The Long Road to Equality: A Meta-analysis of the African American Test Score Gap

Nick Huntington-Klein

California State University, Fullerton

Elizabeth Ackert

University of Texas at Austin

Author Note

Nick Huntington-Klein, Department of Economics, California State University, Fullerton

Elizabeth Ackert, Population Research Center, University of Texas at Austin

Research funds and support for this project have been provided by Collaborative Researchers for Education Sciences Training, Institute of Education Sciences Grant #R305B090012. Support has also been provided by a Eunice Kennedy Shriver National Institute of Child Health and Human Development grant, R24 HD042828, and from a Shanahan Endowment Fellowship and a Eunice Kennedy Shriver National Institute of Child Health and Human Development training grant, T32 HD007543. We thank Pasita Chaijaroen and Andre Stephens for their research assistance and Ashley Jochim, Sean Reardon, Stew Tolnay, William Zumeta, and session participants at the Association for Education Finance and Policy conference for comments.

Correspondence concerning this article should be addressed to Nick Huntington-Klein, Department of Economics, California State University, Fullerton, 800 N. State College Blvd., Fullerton, CA, 92831. Contact: nhuntington-klein@fullerton.edu.

Abstract

This study performs a meta-regression analysis of over 1,100 regressions in 165 studies to examine the relationship between African American racial status and student achievement scores in K-12 education from 1979 to 2010. The study examines time trends in the black test score gap and estimates the extent to which controls for confounding variables including socioeconomic status and schooling characteristics attenuate the size of the gap. Across the samples in the study, the absolute relationship between Black status and achievement decreased during the 1980s and early 1990s, but has been stagnant since the late 1990s. We estimate that socioeconomic status alone explains more than half of the gap, and this influence does not vary significantly over the time period of interest. Controlling for differences in school characteristics only reduces the gap slightly, but school-level factors explain an increasing proportion of the gap over time.

The Long Road to Equality: A Meta-analysis of the African American Test Score Gap

Introduction

Despite research and policy efforts to lessen the influence of ascriptive background characteristics on student achievement, race continues to be a defining factor shaping patterns of educational inequality among students in U.S. schools. One expression of the influence of race on educational outcomes is the Black-White test score gap, which remains sizeable (Jencks and Phillips 1998; Hedges and Nowell 1999; Magnuson and Waldfogel 2008; Reardon et al. 2013; Reardon, Robinson-Cimpian, and Weathers 2014). Enduring test score inequalities between Black students and their peers will have long-term consequences for racial/ethnic stratification in adult life.

This study seeks to characterize changes in the Black test score gap over time and to estimate the potential contributions of confounding variables, such as household socioeconomic status and school characteristics, to the gap from the period of 1979 to 2010. Previous work shows an overall shrinking trend in the Black test score gap from the mid-1960s through 2010 (Grissmer, Flanagan, and Williamson 1998; Hedges and Nowell 1999; Berends and Peñaloza 2008; Reardon et al. 2013; Reardon, Robinson-Cimpian, and Weathers 2014; Magnuson and Waldfogel 2008). However, questions remain about the rate of gap closure and the existence of periods of stall or widening in the gap. Some studies show a pattern of what Magnuson and Waldfogel (2008) refer to as "steady gains and stalled progress," with a period of gap closure in the 1970s to 1980s, but stagnation or widening in the gap in the late 1980s and 1990s (Grissmer, Flanagan, and Williamson 1998; Magnuson and Waldfogel 2008). Other studies show a slow but steady decline in the gap from the mid-1990s through the first decade of the 21st century (Reardon et al. 2013; Reardon, Robinson-Cimpian, and Weathers 2014).

Scholars also debate the roles that families, schools, public policy, and even genetics play in shaping Black achievement disparities with peers (Herrnstein and Murray 1994; Jencks and Phillips 1998). Prior research has provided estimates of the contribution of socioeconomic factors to the gap (Hedges and Nowell 1999; Herrnstein and Murray 1994; Phillips et al. 1998) and has decomposed time trends in the gap into changes in household versus schooling characteristics (Berends and Peñaloza 2008). Most quantitative datasets used in test score gap research, however, do not allow researchers to simultaneously estimate fine-grained time trends in the gap and quantify the contribution of control variables to the size of the gap over time. The National Assessment of Educational Progress (NAEP) provides yearly estimates of the Black test score gap, but does not have comprehensive measures of student background factors such as household socioeconomic status (Berends and Koretz 1995; National Center for Education Statistics 2012a). Datasets from the National Center for Education Statistics such as the National Educational Longitudinal Study (NELS) and the Educational Longitudinal Study (ELS) allow researchers to measure the relationship between student test scores and covariates such as household socioeconomic status and school composition for cohorts of students, but they do not allow for a continuous treatment of time.

This study extends previous research on African American test score gaps using meta-regression analysis. Meta-regression is an under-utilized analytic approach in sociological studies of educational stratification. We argue that meta-regression can serve as a powerful analytical tool for summarizing information on educational disparities across a wider body of studies than could possibly be found in a typical review of the literature. Meta-regression allows us to comprehensively analyze Black test score gaps by leveraging information from a large body of studies that include control variables for student racial/ethnic status and student

achievement outcomes. We can compare estimated test score gaps from studies that include certain background controls against those that do not include these covariates. In doing so, we can leverage information from a wide range of samples in order to estimate the proportion of the estimated gap that can be explained by observable confounding factors. We can also investigate whether confounding factors have had a constant or variable influence on the size of the test score gap over time.

Our study has three main goals: (a) We analyze time trends in the published African American test score gap from 1979 to 2010, and compare our results with previous studies; (b) we evaluate the extent to which the gap is explained by factors that are correlated with Black racial status, such as socioeconomic status, as well as other variables that may shape student achievement patterns, such as school and teacher characteristics, and; (c) we assess how the relationship between confounding variables and the Black test score gap changes over time.

Before proceeding to our analysis, we review the prior literature on the Black-White test score gap, highlighting prior research on time trends in the Black-White test score gap and previous estimates of the proportion of the gap that can be explained by observed factors such as household socioeconomic status and schooling contexts.

Background

Time Trends in the Black Test Score Gap

Several previous studies have examined time trends in the Black-White test score gap (Berends and Peñaloza 2008; Grissmer, Flanagan, and Williamson 1998; Herrnstein and Murray 1994; Jencks and Phillips 1998; Magnuson and Waldfogel 2008; Reardon, Robinson-Cimpian,

and Weathers 2014; Reardon et al. 2013). The Black test score gap with peers is persistent over time, but has declined relative to the 1960s. However, there is disagreement in the literature regarding the rate of gap closure over time and whether there have been periods of stall or widening in the gap, particularly in the 1990s and 2000s (Grissmer, Flanagan, and Williamson 1998; Magnuson and Waldfogel 2008).

A number of studies have used data on nationally representative cohorts of students in datasets from the U.S. Department of Education to examine time trends in the Black-White test score gap. Hedges and Nowell (1999) analyzed changes in the Black-White achievement gap from 1965 to 1996 using seven nationally representative datasets, including the National Longitudinal Study of 1972 (NLS-72), the High School and Beyond (HS&B) Study of 1982, and NAEP. They found a persistent Black-White test score gap over the datasets analyzed, with a mean effect size of .9 standard deviation units. The Black-White difference declined from 1969 to 1992, by approximately .14 standard deviation units.

Similarly, Berends and Peñaloza (2008) used nationally representative data on four cohorts of high school seniors to analyze changes in the Black-White math achievement gap from 1972 to 2004. Berends and Peñaloza used data on seniors in high school from the NLS-72, the NLS-92, ELS-04, NAEP, and the 1982 HS&B, and showed that the Black-White test score gap declined from 1.01 standard deviations among the senior cohort of 1972 to .81 standard deviations among the senior cohort of 1992. However, the gap stagnated between the 1992 and 2004 cohorts. Additionally, for math achievement, they found that the bulk of decline in the Black-White gap occurred between 1972 and 1982, with slow decline thereafter. Studies using NAEP, a norm-referenced achievement test administered at regular intervals to a nationally representative sample of students across the United States since 1969, demonstrate that Black-

White test score gaps are persistent across all subject areas and student ages, but have been decreasing slowly over time. Grissmer et al. (1998) examined time trends in the gap among cohorts of 9, 13, and 17 year-olds from 1971 to 1996. They found that gap closure in math and reading occurred most rapidly for adolescents (13 and 17 year-olds) in the 1980s, but detected a period of stagnation or widening (in the case of 17 year-old reading) in achievement disparities from approximately 1988 through 1996. Magnuson and Waldfogel (2008) extended the analysis through 2004, and found that convergence in test scores between White and Black students (9, 13, and 17 year-olds) was most apparent in the 1970s and 1980s, with stagnation in the 1990s. They found some evidence of renewed gap closure for 9 and 13 year olds from 1996-2004. Both sets of studies suggest that the rate of gap closure was highest in the 1970s and the 1980s, but the gap stagnated or widened in the 1990s.

Using State NAEP data on kindergarten cohorts, Reardon, Greenberg, Kalogrides, Shores, and Valentino (2012a) showed an overall decline in the gap from 1980 to 2005. They found slow but steady narrowing of the gap from the 1990s through 2005. Similarly, Reardon et al. (2014) found an overall narrowing of the Black-White gap in NAEP reading achievement for 9, 13, and 17 year-olds, and in NAEP math achievement for 9 and 13 year-olds from the late 1990s through 2012.

Scholars have relied heavily on NAEP and other nationally representative datasets to evaluate how the Black-White test score gap has changed over time. One advantage of tests given to nationally representative samples of students is that they are not subject to the same degree of political manipulation as state standardized tests, which may potentially affect the test score gap. The school accountability movement, however, has spurred the development and implementation of a myriad of tests to measure achievement, and the magnitude of educational

gaps may vary according to the assessment administered. Work by Herrnstein and Murray (1994) found that the distribution of the Black-White test score gap across 156 studies of score differences ranged from 0.0 to 2.0 standard deviations, with a mean of 1.0. The rate of change in the gap may also vary according to sample characteristics. Reardon et al. (2013) also demonstrated that the rates of change in racial achievement gaps varies across states, which is likely due in part to compositional differences among samples of students between states.

This study widens the scope of tests that are used to evaluate Black test score gaps. Importantly, we extend the time horizon for potential gap closure through 2010, the end of the first decade of the implementation of NCLB. An explicit goal of the No Child Left Behind Act of 2001 is to close racial/ethnic achievement gaps through regular testing and performance accountability by subgroup. Some scholars find that accountability threats (i.e. sanctions) have had positive effects on student achievement (Dee and Jacob 2011; Lauen and Gaddis 2012). However, other scholars show that accountability measures can raise overall student levels of achievement while failing to close the Black-White achievement gap (Hanushek and Raymond 2005). Previous research using NAEP also indicates that Black-White achievement gaps narrowed slowly but steadily from the mid-1990s through the first decade of the 2000s (Reardon et al., 2012a, 2014), suggesting that NCLB has had little influence on pre-existing gap trends. While not a policy analysis, our study provides information about the overall efficacy of NCLB indirectly by examining broad trends in the Black test score gap through 2010.

Confounding Variables and the Black Test Score Gap

A student's racial status is correlated with other characteristics that have been shown to influence student achievement. Many Black students face challenges in their home and schooling environments that could influence the magnitude of the Black test score gap and

changes in the gap over time, including parental unemployment, single parent status, and segregation in neighborhoods and schools.

The educational attainment and income profile of African American adults, the potential parents of African American students, has improved steadily since the mid-twentieth century. The percentage of African Americans with a high school degree or higher rose from 21.7% in 1960 to 85.7% in 2012,[1] and the percentage of African American adults with a B.A. or higher increased from 3.5% to 21.4% over this time period (National Center for Education Statistics 2012b). Mean incomes have also increased since the mid-1960s among all income quintiles within the African American population (U.S. Census Bureau 2012). These observed educational and socioeconomic improvements, however, may be overestimated due to the omission of the African American male incarcerated population from major demographic and socioeconomic surveys. When the characteristics of incarcerated Black men are added back into Black non-incarcerated population estimates, the educational and socioeconomic profile of Black men appears to be relatively stagnant from 1980 to 2005 (Pettit 2012).

The employment stability of many African American adults, particularly of young black men in central cities, has deteriorated since the 1970s (Wilson 2012). Unemployment rates have remained higher for African American males than white males over the last four decades (Fairlie and Sundstrom 1999). Since the 1960s, the proportion of Black children living in single-parent families has also increased. In 2013, approximately 21% of non-Latino White children lived in a single-parent household, compared with 55% of African American children, up from 30% in 1960 (Vespa, Lewis, and Kreider 2013; Ruggles 1994). Mass incarceration from the 1980s onward has also disproportionately affected Black children. Approximately 1 in 7 Black children born in 1978, and 1 in 4 Black children born in 1990, have experienced parental

imprisonment (Wildeman 2009). The mass incarceration of African American males has two implications for our understanding of the Black-White test score gap.

Differences in housing and schooling contexts may also influence the Black test score gap with peers. Neighborhood and school segregation among African American children has improved relative to the 1960s, but Black children remain highly isolated in U.S. schools. Black-White residential segregation, as measured by the index of dissimilarity, declined in most major metropolitan areas between 1980 and 2010 by around 5% each decade (Logan, Stults, and Farley 2004; Logan and Stults 2011). Commensurate with declining residential segregation the percent of Black children attending schools with 90-100% minority students decreased from 64% to 38% from 1968 to 2010 (Orfield, Kucsera, and Siegel-Hawley 2012). Nonetheless, African American neighborhood segregation from Whites in most metropolitan areas remains high. In 2010, approximately 59% of African Americans would need to move to a different neighborhood to ensure an even distribution of Blacks and Whites across cities (Logan and Stults 2011).

Black-White school segregation levels are even higher than Black-White neighborhood segregation levels; 67% would need to switch schools to create an even distribution (Orfield, Kucsera, and Siegel-Hawley 2012). Recently, Black students have become resegregated from Whites as school districts have been released from court-mandated desegregation orders, and as courts have struck down the use of race in school assignment policies (Orfield & Eaton, 1996; Orfield et al., 2012; Reardon, Grewal, Kalogrides, & Greenberg, 2012b). Differences in exposure to minority and poor students in schools between White and Black high school seniors were greater in 2004 than in 1972 (Berends and Peñaloza 2008).

