levels. This evidence is consistent with national trends in racial differences (see Figure 1), and gender gaps present in NAEP data, as well as findings from previous studies using STAR data (Dee 2005; Krueger 1999). These gaps are also reflected in the mean score gaps presented in Table 5. Once we introduce student-teacher racial pairings in Model 1', it is clear how large score gaps are for black students under the instruction of white teachers, in each grade level and across subject tests, black students with white teachers have racial score gaps as wide as the observed gaps for the total population. Overall, black students with black teachers perform about 2 to 3 percentage points better relative to their white peers (with white teachers). This is not, however, the "true" effect of race of a teacher on black student achievement because having a black teacher is highly correlated with attending metropolitan low income schools.

How much of the race gap is due to the correlation between race and socioeconomic status (family income and residence) and how does SES affect the relationship of race of teacher on student achievement (Model 2)

Comparing Model 1' and Model 2, with the latter taking into account differences in family income and neighborhood poverty, reveals that the test score gap is greatly reduced across reading and math tests, and across grade levels. These reductions appear slightly greater for black students who are taught by black teachers relative to black students taught by white teachers (with white students with white teachers as a referent for both). If black and white students had the same SES distributions (poverty and residence), the gap in reading scores would drop from 21 points to only 8 points for black students with white teachers and from 18 points to only 2 points (only marginally significant) for black students with black teachers. The impact is only slightly less for math scores. This finding, that social class accounts for a substantial portion of the racial achievement gap is consistent with findings from a number of studies of the achievement gap outlined earlier and these results are consistent with estimates from previous studies using these data (Dee 2005; Krueger 1999). Interestingly, this study finds that in the first grade, SES and the race of teacher collectively explain nearly all of the black-white test score gap; however in second and third grade, racial gaps persist across subject tests net of these indicators of SES.

Given the salience of racial segregation which causes black children to live in circumstances that preclude upward mobility and causes black students to attend schools that are often without resources that are important to educational achievement and attainment, this apparent reduction in the achievement gap once SES is introduced into the model is a reflection of how heavily these social circumstances impose themselves on the achievement of black students. Given these extramural conditions that impose themselves on student achievement

within schools, how much do the factors over which schools exert an enormous amount of control, like the hiring of quality teachers, succeed in equalizing black-white test scores?

How much racial inequality is due to measureable attributes of teacher quality? (Model 3)

There is inconsistent support for the idea that teachers with less than two years of teaching experience or teachers with less than a master's degree have a different impact on student test scores than their more experienced or educated peers. The changes in these significance levels across tests and grade levels could be a product of some unobserved school organizational feature or practice that benefits more experienced and educated teachers, for example, these "higher quality" teachers might be assigned to classes of higher performing students. According to BIC and likelihood-ratio test statistics, the models that introduce measureable aspects of teacher quality do not fit significantly better than the models without the teacher quality indicators. Further, across grade levels and test subjects, these measures do not explain a significant portion of the black-white test score gap. This finding is consistent with previous studies of teacher quality that find the impact of teacher experience and teachers' education levels on student test scores negligible, in part because these data do not capture whether or not teachers are teaching subjects that are aligned with the fields in which their advanced degrees where obtained (Hanushek & Welch, 2006). This finding is consistent with the inconsistent influence of these attributes of teacher quality referenced in Table 5. One would expect that teacher quality would offer a relatively stable influence on achievement, but the erratic fluctuations observed in Table 5 indicate there are likely important differences in teacher quality that are only partially, if at all, reflected in the measures included in these data.

Other than the years of experience and education levels of the employees hired by schools, class sizes are also subject to school control.

How much is racial inequality reduced by reducing class size? (Model 4)

When we add class size to our model, model fit improves according to BIC and the likelihood ratio test values, and it appears that all students, both black and white, in smaller class sizes achieve test scores around 5 percentile points higher than those in normal class sizes across grade levels and subject tests. Our class size variable also reveals that although there is a significant difference between having a full-time aide relative to a part-time aide in students' first and second grade classes, students under the care of a full-time aide only perform between 1 and 2 percentage points better than their peers with a part-time aide, and in the third grade year there is no apparent benefit. The similarity of student performance across these normal sized classes, with and without full-time aides, is consistent with previous studies that employ these STAR data (Chetty 2011; Dee 2005; Krueger 1999). In addition, the introduction of these class size variables offer no significant reduction in the black-white test score gap, as smaller classroom sizes are positive for both black and white students. Given that the research design of the STAR experiment effectively randomly assigned students to different class sizes, it is no surprise that the introduction of this class size variable did not affect other measureable student characteristics.

Do unobserved differences between the schools affect these estimates? (Model 5)

Since the design of the STAR experiment dictated that random assignment happen within schools, the true benefits of randomization are present in estimates within schools, the benefits of

randomization are present in estimates within schools, not necessarily between them as there are numerous differences in school resources, quality of teachers and students, and other characteristics of schools that are not captured in the model and might bias our estimates. As a result, introducing controls for school effects in the form of a separate dummy variable for each school (α_j), absorbs unobserved differences between schools. Conceptually, this means that all differences between schools are controlled, not just those measured with the covariates in Models 2 through 4. The remaining black white gap could be due to within school difference in family background or within school differences that differentially affect black and white students.

According to a comparison of Model 4 and Model 5, this fixed effects model is an improvement over previous models, according to the BIC, likelihood ratio test, and the amount of variance explained. In addition, it is clear that smaller class sizes benefit all students and total school effects reduce the gap, with the exception of first grade math tests where the gap increased for black students with both white and black teachers, and third grade math tests where the gap increased only for black students with black teachers. This model also reveals that score gaps are largest in classrooms with white teachers, with the exception of our within-school estimate of match performance where gaps appear largest for black students with black teachers. This model, as did the previous models, reveals a persistent black-white test score gap, a large and persistent family income gap, and a persistent benefit of being in a smaller class relative to a larger class sizes, while there is no evidence of the influence of teacher quality on test scores.

These findings are consistent with previous models, although the magnitude of the racial test score gap varies very slightly across years and subject tests which suggests that between school differences accounted for some variation in the estimates. One potential explanation for

the remaining gap could be "tracking" of students or other features of school organization that have an adverse impact on black student learning relative to their white peers.

Overall, I find a persistent racial test score gap net of measures of family and neighborhood poverty, teacher quality, class size, and any unobserved differences between schools. Given this persistent inequality, in subsequent models I ask, how much of this gap is due to the current paradigm in which teachers are predominantly white.

What is the benefit of having a black teacher?

Table 8 provides an overall summary of the influence of black teachers on the achievement of black and white students. As discussed earlier, the impact of black teachers on the achievement of white students is observed directly as β_{01}, given that the racial pairing of white students with white teachers is our omitted category.

[Table 8 about here]

According to Table 8, our estimates indicate that black teachers significantly reduce white performance in all grades relative to their performance in the classrooms of white teachers, although this reduction is not significant in third grade reading assessments, or first grade reading scores when controlling for between school differences. Given the extremely low numbers of white students who are taught by black teachers, and the regions in which these pairings are clustered, it is difficult to interpret the negative impact of black teachers on white students. Perhaps the 5 percent of white students with black teachers are selective on some observed characteristics that explain their lower achievement.

Black teachers significantly improve the performance of black students relative to white teachers in first grade, although the significance of the influence of black teachers is much less in

second and third grade across subject tests. In the second and third grade samples presented in Table 8, it is clear that the impact of black teachers relative to white teachers on math performance is small and statistically insignificant, a finding that, in these two grades is robust to model specification. In these same grade-years, although black teachers appear to improve student performance on reading tests, our estimates are relatively small and largely insignificant. These results suggest that in our second and third grade replications, there may not exist a significant advantage for black students who have a black teacher as compared to black students with a white teacher.

What is the net benefit of having a black teacher?

Ultimately, these results indicate that the benefit of having a black teacher for black students is largely eclipsed by the negative impact of black teachers on white student performance. The exceptions to this finding are in our estimates of first grade math performance, and the models controlling for school effects for first and third grade reading performance. In these classrooms, black students in first grade achieve at greater levels than white students are disadvantaged, and black students in third grade seem to reap advantages equal to white students' disadvantage, according to significance levels.

Although black students perform better with black teachers relative to white teachers, it is clear that introducing policy that bolsters the number of black teachers in all classrooms may not be a solution for the reduction of inequality that prizes both black and white student achievement.

VIII. Conclusions

This study provides substantial support for the thesis that black students benefit from black teachers. For black students, black teachers may provide a buffer from the low expectations of white teachers, and may also provide more encouragement for black students than that provided by white teachers. Given the legacy of discrimination, the logic seems plausible.

