Running head: BLACK MALES' MATH ACHIEVEMENT TRAJECTORIES

Math Achievement Trajectories among Black Male Students in the Elementary and Middle
School Years

Abstract

In this paper, we analyze the variation in math achievement trajectories of Black male students to understand the different ways these students successfully or unsuccessfully navigate schools and the school characteristics that are associated with their trajectories. Using longitudinal student-level data from a large urban US city (n = 7,039), we analyze Black male students from one cohort to identify trajectories. We find a lack of growth in standardized math scores suggesting that, on average, math proficiency among Black male students in our sample is declining over time. We found that the 4th grade standardized math scores of subsidized-lunch students were somewhat lower than those of non-subsidized students and those of retained students were substantially lower than their counterparts. The average math score of a Black male student's cohort appears to be the only variable amenable to policy manipulation that has a sizeable association with the growth of their standardized math scores, suggesting that putting Black male students in more challenging learning environments may be the best way to increase math proficiency over time. By themselves, other policy decisions (reducing student mobility, teacher turnover, or special education classification; increasing attendance or spending on after-school programming; or hiring more qualified or experienced teachers) all appear to have no or negligible associations with growth in math scores.

Keywords: Black education; Urban education; Quantitative research

Introduction

Black males are underrepresented in most categories associated with academic success and overrepresented in categories associated with failure. On standardized measures of achievement, Black students score below White students in all core academic areas including science, mathematics, reading and writing (Fashola & Cooper, 1999). The achievement gap for Black males is even wider than for other groups.

For example, whereas science and math achievement continues to be higher for White, Asian and Latino males than for females from these groups, Black males perform at a lower rate than Black females in these subject areas (Metro Center, 2009; "Author", 2008). Black males are severely underrepresented in the most rigorous academic programs including gifted programs, honors courses, and Advanced Placement courses while being vastly overrepresented in remedial academic tracks (Darling-Hammond, 2000). Additionally, Black males have the highest rates of special education classification including classifications as mildly mentally retarded, emotionally disturbed, and/or learning disabled (Klingner, 2004). Furthermore, of all demographic groups, Black males have the highest rates of expulsions, suspensions, and detentions ("Author", 2010).

Research offers only general lessons typically focused on the over-representation of Black males in categories that predict poor performance and likelihood of dropping out, such as: rates of suspension, expulsion from school, and arrests in school settings (Skiba, 2002); special education classifications of mildly mentally retarded, emotional disturbance, and/or learning disabled (Orfield, 2004); and enrollment in poorly resourced, poor performing schools (Rothstein, 2004) that are highly segregated by race and class (Orfield, & Lee, 2006).

Few prior studies take account of possible variations in educational trajectories within these broad groupings. We know a lot about the young men with the worst outcomes, but very

little about when the signs of trouble first become manifest. We also know relatively little about those who succeed other than that they managed to avoid failure. The absence of a more nuanced analysis of the evolution of achievement of Black males and the prevalence of an analytical perspective that reifies existing stereotypes of academic failure among Black males contributes to the notion that Black males have monolithic experiences and outcomes. To make targeted decisions about how, when, and where to intervene, we need to know more about the complex pathways and educational experiences that shape outcomes for Black male students in the public school system. Longitudinal data on the performance of Black male students in a large urban setting, including the schools they attended, are examined in this paper and yield new insights about these complex pathways.

In 2004, African American students were under-represented in the top achievement categories. Only four percent of African American high school seniors scored proficient compared with 24 percent of Whites and 19 percent of Asians (Camara & Schmidt, 1999). Since 1973, a substantial gap exists between White and African-American children in reading, mathematics, and science during the middle and high school years (Fryer & Levitt, 2004). Recent findings on gaps in academic achievement suggest the persistence of these inequalities despite years of reform. For example, in New York City, the largest public school system in the nation, a little over half of its Black male students from the 2010 cohort graduated with a Regents or college-bound diploma after four years of high school. Black males confront a wide array of obstacles, including disproportionate placement in special education classes (Metro Center, 2009) and disproportionate rates of discipline (Foley, Mishook, Thompson, Kubiak, & Supovitz, 2008). Despite numerous reform initiatives carried out over the last several years, there is little evidence that these reforms have mitigated the severe challenges faced by this

population. Other school districts in urban areas with high Black male enrollment face similar challenges in reducing the graduation gap between White and Black male students (Schott Foundation, 2012). In some cities, the graduation rate for Black males is less than the national average for these students.

This study presents findings from a longitudinal study on the variation in the educational pathways among Black male students for one graduation cohort and explores the factors, amenable to policy intervention, which may influence their success or failure. We use data from one school district in a large, urban American city; the U.S. Department of Education requires we anonymize the city location. Our central research questions are: 1) What are the patterns of educational trajectories of Black males in a large urban education system?; and 2) How are these educational trajectories associated with factors that may determine academic achievement?

We focus specifically on the early years—from elementary school through middle school—because several studies show that indicators from these years are strong predictors of which students are likely to drop out later (Balfanz and Herzog, 2005; Parthenon Group, 2005). Though several indicators are important—including low attendance, teacher reports of poor behavior, and failure in math or English courses—we focus on math scale scores, because performance in math is a key predictor of academic performance. Additionally, this focus on math achievement allowed us to include limited English proficient students who may have been excluded from the state English exams. We then explore associations between family background, student and school level characteristics and growth in math scale scores.