While researchers studying racial/ethnic educational inequality agree that test score gaps are partially due to confounding factors, they disagree about the proportion of the Black-White

test score gap that can be attributed to these observable factors. The role of socioeconomic status in explaining the Black-White test score gap represents a case in point. Hedges and Nowell (1999), Herrnstein and Murray (1994), and Phillips, Brooks-Gunn, Dunkan, Klebanov, and Crane (1998) have estimated that household socioeconomic characteristics such as parental education and income explain about one-third of the Black-White test score gap. Phillips et al. (1998), however, have concluded that a wider set of socioeconomic indicators, such as grandparents' educational attainment, mothers' school quality, birth weight, and parenting practices, explained about two-thirds of the Black-White test score gap.

There is also debate about the impact of schooling characteristics on the Black-White test score gap. For instance, there is not consensus in the literature about the contribution of school segregation to Black test score gaps. There is a limited amount of causal research on the impact of school segregation on student outcomes (Reardon and Owens 2014) and a dearth of research directly linking school segregation to the Black-White test score gap and changes in the Black-White test score gap over time (Vigdor & Ludwig, 2008). Some studies suggest that persistent Black-White school segregation is associated with unequal achievement outcomes. Using a meta-analytic approach similar to the one employed in this paper, Mickelson, Bottia, and Lambert (2013) showed that attending a racially segregated school had a negative association with mathematics achievement. Similarly, Hanushek, Kain, and Rivkin (2009) found that Black achievement decreased as Black students gained exposure to higher percentages of Black students in schools. A number of studies have also demonstrated that both residential and school segregation can have negative impacts on Black achievement growth, and can exacerbate Black-White test score gaps (Rumberger and Willms 1992; Rumberger and Palardy 2005; Card and Rothstein 2007; Condron et al. 2013). However, other scholars argue that contextual factors

such as neighborhood conditions may not affect African American achievement to the degree that is posited by the "neighborhood effects" literature (Johnson Jr. 2010).

Most previous studies have not assessed whether observable measures such as household background and schooling contexts have constant or varying influences on the magnitude of the Black test score gap over time. The foremost dataset used to analyze achievement gaps over time, NAEP, lacks a wide range of background measures that are correlated with racial/ethnic status, particularly those related to household socioeconomic status (Berends and Peñaloza 2008; National Center for Education Statistics 2012a). Berends & Peñaloza (2008), however, combined multiple NCES datasets to decompose changes in the test score gap into components due to changes in student characteristics and/or school characteristics. Their work demonstrates that changes in family background characteristics from 1972 to 2004 accounted for a 35-62% reduction in the Black-White test score gap during this time period, whereas differences in the minority composition of the school increased the gap by approximately 65%. However, Berends and Peñaloza do not evaluate whether the relationship between background factors such as socioeconomic status and the black test score gap remains constant or varies over time.

Our sample includes coefficients from a large number of regressions of test scores on Black racial status for each year from 1979 to 2010, and contains variability in the types of control measures that are present or absent within these regressions. We leverage our unique dataset to assess the contribution of observable confounding variables to the gap, and can evaluate whether confounding variables have a constant or time-varying influence on the gap.

Research Questions

Our research answers the following research questions: (a) How has the published relationship between African American racial status and achievement changed in recent decades?

(b) How does the relationship between African American racial status and achievement change when statistical models control for observable student characteristics? And; (c) How does the ability of observable student characteristics to explain the relationship between African American racial status and achievement change over time?

Methods

Data Collection and Sample

We collected information on 1,196 regression coefficients published in 165 journal articles.[2] Our data collection process consisted of (a) establishing inclusion criteria, (b) creating an article search pool based on inclusion criteria, (c) screening studies in the article search pool for inclusion in the study, and (d) coding qualifying articles for selected attributes.

Our phenomenon of interest is the relationship between African American status and standardized test scores. Accordingly, our inclusion criteria began with the requirement that an article must have had at least one statistical analysis with individual-level measures of standardized test scores as the outcome variable, and must include a sample of K-12 students in the U.S. education system. For inclusion in the analysis, the study also needed to report at least one regression coefficient on African American racial status, or provide sufficient information such that the coefficient could be calculated. The standardized tests in the final sample include state high-stakes achievement tests administered for evaluative purposes, such as the California Achievement Test as well as tests administered for research purposes, such as NAEP.

Importantly, we have included studies that did not specifically seek to examine differences in racial/ethnic achievement outcomes, but the test score gap was incidentally calculated because the regression included a standardized test score outcome and African American racial status as a control. We include these studies in order to mimic ideal meta-

analytic studies that include "gray literature"—studies that have not yet been published—to account for the possibility that the publication of a result is correlated with the size or significance of that result, known as publication bias (McAuley et al. 2000). Studies that focus on race are likely to be evaluated on the basis of whether the racial effects they find accord with prior literature. In non-race-focused studies, the racial test score gap is not the estimate of interest, and so standard publication bias pressures are unlikely to apply to those estimates.[3]

Because thousands of published articles contain quantitative analyses with test scores as a dependent variable and African American racial status as a control, we narrowed our inclusion criteria to only include articles published between January 1990 and June 2013 in one of the top 50 journals in each of economics, sociology, and education, 150 journals in total.[4] These fields commonly make use of regression analysis, rather than experiments that do not require statistical controls for race. The time frame began in 1990 to ensure that most datasets would have been collected in recent decades, and ended in June 2013 when candidate articles were selected.

We used Google Scholar to create the initial article search pool. Our 94 search terms (listed in the Web Appendix) related to standardized exams in general, as well as the specific names of national and state standardized exams (NAEP, CBEST, etc.). We identified 9,062 articles in our target publications containing our search terms. We screened the 9,062 articles in a randomized order to determine whether they meet the specific criteria for inclusion in our sample. We set a goal of finding 165 eligible studies before screening any articles, in order to have acceptable statistical power for estimating primary effects,[5] although the estimation of all potential interactions is not feasible at this sample size. We found 165 eligible articles after screening 6,730 articles. We then returned to the eligible articles and recorded information about the study, information about the coefficient on African American racial status, and other

information about the regression, including the number of control variables, sample size, student age, year of test administration, and the presence of controls for particular confounding variables.

There was high agreement between our two primary screeners (one author and one assistant) as to whether particular articles satisfied the inclusion criteria. In order to assess inter-rater reliability, we created a random subsample of 100 studies to re-evaluate without knowledge of the other rater's conclusions. On this subsample there was 100% agreement. An additional random subsample of 100 studies was generated and re-evaluated by the other author, not a primary screener. The Cohen's kappa measure of agreement with the outside screener was .753, in a range commonly considered good or excellent for inter-rater reliability (Gwet 2012).

The total sample consisted of 165 articles. Per our search process, publication dates are between 1990 and 2013. Qualifying articles are found in 38 of the 150 candidate journals. Forty percent of the sample comes from five journals: *Educational Evaluation and Policy Analysis* (18 articles), *American Educational Research Journal* (14), *Sociology of Education* (13), *Social Forces* (12), and *Economics of Education Review* (9). The number of articles in the sample from each journal is listed in the Web Appendix.

Most of the articles in the sample did not seek to analyze racial differences in achievement. We recorded an article as race-focused if the article mentioned race as a motivating topic in the title, abstract, or introduction. Based on these criteria, only 15.8% of the articles in our sample focused specifically on topics of race. The rest of the articles included African American status as a control variable, and may have briefly mentioned racial differences in achievement when summarizing results, but were not explicitly about race-focused research.

Observed regressions use student data collected between 1960 and 2010. However, only 34 regression coefficients are from samples of students before 1979, a year that is well

represented (26 regression coefficients). We did not expect our results to have validity far outside of the range of 1979-2010, so we dropped any coefficient using a sample from before 1979. A full description of the number of observations by year is available in the attached Appendix. Five further outlier coefficients are dropped because the published standard errors appear to be erroneous, leading to anomalously large achievement gaps.

Table 1 describes the features of the included observations. The final sample contains 1,157 observations from 165 studies. Each of the 1,157 observations represents a different statistical analysis. The median sample size across observations is 6,818 students, with a mean of 13.2 total covariates in the estimated model. Sampled students cover the entire K-12 range, with a mean sample age of about 13 years. Older students are slightly better represented in our sample, with about 45% of all regressions covering test scores elicited in high school. In over three-quarters of the observations, the coefficient on Black achievement relative to the achievement of the reference group is significant and negative. Regressions typically compare Black students to White students. However, 14.1% of studies compare Black students to other groups. In these cases, African American status is the only included racial control variable, and the reference group includes both White and non-White students. In the analysis, we control for whether a White or "non-Black" comparison group is used. Results are robust if the studies using "non-Black" comparison groups are dropped.

In addition to examining changes in the size of the test gap over time, we are also interested in evaluating how confounding variables and other aspects of the regressions explain variation in the test score gap. The regression analyses we observe often include control variables to account for the correlation between African American status and the confounding variable. Given the large number of regressions, it is not possible to record each type of control

variable that is ever present in a model. As such, we focus on recording the presence or absence of controls for confounding variables that are discussed widely in the literature on racial achievement gaps.

We refer to all regression characteristics recorded as "control" variables. However, we refer to a specific set of control variables that could increase or decrease the magnitude of the test score gap as "confounding" variables. We focus on the presence or absence of different types of confounding variables. The presence of controls for socioeconomic status is a central focus of our study. In an attempt to make our results comparable to other studies, we distinguish between two types of controls for socioeconomic status: Parental education and all other forms of socioeconomic status. For simplicity, we will refer to this type of confounding variable simply as SES or socioeconomic status, keeping in mind that parental education is also an indicator of socioeconomic status. The SES measure is widely inclusive, and encompasses variables such as direct measures of income as well as broad measures like free and reduced-price lunch eligibility. We also code whether the study controls for certain other confounding factors that have been shown to influence Black-White achievement gaps: parental marriage, urbanicity, English fluency, prior academic achievement (prior grades or test scores),[6] school characteristics (school composition, school sector, school location, etc.), and teacher characteristics (years of experience, teacher race, etc.).

As Table 1 indicates, controls for SES are the most prevalent across regressions in our sample, but controls for prior academic achievement or parental education are less common. Few studies include controls for urbanicity, English fluency, or school or teacher characteristics. We also code the subject area of the exam administered. Most exam scores reported are for math or reading/English exams. About 14% of regressions use composite exams as a dependent

variable, which include exams that cover both math and English, as well as exams that cover a wide range of standard academic subjects. The Other category primarily consists of spatial-mechanical reasoning tests. Results are not affected if the analysis is limited to only reading, math, and composite tests.

We additionally record whether the studies use exams that are administered on a nationally representative scale or not. Non-nationally representative exams include state and local assessments. Typically, exams given to nationally representative samples of students are low-stakes for students and teachers. In a nationally representative exam, teachers may not know what exactly will be covered and/or may have no incentive to "teach to the test" in a way that may alter the racial test score gap. 56.8% of our sample uses data from nationally representative exams.

Estimation

Following Ringquist (2013), we calculate a generalized partial correlation *r* between African American racial status and standardized test scores. The partial correlation allows for a general measure of the racial test score gap that accounts for differences in testing levels and the spread of exam scores. We then evaluate how this partial correlation changes over time, and assess how the partial correlation changes in the presence of controls for confounding variables and other aspects of the regression analysis. We use partial correlations rather than other measures of effect size, for example *d*-indices, because they allow for the gap to be adjusted for the presence of confounding variables, provide an easily interpretable estimate of the gap that is comparable across studies, and can be computed from nearly any regression table without further information such as the standard deviation of the variables. The estimated partial correlation is comparable to a standardized regression coefficient. The partial correlation measure is thus

helpfully amenable to meta-regression analysis. This is the same approach taken by Mickelson et al. (2013), who look at the relationship between race and student performance, but with a focus on school composition as a confounding variable.

Each regression coefficient *i* in an original study *s* contributes a single partial correlation r_{is}, which measures the degree of association between the standardized test score and African American racial status, holding all other variables in the regression constant. These partial correlations can be extracted from linear and binary regression results. When the *t*-score for the coefficient on African American status is reported or can be calculated, the partial correlation coefficient is

$$r_{is} = t_{is}\sqrt{1/(t_{is}^2 + df_{is})}$$

where t_{is} is the reported *t*-score in regression *i* and df_{is} is the number of degrees of freedom in that regression. When a *Z*-score is reported, the partial correlation coefficient is

$$r_{is} = Z_{is} / \sqrt{n_{is}}$$

where Z_{is} is the reported *Z*-score and n_{is} is the number of observations in regression *i*.[7] Using these calculations, r_{is} is not normally distributed and is bounded by -1 and1. As such, we use a Fisher transformation (Fisher 1921; Hotelling 1953) to generate the transformed partial correlation Θ_{is} that we use in analysis:

$$\Theta_{is} = \frac{1}{2}\ln\frac{1+r_{is}}{1-r_{is}}$$

Θ_{is} is normally distributed with individual variance $v_{is} = 1/(n_{is} - 3)$. Θ_{is} can be used to estimate the equation of interest

$$\Theta_{is} = \alpha + F(\gamma, t_{is}) + \beta X_{is} + \varepsilon_{is} \quad (1)$$

where $F(\gamma, t_{is})$ is a function of the year in which the regression sample was collected t_{is} and a parameter vector γ, and X_{is} is a vector describing the original regression, such as what other variables were controlled for (e.g. prior academic achievement).

The intercept in the model is adjusted such that it can be interpreted as the estimated effect size at the mean of the data X_{is} and t_{is}. The presence or absence of confounding variables X_{is} help to explain variation in the partial correlation Θ_{is}. For example, in a regression that does not feature a control for the confounding variable of parental education, Θ_{is} contains the true test gap as well as the omitted variable bias associated with unobserved parental education. The coefficient on "includes a control for parental education" accounts for and measures the difference in partial correlations for which this omitted variable bias is responsible.