Although it appears that white students also benefit from having a teacher of the same race, there is no clear theoretical expectation for this interpretation. In fact, very few white students have black teachers – only about one in 20 students – and as white students are clustered into predominately white schools, black teachers are teaching schools with mostly black students. Assuming that most white families avoid black teachers for their children by sending them to white schools with few black teachers, the few white students in classrooms with black teachers may be disadvantaged in other ways that would curtail educational achievement.

On the surface, it appears that teachers themselves are responsible for learning differences, but, as is the case with any interactive enterprise, learning is a reciprocal process, one that relies on the collective effort of students and teachers. As a result, it is plausible that the reduced achievement of white students is a consequence of their own efforts as well.

It seems possible that white students who end up in classrooms with black teachers are those that would perform poorly under the instruction of any teacher. Given the level of segregation present in the sample, the 3% of white kids living in predominantly black neighborhoods might be those who are under the direst of circumstances, those devoid of supportive educational resources.

Perhaps schools contribute to the underachieving relationship of black teachers and white students. In line with theories of minority teacher "burnout," perhaps black teachers teaching in schools with predominantly white students are unsupported or undermined by their non-black colleagues. Or perhaps the racial biases of their colleagues or the parents of their students conspire to discredit black teachers in the eyes of white students.

An alternative explanation might be that schools with high concentrations of black students are only able to recruit ineffective, white teachers and high quality black teachers, while predominantly white schools are able to recruit higher quality white teachers and ineffective black teachers.

There are many potential explanations for the reduced achievement of white students under the instruction of black teachers, and fortunately most of these explanations could be proved or disproved with better data, in which there are more observations of students and teachers.

REFERENCES

Agresti, A., & Finlay, B. (2008). Statistical methods for the social sciences. Upper Saddle River, NJ: Prentice Hall

Ainsworth, J. W. (2002). Why does it take a village? The mediation of neighborhood effects on educational achievement. *Social Forces*, *81*(1), 117-152.

Alexander, K.L., Entwistle, D.R., & Thompson, M.S. (1987). School performance, status relations, and the structure of sentiment: Bring the teacher back in. American Sociological Review, 52, 665-682.

Alexander, K. L., Entwisle, D. R., & Horsey, C. S. (1997). From first grade forward: Early foundations of high school dropout. *Sociology of education*, 87-10.
Anderson, E. (1999). *Code of the Street* (pp. 107-141). New York: Norton.

Bates, L. A., & Glick, J. E. (2013). Does it matter if teachers and schools match the student? Racial and ethnic disparities in problem behaviors. *Social science research*, *42*(5), 1180-1190.

Brooks-Gunn, J., & Duncan, G. J. (1997). The effects of poverty on children. *The future of children*, 55-71.

Brophy, J. E., & Evertson, C. M. (1974). Process-Product Correlations in the Texas Teacher Effectiveness Study: Final Report.

Brophy, J. E., & Good, T. L. (1984). *Teacher behavior and student achievement* (No. 73). Institute for Research on Teaching, Michigan State University.

Carlson, M. J., & Corcoran, M. E. (2001). Family structure and children's behavioral and cognitive outcomes. *Journal of Marriage and Family*, *63*(3), 779-792.

Casteel, C. A. (1998). Teacher–student interactions and race in integrated classrooms. *The Journal of Educational Research*, *92*(2), 115-120.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Schanzenbach, and Danny Yagan. (2011). How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR Quarterly Journal of Economics 126(4): 1593-1660, 2011.

Clewell, B. C., & Villegas, A. M. (1998). Diversifying the Teaching Force To Improve Urban Schools: Meeting the Challenge. Introduction. *Education and Urban Society*, *31*(1), 3-17.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. (2005). Who teaches whom? Race and the distribution of novice teachers. *Economics of Education review*, *24*(4), 377-392.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review*, *26*(6), 673-682.

Coates, B. (1972). White adult behavior toward black and white children. *Child Development*, 143-154.

Cole, B. P. (1986). The black educator: An endangered species. *Journal of Negro Education*, 326-334.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. (1966). Equality of educational opportunity. Washington, DC. 1066-5684.

Condron, D. J., & Roscigno, V. J. (2003). Disparities within: Unequal spending and achievement in an urban school district. *Sociology of Education*, 18-36.

Condron, D. J. (2007). Stratification and educational sorting: Explaining ascriptive inequalities in early childhood reading group placement. *Social Problems*, Vol. 54, No. 1 (February 2007), pp. 139-160.

Danziger, S., & Stem, J. (1990). The causes and consequences of child poverty in the United States, University of Michigan Population Studies Center Research Report No. 90-194. September.

Dee, T. S. (2004). Teachers, race, and student achievement in a randomized experiment. *Review of Economics and Statistics*, *86*(1), 195-210.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter?. *American Economic Review*, 158-165.

Downey, D. B., & Pribesh, S. (2004). When race matters: Teachers' evaluations of students' classroom behavior. *Sociology of Education*, *77*(4), 267-282.

Duncan, G. J., Brooks-Gunn, J., & Klebanov, P. K. (1994). Economic deprivation and early childhood development. *Child development*, *65*(2), 296-318.

Duncan, G., & Brooks-Gunn, J. (Eds.). (1997). Consequences of growing up poor. New York: Russell Sage Foundation.

Duneier, M., & Carter, O. (1999). *Sidewalk*. Macmillan.

Eaves, R. C. (1975). Teacher race, student race, and the behavior problem checklist. *Journal of Abnormal Child Psychology*, *3*(1), 1-9.

Ehrenberg, R. G., Goldhaber, D. D., & Brewer, D. J. (1995). Do teachers' race, gender, and ethnicity matter? Evidence from the National Educational Longitudinal Study of 1988. *Industrial & Labor Relations Review*, *48*(3), 547-561.

Entwisle, D. R., & Alexander, K. L. (1992). Summer setback: Race, poverty, school composition, and mathematics achievement in the first two years of school. *American Sociological Review*, 72-84.

Ethridge, S. (1979). Impact of the 1954 Brown vs. Topeka Board of Education decision on Black educators. The Negro Educational Review, 30 (4), 217-232.

Farkas, G., Grobe, R. P., Sheehan, D., & Shuan, Y. (1990). Cultural resources and school success: Gender, ethnicity, and poverty groups within an urban school district. *American Sociological Review*, 127-142.

Featherman, D. L., & Hauser, R. M. (1978). Opportunity and change. *Studies in Population*. Academic Press. New York NY United States 1978..

Ferguson, R. F. (2003). Teachers' perceptions and expectations and the Black-White test score gap. *Urban Education*, *38*(4), 460-507.

Fryer Jr, R. G., & Levitt, S. D. (2004). Understanding the black-white test score gap in the first two years of school. *Review of Economics and Statistics*, *86*(2), 447-464.

Fryer, R. G., & Levitt, S. D. (2006). The black-white test score gap through third grade. *American Law and Economics Review*, *8*(2), 249-281.

Fultz, M. (2004). The displacement of Black Educators post-Brown: An overview and analysis. *History of Education Quarterly*, *44*(1), 11-45.

Garfinkel, I., & McLanahan, S. (1986). *Single mothers and their children: A new American dilemma*. Washington, DC: Urban Institute Press.

Graham, P. A. (1987). Black teachers: A drastically scarce resource. *Phi Delta Kappan*, 598-605.

Hallinan, M. T. (1996). Track mobility in secondary school. *Social Forces*, *74*(3), 983-1002.

Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of political economy*, 84-117.

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational evaluation and policy analysis*, *19*(2), 141-164.

Hanushek, E. A. (2003). The Failure of Input-based Schooling Policies*. *The economic journal*, *113*(485), F64-F98.

Hanushek, E. A., & Welch, F. (2006). *Handbook of the Economics of Education. Volume 2*. Elsevier. 11830 Westline Industrial Drive, St. Louis, MO 63146.

Hauser, Robert M., and David L. Featherman. "The Process of Stratification: Trends and analysis." *New York: Academic* (1977).

Hedges, L. V., & Nowell, A. (1999). Changes in the Black-White gap in achievement test scores. *Sociology of education*, 111-135.

Hess, F. M., & Leal, D. L. (1997). Minority Teachers, Minority Students, and College Matriculation: a New Look at the Role-Modeling Hypothesis. *Policy Studies Journal*, *25*(2), 235-248.

Irvine, J. J. (1989). Beyond role models: An examination of cultural influences on the pedagogical perspectives of Black teachers. *Peabody Journal of Education*, *66*(4), 51-63.

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*, 281-388.

Kauchak, D. (1984). Testing Teachers in Louisiana: A Closer Look. *Phi Delta Kappan*, *65*(9), 626-28.