We organize the study as follows. First, we review an extensive literature that informs our work. Next, we describe the data and methodology we use to estimate educational trajectories. Next, we present descriptive analyses of the entire body of Black male students

included in our study, who were in 4th grade in 1998 and were expected to graduate in 2007. We also describe the data on the schools they attended before high school. Next, we present findings from our growth-model analysis (HLM), which shows how individual performance on standardized math exams changed over time and identifies the factors that are associated with math achievement before entering 9th grade. Finally, we summarize our results and present implications for policy and further research.

Literature Review

Prior research points to factors in three domains that affect academic achievement:1) Family Background Characteristics 2) Individual Student-Level Characteristics; and 3) School Characteristics. Except for the first domain, each domain includes factors that educational policymakers can change in an effort to improve academic achievement. For example, two widely used school policies, which we observe as student-level characteristics, involve retention and special education classification. Both may occur in elementary school, but there is much controversy about their validity and subsequent impact on academic achievement, attainment, and earnings (Roderick, Nagaoka, & Allensworth, 2005; "Author", 2008; Oswald, Best, & Coutinho, 2006). Mobility is a third student-level characteristic, which can be affected by policy decisions. All of the School Level-Characteristics are amenable to policy change and, in turn, can be organized into four categories: 1) School enrollment; 2) School resources; 3) School climate 4) School type enrollment. We use this typology to organize our review of the literature, focusing especially on policy levers.

Family Background

Prevailing differences in family resources and parenting behaviors/practices account for a large share of race/ethnic differences in cognitive development in early childhood. Many family

background characteristics are associated with academic achievement. These include family socioeconomic status (SES), (i.e. a composition of family income, parental education and employment/occupation); parental marital status, home environment, language use, parent-child interaction, parental warmth, discipline and mental health (Duncan & Brooks-Gunn, 1997). Because these characteristics are highly correlated, it is difficult to disentangle the independent contributions of each.

For example, family SES is highly associated with family structure, which, in turn, is associated with academic achievement. Many single women trying to raise children without the financial, emotional and other support of the father have difficulty maintaining their children's emotional and social well-being and their children's cognitive development (Amato, 2005). Studies also associate the presence of family resource with the quality of parental investment in their children (Brooks-Gunn & Duncan, 1996). Reduced parental involvement is associated with delinquent behavior, drug addiction, childhood injuries, accidents and poor health (Barber & Delfabbro, 2000). Additionally, family SES and family structure are highly associated with parental mental health, which in turn is linked to effective parenting and negative child behavioral outcomes (Frank & Meara, 2009; Almond & Currie, 2011).

Student-Level Characteristics

Three student-level characteristics have important associations with academic achievement: retention, special education status, and mobility. The relationship between these characteristics and academic achievement is bi-directional and may be amenable to policy changes.

Retention.

There is probably no bigger disagreement between teachers and educational policymakers on the one hand, and researchers on the other than the disagreement involving the use of retention to improve academic achievement among low-achieving students (Winsler et al., 2012; Roderick et al., 2005). Proponents argue that retention encourages students, parents, and teachers to change their behavior in ways that improve academic achievement, because students would be required to receive additional instruction, unless their performance meets minimal standards. Researchers, by contrast, find no evidence that retention improves the academic achievement of low-achieving students. Although some recent studies find that low-achieving students who are retained may have fewer behavioral problems than their counterparts who are promoted. Other recent studies do not support the beneficial effects of retention on adjustment problems and find that any such gains would be outweighed by the higher probability that retained students dropout (Willson & Hughes, 2006; Deming & Dynarski, 2008)

Special education classification.

A longer standing controversy involves the overrepresentation of Black male students among students with special education classification. This practice continues after decades of criticisms of the practice by studies, legal rulings and commissions have on several grounds. First, the practice is based upon discriminatory and unscientific standards, and a failure of teachers to understand the meaning of behavior, language, and learning styles of Black males (Ford, 2012). Second, it unnecessarily stigmatizes Black males, increases their dropout rates, relegates them to mediocre educational experiences, and diminishes their future economic prospects (Judge & Watson, 2011; "Author", 2008).

Mobility.

While student mobility is often associated with changes in students' residence, research has also found that between 30-40 percent of school changes are associated with other factors such as school overcrowding, class size reduction policies, suspension and expulsions, and schools' academic and social climate (Rumberger, 2003). Studies also find that high mobility rates among elementary students are associated with negative academic achievement and academic outcomes such a dropouts (Voight, Shinn, & Nation, 2012)

However, the association between achievement and mobility may arise because both are associated with other factors that affect student achievement. Some of these other factors are observable in survey data; others are not. For example, poorer families also tend to have the highest levels of mobility due to the difficulties their families may experience in finding affordable housing or stable employment. Both factors (poverty and mobility) are also correlated with lower academic performance (Obradovic et al., 2009). On the other hand, personal and family problems (e.g., substance abuse), which is rarely available in survey data, can affect student's mobility and achievement.