In estimates of our equation of interest, observations cannot be considered independent. Measured test score gaps are likely to be correlated within studies, which often use similar models and the same data. We assume that the covariance matrix of ε_{is} takes the form

$$\varepsilon_{is}\varepsilon_{is}' = \begin{pmatrix} \mathbf{V}_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{V}_S \end{pmatrix} \text{ where } \mathbf{V}_s = \begin{pmatrix} \varphi_{1s} & \rho_s & \rho_s & \cdots & \rho_s \\ \rho_s & \varphi_{2s} & \rho_s & \cdots & \rho_s \\ \rho_s & \rho_s & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \rho_s \\ \rho_s & \rho_s & \cdots & \rho_s & \varphi_{I_s s} \end{pmatrix} \quad (2)$$

where S is the total number of studies, I_s represents the number of regressions in study s, and φ_{is} is the sum of the common variance parameter across all studies τ^2, the study-specific error variance for study σ_s^2, and the observation-specific variance v_{is}. ρ_s is the correlation among observations in study s. We do not model dependence between studies, which may be nonzero if the same dataset or sample setting is studied multiple times. Between-study correlation exists

but likely provokes only a small bias; overall 114 different samples were used and most repeated samples studied outcomes on different exams or at different times, used meaningfully different sample restrictions, or incorporated additional data sources. Generally, differences between analyses within a study are solely due to changes in specification, while differences between studies are more meaningful. Results are robust to clustering by sample rather than by study.

To account for within-study dependence, we estimate equation 1 using random effects in a Generalized Estimating Equations setting (Liang and Zeger 1986). Generalized Estimating Equations (GEE) offer an attractive estimator in the case of meta-regression, as they adequately account for the fact that the number of observations differs by study. We estimate $\hat{\tau}^2$ empirically using weighted least squares where observations are weighted by $v_{is}^{-1/2}$, which allows GEE to incorporate differences in the number of observations by study into parameter and variance estimates (Ringquist 2013).

Since we use many different exam scores as outcomes, a potential concern about our approach is in the comparability of the partial correlations Θ_{is}. However, our estimate of interest is a generalized measure of the racial test score gap. Meta-regression gives us the tools to handle differences between studies by using partial correlation to handle differences in scoring levels and spreads between exams and including indicators for the presence of confounding variables as controls in the regression models. Additionally, our approach allows for the fact that the exams given to students have changed over time. The test gap we estimate more closely represents the gap as it exists in educational practice.

Results

The Test Score Gap over Time

We begin with a meta-regression model in which the test score gap is only allowed to vary over the sampling year, to examine how the published generalized test score gap has changed over time. The plotted values in Figure 1 are the predicted partial correlations between African American status and test score, based on the year in which the sample was generated. A negative partial correlation values indicates that African American scores are lower than those of the reference group. Values closer to zero represent a smaller partial correlation between Black racial status and test scores, and thus a narrower achievement gap. We explore several different functional forms for sampling year.[8] The regression coefficients that we use to generate Figure 1, as well as a scatterplot of the individual partial correlations, are presented in the Appendix Table A1.

In Figure 1, two of the time trend lines take year as continuous, and estimate a quadratic and a cubic form of year. The horizontal dummy lines represent averages of the partial correlations for samples of students within the indicated time spans, in groups of three and five years each. Finally, we introduce a linear term for year which is allowed to vary before and after 1995, the midpoint for the years of interest in our study.

Figure 1 suggests non-linear change in the racial test gap from 1979 to 2010. Significant improvements in the test gap were made in the early 1980s, but all three non-continuous measures of time suggest hardly any decrease in the magnitude of the gap after 1995. The linear rate after 1995 is negative and not statistically distinguishable from zero, but distinguishable from the positive slope before 1995 at the 95% confidence level. The three- and five-year dummy results confirm that closure of the gap has slowed considerably over time. There is little progress in gap closure in recent years.

No Child Left Behind (NCLB) sought to close achievement gaps between Black students and their peers. The time span dummies allow for a test of the change in the estimated gap after the implementation of NCLB, although we cannot distinguish between changes that arise due to the policy, changes that arise at the same time for other reasons, or differences in the populations or exams included in our studies before and after the introduction of NCLB. Our results show that most change in the racial test gap occurred before NCLB was put into effect. We can only weakly identify a statistical difference between the residual time effect in 1998-2001 (before implementation), 2002-2004 (during implementation), or 2005-2007 and 2008-2010 (after implementation). The widest difference between these time spans is only significant with p = .118, between 2002-2004 and 2005-2007.

We cannot conclude that NCLB has had no effect on the test score gap. It is possible that NCLB altered the size of the racial test gap but was countervailed by unobserved changes in families and schools that reversed the effect. However, it is clear that NCLB has not been able to supersede unobserved factors that could potentially widen the gap.

Visual analysis of Figure 1 indicates that the cubic functional form is an adequate approximation of the three-year dummy growth path, and the quadratic nearly so. The use of a continuous path allows for a simpler treatment of time as we analyze the impact of confounding variables on the measured test score gap in the next section. For the subsequent analysis of confounding variables, we use the cubic specification for year. The use of a quadratic specification leads to similar results.

The Test Gap and Confounding Variables

A main goal of our paper is to determine how the presence of controls for confounding variables alters the size of the African American test score gap. A regression of test scores on

African American status alone would produce an estimate of the test gap that is biased by omitted variables such as socioeconomic background. In this section, we use meta-regression to examine how confounding variables and other features of the regression models in the studies in our sample change the observed size of the test gap. The studies in our sample included a race/ethnicity control under the impression that it would affect the outcome of interest—standardized test scores. These studies also included other control variables for confounding factors that might influence student test scores, such as socioeconomic status, parent marital status, and classroom and school characteristics. The meta-regression in this analysis examines how the partial correlation between African American status and test scores varies across models that include or do not include different types of confounding variables.

Table 2 displays the results of three meta-regressions. In each case, the dependent variable is the partial correlation between African American status and exam scores. Changes in terms of the partial correlation can be taken as a change in the gap between Black students and their peers that is roughly three times larger in standard deviation units.[9] A coefficient of .05 suggests that the inclusion of that control variable in a regression models shrinks the published gap by .05 in partial correlation units, or about .15 in terms of the standard deviation of the test score. The inclusion of confounding variables associated with gap reduction will have positive coefficients.

The main model in column 1a indicates that the confounding variables that most significantly reduce the test gap are parental education, SES, and prior academic achievement.[10] Surprisingly, controls for English fluency and teacher characteristics increase the estimated test gap. The negative sign on English fluency may be due to a higher proportion of non-fluent English speakers in the non-Black population than in the Black population. The negative sign on

teacher characteristics appears to be due to an association with school characteristics and socioeconomic status; with these controls dropped, the coefficient on the teacher characteristics control becomes positive, but is not statistically significant.

There are some noteworthy null results in Model 1a. There is not a significant level of difference in the test gap between nationally representative and non-nationally representative exams or across different exam subjects, with the exception of social science exams, which have a smaller gap. The coefficient on the presence of a control for parental marriage status is also small and non-significant.

The results in Model 1a show that household background characteristics and prior academic trajectories have heavy influences on the African American test score gap. These results are consistent with prior studies (Phillips et al. 1998; Hedges and Nowell 1999). Interestingly, school and teacher controls have little influence on the magnitude of the gap, net of other confounding variables recorded.

Situating our results in the literature, we are interested in estimating the proportion of the gap that is explained by observable confounding variables. We compare the predicted gap with no controls for confounding variables and the predicted gap with one control included. We calculate a baseline gap by predicting the gap for a study that includes no confounding variables of interest. In this baseline model, the predicted partial correlation between African American status and test scores is -.172. Column 1b in Table 2 displays the percentage of the gap that is explained when each control variable is added to the baseline model. Relative to the baseline, the inclusion of, for example, socioeconomic status as a control variable explains .075/.172 = 43.6% of the gap.

Parental education and other forms of socioeconomic status explain 30.8% and 43.6% of achievement gaps between Black students and their peers, respectively. To predict the share of the gap explained by socioeconomic status as a whole, we estimate a new model with an interaction term to account for any overlap in the portion of the gap explained by parental education and by other forms of socioeconomic status. In this model, socioeconomic status as whole explains 74.5% of the gap.

Prior achievement also plays an important role in driving achievement disparities between Black students and their peers, and explains over one-quarter of the African American test score gap. This finding reinforces the need to understand the Black test score gap in the context of student achievement trajectories. In contrast, parental marriage, urbanicity, school, and teacher characteristics individually explain less than 15% of the achievement gap. Note that the total explained gap can add up to more than 100% because we do not account for interactions between all confounding variables.

Model 2 in Table 2 includes only studies that use nationally representative exams such as NAEP, which are most often analyzed in prior work on the black test score gap. In nationally representative studies, controlling for parental education and prior achievement leads to a larger reduction in the measured gap than in studies in the full sample. Controlling for SES in nationally representative studies reduces the gap to a lesser extent than controlling for this variable in the full sample of studies. This disparity may be due to the wider range of socioeconomic measures observed in the full sample.

In Model 2, the coefficients on the inclusion of English fluency, school, and teacher controls, and the White comparison group control, are non-significant, but this largely reflects imprecision in the coefficients, since relatively few studies using nationally representative data

include these controls. Overall, the results in Model 2 demonstrate that nationally representative studies are capturing certain parts of the African American test gap, but they do not capture the same variation that is found in a broader sample of studies.

A common finding in the test gap literature is that the racial gap is larger for older students than for younger students (e.g. Carneiro and Heckman 2003; Mickelson, Bottia, and Lambert 2013). In Model 3 we also find a larger gap for older students, but the difference is not statistically significant. Disregarding statistical significance, we find that the partial correlation between African American status and test scores becomes .025 more negative from age 13 to age 18.

The estimates in Table 2 assume a constant influence of confounding variables on the test score gap over time. Under the strong assumption that these control variable coefficients are constant across time, the inclusion of these controls explains a larger proportion of the gap in later years (when the gap is smaller) than in earlier years (when the gap is larger).

The Influence of the Confounding Variables on the Gap over Time

In Table 3 we relax the assumption of constant coefficients across time to evaluate how the influence of confounding variables on the size of the gap may change from 1979 to 2010. We use a linear specification because the sample size is not large enough to precisely estimate, for example, confounding variables interacted with three-year time span dummies.

The influence of several confounding variables on the gap changes significantly over time, with differences in urbanicity, English fluency, and teacher characteristics having a weaker influence on the gap over time, and school characteristics having a stronger influence. We may expect the influence of English fluency on the gap to weaken over time as the proportion of non-Black Hispanic ELL students increases relative to ELL Black students. Similarly, we may

expect school characteristics to have a greater impact over time as Black students increasingly attend different schools than their peers, and/or if the variance in school quality in general grows.

We do not find statistically significant changes within the studies over time for the impact of socioeconomic status, parental marriage, or prior academic achievement on the test score gap. However, given the linear specification and the relatively wide time window (1979 to 2010), the estimated effects are still meaningfully large.

The lack of significance for the interactions between year and either parental education or SES is at odds with recent work by Reardon (2011), which shows that both parental education and income have had a stronger relationship with student achievement levels over time. Reardon (2011), however, is interested in explaining time trends in socioeconomic achievement gaps rather than racial achievement gaps, and does not assess whether socioeconomic status has had a stronger influence on Black test score gaps over time.

We further investigate how the relationship between the published Black test score gap and parental education/SES changes over time. Figure 2 displays the proportion of the gap that is explained by the introduction of a control for either SES, parental education, or both. We allow this proportion to change over time, using both a quadratic specification for time and time dummies that span a three-year window. Figure 2 suggests that the proportion of the gap which is explained by socioeconomic status is rising. However, the growth is slow and is not statistically significant, as can be seen in the confidence bands on the three-year dummies.

Discussion

This study sheds light on Black test score gaps with peers by using meta-regression to summarize information on the achievement gap from samples of students from 1979 to 2010 in over 1,100 regressions published in 165 peer-reviewed journals. We find an overall decrease in

the Black achievement gap during this time period, with the reduction mainly concentrated in the 1980s. This finding aligns with other studies that show the gap narrowing directly after the major reforms of the Civil Rights movement in the 1970s and 1980s (Berends and Peñaloza 2008; Grissmer, Flanagan, and Williamson 1998; Hedges and Nowell 1999), but stalling in the period of the 1990s and early 2000s (Magnuson and Waldfogel 2008; Reardon et al. 2013).

No Child Left Behind (NCLB) is the presiding national policy concerning racial achievement gaps. We fail to find any improvement in racial achievement disparities during the era of increased school accountability pressures associated with the implementation of NCLB (the late 1990s to 2010). We cannot provide a causal estimate of the effect of NCLB, since we do not estimate counterfactual outcomes. However, if NCLB reduced the gap, these positive effects must have been offset by other unobserved changes occurring at the same time.

Importantly, our study highlights how a substantial portion of Black achievement gaps with peers can be explained by observable confounding variables that are correlated with racial status and student achievement. Our results suggest that prior research has understated the role that socioeconomic differences play in fueling Black test score gaps. Previous studies have found that differences in socioeconomic status only account for approximately one-third of the Black-White test score gap (Herrnstein and Murray 1994; Phillips et al. 1998; Hedges and Nowell 1999). In contrast, we show that measures of socioeconomic status explain almost three-quarters of the partial correlation between Black racial status and achievement. We contend that the use of broad measures of socioeconomic status allow us to explain a much larger proportion of the gap than has been reported previously, consistent with Phillips et al. (1998). However, we do not find significant evidence that most control variables have a changing influence on the test score gap over time.

Despite an increasing emphasis on school and teacher accountability in education reform movements, we do not find that a significant proportion of the gap is explained by school and teacher characteristics from 1979 to 2010. Rather, controls for socioeconomic status appear to be the main mediators of the published Black test score gap. These results do not imply that "schools don't matter." Given that prior achievement explains a noticeable portion of the gap, it is possible that school- and classroom-level factors in early childhood influence the test score gap via academic development trajectories. We also find that the influence of schools on the gap is increasing over time, which could be the result of increased accountability pressures. Nonetheless, our work suggests that socioeconomic status as a determinant of Black-White test score gaps is underappreciated in the current scholarly literature and in policy discussions about the best means to close Black-White achievement gaps.

Our results are based on the use of meta-analysis. The limitations of meta-analysis are well-documented by methodologists in many fields (Cooper, Hedges, and Valentine 2009; Franke 2001; Hedges 1992). One limitation is that our results only generalize to the population of studies in our sample. If these studies are consistently biased representations of reality, so too will be our results. Other limitations include publication bias, varying quality in the studies included in the sample, comparability of outcome measures across studies, variability in the precision of findings across studies, and the non-independence of observations within studies. Our study design attempts to address these weaknesses in several ways.