Kane, Andrea and Isabel V. Sawhill, "Preventing Teen Childbearing" in *One Percent for the Kids,* edited by Sawhill (Brookings, 2003), pp. 56–75.

King, S. H. (1993). Why did we choose teaching careers and what will enable us to stay?: Insights from one cohort of the African American teaching pool. *Journal of Negro Education*, 475-492.

King, S. H. (1993). The limited presence of African-American teachers. *Review of Educational Research*, *63*(2), 115-149.

Krueger, A. B. (1997). *Experimental estimates of education production functions*. Quarterly Journal of Economics, Vol. 114 (2), 497-532.

Kuklinski, M. R., & Weinstein, R. S. (2001). Classroom and developmental differences in a path model of teacher expectancy effects. *Child development*, *72*(5), 1554-1578.

Lareau, A. (2002). Invisible inequality: Social class and childrearing in black families and white families. *American sociological review*, 747-776.

Lareau, A. (2003). Unequal Childhoods. Berkeley.

Lee, V. E., & Burkam, D. T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. Economic Policy Institute, Washington, DC.

Leventhal, T., & Brooks-Gunn, J. (2000). The neighborhoods they live in: the effects of neighborhood residence on child and adolescent outcomes. *Psychological bulletin*, *126*(2), 309.

Lucas, S. R. (2001). Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects1. *American journal of sociology*, *106*(6), 1642-1690.

Madon, S., Jussim, L., Keiper, S., Eccles, J., Smith, A., & Palumbo, P. (1998). The accuracy and power of sex, social class, and ethnic stereotypes: A naturalistic study in person perception. *Personality and Social Psychology Bulletin*, *24*(12), 1304-1318.

McGroder, S. M., Moore, K. A., & Zaslow, M. J. (2000). *National Evaluation of Welfare-to-work Strategies: Impacts on Young Children and Their Families Two Years After Enrollment: Findings from the Child Outcomes Study: Summary Report*. US Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation, Administration for Children and Families.

McKown, C., & Weinstein, R. S. (2002). Modeling the role of child ethnicity and gender in children's differential response to teacher expectations1. *Journal of Applied Social Psychology*, *32*(1), 159-184.

McKown, C., & Weinstein, R. S. (2008). Teacher expectations, classroom context, and the achievement gap. *Journal of school psychology*, *46*(3), 235-261.

McLanahan, S. (1997). Parent absence or poverty: Which matters more. *Consequences of growing up poor*, 35-48.

McLanahan, S., & Sandefur, G. (1994). *Growing Up with a Single Parent. What Hurts, What Helps*. Harvard University Press, Cambridge, MA.

McLoyd, V. C., & Wilson, L. (1990). Maternal behavior, social support, and economic conditions as predictors of distress in children. *New Directions for Child and Adolescent Development*, *1990*(46), 49-69.

Menacker, J. (1990). Equal educational opportunity: Is it an issue of race or socioeconomic status? Urban Education, 25 (3), 317–325.

Morris, P. A., Huston, A. C., Duncan, G. J., Crosby, D. A., & Bos, J. M. (2001). How Welfare and Work Policies Affect Children: A Synthesis of Research.

Murnane, R. J., & Phillips, B. R. (1981). Learning by doing, vintage, and selection: Three pieces of the puzzle relating teaching experience and teaching performance. *Economics of education review*, *1*(4), 453-465.

Oates, G. L. S. C. (2009). An empirical test of five prominent explanations for the black–white academic performance gap. *Social Psychology of Education*, *12*(4), 415-441.

Orfield, G., Monfort, F., & Aaron, M. (1989). Status of school desegregation: 1968-1986. *Alexandria, VA: National School Boards Association*.

Orr, Larry, Judith Feins, Robin Jacob, Eric Beecroft, Lisa Sanbonmatsu, Lawrence F. Katz, Jeffrey B. Liebman, and Jeffrey R. Kling.(2003).*Moving to opportunity for fair housing demonstration program: Interim impacts evaluation*. US Department of Housing and Urban Development, Office of Policy Development and Research.

Phillips, Meredith. (2000). "Understanding Ethnic Differences in Academic Achievement: Empirical Lessons from National Data." In David W. Grissmer and J. Michael Ross (eds.), Analytic Issues in the Assessment of Student Achievement (pp. 103-32). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Phillips, M., Crouse, J., & Ralph, J. (1998). Does the Black-White test score gap widen after children enter school. *The Black-White test score gap*, 229-272.

Pribesh, S., & Downey, D. B. (1999). Why are residential and school moves associated with poor school performance? *Demography*, *36*(4), 521-534.

Rist, R. (1970). Student social class and teacher expectations: The self-fulfilling prophecy in ghetto education. *Harvard educational review*, *40*(3), 411-451.

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

Rouse, C. E., Brooks-Gunn, J., & McLanahan, S. (2005). The Future of Children: School Readiness: Closing Racial and Ethnic Gaps. *Brookings and Woodrow Wilson School of Public and International Affairs, Princeton University*.

Rosenthal, J. O. (1957). Negro Teachers' Attitudes Toward Desegregation. *Journal of Negro Education*, 63-71.

Rothstein, R. (2002). *The class size debate*. Economic Policy Institute.

Rumberger, R. W., & Larson, K. A. (1998). Student mobility and the increased risk of high school dropout. *American Journal of Education*, 1-35.

Rumberger, R. W., Ghatak, R., Poulos, G., Ritter, P. L., & Dornbusch, S. M. (1990). Family influences on dropout behavior in one California high school. *Sociology of education*, 283-299.

Sandefur, G. D., McLanahan, S., & Wojtkiewicz, R. A. (1992). The effects of parental marital status during adolescence on high school graduation. *Social Forces*, *71*(1), 103-121.

Simpson, A. W., & Erickson, M. T. (1983). Teachers' verbal and nonverbal communication patterns as a function of teacher race, student gender, and student race. *American Educational Research Journal*, *20*(2), 183-198.

Spruill, A. W. (1960). The Negro Teacher in the Process of Desegregation of Schools. *Journal of Negro Education*, 80-84.

Tillman, L. C. (2009). *The sage handbook of African American education*. Sage.

Tillman, L. C. (2004). (Un)intended consequences? The impact of the Brown v. Board of Education decision on the employment status of black educators. *Education and Urban Society*, *36*(3), 280-303.

Turley, R. N. L. (2003). When do neighborhoods matter? The role of race and neighborhood peers. *Social Science Research*, *32*(1), 61-79.

Waters, M. M. (1989, September). An agenda for educating black teachers. In *The Educational Forum* (Vol. 53, No. 3, pp. 267-279). Taylor & Francis Group.

Williams Jr, Robin M., and Gerald David Jaynes, eds. (1989). *A Common Destiny: Blacks and American Society*. Washington, DC: National Academy Press.

Wilson, W. J. (1987). *The truly disadvantaged: The inner city, the underclass, and public policy*. University of Chicago Press.

Word, E. R. (1990). The State of Tennessee's Student/Teacher Achievement Ratio (STAR) Project: Technical Report (1985-1990).

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1978-2012 Long-Term Trend Reading and Mathematics Assessments.

TABLES AND FIGURES

Table 1. Selected Population Characteristics: Sample, State, and National Comparison

Characteristics	STAR Sample	Tennessee	United States
Average Composite ACT Score	19.2	19.7	21.1
Percentage of Black Students	34.3%	23.0%	15.7%
Percentage of Children below the poverty level	49.4%	24.4%	21.0%
Percentage of Black Teachers	20.3%	5.2%	6.8%
Percentage of Teachers with a Master's degree or higher	37.6%	60.5%	56.4%
Percentage of Teacher with more than three years of teaching experience	87.8%	89.4%	91.0%

Notes: Project STAR characteristics were calculated from the sample directly, while national and statewide characteristics are from a number of sources gathered in 2012. ACT scores for Tennessee and the United States in 2012 are from ACT Inc. National and statewide percentages of black students in 2012 are from the US Department of Education's Common Core Data (CCD). National and statewide percentages of children living below the poverty level in 2012 are from the US Census Bureau's American Community Survey. National and statewide percentages in 2012 of black teachers, teachers with a Master's degree or higher, and teachers with 3 years or more experience are the US Department of Education's Schools and Staffing Survey (SASS).