Such spurious correlations have made it difficult for studies to determine if mobility was a symptom or cause of low achievement. Research has found the relationship between student mobility and test scores to be largely accounted for by past academic performance and characteristics that pre-dated school changes (Temple & Reynolds, 2000). This means that mobility may be more amenable to policy changes at several levels. For example, by improving the quality of a particular school or by establishing a common curriculum, administrators may reduce parental incentives to move their children elsewhere.

A third option policymakers have used to reduce the adverse effects of mobility is to change when students transition between schools as students move through grade levels.

Alspaugh (1999) found that students who transitioned to high school in the 7th grade – i.e., students attending 7-12 schools – dropped out of high school at a significantly lower rate than students who transitions to high school in the 9th grade or 10th grade. The developmental needs of student in the middle grades may account for these results (e.g. Jenkins & McEwin, 1992). Alspaugh (1999) posits that the high dropout rate attributed to students making later high-school transitions may be related to the achievement loss experienced by many students during a transitional year.

School-Level Characteristics

School-level characteristics in four domains (student enrollment, school resources, school climate, and school type) have important effects on academic achievement, which are often difficult to quantify, because of collinearity and endogeneity.

Student enrollment.

Enrollment characteristics play an important role in shaping the educational outcomes of its students in part because variables such as the racial and socioeconomic status of students are often highly inter-correlated and highly correlated with other educational opportunities. As a result, students who attend schools with high concentrations of poor, minority, and underachieving students may be at increased risk for academic failure (Balfanz, 2009). Attending schools with high concentration of low-income Black students is associated with lower achievement for individual Black students– particularly high achieving Black students (Hanushek, Kain, & Rivkin, 2009). In particular, some studies find Black students are three times more likely than White students to be in enrolled in high poverty schools (Balfanz, 2009). Moreover, because of the high correlation between race and poverty, especially in large urban areas, many students attend schools where 90-100 percent of the students are Black and Latino

students (Fry, 2005). This makes high poverty concentration synonymous with race and ethnic segregation, both of which are adversely associated with academic achievement (Van Maele & Van Houtte, 2010).

Research on the concentration of "high need" minority students has shown that schools often become overwhelmed by the broad array of needs that such students bring with them to school. Such schools are also more likely to have: remedial courses; high teacher turnover; inexperienced or unqualified teachers; and fewer demanding pre-collegiate courses (Lee, 2012; Goldhaber, Choi, & Cramer, 2007). These schools tend to have crowded classroom space and inadequate supplies of textbooks and materials (Crosnoe, 2005). Schools with these characteristics often fail to provide adequate educational resources and a supportive school climate (Griffith, 2000).

School Resources.

Researchers offer mixed results on whether school resources have beneficial effects on student achievement. Some studies posit that schools' pedagogical and social resources have little to no positive impact on students' academic outcomes (Klugman, 2012). On the other hand, Greenwald, Hedges, and Laine's (1996) meta-analysis of the effect of school resources on student achievement concludes that school resources (e.g., counselors, social workers, and technology experts) have a positive effect on student achievement. Unfortunately, Black males are disproportionately enrolled in under resourced, poorly performing schools (Rothstein, 2004). In some urban areas, high need Black males— special education students and under-performing students— are also more likely to be concentrated in under-resourced schools in low-income neighborhoods (Schott Foundation, 2012).

Teacher quality.

Access to highly qualified and effective teachers is an issue of great concern because many studies have shown that teachers have a significant impact on student achievement (e.g. Goldhaber, 2008; Rockoff, 2004). Teacher quality is also central to efforts to understand and reduce achievement gaps between students of color and their White and Asian peers (Ferguson, 2003).

One measure of teacher quality is teacher certification. Findings on the effects of teacher certification are mixed. Proponents of teacher certification (e.g., Sass, Hannaway, Xu, Figlio, & Feng, 2012) claim that the teacher certification process ensures that teachers have both the pedagogical and content area knowledge, necessary to increase teachers' effectiveness in teaching at-risk students. Evidence suggests a positive association between teachers with full certification and math and reading exams and a negative link between uncertified teachers and student achievements on these tests (Darling-Hammond, 1999).

Among the more robust indicators of teacher quality, three or more years of teacher experience has been shown to correlate positively with student performance (Goldhaber, 2008). Experienced teachers, however, are not equally distributed across low and high poverty schools. Goldhaber et al. (2007) demonstrate that experienced teachers are drawn to schools with low concentrations of poor and minority students and high levels of student achievement. There is also evidence that high teacher turnover inhibits low-performing schools from raising student achievement (Scafidi, Sjoquist, & Stinebrickner, 2007).

After school programming.

Throughout the last two decades, there have been numerous studies documenting benefits of participation in out of school time (OST) activities (Grossman, Lind, Hayes, McMaken, & Gersick, 2009). Participation has been associated with higher (more, or improved): school

attendance, academic achievement, positive attitudes towards schoolwork (Darling, 2005) and homework completion, aspirations for college, work habits and interpersonal skills (Durlak & Weissberg, 2007). Besides benefits, studies have shown that participation in OST activities are linked to decreases in negative behaviors such as juvenile arrests, drug activity, and teenage pregnancy (Barnes, Hoffman, Welte, Farrell, & Dintchef, 2007; Durlak & Weissberg, 2007). Emerging evidence suggests at-risk students, especially boys benefit most from OST activities (Hoffman, 2006; Lauer et al. 2006; Pierce, Bolt, & Vendell, 2010).