Publication bias arises when studies of a particular focal relationship between an independent and dependent variable that are unpublished yield different results than those that are published. Significant structural differences across studies make a formal test for publication bias difficult in our setting. However, our sample includes a variant of gray literature -

observations from studies that did not purposefully seek to examine racial achievement gaps. We argue that our aggregated analysis minimizes the influence of publication bias. The inclusion of low quality studies can also bias the conclusions of a meta-analysis (Moher et al. 1998). We minimize heterogeneity in study quality within our sample by focusing on studies in top peer-reviewed journals and including many studies, reducing the impact of any one low-quality paper.

Our results point to a number of avenues for future research. Studies of racial/ethnic gaps in education often focus on school and teacher characteristics. There is value in this approach, as school and teacher characteristics are often viewed as "malleable factors" that are under the jurisdiction of policymakers, whereas household characteristics are often seen as the result of private decision-making. We find that the importance of school characteristics is increasing over time, and schools should not be ruled out as targets for intervention that could close test score gaps.

However, our results suggest that research should re-direct attention to socioeconomic status as a main driver of the Black-White test score gap. Efforts in this vein could be considerably improved by the use of more detailed socioeconomic status measures in yearly NAEP exams (National Center for Education Statistics 2012a). Both researchers and policymakers should scrutinize the reasons for the large and time-invariant relationship between student socioeconomic status and test scores, and should identify factors that attenuate this linkage. This research and policy inquiry would lead to further understanding of not only how background factors explain the gap, but also how available policy levers could affect the gap and its relation with background factors.

We also see the potential for using meta-analysis in future sociological research to understand educational inequalities in other domains. Our approach takes advantage of a large number of research results that would be ignored in a typical literature review or a traditional meta-analysis. Any study that includes an achievement outcome and a coefficient for African American racial status can provide information about the Black test score gap. This allows us to gain a sense of the gap over a wider array of contexts than is typically possible in a traditional meta-analysis. Our approach could be expanded to the study of gaps in other outcome variables, such as dropout, high school graduation, college enrollment, or income, and could be used to quantify disparities in outcomes between other groups, such as between students in intact versus non-intact families or between immigrant and non-immigrant students.

Overall, our results suggest that African American students face a "long road" to achieving equal test score outcomes with their peers. We see avenues for improvement and ways in which policy could help, but recent improvements have been hard to come by. Our findings suggest that policies seeking to ameliorate the Black-White test score gap need to target multiple domains of African American students' lives. Closing Black-White achievement gaps will be difficult, however, given the extent to which the gap is driven by hard-to-change factors like socioeconomic status.

Endnotes

[1] These numbers are not directly comparable; early figures do not include alternative high school certifications such as the GED. We still suggest that the change from 21.7% to 85.7% represents a significant increase.

[2] This dataset is available upon request from the authors.

[3] In work available from the authors, we find that, controlling for observed structural differences - race-focused studies included fewer controls and were more likely to use national samples - racial gaps were 27% larger in race-focused studies. The likelihood of unobserved structural differences precludes a formal test of publication bias. Instead, we minimize the influence of publication bias on our results by including both kinds of studies in analysis.

[4] Top journals were determined by the quality-weighted citation index at Journal-Ranking.com on May 2, 2013. As of this writing (September 13, 2015), Journal-Ranking.com is not accessible. However, a full list of journals included in our study is in the Web Appendix, and the Journal-Ranking.com methodology for ranking journals is in Lim et al. (2007).

[5] Making the assumption that $R^2 = .2$ in the full model and $R^2 = .15$ in a reduced model, we require 475 observations to achieve power of .9 for our planned confounding variables and several interaction terms. Assuming that each study would contribute on average 5 observations, this initial power analysis suggests a sample size of 95 studies. Since our assumptions may be optimistic, we increase the study size to the maximum number feasible given our resources.

[6] The inclusion of a control for prior academic achievement separates raw achievement gaps from gaps in achievement growth. We include studies that control for prior achievement and record whether this control is present in the model to situate the gap properly as a part of a developmental and cumulative process, and then control for the presence of the confounding variable in the regression.

[7] An additional case is that the results are published as standardized regression coefficients. In that case, r_{is} is taken to be equal to the published standardized coefficient.

[8] These estimates are corrected for sample size and the intra-correlation of effects within studies, as outlined in the methods section. Figure 1 does not take into account the presence or

absence of controls for confounding variables. Results are robust to the inclusion of other confounding variables. For the results in this section and the next, the use of the Fisher transformation to calculate adjusted partial correlations means that the coefficients are transformed back in order to be interpreted as the ability of a confounding variable to explain the partial correlation between African American status and achievement.

[9] To see this, take the example of a reported standardized coefficient. In that case, the partial correlation records the effect of a one-standard deviation change in African American status on standardized exam scores. Dividing the correlation by the standard deviation of African American status produces the gap between Black students and others in standard deviation units. Taking the typical proportion of Black students in the sample as about .1, this suggests multiplying by roughly $(.1 \times .9)^{.5} / .1 = 3$.

[10] Results are robust to the exclusion of any single study. Cook's Distance was above the standard $4 / n$ cutoff for two studies. Removing these two studies from the sample does not change any substantive results.

[11] The proportion is calculated using a model in which “controls for parental education, SES, or both” is included as a predictor. The total proportion is consistently smaller than the 74.5% reported in the previous section due to the use of a single time-varying indicator, as opposed to including both measures and their interaction independently. This approach is taken to avoid including too many interaction terms for the sample size to appropriately estimate.

References

Berends, Mark, and Daniel M. Koretz. 1995. "Reporting Minority Students' Test Scores: How Well Can the National Assessment of Educational Progress Account for Differences in Social Context?" *Educational Assessment* 3 (3): 249–85.

Berends, Mark, and Robert V. Peñaloza. 2008. "Changes in Families, Schools, and the Test Score Gap." In *Steady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap*, edited by Katherine Magnuson and Jane Waldfogel, 66–109. New York: Russell Sage Foundation.

Card, David, and Jesse Rothstein. 2007. "Racial Segregation and the Black–White Test Score Gap." *Journal of Public Economics* 91 (11): 2158–84.

Carneiro, Pedro M., and James J. Heckman. 2003. "Human Capital Policy." In *Inequality in America: What Role for Human Capital Policies?*, edited by Benjamin Friedman, 77–239. NBER Working Paper. Cambridge, MA: MIT Press.

Condron, Dennis J., Daniel Tope, Christina R. Steidl, and Kendralin J. Freeman. 2013. "Racial Segregation and the Black/White Achievement Gap, 1992 to 2009." *The Sociological Quarterly* 54 (1): 130–57.

Cooper, Harris, Larry V. Hedges, and Jeffrey C. Valentine. 2009. *The Handbook of Research Synthesis and Meta-Analysis*. New York City, NY: Russell Sage Foundation.

Dee, Thomas S., and Brian Jacob. 2011. "The Impact of No Child Left Behind on Student Achievement." *Journal of Policy Analysis and Management* 30 (3): 418–46.

Fairlie, Robert W., and William A. Sundstrom. 1999. "The Emergence, Persistence, and Recent Widening of the Racial Unemployment Gap." *Industrial and Labor Relations Review* 52 (2): 252–70.

Fisher, Ronald Aylmer. 1921. "On the 'Probable Error' of a Coefficient of Correlation Deduced from a Small Sample." *Metron* 1: 3–32.

Franke, George R. 2001. "Applications of Meta-Analysis for Marketing and Public Policy: A Review." *Journal of Public Policy & Marketing* 20 (2): 186–200.

Grissmer, David, Ann Flanagan, and Stephanie Williamson. 1998. "Why Did the Black-White Score Gap Narrow in the 1970s and 1980s?" In *The Black White Test Score Gap*, edited by Christopher Jencks and Meredith Phillips, 182–228. Washington, D.C.: The Brookings Institution.

Gwet, KL. 2012. *Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Multiple Raters*. 3rd ed. Gaithersburg, MD: Advanced Analytics Press.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2009. "New Evidence about Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement." *Journal of Labor Economics* 27 (3): 349–83.

Hanushek, Eric A., and Margaret E. Raymond. 2005. "Does School Accountability Lead to Improved Student Performance?" *Journal of Policy Analysis and Management* 24 (2): 297–327.

Hedges, Larry V. 1992. "Meta-Analysis." *Journal of Educational and Behavioral Statistics* 17 (4): 279–96.

Hedges, Larry V., and Amy Nowell. 1999. “Changes in the Black-White Gap in Achievement Test Scores.” *Sociology of Education* 72 (2): 111–35..

Herrnstein, Richard J., and Charles Murray. 1994. *The Bell Curve: The Reshaping of American Life by Differences in Intelligence*. New York City: Free Press.

Hotelling, Harold. 1953. “New Light on the Correlation Coefficient and Its Transforms.” *Journal of the Royal Statistical Society. Series B (Methodological)* 15 (2): 193–232.

Jencks, Christopher, and Meredith Phillips. 1998. *The Black-White Test Score Gap*. Washington, D.C.: Brookings Institution Press.

Johnson Jr., Odis. 2010. “Assessing Neighborhood Racial Segregation and Macroeconomic Effects in the Education of African Americans.” *Review of Educational Research* 80 (4): 527–75.

Lauen, Douglas L., and S. Michael Gaddis. 2012. “Shining a Light or Fumbling in the Dark? The Effects of NCLB’s Subgroup-Specific Accountability on Student Achievement.” *Educational Evaluation and Policy Analysis* 34 (2): 185–208.

Liang, Kung-Yee, and Scott L. Zeger. 1986. “Longitudinal Data Analysis Using Generalized Linear Models.” *Biometrika* 73 (1): 13–22.

Lim, Andrew, Hong Ma, Qi Wen, Zhou Xu, Brenda Cheang, Bernard Tan, and Wenbin Zhu. 2007. “Journal-Ranking.com: An Online Interactive Journal Ranking System.” In *Proceedings of the National Conference on Artificial Intelligence*. Cambridge, MA: AIII Press.

Logan, John R., and Brian Stults. 2011. “The Persistence of Segregation in the Metropolis: New Findings from the 2010 Census.” *Census Brief Prepared for Project US2010*. Washington, D.C.

Logan, John R., Brian J. Stults, and Reynolds Farley. 2004. “Segregation of Minorities in the Metropolis: Two Decades of Change.” *Demography* 41 (1): 1–22.

Magnuson, Katherine, and Jane Waldfogel. 2008. “Introduction.” In *Steady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap*, 1–29. New York City: Russell Sage Foundation.

McAuley, Laura, Ba’ Pham, Peter Tugwell, and David Moher. 2000. “Does the Inclusion of Grey Literature Influence Estimates of Intervention Effectiveness Reported in Meta-Analyses?” *The Lancet* 356 (9237): 1228–31.

Mickelson, Roslyn A., Martha C. Bottia, and Richard Lambert. 2013. “Effects of School Racial Composition on K-12 Mathematics Outcomes: A Metaregression Analysis.” *Review of Educational Research* 83 (1): 121–58.

Moher, David, Ba’ Pham, Alison Jones, Deborah J. Cook, Alejandro R. Jadad, Michael Moher, Peter Tugwell, and Terry P. Klassen. 1998. “Does Quality of Reports of Randomised Trials Affect Estimates of Intervention Efficacy Reported in Meta-Analyses?” *The Lancet* 352 (9128): 609–13.

National Center for Education Statistics. 2012a. “Improving the Measurement of Socioeconomic Status for the National Assessment of Educational Progress: A Theoretical Foundation.” Washington, D.C.

———. 2012b. “Table 8. Percentage of Persons Age 25 and Over with High School Completion or Higher and a Bachelor’s or Higher Degree by Race/Ethnicity and Sex: Selected Years, 1910 through 2012.” *Digest of Education Statistics*.

Orfield, Gary, and Susan E. Eaton. 1996. *Dismantling Desegragation: The Quiet Reversal of Brown v. Board of Education*. New York City, NY: The New Press.

Orfield, Gary, John Kucsera, and Genevieve Siegel-Hawley. 2012. "'E Pluribus'... Separation: Deepening Double Segregation for More Students." Los Angeles, CA: Civil Rights Project / Proyecto Derechos Civiles.

Pettit, Becky. 2012. *Invisible Men: Mass Incarceration and the Myth of Black Progress*. New York City, NY: Russell Sage Foundation.

Phillips, Meredith, Jeanne Brooks-Gunn, Greg J. Duncan, Pamela Klebanov, and Jonathan Crane. 1998. "Family Background, Parenting Practices, and the Black-White Test Score Gap." In *The Black White Test Score Gap*, edited by Christopher Jencks and Meredith Phillips, 103–50. Washington, D.C.: Brookings Institution Press.

Reardon, Sean F. 2011. "The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations." In *Whither Opportunity?: Rising Inequality, Schools, and Children's Life Chances*, edited by Greg J. Duncan and Richard J. Murnane, 91–115. New York City, NY: Russell Sage.

Reardon, Sean F., Erica H. Greenberg, Demetra Kalogrides, Kenneth A. Shores, and Rachel A. Valentino. 2012. "Trends in Academic Achievement Gaps in the Era of No Child Left Behind." In *Society for Research on Educational Effectiveness*.

Reardon, Sean F., Elena Tej Grewal, Demetra Kalogrides, and Erica Greenberg. 2012. "Brown Fades: The End of Court-Ordered School Desegregation and the Resegregation of American Public Schools." *Journal of Policy Analysis and Management* 31 (4): 876–904.

Reardon, Sean F., and Anne Owens. 2014. "60 Years After Brown: Trends and Consequences of School Segregation." *Annual Review of Sociology* 40: 199–218.

Reardon, Sean F., Joseph P. Robinson-Cimpian, and Ericka S. Weathers. 2014. "Patterns and Trends in Racial/Ethnic and Socioeconomic Academic Achievement Gaps." In *Handbook*

of Research in Education Finance, edited by Helen A. Ladd, Margaret E. Goertz, and Lawrence Erlbaum. New York City, NY: Routledge.

Reardon, Sean F., Rachel A. Valentino, Demetra Kalogrides, Kenneth A. Shores, and Erica H. Greenberg. 2013. "Patterns and Trends in Racial Academic Achievement Gaps Among States, 1999-2011."

Ringquist, Evan. 2013. *Meta-Analysis for Public Management and Policy*. San Francisco, CA: Jossey-Bass.