Table 2. Sample Construction

	First Grade		Second Grade		Third Grade	
	N	Missing	N	Missing	N	Missing
Initial number of respondents	6829	0	6840	0	6802	0
Less those missing values on reading and math SATs	-452	6.62%	-791	11.56%	-835	12.28%
Less those students missing values for race	-2	0.42%	-20	1.30%	-5	0.21%
Less those students who were neither black nor white	-43	1.17%	-32	1.89%	-29	0.68%
Less those students with teachers who were missing values for race	0	0.28%	-45	0.88%	-49	0.75%
Less those students with a teacher who was neither black nor white	0	0.28%	0	0.88%	-14	0.96%
Less those students missing values for teacher's educational attainment	-21	0.60%	-34	1.40%	-14	0.97%
Less those students with teachers missing values for years of experience	0	0.28%	-35	1.48%	0	0.75%
Less those students missing values for free or reduced price lunch receipt	-149	2.62%	-241	3.57%	-173	4.15%
Sample after missing data are removed(Total Percentage Missing)	6162	(9.8%)	5642	(17.51%)	5683	(15.88%)

Note: Total percentage missing is reflective of the percent missing from the final sample N, relative to the intial number of respondents, not the sum of missing values

Source: Health & Education Research Operative Service, the State of Tennessee's Student/Teacher Achievement Ratio (STAR) Project (1985-1989).

Table 3. Characteristics of Students in the Star Sample by Grade Level

	GRADE 1		Grade 2		Grade 3	
Dependent Variable	Mean/%	SD	Mean/%	SD	Mean/%	SD
Reading Index (Percentile)	50.915	28.49	50.94	28.6	51.24	28.92
Mathematics Index (Percentile)	51.28	28.67	50.61	28.45	50.73	28.62
Background Variables						
Race of Student						
Black	33.8%		35.2%		33.7%	
White	66.2%		64.8%		66.3%	
Gender						
Female	48.4%		48.3%		48.6%	
Male	51.6%		51.7%		51.4%	
Family Income						
Poor (Receiving Free/Reduced Price Lunch)	50.6%		48.8%		48.7%	
Non-Poor	49.4%		51.2%		51.3%	
Percent of Students Residing in						
Rural	45.8%		46.7%		47.1%	
Non-Metropolitan Urban	9.3%		5.8%		7.1%	
Metropolitan Low-Income	23.4%		24.5%		24.5%	
Metropolitan Non Low-Income	21.4%		22.9%		21.3%	
Schooling Variables						
Race of Teacher						
Percent of Students with a Black Teacher	17.9%		21.0%		21.8%	
Percent of Students with a White Teacher	82.1%		79.0%		78.2%	
Perecent of Students with Teachers who have						
Master's Degree or above	34.1%		35.8%		42.9%	
Bachelor's Degree Only	65.9%		64.2%		57.1%	
Two years or more of teaching experience	90.5%		92.7%		95.0%	
Less than two years of experience	9.5%		7.3%		5.0%	
Perecent of Students in classes that are						
Small	28.7%		29.9%		32.0%	
Normal with a part-time aide	38.4%		34.3%		29.9%	
Normal with a full-time aide	32.9%		35.8%		38.2%	
Sample N	6162		5642		5683	

Source: Health & Education Research Operative Service, the State of Tennessee's Student/Teacher Achievement Ratio (STAR) Project (1985-1989).

Table 4a. Differences in Characteristics of Students, including Teacher Attributes and Classroom Size by Grade Level

	GRADE 1			GRADE 2			GRADE 3		
Gender	Black (N=2084)	White (N=4078)	Difference	Black (N=1987)	White (N=3655)	Difference	Black (N=1916)	White (N=3767)	Difference
Male	50.8%	52.1%	-1.3%	50.8%	52.1%	-1.3%	50.4%	51.8%	-1.4%
Female	49.2%	47.9%		49.2%	47.9%		49.6%	48.2%	
Family Income									
Poor (Free/Reduced Lunch)	81.7%	34.7%	46.9%	79.6%	32.1%	47.5%	80.3%	32.6%	47.7%
Non Poor	18.3%	65.3%		20.4%	67.9%		19.7%	67.4%	
Residence									
Rural	9.4%	64.4%	-55.0%	8.7%	67.3%	-58.67%	8.8%	66.6%	-57.8%
Non-Metropolitan Urban	3.8%	12.2%	-8.4%	2.3%	7.8%	-5.5%	1.9%	9.7%	-7.8%
Metropolitan Low-Income	64.3%	2.5%	61.7%	65.6%	2.2%	63.3%	66.4%	3.2%	63.2%
Metropolitan Non Low-Income	22.6%	20.9%	1.7%	23.5%	22.6%	0.9%	22.9%	20.5%	2.4%
Teacher Quality									
Master's Degree or above	29.9%	36.3%	-6.4%	33.3%	37.2%	-3.9%	47.5%	40.6%	6.9%
Bachelor's degree only	70.1%	63.7%		66.7%	62.8%		52.5%	59.4%	
Two years of experience or less	12.6%	7.8%	4.7%	9.5%	6.1%	3.4%	4.6%	5.1%	-0.5%
More than two years of teaching experien	87.4%	92.2%		90.5%	93.9%		95.4%	94.9%	
Class Size									
Small	27.3%	29.5%	-2.2%	28.5%	30.6%	-2.1%	29.6%	33.2%	-3.5%
Normal	43.3%	35.9%	7.4%	35.9%	33.5%	2.4%	31.3%	29.1%	2.2%
Normal with aide	29.4%	34.7%	-5.3%	35.6%	35.9%	-0.3%	39.0%	37.7%	1.3%

Source: Health & Education Research Operative Service, the State of Tennessee's Student/Teacher Achievement Ratio (STAR) Project (1985-1989).

Table 4b. Percentages of Black and White Students by Race of Teacher and Residence

	Grade 1				Grade 2			
	Black Students (N=2084)		White Students (N=4078)		Black Students (N=1987)		White Students (N=3655)	
	with Black Teachers	with White Teachers	with Black Teachers	with White Teachers	with Black Teachers	with White Teachers	with Black Teachers	with White Teachers
Percentage of Students	44.1%	55.9%	4.6%	95.4%	43.1%	56.9%	8.9%	91.1%
Residence								
Rural	0.9%	16.2%	36.6%	65.8%	3.2%	12.8%	49.7%	69.1%
Non-Metropolitan Urban	0.9%	6.1%	8.6%	12.4%	0.1%	3.9%	3.4%	8.2%
Metropolitan Low-Income	76.7%	54.4%	16.7%	1.8%	73.2%	59.8%	11.3%	1.4%
Metropolitan Non Low-Income	21.5%	23.4%	38.2%	20.0%	23.5%	23.5%	35.6%	21.4%
N	920	1164	186	3892	856	1131	326	3329

	Grade 3			
	Black Students (N=1916)		White Students (N=3767)	
Percentage of Students	with Black Teachers	with White Teachers	with Black Teachers	with White Teachers
	51.9%	48.1%	6.6%	93.4%
Residence				
Rural				
Non-Metropolitan Urban	1.2%	16.9%	34.0%	68.9%
Metropolitan Low-Income	0.0%	4.0%	0.0%	10.4%
Metropolitan Non Low-Income	76.0%	56.1%	21.1%	1.9%
N	22.8%	23.0%	44.9%	18.8%
	994	922	247	3520

Source: Health & Education Research Operative Service, the State of Tennessee's Student/Teacher Achievement Ratio (STAR) Project (1985-1989).

Table 5a. Differences in Mean Reading Scores between Black and White Primary School Students by Race, Gender, Family Income, Residence, and Teacher Quality