School climate.

School climate is the social atmosphere of a setting or learning environment that affects the behavior and attitudes of teachers and students (Moos, 1979). Positive school climates are correlated with positive student educational outcomes; whereas "negative" school climates are associated with obstacles that can inhibit students learning (Brand, Felner, Shim, Seitsinger, & Dumas, 2003). Few studies have developed measures of school climate, but many have used disciplinary and attendance records and graduation rates as proxies (Spera, Wentzel, & Matto, 2009). Several such studies have found that academic achievement is highly associated with attendance, classroom behavior and perceptions of safety and order (Jia et al. 2009; Brand et al., 2003).

School Type: Small Schools

There has been considerable research on the effects of school type, especially school size, on academic (and related) outcomes. Most such studies find that small schools produce better academic outcomes compared to large comprehensive schools, particularly for disadvantaged students (Leithwood & Jantzi, 2009). Other studies show that when compared with large schools, small schools have: lower dropouts rates (Rumberger & Palardy, 2005), higher student

engagement (Crosnoe, Johnson, & Elder, 2004), higher attendance rates (Kuziemko, 2006); lower levels of violence and discipline problems (Darling-Hammond, Ancess, & Ort, 2002).

However, the evidence on small schools is not universally favorable. "Author" (2003) found that several of the new small schools in Boston lacked a clear academic focus and failed to develop strategies to provide support for students. Conchas and Rodriguez (2007) found that many small schools offer a limited curriculum compared to larger comprehensive schools and were unable to provide the support services that students with learning disabilities and English language learners require.

Data and Methods

Data

Our data was comprised of two merged data sources: 1) family background and student-level data and 2) school-level data. We obtained the family background and student level data from a school district in one large, urban US city with high Black male enrollment. We linked these data to publicly available school level data and the U.S. Department of Education (DOE) Common Core Data. The result was a large, rich, and rare longitudinal sample of a cohort of Black male students and the schools they attended. Data included family background and detailed information on student and school characteristics. This sample included all the Black male students in the city who began high school in 2003-2004, and on whom DOE based their high school graduation reports for the cohort of 2007.

The initial sample consisted of 11,803 Black males who were in the system between fall 1998 and fall 2003 or entered the system during that time period. Because we sought to track the students from elementary school through high school, we decided to narrow our sample to those students who were present in the system beginning in 1998 (expected 4th grade) and for whom

we had performance data over the entire period of interest. We decided to focus on math performance, because some of the students in our sample were English Language Learners, and, during this sample period, they were not required to take the same state or city ELA exam as native speakers of English.

To study the relationship between individual students and the school they attended over time, we also had to ensure that we included students that we could link to the schools they attended over time. Finally, outliers in terms of age and grade in the first year of the study were removed. Our specific criteria for inclusion in our study were: 1) scores for at least 4 state/district math exams between 1998 and 2003 (expected grades 4 through 8), 2) identified school settings that each student attended for all years between 1998 and 2003, and 3) within 2 standard deviations of the expected age for a 4th grader in 1998 (roughly 8-10 years). Based on these criteria an analytic sample of 7,039 students was used in our trajectory analyses.

Dependent variable: Math scale scores

Our outcome variable is the math scale score for each student in each year of 1998-2003. Between 4th and 8th grade, when students are usually between the ages of 8-14, students in our sample take standardized exams in math and English Language Arts. While NCLB now requires states to conduct annual testing for grades 3-8, during the period of our study, the federal government only required standardized testing in grades 4 and 8. During the interim years, the city from which the schools were sampled administered annual district exams with the same underlying design as the state exams. These exams formed the basis of our assessment of performance for our sample from 1998-2003.

Raw math scores cannot be compared year to year; thus such scores were converted to *scale scores* by the DOE to allow for cross-grade comparisons and to assess growth in

performance over time. The scale score shows the level of proficiency a student has in Math. Scale scores are arrived at through a standard setting study conducted with an expert panel of teachers at the state level. The Math scale scores use a numerical scale that runs continuously from beginning skills to advanced skills. However, these scales scores were not benchmarks of what is considered to represent *math proficiency*. For that purpose, scale scores were further converted into a Likert scale, with a range of 1-4 levels. Level 1 indicates not meeting learning standards (scores range 448-601), Level 2 is below proficiency (scores range 602-636), Level 3 represents proficiency (scores 637-677), and Level 4 indicates learning standards with distinction (scores range 678-810). A student's performance level provides information on a student's abilities in relation to the state Math Learning Standards. Teachers at the state level established descriptions of the skills and knowledge required in each performance level. To see how students performed, we examined their growth both in terms of performance and proficiency over time.

Independent variables

Our data includes only two measures of family background characteristics: nativity and SES. Following most education research, we use qualification for the federal subsidized school lunch program (either free or at a reduced rate) to measure SES, because this measure is linked to reported income and qualification for other government subsidies (Hanushek, Kain, & Rivkin, 2009). Student-level characteristics included special education status and transitioning to schools (Klingner, 2004; Voight, Shinn, & Nation, 2012). School-level characteristics, following from previous studies, include percent of racial/ethnic minority students, percent of student qualifying for subsidized lunch, average teacher experience, attendance rates and teacher turnover rate as proxies for school climate, and school size (Balfanz, 2009; Lee, 2012). There were no data available on classroom characteristics.