Ruggles, Steve. 1994. "The Origins of African-American Family Structure." *American Sociological Review* 59 (1): 136–51.

Rumberger, Russell W., and Gregory J. Palardy. 2005. "Does Segregation Still Matter? The Impact of Student Composition on Academic Achievement in High School." *Teachers College Record* 107 (9): 1999–2045.

Rumberger, Russell W., and J. Douglas Willms. 1992. "The Impact of Racial and Ethnic Segregation on the Achievement Gap in California High Schools." *Educational Evaluation and Policy Analysis* 14 (4): 377–96.

U.S. Census Bureau. 2012. "Table H-2. Share of Aggregate Income Received by Each Fifth and Top 5 Percent of Black Households: 1967 to 2012." *Historical Income Tables: Income Inequality*.

Vespa, Jonathan, Jamie M. Lewis, and Rose M. Kreider. 2013. "America's Families and Living Arrangements: 2012." Washington, D.C.: U.S. Census Bureau.

Vigdor, Jacob L., and Jens Ludwig. 2008. "Segregation and the Test Score Gap." In *Steady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap*, edited by Katherine Magnuson and Jane Waldfogel, 181–211. New York: Russell Sage Foundation.

Wildeman, Christopher. 2009. “Parental Imprisonment, the Prison Boom, and the Concentration of Childhood Disadvantage.” *Demography* 46 (2): 265–80.

Wilson, William J. 2012. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago, IL: University of Chicago Press.

Table 1
Summary Statistics

Variable	Mean	Std. Dev.
Adjusted Partial Correlation Θ_{is}	-.103	.204
Coefficient on African American is:		
Negative and significant at α = .95	.765	
Insignificant at α = .95	.204	
Positive and significant at α = .95	.031	
Confounding variable controlled for:		
Parental education	.362	
Socioeconomic status	.742	
Parental marriage	.264	
Urbanicity	.143	
English fluency	.221	
Prior academic achievement	.444	
School characteristics	.175	
Teacher characteristics	.086	
Sample size (median)	6,818	
Number of regression covariates	13.204	15.144
Age of students	13.198	4.305
White comparison group	.859	
Non-Black comparison group	.141	
Exam types:		
Math	.347	
Reading/English	.356	
Science	.114	
Social science	.029	
Composite	.135	
Other	.019	
Nationally representative	.568	

Table 2
Effects of Confounding Variables on the Black Test Score Gap

Model:	(1a)		(1b)	(2)		(3)	
	Coef.	(s.e.)	Pct.	Coef.	(s.e.)	Coef.	(s.e.)
Controls included:							
Parental ed.	.053***	(.017)	30.8%	.077***	(.019)		
SES	.075***	(.020)	43.6%	.036***	(.013)		
Parental mrg.	-.009	(.017)	5.2%	.006	(.012)		
Urbanicity	-.023	(.017)	13.4%	-.015*	(.009)		
English fluency	-.037***	(.013)	21.5%	-.022	(.016)		
Prior achieve.	.048***	(.013)	27.9%	.067***	(.014)		
School chars.	.021	(.013)	12.2%	.001	(.009)		
Teacher chars.	-.009*	(.005)	5.2%	.010	(.007)		
Exam types:							
Composite (omitted)							
Math	-.002	(.010)		.003	(.011)		
Reading/English	-.001	(.010)		.004	(.014)		
Science	-.006	(.006)		-.001	(.007)		
Social science	.046*	(.026)		.057**	(.027)		
Other	-.003	(.015)		.004	(.012)		
Year	.004	(.005)		.002	(.004)	.002	(.004)
$Year^2$	-.0001	(.0002)		-.0003*	(.0002)	-.0001	(.0001)
$Year^3$	$-1.48 * 10^{-7}$	$(2.37 * 10^{-5})$		$-2.94 * 10^{-6}$	$(4.65 * 10^{-5})$	$5.68 * 10^{-6}$	$(1.92 * 10^{-5})$
Age						-.003	(.002)
Age^2						-.0004	(.0005)
Comparison = White	-.053**	(.024)		.046	(.036)	-.023	(.022)
Nat. Rep. Exam	.012	(.019)				.025	(.023)
Constant	-.088***	(.017)		-.067***	(.013)	-.071***	(.017)
n	1,157			657		1,157	

*/**/*** indicates statistical significance at the 10%/5%/1% level.

Table 3
Influence of Background Characteristics over Time

	Coefficient	(s.e.)
Controls included:		
Parental education	.039**	(.017)
Socioeconomic status	.070***	(.016)
Parental marriage	-.001	(.019)
Urbanicity	-.028**	(.011)
English fluency	-.029**	(.013)
Prior academic achievement	.048***	(.012)
School characteristics	.027**	(.013)
Teacher characteristics	-.008**	(.004)
Year interaction:		
Parental education * Year	.003	(.002)
Socioeconomic status * Year	.002	(.002)
Parental marriage * Year	.003	(.002)
Urbanicity * Year	-.004**	(.002)
English fluency * Year	-.003**	(.002)
Prior academic achievement * Year	-.002	(.001)
School characteristics * Year	.005*	(.003)
Teacher characteristics * Year	-.001**	(.0003)
Comparison = White	-.057**	(.026)
Nationally Rep. Exam	.010	(.018)
Constant	-.086***	(.019)
Exam type controls	Yes	
Cubic year controls	Yes	

*/**/*** indicates statistical significance at the 10%/5%/1% level.

Table A1
Observations Per Year

Year	Obs.	Year	Obs.	Year	Obs.
1960	6	1986	39	1999	99
1965	12	1987	12	2000	76
1971	1	1988	38	2001	8
1972	12	1989	68	2002	28
1975	1	1990	115	2003	27
1977	1	1991		2004	42
1978	1	1992	129	2005	15
1979	26	1993	30	2006	5
1980	64	1994	33	2007	32
1981	8	1995	8	2008	21
1982	22	1996	12	2009	6
1983	4	1997	93	2010	6
1984	9	1998	73		

Table A2
Coefficients for Figure 1

Variable	Coefficient	(s.e.)	Variable	Coefficient	(s.e.)
Three-year dummies			Five-year dummies		
1979-1980	-.163***	(.026)	1979-1984	-.150***	(.024)
1981-1983	-.123***	(.023)	1985-1989	-.101***	(.024)
1984-1986	-.010***	(.030)	1990-1994	-.096***	(.028)
1987-1989	-.121***	(.040)	1995-1999	-.052***	(.027)
1990-1992	-.090***	(.024)	2000-2004	-.067***	(.015)
1993-1995	-.095*	(.051)	2005-2010	-.053***	(.012)
1996-1998	-.046**	(.023)	Cubic		
1999-2001	-.060	(.042)	Year	.002	(.003)
2002-2004	-.079***	(.017)	$Year^2$	-.0002	(.0001)
2005-2007	-.039**	(.019)	$Year^3$	$4.91 * 10^{-6}$	$(1.83 * 10^{-5})$
2008-2010	-.054**	(.025)	Constant	-.446***	(.085)
Linear pre/post 1995			Quadratic		
Year (pre 1995)	.005***	(.001)	Year	.003*	(.002)
Year (post 1995)	-.002	(.004)	$Year^2$	-.0002	(.0001)
Post 1995	.264	(.352)	Constant	-.438***	(.087)
Constant	-.485***	(.183)			

*/**/*** indicates statistical significance at the 10%/5%/1% level.

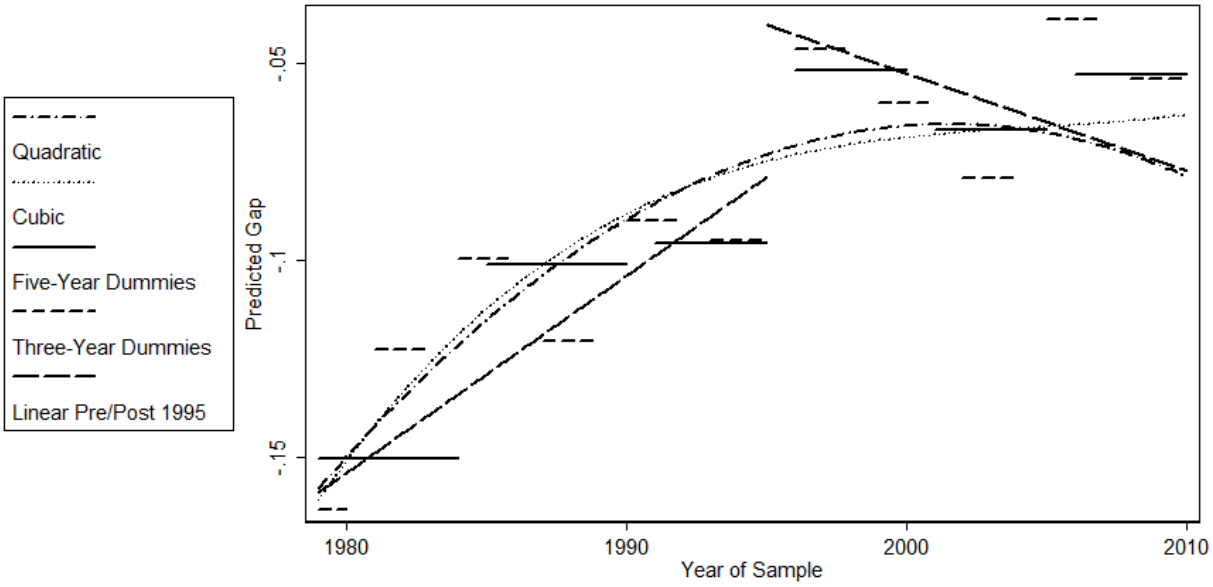

Figure 1.
Predicted Black Test Score Gap, 1979-2010, With Differing Functional Forms for Time

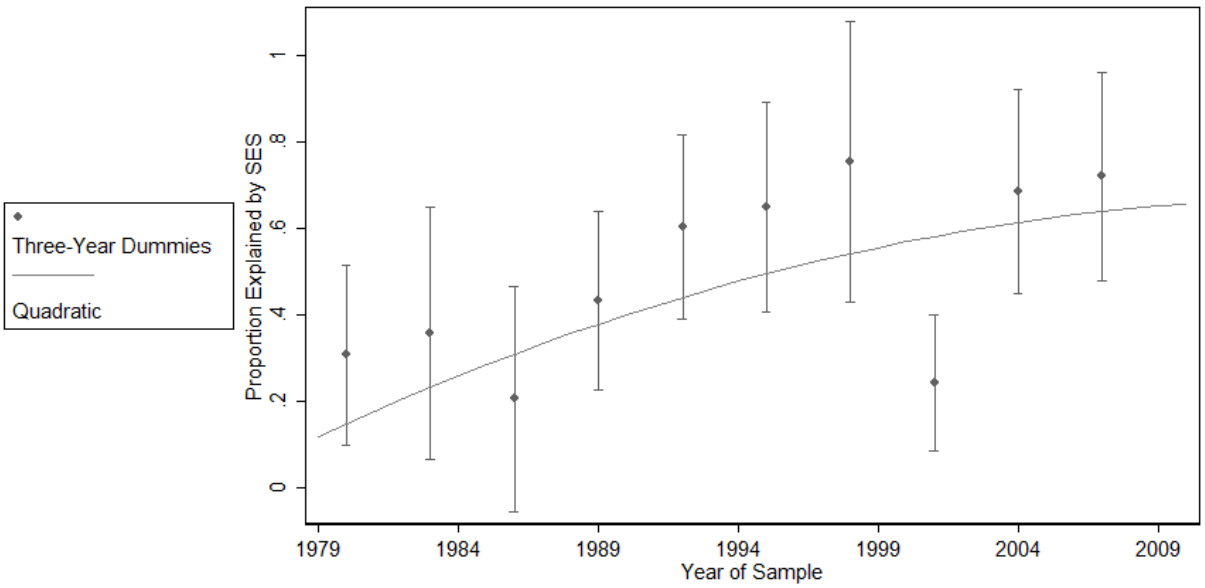

Figure 2.
Proportion of Black Test Score Gap Explained by Parental Education and SES over Time

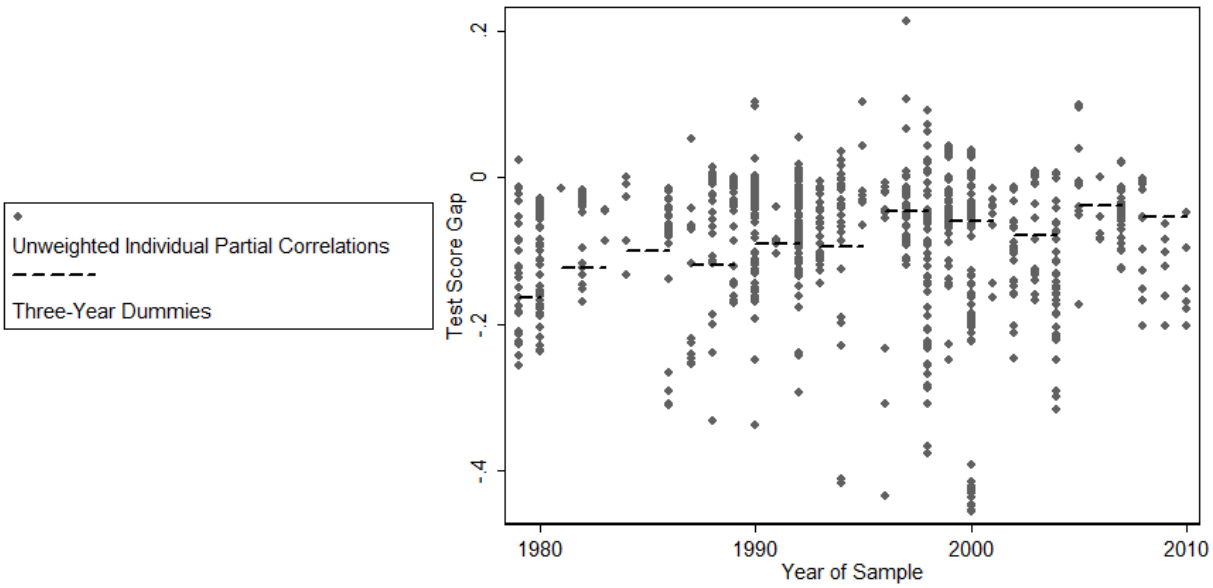

Figure 3.
Scatterplot of Partial Correlations Alongside Estimated Three-Year Dummies

Appendix A: Auxiliary Figures and Tables

This appendix lists the number of regressions in the sample for each year in which student exam data were collected, and the coefficients used to generate Figure 1.