	GRADE 1					GRADE 2					GRADE 3				
	Black Students		White Students		Mean Gap	Black Students		White Students		Mean Gap	Black Students		White Students		Mean Gap
Gender	Mean	SD	Mean	SD		Mean	SD	Mean	SD		Mean	SD	Mean	SD	
Male	35.2	24.9	54.4	28.3	-19.2	33.9	24.5	55.6	28.1	-21.7	36.0	26.1	54.7	28.6	-18.7
Female	41.4	25.9	60.7	26.7	-19.2	39.8	24.7	61.9	26.7	-22.1	42.7	27.0	60.2	27.3	-17.5
Gap:	-6.2		-6.3			-5.9		-6.3			-6.7		-5.5		
Family Income															
Poor (Free/Reduced Lunch)	34.8	24.4	47.2	28.1	-12.4	33.7	23.4	49.5	28.7	-15.7	36.4	25.9	48.0	27.9	-11.6
Non Poor	53.6	25.3	62.8	25.9	-9.2	48.9	26.1	63.0	26.0	-14.0	51.3	27.0	61.8	27.1	-10.5
Gap:	-18.8		-15.6			-15.2		-13.5			-14.9		-13.9		
Residence															
Rural	44.0	26.4	57.6	27.8	-13.6	45.9	25.4	58.6	27.8	-12.7	42.7	23.9	57.2	27.9	-14.5
Non-Metropolitan Urban	44.7	27.4	55.6	29.2	-10.9	47.4	25.7	59.6	27.5	-12.3	41.7	27.6	54.5	29.2	-12.8
Metropolitan Low-Income	33.5	23.7	42.5	28.2	-9.0	33.0	23.4	38.6	24.1	-5.6	36.1	26.1	41.8	27.4	-5.7
Metropolitan Non Low-Income	48.3	26.6	59.6	25.9	-11.3	43.0	25.9	60.2	26.6	-17.2	47.2	27.7	61.5	27.1	-14.3
Teacher Quality															
Master's Degree	38.6	26.2	58.3	27.0	-19.8	36.0	24.6	58.1	27.9	-22.1	42.8	26.8	57.0	28.2	-14.2
No Master's degree	38.1	25.3	56.8	28.1	-18.7	37.2	24.8	58.9	27.4	-21.7	36.2	26.3	57.5	28.1	-21.3
Gap:	0.4		1.5			-1.2		-0.8			6.6		-0.5		
Two years of experience or less	38.4	26.3	56.2	28.5	-17.9	31.3	25.1	54.5	28.1	-23.2	35.2	23.5	56.3	28.3	-21.1
More than two years of teaching experience	38.2	25.5	57.5	27.7	-19.2	37.4	24.7	58.9	27.6	-21.5	39.5	26.9	57.4	28.1	-17.9
Gap:	0.1		-1.2			-6.1		-4.3			-4.3		-1.0		

Table 5b. Differences in Mean Math Scores between Black and White Primary School Students by Race, Gender, Family Income, Residence, and Teacher Quality

	GRADE 1					GRADE 2					GRADE 3				
	Black Students		White Students		Mean Gap	Black Students		White Students		Mean Gap	Black Students		White Students		Mean Gap
Gender	Mean	SD	Mean	SD		Mean	SD	Mean	SD		Mean	SD	Mean	SD	
Male	37.7	26.6	58.1	27.6	-20.3	37.1	25.6	57.6	27.7	-20.6	38.8	26.9	56.6	28.4	-17.8
Female	39.1	26.4	57.7	27.3	-18.6	37.9	25.7	57.9	26.9	-20.0	39.7	26.6	56.6	27.1	-16.9
Gap:	-1.3		0.4			-0.8		-0.3			-0.9		-0.1		
Family Income															
Poor (Free/Reduced Lunch)	35.8	25.9	48.9	28.0	-13.1	35.2	25.0	50.5	28.3	-15.3	36.3	25.8	47.5	27.6	-11.2
Non Poor	49.9	26.3	62.6	25.9	-12.7	46.4	26.0	61.2	26.2	-14.8	50.9	27.3	61.0	26.8	-10.1
Gap:	-14.2		-13.7			-11.2		-10.7			-14.6		-13.5		
Residence															
Rural	40.4	26.1	58.4	27.3	-18.0	42.6	25.3	58.8	27.4	-16.2	40.3	25.6	56.8	27.9	-16.5
Non-Metropolitan Urban	35.4	24.5	54.8	29.9	-19.4	32.8	22.6	54.8	27.2	-21.9	36.0	25.1	53.5	27.7	-17.6
Metropolitan Low-Income	35.6	25.9	45.2	24.5	-9.6	36.2	25.9	37.5	23.0	-1.3	36.6	26.4	37.7	24.1	-1.1
Metropolitan Non Low-Income	45.9	27.2	59.6	26.3	-13.8	39.6	25.0	57.8	26.7	-18.2	46.6	27.1	60.2	26.7	-13.6
Teacher Quality															
Master's Degree	38.5	26.6	59.3	27.6	-20.7	35.8	26.2	56.9	27.6	-21.1	42.4	26.9	57.4	27.9	-14.9
No Master's degree	38.3	26.5	57.1	27.4	-18.8	38.3	25.3	58.3	27.2	-20.0	36.3	26.2	56.1	27.7	-19.8
Gap:	0.2		2.2			-2.4		-1.4			6.1		1.3		
Two years of experience or less	45.1	29.0	56.2	27.8	-11.1	42.0	27.1	56.4	27.5	-14.4	37.7	25.4	57.4	27.9	-19.7
More than two years of teaching experience	37.4	26.0	58.0	27.4	-20.6	37.0	25.4	57.9	27.3	-20.9	39.3	26.8	56.5	27.8	-17.3
Gap:	7.7		-1.8			5.0		-1.4			-1.6		0.9		

Source: Health & Education Research Operative Service, the State of Tennessee's Student/Teacher Achievement Ratio (STAR) Project (1985-1989).

Table 7a. First Grade Models of Student Performance by Student-Teacher Race Interaction, Gender, Socioeconomic Background and Class Size

First Grade Reading

Background Variables Race	Model 1	Model 1'	Model 2	Model 3	Model 4	Model 5
Black Students (White Students referent)	-19.21*** (.72)					
White Students & White Teacher		--	--	--	--	--
White Students & Black Teacher		-10.88*** (2.00)	-6.50** (1.94)	-6.36** (1.94)	-5.94** (1.94)	-3.16 (2.21)
Black Students & White Teacher		-20.98*** (.89)	-8.24*** (1.11)	-8.16*** (1.11)	-7.57*** (1.11)	-6.19*** (1.27)
Black Students & Black Teacher		-18.10*** (.98)	-2.39+ (1.34)	-2.21 (1.34)	-2.40+ (1.34)	.56 (1.63)
Gender						
Male	--	--	--	--	--	--
Female	6.26*** (.68)	6.30*** (.68)	6.31*** (.65)	6.32*** (.65)	6.31*** (.65)	6.06*** (.61)
Family Income						
Poor (Free/Reduced Lunch)			-15.10*** (.75)	-15.08*** (.75)	-15.00*** (.75)	-14.20*** (.77)
Non Poor			--	--	--	--
Residence						
Rural			.96 (1.17)	1.03 (1.17)	1.20 (1.16)	--
Non Metro Urban			--	--	--	--
Metro Low Income			-10.09*** (1.52)	-10.10*** (1.52)	-9.98*** (1.51)	--
Metro non Low Income			-.07 (1.31)	-.04 (1.31)	.10 (1.31)	--
Teacher Quality						
MA or higher				1.15 (.70)	.99 (.70)	-.33 (.76)
BA Only				--	--	--
Two years of experience or less				--	--	--
More than two years of teaching experience				.02 (1.14)	-.40 (1.14)	2.06 (1.31)
Class Size						
Small					6.51*** (.80)	6.80*** (.78)
Normal					--	--
Normal with full-time aide					2.64** (.78)	2.06** (.76)
School Fixed Effects	no	no	no	no	no	yes
BIC	58054.97	58037.23	57513.34	57527.97	57480.18	56686.39
log likelihood ratio	-29014.39	-28996.8	-28717.4	-28715.99	-28683.37	-28299.56
R-squared	0.1130	0.1181	0.1945	0.1949	0.2034	0.2967
N	6162	6162	6162	6162	6162	6162

Notes:
-- represents omitted category
standard errors are reported in parentheses
+ P<.10
* P<.05
** P<.01
*** P<.001

Model 1 includes race and gender
Model 2 adds family income and residence
Model 3 adds teacher quality
Model 4 adds class size
Model 5 adds school fixed effects

First Grade Math

Background Variables Race	Model 1	Model 1'	Model 2	Model 3	Model 4	Model 5
Black Students (White Students referent)	-19.49*** (.73)					
White Students & White Teacher		--	--	--	--	--
White Students & Black Teacher		-11.58*** (2.02)	-8.21*** (1.99)	-8.07*** (1.99)	-7.49*** (1.98)	-6.07** (1.99)
Black Students & White Teacher		-23.13*** (.90)	-13.46*** (1.14)	-13.32*** (1.14)	-12.85*** (1.14)	-14.96*** (1.26)
Black Students & Black Teacher		-16.08*** (.99)	-4.44** (1.37)	-4.35** (1.38)	-4.51*** (1.37)	-6.68*** (1.64)
Gender						
Male	--	--	--	--	--	--
Female	.19 (.69)	.27 (.68)	.28 (.66)	.28 (.66)	.25 (.66)	-.00 (.61)
Family Income						
Poor (Free/Reduced Lunch)			-12.93*** (.77)	-12.91*** (.77)	-12.79*** (.77)	-11.50*** (.76)
Non Poor			--	--	--	--
Residence						
Rural			3.16** (1.20)	3.28** (1.20)	3.41** (1.19)	--
Non Metro Urban			--	--	--	--
Metro Low Income			-4.22** (1.56)	-4.32** (1.56)	-4.27** (1.55)	--
Metro non Low Income			2.17 (1.35)	2.15 (1.35)	2.19 (1.34)	--
Teacher Quality						
MA or higher				1.82* (.72)	1.75* (.71)	1.31+ (.76)
BA Only				--	--	--
Two years of experience or less				--	--	--
More than two years of teaching experience				-2.89* (1.17)	-3.20 (1.17)	1.36 (1.26)
Class Size						
Small					6.66*** (.82)	7.46*** (.78)
Normal					--	--
Normal with full-time aide					1.07 (.79)	.50 (.75)
School Fixed Effects	no	no	no	no	no	yes
BIC	58197.13	58147.22	57816.66	57823.81	57769.5	56665.6
log likelihood ratio	-29085.47	-29051.79	-28869.06	-28863.91	-28828.03	-28289.17
R-squared	0.1035	0.1132	0.1643	0.1657	0.1753	0.3077
N	6162	6162	6162	6162	6162	6162