Methods

We explored associations between math scale scores and family background, student-level, and school-level characteristics by estimating individual patterns of math scores over time and identifying factors that differentiated students who displayed different patterns. We fit two models to the longitudinal data to test the association between student math exam scores and substantive predictors. The first model uses time to predict math exam scores over time; the second model adds all other student and school level predictors. We fit these models to the data using growth curve analysis.

Growth curve methods permit the examination of individual performance patterns and the identification of factors related to those patterns by distinguishing between *within-individual* performance patterns and *between-individual* differences in those patterns. Hierarchical linear models (HLM) or multi-level models (MLM) estimate individual growth curves and identify school level characteristics related to patterns of development. Individual and population growth curves are estimated simultaneously to describe the change we expect each member of the population to experience between the fourth and eighth grades (individual) and to identify characteristics that vary across individuals that are associated with those patterns of change over time (population). Specifically, HLM analyses estimates individual and population growth curves from the fixed and random effect variables specified in the models. The iterative estimation process is divided into the estimation of the individual (within-subject) growth curves and the population (i.e. between-subjects) growth curves.

The first equation below (1) describes individual change over time, also known as level 1. The individual growth model describes the temporal dependence of individual status on time. The two *individual growth parameters*, whose values determine the trajectory of individual

change over time, are (π_0) and (π_1). The first parameter is the intercept, representing student i's math exam score in grade 4 (initial level). The second parameter is the slope and represents the rate at which student i's math scores are growing over time (that is, the linear rate of change per unit increase in grade).

Model 1:

$$Y_{ij} = \pi_{0i} + \pi_{1i}Time_{ij} + \epsilon_{ij} \quad (1)$$

$$\pi_{0i} = \gamma_{00} + \theta_{0i} \quad (2)$$

$$\pi_{1i} = \gamma_{10} + \theta_{1i} \quad (3)$$

With the level 1 model we can ask: Do individuals differ in the changes they experience in math scores between grades 4 and 8, and if so how? More specifically: Do their initial math scores differ (intercept)? Does the rate of change (slope) in their math scores differ over this period? Besides these questions, we can ask questions about the average pattern of change, for example: What is the average initial math score in the sample? What is the average rate of change in math scores in the sample?

The third term in equation 1 (ϵ_{ij}) represents the residual, that part of student's observed math score that is not predicted by the student's age. It may reflect the influence of other predictors of math scores that are not included in the level 1 model, measurement error, or sampling error.

Once a level-1 model has been established, everyone in the sample is assumed to have the same form for their growth. However, different people can have different values of the individual growth parameters. For example, individuals may differ in both their intercepts (π_0) and slopes (π_1).

Equations 2 and 3 represent level 2 of the model which describes inter-individual differences in the growth parameters. In the level-2 submodel, γ_{00} and γ_{10} represent the population average initial status and rate of change, respectively. These coefficients are also known as fixed effects. The level 2 residuals, θ_{0i} and θ_{1i}, represent those portions of the initial status or rate of change that are unexplained. They represent deviations of the individual trajectory around the group average.

An alternative specification that is more parsimonious is one that collapses level 1 and level 2 into a composite model. Specifically, since the first level parameters (π_0 and π_1) are the outcomes of the level 2 submodel, we can substitute for π_0 and π_1 from the level 2 model into the level 1 model. This results in:

$$Y_{ij} = \gamma_{00} + \gamma_{10}Time_{ij} + \theta_{0i} + \theta_{1i}Time_{ij} + \epsilon_{ij} \ (4)$$

The two specifications are mathematically equivalent. Each stipulates a relationship between an outcome (Y_{ij}) and a predictor (*Time*). The benefit of a level 1/level 2 representation is that it is conceptually intuitive: we first describe individual change and then we examine interindividual differences in change. We can easily focus on the parameters that describe interindividual differences in initial status (γ_{00}, fixed effect) and parameters that describe differences in change (γ_{10}, fixed effect). The composite model is advantageous because it reflects the statistical model that is fit to the data.

Our second model incorporates school characteristics for each student at every measurement occasion (4th to 8th grades). These predictors are known as time-varying predictors since their values may differ over time. For example, a student's cohort average math score, the average of the math scores for all students in his cohort, may differ from when he is in 4th grade and when he is in 5th grade. The easiest way to understand how to include time varying

predictors is to use the composite specification of the HLM model. We can begin with equation 4

from above and add the main effects of all the predictors that are available to us. This renders a

new composite model:

Model 2:

$$Y_{ij} = \gamma_{00} + \gamma_{01}Time_{ij} + \gamma_{02}Special\ Ed\ Status_{ij} + \gamma_{03}Fifth\ grade\ retention_{ij}$$

$$+\ \gamma_{04}Foreign\ Born_{ij} + \gamma_{05}Ever\ free\ lunch_{ij} + \gamma_{06}School\ Mobility_{ij}$$

$$+\ \gamma_{07}Special\ Ed\ Status\ x\ Time_{ij} + \gamma_{08}Foreign\ Born\ x\ Time$$

$$+\ \gamma_{09}Retention\ x\ Time + \gamma_{10}Lunch\ status\ x\ time$$

$$+\gamma_{11}\%\ After\ School\ programming$$

$$+\gamma_{12}\%\ Teachers\ outside\ of\ certification$$

$$+\gamma_{13}\%\ Teachers\ less\ than\ 3yrs\ experience + \gamma_{14}Teacher\ Retention$$

$$+\gamma_{15}Attendance\ Rate + \gamma_{16}School\ lunch\ status\ rate$$

$$+\gamma_{17}Cohort\ Avg\ Math + \gamma_{18}Small\ School + \theta_{0i} + \theta_{1i}Time_{ij} + \epsilon_{ij}\ (5)$$

The two subscripts (*i,j*) on each predictor signify that it is time varying. Here we assume that

student *i*'s value of Y at time *j* depends on the contemporaneous values of the above school

characteristics and the three person-specific residuals, $\theta_{0i} + \theta_{1i}Time_{ij} + \epsilon_{ij}$. Again, the model

has two types of effects. The fixed effects provide estimates of average math scores (γ_{00}),

average change in math scores (γ_{01}), and how covariates, such as, percent of the budget toward

after school programming (γ_{11}), the percent of teachers with fewer than 3 years of experience

(γ_{13}), and others are related to math scores across all subjects. In contrast, random components

($\theta_{0i} + \theta_{1i}$) represent effects for each individual; we include these components since we believe

that the grade-related mean level and change of the math scores vary significantly between

individuals. Thus, a single regression equation is estimated which interrelates student and school level characteristics.

Methodologically, these models are appropriate given our repeated measures design. Hierarchical linear models permit the inclusion of predictors that change over time. Thus, the student's school type each year can be related to his math scores longitudinally as a time varying covariate. Second, models are estimated using maximum likelihood (ML). ML estimates have several desirable properties. As the sample size increases ML estimates: (1) converge to the true values of population parameters (π_{0i} and π_{1i}), (2) sampling distributions of these estimates are normally distributed, and (3) the estimates are more precise than those derived from other estimators. Third, other more traditional multivariate models drop missing cases in any wave of data; in contrast, since the person-year is the unit of analysis in hierarchical models, all available data is utilized.

All models were run using Stata 12.0 (Stata, 2001).

Results

Figure 1 shows the math scale scores of our sample in 4th and 8th grade, compared to the minimum scale score needed in each year to be considered a Level 3 or proficient. In 1999, the sample performed, on average, just below proficiency. Although mean scores increased in 2003, the sample in terms of proficiency fell behind.

<Figure 1 about here>

In Figure 2, we add the minimum scale scores needed to perform at level 2, or the "almost proficient" level. Over time, the trend in the sample's mean scale score begins below proficiency and ends closer to the lower limit for Level 2. Figures 1 and 2 elucidate the sample's lack of progress in math achievement between the 4th and 8th grades.

<Figure 2 about here>

Looking more closely at students' proficiency levels on the math exam, we see a rather acute downward shift in proficiency levels between 1999 and 2003 (see Figure 3). Overall, in 1999, only 43% of the sample obtained a score that met proficiency or performed better (Level 3 or 4). More disturbing, in 2003, this percentage decreased by 17 percentage points, with only 26% of the sample obtaining proficiency or above. Most notable was the change in performance for the students in the sample that were the highest performers in 1999. While only 7.7% obtained advanced proficiency (Level 4) on the math exam in 1999, by 2003 this small percentage had decreased dramatically to 1.9% of the total sample that obtained advanced proficiency.

<Figure 3 about here>

The vast majority of students in our sample were born in the United States (91.0 percent). The vast majority of the students in our sample (89.2 percent) qualified for the federal subsidized lunch program, suggesting that an overwhelming number of students in our sample were from low-income families.

<Table 1 about here>

Nearly a fifth of the students in the sample were referred for special education services at some point during their academic career. Of those referred, 11.6% spent at least one year in integrated settings between 1998 and 2003, almost half (48.9%) were assigned at least one school year during the same time period to settings where they were isolated from the general school population, either in special classrooms or programs within schools or in specialized school settings (data not shown). By contrast, retention by the 5th grade, which could have occurred in the 3rd or 4th grade, was rare, involving less than 1 percent of the sample.

There was much mobility in the sample, but some mobility was to be expected as students transitioned from elementary to middle school. Most of the sample changed schools between the 5th and 8th grades. If a student had attended a traditional elementary school (grades K-5) and middle school (grades 6-8), we would expect them to move once; almost three-quarters of our sample followed this expected pattern.

This section describes the characteristics, summarized in Table 2, of the students and schools in the elementary and middle schools attended by the sample during their expected 4th grade year (1998-1999) and expected 8th grade year (2002-2003).

<center>**<Table 2 about here>**</center>

During the elementary and middle-school years, students attended predominantly minority schools, in which Black students made up the largest enrollment grouping (62.8 percent and 59.3 percent respectively). As students transitioned from their elementary to middle school years fewer qualified for subsidized lunch. The mean math scale scores increased from the 4th grade to the 8th grade.