Table A1 reports the number of regressions coded for the sample from the full range of recorded data - from 1960 to 2010, to motivate the dropping of observations from before 1979. Observations dropped for other reasons, such as apparent miscoding in the original article, are not included here.

Table A2 reports the regression coefficients used to generate the plots in Figure 1. Figure A1 presents a scatterplot of individual partial correlations over time, similar to the analysis presented in Figure 1. Overlaid on the scatterplot are the three-year dummies from Figure 1. Figure A1 offers an idea of the variance in reported partial correlations in each year. However, the figure does not adjust the observations for either of two levels of weighting outlined in the Estimation section. As such, although the figure gives a sense of the data at a fine-grained level, the overall picture is somewhat misleading and should be interpreted cautiously.

Appendix B: Web Appendix for Online Publication Only

This web appendix provides a full list of all journals searched and a bibliography of each article included in the final sample.

Journals were chosen as the top 50 journals in each of the "Economics," "Sociology," and "Education and Educational Research" common ranking categories as reported on Journal-Ranking.com on May 2, 2013. Note that the journal "Economics of Education Review" is listed as an education journal rather than in economics - the journal is classified as both but did not make the top 50 cutoff for economics. Due to some overlap between top 50 education and sociology journals, there are only 49 reported education journals, and the journal "Demography" is added to the sociology list as a 50th entry.

Each journal which provided at least one article for the final sample is followed by a number indicating the number of articles from that journal which appear in the final sample.

List of Journals Searched and Number of Articles per Journal Included in the Final Sample

Economics journals:

- American Economic Review
- Brookings Papers on Economic Activity
- Econometric Theory
- Econometrica
- Economic Journal, 1
- Economic Policy
- Economica
- European Economic Review

- Experimental Economics
- Games and Economic Behavior
- International Economic Review
- International Journal of Game Theory
- Journal of Accounting & Economics
- Journal of Applied Econometrics
- Journal of Business & Economic Statistics
- Journal of Development Economics
- Journal of Econometrics
- Journal of Economic Growth
- Journal of Economic History
- Journal of Economic Literature
- Journal of Economic Perspectives
- Journal of Economic Theory
- Journal of Economics & Management Strategy
- Journal of Environmental Economics and Management
- Journal of Finance
- Journal of Financial and Quantitative Analysis
- Journal of Financial Economics
- Journal of Health Economics, 1
- Journal of Human Resources, 7
- Journal of Industrial Economics
- Journal of International Economics

- Journal of Labor Economics, 3
- Journal of Law & Economics
- Journal of Law Economics & Organization
- Journal of Mathematical Economics
- Journal of Monetary Economics
- Journal of Money Credit and Banking
- Journal of Political Economy
- Journal of Public Economics
- Journal of Risk and Uncertainty
- Journal of the European Economic Association, 1
- Journal of Urban Economics, 1
- Mathematical Finance
- Quarterly Journal of Economics, 1
- Rand Journal of Economics
- Review of Economic Dynamics
- Review of Economic Studies
- Review of Economics and Statistics, 3
- Review of Financial Studies
- World Bank Economic Review

Education journals:

- Academic Psychiatry
- Advances in Health Sciences Education
- Aids Education and Prevention

- American Educational Research Journal, 14
- American Journal of Education, 4
- Anthropology & Education Quarterly
- Applied Measurement in Education, 3
- Comparative Education Review
- Curriculum Inquiry
- Early Childhood Research Quarterly, 7
- Economics of Education Review, 9
- Educational Administration Quarterly
- Educational Evaluation and Policy Analysis, 18
- Educational Policy, 4
- Elementary School Journal, 5
- Harvard Educational Review, 3
- Health Education Research
- Instructional Science
- International Journal of Science Education
- Journal for Research in Mathematics Education, 3
- Journal of American College Health
- Journal of College Student Development
- Journal of Economic Education, 5
- Journal of Educational and Behavioral Statistics, 3
- Journal of Educational Research, 8
- Journal of Experimental Education

- Journal of Higher Education
- Journal of Literacy Research
- Journal of Research in Reading
- Journal of Research in Science Teaching, 3
- Journal of School Health, 2
- Journal of Teacher Education, 2
- Journal of the Learning Sciences
- Language Learning
- Learning and Instruction
- Physical Review Special Topics-Physics Education Research
- Reading and Writing
- Reading Research Quarterly
- Research in Higher Education
- Research in the Teaching of English
- Review of Educational Research
- Review of Higher Education
- Review of Research in Education
- Science Education, 1
- Scientific Studies of Reading
- Second Language Research
- Teachers College Record, 6
- Tesol Quarterly
- Theory Into Practice

Sociology journals:

- Acta Sociologica
- American Journal of Sociology, 1
- American Sociological Review, 4
- Annual Review of Sociology
- British Journal of Sociology
- Comparative Studies in Society and History
- Demography, 3
- Discourse & Society
- Economy and Society
- Ethnic and Racial Studies
- European Sociological Review
- Gender & Society
- Human Ecology
- International Journal of Intercultural Relations
- Journal for the Scientific Study of Religion, 1
- Journal of Contemporary Ethnography
- Journal of Leisure Research
- Journal of Marriage and the Family, 4
- Journal of Mathematical Sociology
- Language in Society
- Law & Society Review
- Mobilization

- Poetics
- Politics & Society
- Population and Development Review
- Rationality and Society
- Review of Religious Research
- Rural Sociology, 1
- Social Forces, 12
- Social Networks
- Social Problems
- Social Science Research, 5
- Society & Natural Resources
- Sociologia Ruralis
- Sociological Forum
- Sociological Inquiry, 1
- Sociological Methodology
- Sociological Methods & Research, 1
- Sociological Perspectives
- Sociological Quarterly, 1
- Sociological Review
- Sociological Theory
- Sociology of Education, 13
- Sociology of Health & Illness
- Sociology of Religion

- Sociology of Sport Journal
- Sociology - The Journal of the British Sociological Society
- Symbolic Interaction
- Theory and Society
- Work and Occupations
- Youth & Society

Search Terms

- “achievement exam”
- “achievement test”
- “achievement test”
- “ACT Exam”
- “ACT EXPLORE”
- “aptitude test”
- “Arizona's Instrument to Measure Standards”
- “Assessments for Wyoming Students”
- “Commonwealth Accountability”
- “Connecticut Mastery”
- “DC-CAS”
- “entrance exam”
- “exit exam”
- “High School Assessment”
- “Illinois Standards”
- “Indiana Statewide Testing”

- “Iowa Tests”
- “Louisiana Educational Assessment”
- “Maine Educational Assessment”
- “Maryland School Assessment”
- “MCA-II”
- “Missouri Assessment Program”
- “norm-referenced”
- “norm-references”
- “Oregon Assessment”
- “Palmetto Assessment”
- “Preliminary ACT”
- “Regents Exam”
- “standardized exam”
- “standardized test”
- “Standards of learning”
- “Standards-based assessment”
- “Stanford Achievement Test”
- “STAR Exam”
- “STAR Test”
- “THE SAT”
- “Utah Tests”
- AHSGE
- ARMT

- ASVAB
- CAHSEE
- CAPT
- CRCT
- CSAP
- DCAS
- DSTEP
- EOCS
- EOCT
- EOGS
- FCAT
- GAA
- GEPA
- GHSGT
- HSA
- HSAA
- HSAP
- HSGQE
- HSPA
- HSPE
- HSPT
- ILEAP
- ISAT

- ISEE
- ITBS
- MCAS
- MCT
- MEAP
- MFLE
- MHSA
- MME
- MONTCAS
- NAEP
- NECAP
- NECAP
- NJASK
- NMAPA
- NMSBA
- NMSQT
- NPEP
- OAA
- OCCT
- OGT
- PASA
- PSAE
- PSAT

- PSSA
- SSAT
- STAAR
- TCAP
- TerraNova
- WASL
- WESTEST
- WKCE
- WorkKeys

Sample References

The references that follow contain all 165 studies included in the sample.

Ainsworth, James W. 2002. “Why Does It Take a Village? The Mediation of Neighborhood Effects on Educational Achievement.” *Social Forces* 81 (1): 117–52.

Argys, Laura M., H. Elizabeth Peters, Jeanne Brooks-Gunn, and Judith R. Smith. 1998. “The Impact of Child Support on Cognitive Outcomes of Young Children.” *Demography* 35 (2): 159–73.

Arum, Richard. 1996. “Do Private Schools Force Public Schools to Compete?” *American Sociological Review* 61 (1): 29–46.

Aughinbaugh, Alison. 2000. “Does Head Start Yield Long-Term Benefits?” *Journal of Human Resources* 36 (4): 641–65.

Bankston III, Carl L., and Stephen J Caldas. 2014. “The American School Dilemma: Race and Scholastic Performance.” *The Sociological Quarterly* 38 (3): 423–29.

Bankston III, Carl L., and Stephen J. Caldas. 1996. “Majority African American Schools and Social Injustice: The Influence of De Facto Segregation on Academic Achievement.” *Social Forces* 75 (2): 535–55.

———. 1997. “Family Structure, Schoolmates, and Racial Inequalities in School Achievement.” *Journal of Marriage and the Family* 60 (3): 715–23.

Bartik, Timothy J., William Gormley, and Shirley Adelstein. 2012. “Earnings Benefits of Tulsa’s Pre-K Program for Different Income Groups.” *Economics of Education Review* 31 (6): 1143–61.

Baum, Charles L. 2003. “Does Early Maternal Employment Harm Child Development? An Analysis of the Potential Benefits of Leave Taking.” *Journal of Labor Economics* 21 (2). University of Chicago Press: 381–408.

Benson, James, and Geoffrey Borman. 2009. “Family, Neighborhood, and School Settings across Seasons: When Do Socioeconomic Context and Racial Composition Matter for the Reading Achievement Growth of Young Children?” *Teachers College Record* 112 (5): 1338–90.

Blackburn, McKinley L. 2004. “The Role of Test Scores in Explaining Race and Gender Differences in Wages.” *Economics of Education Review* 23 (6): 555–76.

Blau, David. 1999. “The Effect Of Income On Child Development.” *The Review of Economics and Statistics* 81 (2): 261–76.

Booker, Kevin, Tim R. Sass, Brian Gill, and Ron Zimmer. 2011. “The Effects of Charter High Schools on Educational Attainment.” *Journal of Labor Economics* 29 (2): 377–415.

Borg, Mary O’Malley, J. Patrick Plumlee, and Harriet A. Stranahan. 2007. “Plenty of Children Left Behind: High-Stakes Testing and Graduation Rates in Duval County, Florida.” *Educational Policy* 21 (5): 695–716.

Borman, Geoffrey D., James G. Benson, and Laura Overman. 2008. “A Randomized Field Trial of the Fast ForWord Language Computer-Based Training Program.” *Educational Evaluation and Policy Analysis* 31 (1): 82–106.

Brennan, Robert T., Jimmy Kim, Melodie Wenz-Gross, and Gary N. Siperstein. 2000. “The Relative Equitability of High-Stakes Testing versus Teacher-Assigned Grades: An Analysis of the Massachusetts Comprehensive Assessment System (MCAS).” *Harvard Educational Review* 71 (2): 173–216.

Brewer, Dominic J. 1993. “Principals and Student Outcomes: Evidence from U.S. High Schools.” *Economics of Education Review* 12 (4): 281–92.

Briggs, Derek C. 2008. “Using Explanatory Item Response Models to Analyze Group Differences in Science Achievement.” *Applied Measurement in Education* 21 (2): 89–118.

Broh, Beckett A. 2002. “Linking Extracurricular Programming to Academic Achievement: Who Benefits and Why?” *Sociology of Education* 75 (1): 69–95.

Brunsma, David L., and Kerry Ann Rockquemore. 2003. “Statistics, Sound Bites, and School Uniforms: A Reply to Bodine.” *The Journal of Educational Research* 97 (2). Routledge: 72–77.

Buchmann, Claudia, Dennis J Condron, and Vincent J Roscigno. 2014. “Shadow Education, American Style: Test Preparation, the SAT and College Enrollment.” *Social Forces* 89 (2): 435–61.

Buddin, Richard, and Gema Zamarro. 2009. “Teacher Qualifications and Student Achievement in Urban Elementary Schools.” *Journal of Urban Economics* 66 (2): 103–15.

Burkam, David T., and Valerie E. Lee. 1998. “Effects of Monotone and Nonmonotone Attrition on Parameter Estimates in Regression Models with Educational Data: Demographic Effects

on Achievement, Aspirations, and Attitudes." *The Journal of Human Resources* 33 (2): 555–74.

Burkam, David T., Valerie E. Lee, and Becky A. Smerdon. 1997. "Gender and Science Learning Early in High School: Subject Matter and Laboratory Experiences." *American Educational Research Journal* 34 (2): 297–331.

Burkam, David T., Douglas D. Ready, Valerie E. Lee, and Laura F. LoGerfo. 2004. "Social-Class Differences in Summer Learning Between Kindergarten and First Grade: Model Specification and Estimation." *Sociology of Education* 77 (1): 1–31.

Chomitz, Virginia R., Meghan M. Slining, Robert J. McGowan, Suzanne E. Mitchell, Glen F. Dawson, and Karen A. Hacker. 2008. "Is There a Relationship between Physical Fitness and Academic Achievement? Positive Results from Public School Children in the Northeastern United States." *Journal of School Health* 79 (1): 30–37.

Claessens, Amy. 2012. "Kindergarten Child Care Experiences and Child Achievement and Socioemotional Skills." *Early Childhood Research Quarterly* 27 (3): 365–75.

Clark, Christopher, Benjamin Scafidi, and John R. Swinton. 2011. "Do Peers Influence Achievement in High School Economics? Evidence from Georgia's Economics End of Course Test." *The Journal of Economic Education* 42 (1): 3–18.

Clements, Melissa A, Arthur J Reynolds, and Edmond Hickey. 2004. "Site-Level Predictors of Children's School and Social Competence in the Chicago Child–Parent Centers." *Early Childhood Research Quarterly* 19 (2): 273–96.