Notes:
-- represents omitted category
standard errors are reported in parentheses
+ P<.10
* P<.05
** P<.01
*** P<.001

Model 1 includes race and gender
Model 2 adds family income and residence
Model 3 adds teacher quality
Model 4 adds class size
Model 5 adds school fixed effects

Table 7b. Second Grade Models of Student Performance by Student-Teacher Race Interaction, Gender, Socioeconomic Background and Class Size

Second Grade Reading

Background Variables Race	Model 1	Model 1'	Model 2	Model 3	Model 4	Model 5
Black Students (White Students referent)	-21.88***					
	(.73)					
White Students & White Teacher			--	--	--	--
White Students & Black Teacher		-5.91***	-5.57***	-5.57***	-5.81***	-5.95***
		(1.53)	(1.49)	(1.50)	(1.49)	(1.70)
Black Students & White Teacher		-22.71***	-10.73***	-10.81***	-10.68***	-7.74***
		(.90)	(1.19)	(1.19)	(1.18)	(1.40)
Black Students & Black Teacher		-22.00***	-8.74***	-9.15***	-8.81***	-5.37**
		(1.01)	(1.37)	(1.38)	(1.37)	(1.68)
Gender						
Male	--		--	--	--	--
Female	6.15***	6.14***	5.98***	5.94***	5.91***	5.96***
	(.70)	(.70)	(.68)	(.68)	(.67)	(.65)
Family Income						
Poor (Free/Reduced Lunch)			-13.23***	-13.23***	-13.07***	-13.97***
			(.79)	(.79)	(.79)	(.84)
Non Poor			--	--	--	--
Residence						
Rural			-1.20	-1.34	-.67	--
			(1.49)	(1.50)	(1.50)	
Non Metro Urban			--	--	--	--
Metro Low Income			-10.89***	-10.57***	-10.21***	--
			(1.81)	(1.82)	(1.81)	
Metro non Low Income			-3.45*	-3.61*	-3.035+	--
			(1.61)	(1.62)	(1.62)	
Teacher Quality						
MA or higher				-.65	-.59	-.52
				(.72)	(.72)	(.88)
BA Only				--	--	--
Two years of experience or less				--	--	--
More than two years of teaching experience				3.31*	3.24*	3.67*
				(1.35)	(1.35)	(1.62)
Class Size						
Small					5.82***	6.48***
					(.85)	(.84)
Normal					--	--
Normal with full-time aide					2.20**	2.03*
					(.82)	(.80)
School Fixed Effects	no	no	no	no	no	yes
BIC	52997.18	52999.22	52647.1	52658.07	52628.01	52149.97
log likelihood ratio	-26485.63	-26478.01	-26284.68	-26281.53	-26257.86	-26031.79
R-squared	0.1442	0.1465	0.2030	0.2039	0.2106	0.2714
N	5642	5642	5642	5642	5642	5642

Notes:
-- represents omitted category
standard errors are reported in parentheses
+ P<.10
* P<.05
** P<.01
*** P<.001

Model 1 includes race and gender
Model 2 adds family income and residence
Model 3 adds teacher quality
Model 4 adds class size
Model 5 adds school fixed effects

Second Grade Math

Background Variables Race	Model 1	Model 1'	Model 2	Model 3	Model 4	Model 5
Black Students (White Students referent)	-20.31***					
	(.74)					
White Students & White Teacher			--	--	--	--
White Students & Black Teacher		-6.58***	-6.57***	-6.38***	-6.38***	-5.26**
		(1.55)	(1.53)	(1.54)	(1.53)	(1.64)
Black Students & White Teacher		-21.00***	-12.76***	-12.74***	-12.63***	-12.14***
		(.91)	(1.22)	(1.22)	(1.22)	(1.34)
Black Students & Black Teacher		-20.75***	-12.04***	-11.90***	-11.66***	-10.88***
		(1.02)	(1.40)	(1.41)	(1.41)	(1.65)
Gender						
Male	--	--	--	--	--	--
Female	.46	.46	.39	.41	.38	.55
	(.71)	(.71)	(.69)	(.69)	(.69)	(.65)
Family Income						
Poor (Free/Reduced Lunch)			-11.10***	-11.09***	-10.96***	-12.36***
			(.81)	(.81)	(.81)	(.81)
Non Poor			--	--	--	--
Residence						
Rural			4.60**	4.39**	4.94**	--
			(1.53)	(1.54)	(1.54)	
Non Metro Urban			--	--	--	--
Metro Low Income			-.74	-1.18	-.87	--
			(1.85)	(1.87)	(1.86)	
Metro non Low Income			.78	.55*	1.02**	--
			(1.65)	(1.66)	(1.66)	
Teacher Quality						
MA or higher				-.68	-.61	-.40
				(.74)	(.74)	(.81)
BA Only				--	--	--
Two years of experience or less				--	--	--
More than two years of teaching experience				-1.98	-1.98	-.90
				(1.38)	(1.39)	(1.61)
Class Size						
Small					4.65***	5.93***
					(.87)	(.84)
Normal					--	--
Normal with full-time aide					1.51+	-1.77*
					(.84)	(.79)
School Fixed Effects	no	no	no	no	no	yes
BIC	53119.85	53119.03	52933.3	52947.17	52935.24	52079.42
log likelihood ratio	-26546.97	-26537.92	-26427.78	-26426.08	-26411.47	-25996.52
R-squared	0.1163	0.1191	0.1528	0.1534	0.1577	0.2729
N	5642	5642	5642	5642	5642	5642

Notes:
-- represents omitted category
standard errors are reported in parentheses
+ P<.10
* P<.05
** P<.01
*** P<.001

Model 1 includes race and gender
Model 2 adds family income and residence
Model 3 adds teacher quality
Model 4 adds class size
Model 5 adds school fixed effects

Table 7c. Third Grade Models of Student Performance by Student-Teacher Race Interaction, Gender, Socioeconomic Background and Class Size

Third Grade Reading

Background Variables Race	Model 1	Model 1'	Model 2	Model 3	Model 4	Model 5
Black Students (White Students referent)	-18.07*** (.77)					
White Students & White Teacher	--	--	--	--	--	--
White Students & Black Teacher		-3.39+ (1.80)	-2.43 (1.81)	-2.47 (1.81)	-2.60 (1.80)	.67 (2.11)
Black Students & White Teacher		-18.04*** (1.01)	-8.80*** (1.27)	-8.87*** (1.27)	-8.71*** (1.27)	-6.97*** (1.48)
Black Students & Black Teacher		-18.54*** (0.98)	-7.31*** (1.43)	-7.52*** (1.43)	-7.88*** (1.43)	-6.40** (1.84)
Gender						
Male	--		--	--	--	--
Female	5.89*** (.72)	5.90*** (.72)	5.76*** (.70)	5.77*** (.70)	5.69*** (.70)	5.89*** (.69)
Family Income						
Poor (Free/Reduced Lunch)			-12.89*** (.82)	-12.88*** (.82)	-12.76*** (.82)	-12.39*** (.86)
Non Poor			--	--	--	--
Residence						
Rural			1.82 (1.42)	1.95 (1.43)	2.44+ (1.43)	--
Non Metro Urban			--	--	--	--
Metro Low Income			-4.83** (1.81)	-4.58* (1.82)	-3.86* (1.82)	--
Metro non Low Income			2.82+ (1.60)	2.82+ (1.61)	3.45* (1.61)	--
Teacher Quality						
MA or higher				.88 (.73)	1.43+ (.74)	1.48+ (.88)
BA Only				--	--	--
Two years of experience or less				--	--	--
More than two years of teaching experience				.53 (1.67)	.37 (1.66)	.79 (1.90)
Class Size						
Small					4.89*** (.90)	5.24*** (.91)
Normal					--	--
Normal with full-time aide					.19 (.86)	-.28 (.85)
School Fixed Effects	no	no	no	no	no	yes
BIC	53817.89	53831.5	53522.55	53538.11	53516.21	53150.53
log likelihood ratio	-26895.98	-26894.1	-26722.37	-26721.5	-26701.91	-26532.04
R-squared	0.0969	0.0975	0.1504	0.1507	0.1565	0.2054
N	5683	5683	5683	5683	5683	5683