On average, students attended schools in which administrators allocated very little of their budgets to afterschool programming 1999 and 2003 (nearly 1 percent). Further, as students in our sample progressed into their middle school years they were instructed by more teachers with fewer than three years of experience.

Student attendance rates were generally low throughout the elementary school years to middle school years (nearly fourth-fifths). This means that for every one hundred students at least 15 students were absent on any given school day. Our data also includes a proxy for school climate not usually available in other studies: the proportion of teachers who returned to the same school from the previous year. Students attended schools that retained nearly three-quarters

of their teachers in the elementary and middle school years. These findings suggest a lack of a stable environment in many of the schools Black males attended, which we expect to reduce math scale scores.

Finally, less than five percent of the sample attended small schools (<440) in the elementary years. Although, the average small school attendance rate nearly tripled between the elementary and middle school years.

Growth Curve Analysis

<Table 3 about here>

Table 3 presents models predicting math exam scores from time, individual, and school level predictors. In the first column, we present the results of fitting model (1) to this data, where *Year* indicates the grade level at time *j* for student *i*. Since time is centered at grade 4, when time equals 0, we estimate that the average student has a math score of 629 in the fourth grade; over time this level increases linearly at a rate of 15 points per grade level. The variance components for both initial status and rates of change are statistically significant, suggesting the wisdom of exploring the effects of person-specific predictors.

In Figure 4 we can see a collection of individual growth trajectories for a group of 4 randomly selected students (dotted lines) against the estimated model (population average in solid line). There is evidence of heterogeneity in observed change across students—for some, math scores increase with grade levels, for others scores decrease before finally increasing, and for still others math scores increase before decreasing by 8th grade. Despite the differences, these plots show a general, linear average trend in the growth of math scores.

<Figure 4 about here>

Raw exam scores are converted to a common scale in order to compare achievement across grades. Using this conversion, the mean math score for a fourth-grade student assessed to be well below proficiency is 586.9, the mean math score for a fourth-grade student assessed to be below proficiency is 621.9, and the mean math score for a fourth-grade student assessed to have just met proficiency is 653.9. This means the average Black male student is performing below proficiency in the fourth-grade.

With a sample as large as ours, the distribution of math scale scores should be symmetrical, so the mean of the standardized scores should be near zero and these standardized scores should have a unit standard deviation. This suggests that standardizing the math scale scores of Black male students, using the mean of Black male math scores each year, which also remain below proficiency, allows associations between included variables and math performance to be expressed in terms of standard deviations.

Model 2 shows the results of re-estimating the model using standardized scale scores as the outcome variable. The estimate of the sample mean of our fourth grade standardized score is in fact near zero (.05). But more importantly, our estimate of the coefficient of year (grade) is not statistically different from zero, which means that on average, there is no growth in the standardized math scores over time. However, the average math scores of Black male students in the 8th grade were closer to the lower limit of level 2 (below proficiency) than in 4th grade. Thus, the lack of growth in the standardized math scores suggests that, on average, the math proficiency is declining over time.

Model 3 presents the results of adding our measures of family background, individual student characteristics, student-level characteristics and school-level characteristics to our model, using student's standardized math scores as the outcome variable. The standardized average

fourth grade math score is higher in Model 3 than our estimate in Model 2, but on average, we still see no trend in standardized math scores over time. There was no statistically significant difference between the standardized math score of foreign born and native born Black male students, but subsidized-lunch students had standardized math scores one tenth of a standard deviation lower than non-subsidized students in the fourth grade. Like the standardized math scores of the average Black Male student, the scores of foreign born and subsidized-lunch displayed no trend over time.

As compared with the average Black male student, the fourth-grade standardized math scores of special education students were .6 of a standard deviation lower than students who did not have a special education status. Students who were retained in the expected 5th grade year, by being held back one or two years, had standardized math scores nearly three quarters of a standard deviation lower than the average Black male student. Again, there was no trend in standardized math scores for the average Black male student who was retained, but special education students saw their standardized math scores grow by less than tenth of a standard deviation (.03) per year. Mobile students saw their standardized math scores drop by a larger, but still negligible amount (.06 standard deviations per year).

With one exception, student and school level characteristics minimally explain the math scores of Black male students. The proportion of students receiving subsidized lunch was associated with an even smaller reduction in standardized math scores, but a one standard deviation increase in the average standardized math scores of the cohort increases a student's standardized math scores by a quarter of a standard deviation. Teacher retention also has a negligible association with standardized math scores, but none of the other student or school level variables have a statistically significant association with standardized math scores.

Finally, to the extent that variation in individual initial status or rate of change has been explained by the time-varying and time invariant predictors, we need to examine the variance components. The level-1 residual variance, σ_ϵ^2, summarizes the average scatter of an individual's observed outcome values around his own true change trajectory. Comparing Model 1 to Model 3, we find a decline in the variation (from .24 to .20), which suggests the addition of time-varying parameters proved to be profitable. The level-2 variance components quantify the amount of unpredicted variation in the individual growth parameters. σ_0^2 assesses the unpredicted variability in true initial status (the scatter of the π_{oi} around γ_{00}); σ_1^2 assesses the unpredicted variability in true rates of change (the scatter of π_{1i} around γ_{10}). Comparing Model 2 to Model 4, we find a decline in the residual variation in the initial status, which was helped by the addition of time invariant student characteristics to our model. The covariance of the level-2 residuals, σ_{01}, summarizes the magnitude and direction of the association between the intial status and the rate of change, controlling for the given predictors. In our models, they are negligible. And finally, as we examine the relative goodness-of-fit using the AIC and BIC indicators, model 3 fits better due to the smaller information criterion.