Clotfelter, Charles T., Helen F. Ladd, and Jacob Vigdor. 2009. "The Academic Achievement Gap in Grades 3 to 8." *The Review of Economics and Statistics* 91 (2): 398–419.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2006. "Teacher-Student Matching and the Assessment of Teacher Effectiveness." *Journal of Human Resources* 41 (4): 778–820.

Crane, Jonathan. 1995. "Effects of Home Environment, SES, and Maternal Test Scores on Mathematics Achievement." *Journal of Educational Research* 89 (5): 305–14.

DeMars, Christine E. 2000. "Test Stakes and Item Format Interactions." *Applied Measurement in Education* 13 (1): 55–77.

Desimone, Laura, and Daniel A. Long. 2009. "Teacher Effects and the Achievement Gap: Do Teacher and Teaching Quality Influence the Achievement Gap between Black and White and High- and Low-SES Students in the Early Grades?" *Teachers College Record* 112 (12): 3024–73.

Dewey, James, Thomas A. Husted, and Lawrence W. Kenny. 2000. "The Ineffectiveness of School Inputs: A Product of Misspecification?" *Economics of Education Review* 19 (1): 27–45.

Domina, Thurston. 2005. "Leveling the Home Advantage: Assessing the Effectiveness of Parental Involvement in Elementary School." *Sociology of Education* 78 (3): 233–49.

Downey, Douglas B., and James W. Ainsworth-Darnell. 2002. "The Search for Oppositional Culture among Black Students." *American Sociological Review* 67 (1): 156–64.

Dunn, Michelle C., Joseph B. Kadane, and John R. Garrow. 2002. "Comparing Harm Done by Mobility and Class Absence: Missing Students and Missing Data." *Journal of Educational and Behavioral Statistics* 28 (3): 269–88.

Eitle, Tamela McNulty, and David J. Eitle. 2002. "Race, Cultural Capital, and the Educational Effects of Participation in Sports." *Sociology of Education* 75 (2): 123–46.

Engec, Necati. 2005. “Relationship between Mobility and Student Performance and Behavior.” *Journal of Educational Research* 99 (3): 167–78.

Engel, Rozlyn C., Luke B. Gallagher, and David S. Lyle. 2010. “Military Deployments and Children’s Academic Achievement: Evidence from Department of Defense Education Activity Schools.” *Economics of Education Review* 29 (1): 73–82.

Entwisle, Doris R., and Karl L. Alexander. 1995. “Family Type and Children’s Growth in Reading and Math Over the Primary Grades.” *Journal of Marriage and the Family* 58 (2): 341–55.

Entwisle, Doris R., Karl L. Alexander, and Linda Steffel Olson. 1994. “The Gender Gap in Math: Its Possible Origins in Neighborhood Effects.” *American Sociological Review* 59 (6): 822–38.

———. 2005. “First Grade and Educational Attainment by Age 22: A New Story.” *American Journal of Sociology* 110 (5): 1458–1502.

Farkas, George, and Kurt Beron. 2004. “The Detailed Age Trajectory of Oral Vocabulary Knowledge: Differences by Class and Race.” *Social Science Research* 33 (3): 464–97.

Farkas, George, and And Others. 1989. “Cultural Resources and School Success: Gender, Ethnicity, and Poverty Groups within an Urban School District.” *American Sociological Review* 55 (1): 127–42.

Ferguson, Jason L., and Douglas D. Ready. 2011. “Expanding Notions of Social Reproduction: Grandparents’ Educational Attainment and Grandchildren's Cognitive Skills.” *Early Childhood Research Quarterly* 26 (2): 216–26.

Finn, Jeremy D., and Others. 1994. “Disruptive and Inattentive-Withdrawn Behavior and Achievement among Fourth Graders.” *Elementary School Journal* 95 (5): 421–34.

Fleming, Charles B, Kevin P Haggerty, Richard F Catalano, Tracy W Harachi, James J Mazza, and Diana H Gruman. 2005. "Do Social and Behavioral Characteristics Targeted by Preventive Interventions Predict Standardized Test Scores and Grades?" *The Journal of School Health* 75 (9): 342–49.

Fletcher, Jason M., and Steven F. Lehrer. 2011. "Genetic Lotteries Within Families." *Journal of Health Economics* 30 (4): 647–59.

Freedle, Roy O. 2002. "Correcting the SAT's Ethnic and Social-Class Bias: A Method for Reestimating SAT Scores." *Harvard Educational Review* 73 (1): 1–43.

Fryer, Roland G., and Steven D. Levitt. 2004. "Understanding the Black-White Test Score Gap in the First Two Years of School." *The Review of Economics and Statistics* 86 (2): 447–64.

Galindo, Claudia, and Steven B. Sheldon. 2012. "School and Home Connections and Children's Kindergarten Achievement Gains: The Mediating Role of Family Involvement." *Early Childhood Research Quarterly* 27 (1): 90–103.

Gansle, Kristin A., George H. Noell, and Jeanne M. Burns. 2011. "Do Student Achievement Outcomes Differ across Teacher Preparation Programs? An Analysis of Teacher Education in Louisiana." *Journal of Teacher Education* 63 (5): 304–17.

Georges, Annie, and Aaron M. Pallas. 2009. "New Look at a Persistent Problem: Inequality, Mathematics Achievement, and Teaching." *Journal of Educational Research* 103 (4): 274–90.

Gibbs, Benjamin G. 2010. "Reversing Fortunes or Content Change? Gender Gaps in Math-Related Skill Throughout Childhood." *Social Science Research* 39 (4): 540–69.

Glick, Jennifer E., and Michael J. White. 2003. "The Academic Trajectories of Immigrant Youths: Analysis Within and Across Cohorts." *Demography* 40 (4): 759–83.

Goddard, Yvonne L., Roger D. Goddard, and Megan Tschannen-Moran. 2006. "A Theoretical and Empirical Investigation of Teacher Collaboration for School Improvement and Student Achievement in Public Elementary Schools." *Teachers College Record* 109 (4): 877–96.

Goldhaber, Dan D., Dominic J. Brewer, Eric R. Eide, and Daniel I. Rees. 1999. "Testing for Sample Selection in the Milwaukee School Choice Experiment." *Economics of Education Review* 18 (2): 259–67.

Goldhaber, Dan, and Joe Walch. 2012. "Strategic Pay Reform: A Student Outcomes-Based Evaluation of Denver's ProComp Teacher Pay Initiative." *Economics of Education Review* 31 (6): 1067–83.

Goosby, Bridget J., and Jacob E. Cheadle. 2009. "Birth Weight, Math and Reading Achievement Growth: A Multilevel Between-Sibling, Between-Families Approach." *Social Forces* 87 (3): 1291–1320.

Gottfried, Michael A. 2009. "Excused Versus Unexcused: How Student Absences in Elementary School Affect Academic Achievement." *Educational Evaluation and Policy Analysis* 31 (4): 392–415.

———. 2010. "Absent Peers in Elementary Years: The Negative Classroom Effects of Unexcused Absences on Standardized Testing Outcomes." *Teachers College Record* 113 (8): 1597–1632.

———. 2011. "The Detrimental Effects of Missing School: Evidence from Urban Siblings." *American Journal of Education* 117 (2): 147–82.

———. 2012. "Reframing Retention: New Evidence from within the Elementary School Classroom on Post-Retention Performance." *Elementary School Journal* 113 (2): 192–214.

Greenea, Jay P., and Marcus A. Winters. 2009. "The Effects of Exemptions to Florida's Test-Based Promotion Policy: Who Is Retained?" *Economics of Education Review* 28 (1): 135–42.

Grimes, Paul W. 1993. "Public versus Private Secondary Schools and the Production of Economic Education." *Journal of Economic Education* 25 (1): 17–30.

Guo, Guang. 1998. "The Timing of the Influences of Cumulative Poverty on Children's Cognitive Ability and Achievement." *Social Forces* 77 (1): 257–87.

Hallinan, Maureen, and Warren N. Kubitschek. 2012. "A Comparison of Academic Achievement and Adherence to the Common School Ideal in Public and Catholic Schools." *Sociology of Education* 85 (1): 1–22.

Hallinan, Maureen T. 1994. "School Differences in Tracking Effects on Achievement." *Social Forces* 72 (3): 799–820.

Hamilton, Laura S. 1998. "Gender Differences on High School Science Achievement Tests: Do Format and Content Matter?" *Educational Evaluation and Policy Analysis* 20 (3): 179–95.

Hamilton, Laura S., E. Michael Nussbaum, Haggai Kupermintz, Johannes I. M. Kerkhoven, and Richard E. Snow. 1995. "Enhancing the Validity and Usefulness of Large-Scale Educational Assessments: II. NELS:88 Science Achievement." *American Educational Research Journal* 32 (3): 555–81.

Hanushek, Eric A., and Margaret E. Raymond. 2004. "The Effect of School Accountability Systems on the Level and Distribution of Student Achievement." *Journal of the European Economic Association* 2 (2-3): 406–15.

Harris, Robert B., and William C. Kerby. 1996. “Statewide Performance Assessment as a Complement to Multiple-Choice Testing in High School Economics.” *Journal of Economic Education* 28 (2): 122–34.

Heck, Ronald H., and Marian Crislip. 2001. “Direct and Indirect Writing Assessments: Examining Issues of Equity and Utility.” *Educational Evaluation and Policy Analysis* 23 (3): 275–92.

Heck, Ronald H., and Philip Hallinger. 2009. “Assessing the Contribution of Distributed Leadership to School Improvement and Growth in Math Achievement.” *American Educational Research Journal* 46 (3): 659–89.

Hedges, Larry V., and Amy Nowell. 1999. “Changes in the Black-White Gap in Achievement Test Scores.” *Sociology of Education* 72 (2): 111–35.

Heilig, Julian V., and Linda Darling-Hammond. 2008. “Accountability Texas-Style: The Progress and Learning of Urban Minority Students in a High-Stakes Testing Context.” *Educational Evaluation and Policy Analysis* 30 (2): 75–110.

Hill, M. Anne, and June O’Neill. 1993. “Family Endowments and the Achievement of Young Children with Special Reference to the Underclass.” *Journal of Human Resources* 29 (4): 1064–1100.

Hoffer, Thomas B. 1996. “High School Graduation Requirements: Effects on Dropping Out and Student Achievement.” *Teachers College Record* 98 (4): 584–607.

Hofferth, Sandra L. 2006. “Residential Father Family Type and Child Well-Being: Investment Versus Selection.” *Demography* 43 (1): 53–77.

Huynh, Huynh, J. Patrick Meyer, and Dorinda J. Gallant. 2004. "Comparability of Student Performance Between Regular and Oral Administrations for a High-Stakes Mathematics Test." *Applied Measurement in Education* 17 (1): 39–57.

Jackson, Margot I., and Robert D. Mare. 2007. "Cross-Sectional and Longitudinal Measurements of Neighborhood Experience and Their Effects on Children." *Social Science Research* 36 (2): 590–610.

Jacob, Brian A. 2001. "Getting Tough? The Impact of High School Graduation Exams." *Educational Evaluation and Policy Analysis* 23 (2): 99–121.

James-Burdumy, Susanne. 2005. "The Effect of Maternal Labor Force Participation on Child Development." *Journal of Labor Economics* 23 (1): 177–211.

Jeong, Dong Wook. 2009. "Student Participation and Performance on Advanced Placement Exams: Do State-Sponsored Incentives Make a Difference?" *Educational Evaluation and Policy Analysis* 31 (4): 346–66.

Jepsen, Christopher. 2002. "The Effectiveness of Catholic Primary Schooling." *Journal of Human Resources* 38 (4): 928–41.

Jeynes, William H. 2002. "Educational Policy and the Effects of Attending a Religious School on the Academic Achievement of Children." *Educational Policy* 16 (3): 406–24.

Kim, James S., Lauren Capotosto, Ardice Hartry, and Robert Fitzgerald. 2011. "Can a Mixed-Method Literacy Intervention Improve the Reading Achievement of Low-Performing Elementary School Students in an After-School Program?: Results From a Randomized Controlled Trial of READ 180 Enterprise." *Educational Evaluation and Policy Analysis* 30 (2): 183–201.

Knox, Virignia W. 1995. “The Effects of Child Support Payments on Developmental Outcomes for Elementary School-Age Children.” *Journal of Human Resources* 31 (4): 816–40.

Kohlhaas, Kay, Hsin-Hui Lin, and Kwang-Lee Chu. 2010. “Science Equity in Third Grade.” *Elementary School Journal* 110 (3): 393–408.

Konstantopoulos, Spyros, Manisha Modi, and Larry V. Hedges. 2000. “Who Are America’s Gifted?” *American Journal of Education* 109 (3): 344–81.

Krieg, John M. 2006. “Teacher Quality and Attrition.” *Economics of Education Review* 25 (1): 13–27.

Krueger, Alan B. 1999. “Experimental Estimates of Education Production Functions.” *The Quarterly Journal of Economics* 114 (2): 497–532.

Kupermintz, Haggai, and Richard E. Snow. 1997. “Enhancing the Validity and Usefulness of Large-Scale Educational Assessments: III. NELS: 88 Mathematics Achievement to 12th Grade.” *American Educational Research Journal* 34 (1): 124–50.

Lee, Jaekyung, and Kenneth K. Wong. 2004. “The Impact of Accountability on Racial and Socioeconomic Equity: Considering Both School Resources and Achievement Outcomes.” *American Educational Research Journal* 41 (4): 797–832.

Lee, Okhee, Rachael Deaktor, Craig Enders, and Julie Lambert. 2008. “Impact of a Multiyear Professional Development Intervention on Science Achievement of Culturally and Linguistically Diverse Elementary Students.” *Journal of Research in Science Teaching* 45 (6): 726–47.

Lee, Okhee, Jaime Maerten-Rivera, Randall D. Penfield, Kathryn LeRoy, and Walter G. Secada. 2008. “Science Achievement of English Language Learners in Urban Elementary Schools:

Results of a First-Year Professional Development Intervention." *Journal of Research in Science Teaching* 45 (1): 31–52.

Lee, Valerie E., and David T. Burkam. 1996. "Gender Differences in Middle Grade Science Achievement: Subject Domain, Ability Level, and Course Emphasis." *Science Education* 80 (6): 613–50.

Lee, Valerie E., and Susanna Loeb. 2000. "School Size in Chicago Elementary Schools: Effects on Teachers' Attitudes and Students' Achievement." *American Educational Research Journal* 37 (1): 3–31.