Notes:
-- represents omitted category
standard errors are reported in parentheses
+ P<.10
* P<.05
** P<.01
*** P<.001

Model 1 includes race and gender
Model 2 adds family income and residence
Model 3 adds teacher quality
Model 4 adds class size
Model 5 adds school fixed effects

Third Grade Math

Background Variables Race	Model 1	Model 1'	Model 2	Model 3	Model 4	Model 5
Black Students (White Students referent)	17.38*** (.76)					
White Students & White Teacher	--	--	--	--	--	--
White Students & Black Teacher		-6.79*** (1.80)	-6.05** (1.80)	-6.04** (1.80)	-6.17** (1.80)	-4.47* (2.02)
Black Students & White Teacher		-17.16*** (1.01)	-8.54*** (1.27)	-8.74*** (1.27)	-8.61*** (1.27)	-8.33*** (1.47)
Black Students & Black Teacher		-18.44*** (.98)	-8.04*** (1.42)	-8.42*** (1.43)	-8.73*** (1.43)	-10.23*** (1.79)
Gender						
Male	--		--	--	--	--
Female	.35 (.72)	.37 (.72)	.25 (.70)	.26 (.70)	.18 (.70)	.37 (.67)
Family Income						
Poor (Free/Reduced Lunch)			-12.64*** (.82)	-12.61*** (.82)	-12.52*** (.82)	-11.69*** (.84)
Non Poor			--	--	--	--
Residence						
Rural			2.66+ (1.42)	3.02* (1.42)	3.33* (1.43)	--
Non Metro Urban			--	--	--	--
Metro Low Income			-3.16+ (1.81)	-2.56 (1.82)	-2.03 (1.82)	--
Metro non Low Income			3.55* (1.60)	3.79* (1.60)	4.19** (1.61)	--
Teacher Quality						
MA or higher				2.00** (.73)	2.54** (.74)	1.40+ (.85)
BA Only				--	--	--
Two years of experience or less				--	--	--
More than two years of teaching experience				-1.98** (1.66)	-2.04** (1.66)	-.17 (1.83)
Class Size						
Small					3.56*** (.90)	4.02*** (.88)
Normal					--	--
Normal with full-time aide					-.79 (.86)	-.54 (.84)
School Fixed Effects	no	no	no	no	no	yes
BIC	53789.32	53791.36	53502.63	53511.87	53500.81	52875.22
log likelihood ratio	-26881.69	-26874.07	-26712.41	-26708.39	-26694.21	-26394.38
R-squared	0.0824	0.0848	0.1355	0.1367	0.1410	0.2270
N	5683	5683	5683	5683	5683	5683

Notes:
-- represents omitted category
standard errors are reported in parentheses
+ P<.10
* P<.05
** P<.01
*** P<.001

Model 1 includes race and gender
Model 2 adds family income and residence
Model 3 adds teacher quality
Model 4 adds class size
Model 5 adds school fixed effects

Table 8. The Impact of a Black Teacher on Black and White Student Achievement

READING SCORES

	First Grade				Second Grade				Third Grade			
	Model 2	Model 3	Model 4	Model 5	Model 2	Model 3	Model 4	Model 5	Model 2	Model 3	Model 4	Model 5
Impact of a black teacher on white student achievement (β 01)	-6.50**	-6.36**	-5.94**	-3.16	-5.57***	-5.57***	-5.81***	-5.95***	-2.43	-2.47	-2.60	.67
	(.001)	(.001)	(.002)	(.152)	(.000)	(.000)	(.000)	(.000)	(.180)	(.172)	(.149)	(.748)
Impact of a black teacher on black student achievement (β 11 - β 10)	5.85***	5.95***	5.17***	5.63***	1.99+	1.66	1.87	2.37*	1.49	1.35	0.83	0.57
	(.000)	(.000)	(.000)	(.000)	(0.090)	(0.160)	(0.112)	(0.043)	(0.231)	(0.274)	(0.492)	(0.666)
What is the net benefit of having a black teacher? (β 01 + (β 11 - β 10))	-.65***	-.41***	-.77***	2.47***	-3.58***	-3.91***	-3.94***	-3.58***	-.94+	-1.12+	-1.77	1.24
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(0.068)	(0.075)	(0.109)	(0.962)

MATH SCORES

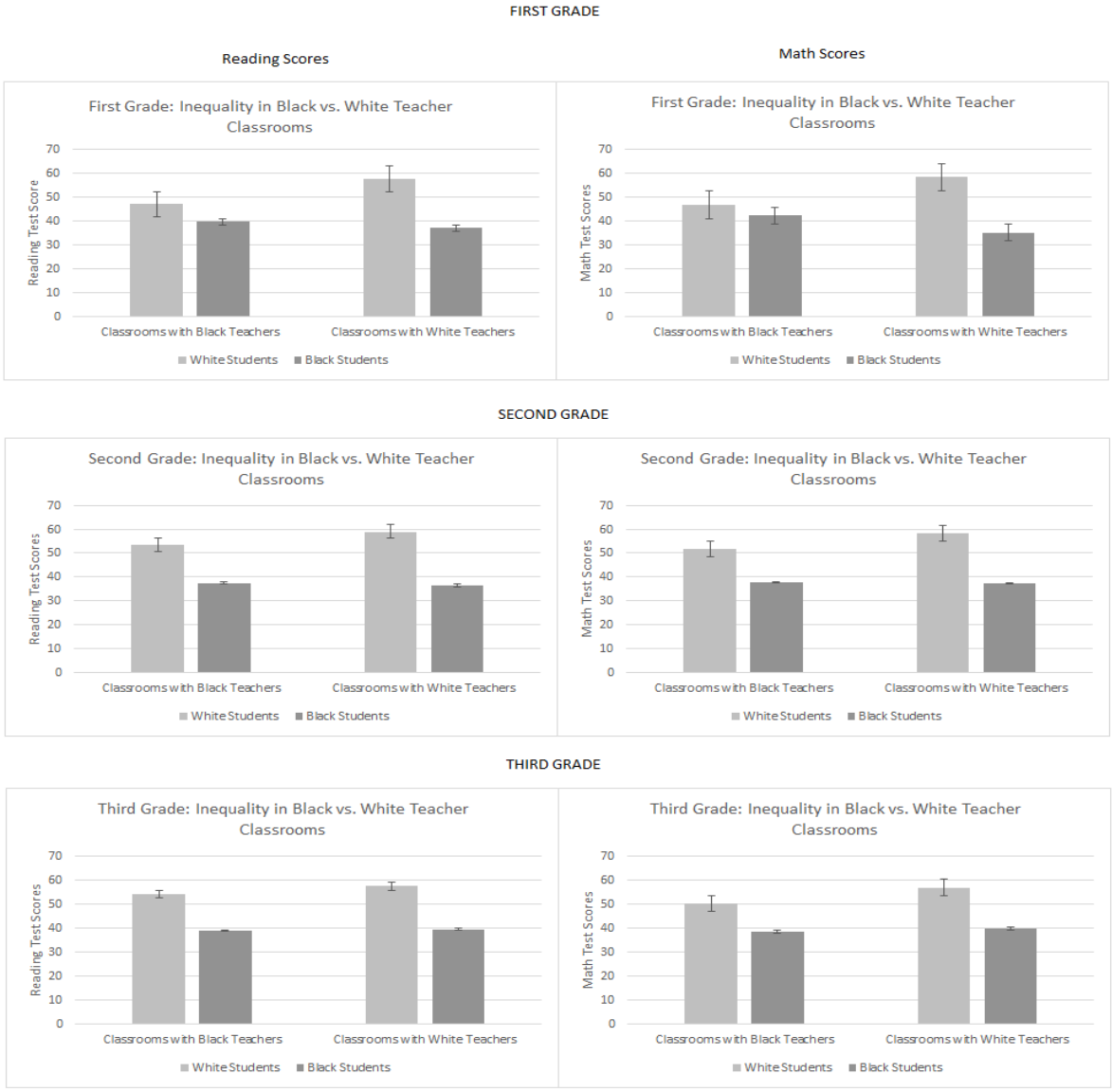
	First Grade				Second Grade				Third Grade			
	Model 2	Model 3	Model 4	Model 5	Model 2	Model 3	Model 4	Model 5	Model 2	Model 3	Model 4	Model 5
Impact of a black teacher on white student achievement (β 01)	-8.21***	-8.07***	-7.49***	-6.07**	-6.57***	-6.38***	-6.38***	-5.26**	-6.05**	-6.04**	-6.17**	-4.47*
	(.000)	(.000)	(.000)	(.002)	(.000)	(.000)	(.000)	(.001)	(.001)	(.001)	(.001)	(.027)
Impact of a black teacher on black student achievement (β 11 - β 10)	9.02***	8.97***	8.34***	8.28***	0.72	0.84	0.97	1.26	0.50	0.32	-0.12	-1.90
	(.000)	(.000)	(.000)	(.000)	(0.547)	(0.503)	(0.431)	(0.299)	(0.687)	(0.800)	(0.922)	(0.148)
What is the net benefit of having a black teacher? (β 01 + (β 11 - β 10))	0.81***	.90***	.85***	2.21***	-5.85***	-5.54***	-5.41***	-4.00***	-5.55**	-5.72**	-6.29**	-6.37
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(0.002)	(0.003)	(0.004)	(0.271)

Notes:

p-values are reported in parentheses
+ P<.10
* P<.05
** P<.01
*** P<.001

Model 6: Race of Student and Race of Teacher Net of Gender, Family Income, and Residence
Model 7: Model 6 plus Teacher Quality
Model 8: Model 7 plus Class Size
Model 9: Model 8 with School Fixed Effects
Model 5 adds school fixed effects

Figure 2. Mean Student Test Scores by Race of Teacher and Grade Level



Source: Health & Education Research Operative Service, the State of Tennessee's Student/Teacher Achievement Ratio (STAR) Project (1985-1989).

APPENDIX A

Why Use Test Scores as Measures of Achievement Differences?

Since the Coleman Report, a large body of research on the apparent achievement gaps between black and white students has emerged. While a number of different measures of achievement have been used, including high school graduation rates, college-enrollment rates, college-completion rates, course grades, dropout rates, measures of absenteeism, and behavioral evaluations; test scores have gained currency both politically and methodologically among educational researchers as a preferred measure of student achievement.

Early studies of achievement revealed that children from families with lower socioeconomic statuses and black students consistently score lower than white students on year-end exams (Phillips 1998). Given the rise of standards-based educational reforms after the reauthorization of the Elementary and Secondary Education Act in the form of the No Child Left Behind Act (2001), standardized test scores have become the main metric for decisions about school funding, school closures, and teacher employment. Specifically, when adequate yearly progress is not met, evaluated according to yearly test score gains of students, schools face restructuring or closure, and teachers are often fired in the pursuit of higher scores. In this way, test scores are enormously consequential to the functioning of schools and the teachers that staff them, as test scores are the main measure of teacher and school effectiveness.

Although there has been considerable debate regarding the accuracy of these student assessments in measuring the academic skills learned during the school year (see Rothstein 2004), and the accuracy of test score gains as metrics for teacher quality (see Baker et al. 2010; Rothstein 2010; Darling-Hammond et al. NAE 2011), test scores have become the most widely used measurement of school and teacher effectiveness among researchers of education. This is due in part to the pervasive use of standardized tests at nearly every grade level of students' scholastic careers. As a consequence of their widespread use, researchers have been able to illustrate that test scores are consequential for a number of later educational and labor market outcomes that are important to social mobility like black-white wage differentials (Neal & Johnson, 1996; O'Neill, 1990), the likelihood of dropping out of high-school or attending college (Alexander et al. 2007), and educational attainment (Enwistle et al. 2005), and in particular, researchers have demonstrated that test scores in the earliest years of schooling are predictive of later academic and labor market outcomes. To summarize, test scores allow researchers to trace differences in achievement over the course of students' educational trajectories, and even link test scores to post-scholastic outcomes.

Table A1a. Differences in Mean Reading Scores between Black and White Primary School Students by Race of Teacher and Residence

	GRADE 1					GRADE 2					GRADE 3				
	Black Students		White Students		Mean Gap	Black Students		White Students		Mean Gap	Black Students		White Students		Mean Gap
Race of Teacher	Mean	SD	Mean	SD		Mean	SD	Mean	SD		Mean	SD	Mean	SD	
Black	39.8	25.5	47.1	28.7	-7.3	37.4	23.1	53.3	28.1	-15.9	39.0	26.9	54.3	29.1	-15.3
White	37.1	25.6	57.9	27.6	-20.8	36.4	25.9	59.1	27.5	-22.8	39.6	26.6	57.5	28.0	-17.9
Gap:	2.7		-10.8			1.0		-5.9			-0.6		-3.2		
Residence and Race of Teacher															
Black:															
Rural	42.9	27.9	55.7	28.2	-12.8	49.9	24.6	54.5	27.4	-4.6	44.5	23.8	55.2	25.4	-10.7
Non-Metropolitan Urban	35.0	33.7	38.5	28.7	-3.5	48.7	0.0	54.5	27.4	-5.9	0.0	0.0	0.0	0.0	0.0
Metropolitan Low-Income	37.1	24.2	37.0	30.6	0.0	36.5	22.7	45.7	23.5	-9.2	36.6	26.1	40.4	28.8	-3.8
Metropolitan Non Low-Income	49.5	27.1	45.2	26.6	4.3	38.4	23.8	54.5	30.5	-16.1	46.7	28.1	60.1	29.9	-13.4
N															
White:															
Rural	44.0	26.5	57.7	27.8	-13.6	45.2	25.6	58.9	27.8	-13.7	42.5	24.0	57.2	28.0	-14.7
Non-Metropolitan Urban	45.8	26.7	56.1	29.1	-10.3	47.3	26.0	60.1	27.5	-12.8	41.7	27.6	54.5	29.2	-12.8
Metropolitan Low-Income	29.6	22.4	44.8	26.9	-15.2	29.8	23.5	32.8	23.2	-3.0	35.3	26.2	42.8	26.4	-7.5
Metropolitan Non Low-Income	47.4	26.3	60.9	25.4	-13.5	46.5	26.9	61.2	25.8	-14.7	47.7	27.4	61.7	26.7	-14.0
N															

Table A1b. Mean Math Scores of Primary School Students by Race of Teacher and Residence

	GRADE 1					GRADE 2					GRADE 3				
	Black Students		White Students		Mean Gap	Black Students		White Students		Mean Gap	Black Students		White Students		Mean Gap
Race of Teacher	Mean	SD	Mean	SD		Mean	SD	Mean	SD		Mean	SD	Mean	SD	
Black	39.8	25.5	47.1	28.7	-7.3	37.4	23.1	53.3	28.1	-15.9	39.0	26.9	54.3	29.1	-15.3
White	37.1	25.6	57.9	27.6	-20.8	36.4	25.9	59.1	27.5	-22.8	39.6	26.6	57.5	28.0	-17.9
Gap:	2.7		-10.8			1.0		-5.9			-0.6		-3.2		
Residence and Race of Teacher															
Black:															
Rural	26.5	22.5	52.9	26.9	-26.4	48.8	25.4	56.0	28.6	-7.2	31.0	19.0	48.4	24.3	-17.4
Non-Metropolitan Urban	18.8	14.9	20.4	17.1	-1.6	31.2	0.0	47.0	25.6	-15.8	0.0	0.0	0.0	0.0	0.0
Metropolitan Low-Income	40.1	27.0	48.7	29.2	-8.7	38.2	26.3	36.6	21.3	1.6	36.6	26.4	36.8	23.9	-0.1
Metropolitan Non Low-Income	51.9	27.6	46.1	25.9	5.9	34.2	21.8	51.2	29.7	-17.0	45.5	27.7	57.9	29.2	-12.5
N															
White:															
Rural	41.0	26.1	58.5	27.3	-17.6	41.4	25.2	59.0	27.3	-17.5	41.1	25.9	57.1	27.9	-16.1
Non-Metropolitan Urban	37.2	24.7	56.0	29.5	-18.7	32.9	22.9	55.1	27.2	-22.2	36.0	25.1	53.5	27.7	-17.6
Metropolitan Low-Income	30.7	23.8	43.7	22.2	-13.0	34.3	25.3	38.3	24.5	-4.0	36.6	26.4	38.4	24.5	-1.8
Metropolitan Non Low-Income	41.4	26.1	60.9	26.0	-19.5	43.6	26.5	58.9	26.1	-15.3	47.8	26.6	60.5	26.3	-12.8
N															

Source: Health & Education Research Operative Service, the State of Tennessee's Student/Teacher Achievement Ratio (STAR) Project (1985-1989).