Lee, Valerie E., and Julia B. Smith. 1999. "Social Support and Achievement for Young Adolescents in Chicago: The Role of School Academic Press." *American Educational Research Journal* 36 (4): 907–45.

Li, Deping, Andreas Oranje, and Yanlin Jiang. 2008. "On the Estimation of Hierarchical Latent Regression Models for Large-Scale Assessments." *Journal of Educational and Behavioral Statistics* 34 (4): 433–63.

Lleras, Christy. 2008. "Race, Racial Concentration, and the Dynamics of Educational Inequality Across Urban and Suburban Schools." *American Educational Research Journal* 45 (4): 886–912.

Lubienski, Sarah Theule, and Christopher Lubienski. 2006. "School Sector and Academic Achievement: A Multilevel Analysis of NAEP Mathematics Data." *American Educational Research Journal* 43 (4): 651–98.

Ludwig, Jens, and Laurie J. Bassi. 1999. "The Puzzling Case of School Resources and Student Achievement." *Educational Evaluation and Policy Analysis* 21 (4): 385–403.

Lynch, Gerald J. 1989. “The Effect of Teacher Course Work on Student Learning: Evidence from the TEL.” *Journal of Economic Education* 21 (3): 287–96.

Ma, Xin, and Jesse L. M. Wilkins. 2007. “Mathematics Coursework Regulates Growth in Mathematics Achievement.” *Journal for Research in Mathematics Education* 38 (3): 230–57.

Mac Iver, Martha Abele, and Douglas J. Mac Iver. 2009. “Urban Middle-Grade Student Mathematics Achievement Growth Under Comprehensive School Reform.” *The Journal of Educational Research* 102 (3): 223–36.

Maerten-Rivera, Jaime, Nicholas Myers, Okhee Lee, and Randall Penfield. 2010. “Student and School Predictors of High-Stakes Assessment in Science.” *Science Education* 94 (6): 937–62.

Maltese, Adam V., and Craig D. Hochbein. 2012. “The Consequences of ‘School Improvement’: Examining the Association Between Two Standardized Assessments Measuring School Improvement and Student Science Achievement.” *Journal of Research in Science Teaching* 49 (6): 804–30.

Matsumura, Lindsay Clare, Helen E. Garnier, Richard Correnti, Brian Junker, and Donna DiPrima Bickel. 2010. “Investigating the Effectiveness of a Comprehensive Literacy Coaching Program in Schools with High Teacher Mobility.” *Elementary School Journal* 111 (1): 35–62.

McCaffrey, Daniel F., Laura S. Hamilton, Brian M. Stecher, Stephen P. Klein, Delia Bugliari, and Abby Robyn. 2001. “Interactions among Instructional Practices, Curriculum, and Student Achievement: The Case of Standards-Based High School Mathematics.” *Journal for Research in Mathematics Education* 32 (5): 493–517.

McClelland, Megan M., Alan C. Acock, and Frederick J. Morrison. 2006. "The Impact of Kindergarten Learning-Related Skills on Academic Trajectories at the End of Elementary School." *Early Childhood Research Quarterly* 21 (4): 471–90.

McNeal, Ralph B. 1999. "Parental Involvement as Social Capital: Differential Effectiveness on Science Achievement, Truancy, and Dropping Out." *Social Forces* 78 (1): 117–44.

Meisels, Samuel J, Donna DiPrima Bickel, Julie Nicholson, Yange Xue, and Sally Atkins-Burnett. 2001. "Trusting Teachers' Judgments: A Validity Study of a Curriculum-Embedded Performance Assessment in Kindergarten to Grade 3." *American Educational Research Journal* 38 (1): 73–95.

Meisels, Samuel J., and Fong-ruey Liaw. 1992. "Failure in Grade: Do Retained Students Catch Up?" *Journal of Educational Research* 87 (2): 69–77.

Mickelson, Roslyn Arlin. 2001. "Subverting Swann: First- and Second-Generation Segregation in the Charlotte-Mecklenburg Schools." *American Educational Research Journal* 38 (2): 215–52.

Milesi, Carolina, and Adam Gamoran. 2006. "Effects of Class Size and Instruction on Kindergarten Achievement." *Educational Evaluation and Policy Analysis* 28 (4): 287–313.

Miller, Raegen T., Richard J. Murnane, and John B. Willett. 2008. "Do Teacher Absences Impact Student Achievement? Longitudinal Evidence From One Urban School District." *Educational Evaluation and Policy Analysis* 30 (2): 181–200.

Miller, Robert J., and Brian Rowan. 2006. "Effects of Organic Management on Student Achievement." *American Educational Research Journal* 43 (2): 219–53.

Moller, Stephanie, Elizabeth Stearns, Judith R. Blau, and Kenneth C. Land. 2006. "Smooth and Rough Roads to Academic Achievement: Retention and Race/Class Disparities in High School." *Social Science Research* 35 (1): 157–80.

Molnar, Alex, Philip Smith, John Zahorik, Amanda Palmer, Anke Halbach, and Karen Ehrle. 1999. "Evaluating the SAGE Program: A Pilot Program in Targeted Pupil-Teacher Reduction in Wisconsin." *Educational Evaluation and Policy Analysis* 21 (2): 165–77.

Mulkey, Lynn M., Robert L. Crain, and Alexander J. C. Harrington. 1992. "One-Parent Households and Achievement: Economic and Behavioral Explanations of a Small Effect." *Sociology of Education* 65 (1): 48–65.

Muller, Chandra. 1994. "Maternal Employment, Parent Involvement, and Mathematics Achievement among Adolescents." *Journal of Marriage and the Family* 57 (1): 85–100.

———. 1997. "Gender Differences in Parental Involvement and Adolescents' Mathematics Achievement." *Sociology of Education* 71 (4): 336–56.

———. 2001. "The Role of Caring in the Teacher-Student Relationship for At-Risk Students." *Sociological Inquiry* 71 (2): 241–55.

Musoba, Glenda Droogsma. 2010. "Accountability Policies and Readiness for College for Diverse Students." *Educational Policy* 25 (3): 451–87.

Muthen, Bengt, Li-Chiao Huang, Booil Jo, Siek-Toon Khoo, Ginger Nelson Goff, John R. Novak, and Jeffrey Chen Shih. 1995. "Opportunity-to-Learn Effects on Achievement: Analytical Aspects." *Educational Evaluation and Policy Analysis* 17 (3): 371–403.

Neild, Ruth Curran, Elizabeth Nash Farley-Ripple, and Vaughan Byrnes. 2009. "The Effect of Teacher Certification on Middle Grades Achievement in an Urban District." *Educational Policy* 23 (5): 732–60.

Newmann, Fred M., and And Others. 1995. "Authentic Pedagogy and Student Performance." *American Journal of Education* 104 (4): 280–312.

Nussbaum, E. Michael, Laura S. Hamilton, and Richard E. Snow. 1997. "Enhancing the Validity and Usefulness of Large-Scale Educational Assessments: IV NELS:88 Science Achievement to 12th Grade." *American Educational Research Journal* 34 (1): 151–73.

Orr, Amy J. 2002. "Black-White Differences in Achievement: The Importance of Wealth." *Sociology of Education* 76 (4): 281–304.

Paino, Maria, and Linda A. Renzulli. 2012. "Digital Dimension of Cultural Capital: The (In)Visible Advantages for Students Who Exhibit Computer Skills." *Sociology of Education* 86 (2): 124–38.

Palardy, Gregory J., and Russell W. Rumberger. 2008. "Teacher Effectiveness in First Grade: The Importance of Background Qualifications, Attitudes, and Instructional Practices for Student Learning." *Educational Evaluation and Policy Analysis* 30 (2): 111–40.

Papay, John P., Richard J. Murnane, and John B. Willett. 2010. "The Consequences of High School Exit Examinations for Low-Performing Urban Students: Evidence From Massachusetts." *Educational Evaluation and Policy Analysis* 32 (1): 5–23.

Parcel, Toby L., and Laura E. Geschwender. 1995. "Explaining Southern Disadvantage in Verbal Facility among Young Children." *Social Forces* 73 (3): 841–74.

Peterson, Norris A. 1991. "The High School Economics Course and Its Impact on Economic Knowledge." *Journal of Economic Education* 23 (1): 5–16.

Phillips, Meredith. 1997. "What Makes Schools Effective? A Comparison of the Relationships of Communitarian Climate and Academic Climate to Mathematics Achievement and

Attendance During Middle School." *American Educational Research Journal* 34 (4): 633–62.

Plecki, Margaret L., Ana M. Elfers, and Yugo Nakamura. 2012. "Using Evidence for Teacher Education Program Improvement and Accountability: An Illustrative Case of the Role of Value-Added Measures." *Journal of Teacher Education* 63 (5): 318–34.

Pong, Suet-Ling. 1996. "Family Structure, School Context, and Eighth-Grade Math and Reading Achievement." *Journal of Marriage and the Family* 59 (3): 734–46.

Powell, Brian, and Lala C. Steelman. 1990. "Beyond Sibship Size: Sibling Density, Sex Composition, and Educational Outcomes." *Social Forces* 69 (1): 181–206.

Raudenbush, Stephen W., Randall P. Fotiu, and Yuk Fai Cheong. 1998. "Synthesizing Results from the Trial State Assessment." *Journal of Educational and Behavioral Statistics* 24 (4): 413–38.

Ready, Douglas D., Laura F. LoGerfo, David T. Burkam, and Valerie E. Lee. 2005. "Explaining Girls' Advantage in Kindergarten Literacy Learning: Do Classroom Behaviors Make a Difference?" *Elementary School Journal* 106 (1): 21-38.

Regnerus, Mark D. 2000. "Shaping Schooling Success: Religious Socialization and Educational Outcomes in Metropolitan Public Schools." *Journal for the Scientific Study of Religion* 39 (3): 363–70.

Reynolds, Arthur J. 1992. "Grade Retention and School Adjustment: An Explanatory Analysis." *Educational Evaluation and Policy Analysis* 14 (2): 101–21.

Riegle-Crumb, Catherine, and Eric Grodsky. 2010. "Racial-Ethnic Differences at the Intersection of Math Course-Taking and Achievement." *Sociology of Education* 83 (3): 248–70.

Roscigno, Vincent J. 1998. “Race and the Reproduction of Educational Disadvantage.” *Social Forces* 76 (3): 1033–61.

Roscigno, Vincent J., and Martha L. Crowle. 2009. “Rurality, Institutional Disadvantage, and Achievement/Attainment.” *Rural Sociology* 66 (2): 268–92.

Roscigno, Vincent J., Donald Tomaskovic-Devey, and Martha Crowley. 2006. “Education and the Inequalities of Place.” *Social Forces* 84 (4): 2121–45.

Rosenbaum, Emily, and Jessie Anne Rochford. 2008. “Generational Patterns in Academic Performance: The Variable Effects of Attitudes and Social Capital.” *Social Science Research* 37 (1): 350–72.

Rouse, Heather L., and John W. Fantuzzo. 2009. “Multiple Risks and Educational Well Being: A Population-Based Investigation of Threats to Early School Success.” *Early Childhood Research Quarterly* 24 (1): 1–14.

Rumberger, Russell W., and Gregory J. Palardy. 2005. “Does Segregation Still Matter? The Impact of Student Composition on Academic Achievement in High School.” *Teachers College Record* 107 (9): 1999–2045.

Rumberger, Russell W., and J. Douglas Willms. 1992. “The Impact of Racial and Ethnic Segregation on the Achievement Gap in California High Schools.” *Educational Evaluation and Policy Analysis* 14 (4): 377–96.

Sander, William. 1998. “Private Schools and Public School Achievement.” *Journal of Human Resources* 34 (4): 697–709.

Schmidt, William H., Leland S. Cogan, Richard T. Houang, and Curtis C. McKnight. 2011. “Content Coverage Differences across Districts/States: A Persisting Challenge for U.S. Education Policy.” *American Journal of Education* 117 (3): 399–427.

Smith, Thomas E. 1992. “Gender Differences in the Scientific Achievement of Adolescents: Effects of Age and Parental Separation.” *Social Forces* 71 (2): 469–84.

Sojourner, Aaron. 2013. “Identification of Peer Effects with Missing Peer Data: Evidence from Project STAR*.” *The Economic Journal* 123 (569): 574–605.

Supovitz, Jonathan A., and Robert T. Brennan. 1996. “Mirror, Mirror on the Wall, Which Is the Fairest Test of All? An Examination of the Equitability of Portfolio Assessment Relative to Standardized Tests.” *Harvard Educational Review* 67 (3): 472–506.

Swanson, Christopher B., and Barbara Schneider. 1998. “Students on the Move: Residential and Educational Mobility in America’s Schools.” *Sociology of Education* 72 (1): 54–67.

Tarr, James E., Robert E. Reys, Barbara J. Reys, Oscar Chavez, Jeffrey Shih, and Steven J. Osterlind. 2008. “The Impact of Middle-Grades Mathematics Curricula and the Classroom Learning Environment on Student Achievement.” *Journal for Research in Mathematics Education* 39 (3): 247–80.

Telese, James A. 2011. “Middle School Mathematics Teachers’ Professional Development and Student Achievement.” *Journal of Educational Research* 105 (2): 102–11.

Thum, Yeow Meng. 2003. “Measuring Progress Toward a Goal: Estimating Teacher Productivity Using a Multivariate Multilevel Model for Value-Added Analysis.” *Sociological Methods & Research* 32 (2): 153–207.

Witte, John F. 1998. “The Milwaukee Voucher Experiment.” *Educational Evaluation and Policy Analysis* 20 (4): 229–51.

Yasumoto, Jeffrey Y., Kazuaki Uekawa, and Charles E. Bidwell. 2000. “The Collegial Focus and High School Students’ Achievement.” *Sociology of Education* 74 (3): 181–209.

Yesil-Dagli, Ummuhan. 2011. “Predicting ELL Students’ Beginning First Grade English Oral Reading Fluency from Initial Kindergarten Vocabulary, Letter Naming, and Phonological Awareness Skills.” *Early Childhood Research Quarterly* 26 (1): 15–29.

Zellman, Gail L., and Jill M. Waterman. 1998. “Understanding the Impact of Parent School Involvement on Children’s Educational Outcomes.” *The Journal of Educational Research* 91 (6): 370–80.