Conclusion

The goal of this analysis was to understand what characteristics of students, schoolmates, and schools attended by Black male students, especially those amenable to policy change, were associated with growth in math proficiency during the elementary and middle school years. We found that the average Black male performed below proficiency in math in the fourth grade, so we decided to focus on standardized, rather than actual math scores, to examine the trend in proficiency. The standardized math scores of Black male students showed no trend; thus, on average, math proficiency was declining over time (no significant growth term). After including

all the theoretically relevant variables, we found that the 4th grade standardized math scores of subsidized-lunch students were somewhat lower than those of non-subsidized students and those of retained students were substantially lower than their counterparts. However, we still found no trend in the standardized math scores of Black male students, though the standardized scores of special education students showed a modest upward trend and the standardized math scores of mobile students showed a somewhat larger, but still negligible, downward trend.

Finally, the average math score of a Black male student's cohort mates appear to be the only variable amenable to policy manipulation that has a sizeable association with the growth of their standardized math scores, suggesting that putting Black male students in more challenging learning environments may be the best way to increase math proficiency over time. By themselves, other policy decisions (reducing student mobility, teacher turnover, or special education classification; increasing attendance or spending on after-school programming; or hiring more qualified or experienced teachers) all appear to have no or negligible associations with growth in math scale scores.

Lastly, we remain cautious about interpreting our results. First, it is difficult to disentangle when the students in our sample are assigned to special education status. We can only gauge if schools assigned this status to students starting in the 6th grade to 12th grade, but schools may start classifying as early as the 4th grade without reporting this assignment. Thus, the early poor performance of special education students in our sample may be a result of being assigned to this status or the lower performance may have predated the special education classification. About a fifth of our sample was classified with special education status, so attempting to sort this out is an important focus of future research. Second, Black male students attend schools that spend about 1 percent of their budgets on after school programming and there

is little variation in spending across schools. Thus, substantial increases in such spending, that would take after school programming to a whole new level, could be associated with improvements in math proficiency, which we could not detect in this sample. Third, some of our independent variables may be insensitive to differences in our sample because our sample is homogenous. For example, the vast majority of schools in our sample had high rates of free or reduced price lunch. Although there is little variation in this variable, it is unwise to conclude that the rate of subsidized lunch has no association with student achievement. Fourth, we were not afforded data on classroom characteristics which would have provided information on teacher motivation, stereotypes, and classroom behavior—all of which may influence math score achievement over time.

Implications for Policy and Future Research

Besides our finding of no trend in math performance for Black male students during the elementary and middle school years, perhaps the most startling realization we have had from our study is that the Black male students in the school district we examined (89 percent of whom come from low SES families, Table 1) attend schools with so many other poor (82 percent, Table 2) and minority (95 percent, Table 2) students. Poverty and minority concentration among Black students is well known and studies show that such high concentrations make for challenging educational environments.

Nevertheless, seeing such high concentrations in our sample has given us a new appreciation for the limited degrees of freedom that exist for improving Black male achievement through marginal changes in the variables we are able to observe. Perhaps the best proxy we have for the learning environment in our data is the average math scores of schoolmates, the only student or school level variable we find to be substantially associated with math performance of

Black male students. How could marginal changes in the variables we are able to observe substantially increase the average math scores in schools with so many poor and minority students? Small wonder that the emerging efforts that appear to be improving the academic achievement of Black male students use radical, not marginal, changes to improve learning environments (e.g. single sex schools, cash incentives, charter schools and promise neighborhoods). Two examples are Family Awards Program in New York City that uses conditional cash transfers for educational and other outcomes to increase academic achievement for high performing students (Riccio, Dechausay et. al, 2010) and the Harlem Children's Zone to parents and whole communities in improving the educational environment for black students resulting in more general gains in academic achievement (Dobbie & Fryer 2009). Recent efforts in school districts with high Black male enrollment have focused on creating personal education plans for students who are performing below proficient in math. These plans provide academic, social, and health supports for struggling students and engage community members to ensure future student success (Schott Foundation, 2012).

References

Almond, D., & Currie, J. (2011). Human capital development before age five. *Handbook of Labor Economics*, *4*, 1315-1486.

Alspaugh, J.W., (1999) The Interaction Effect of Transition Grade to High School With Gender and Grade Level Upon Dropout Rates. Paper prepared for the American Educational Research Association, Montreal, Quebec, Canada.

Amato, P. R. (2005). The impact of family formation change on the cognitive, social, and emotional well-being of the next generation. *The Future of Children*, *15*(2), 75-96.

"Author" (2003).

"Author" (2008).

"Author" (2010).

Balfanz, R. (2009). Can the American high school become an avenue for advancement for all? *Future of Children 19*(1), 17-36.

Balfanz, R., & Herzog, L. (2005, March). *Keeping the middle grades students on track to graduation: Initial analysis and implications.* Paper presented at the second Regional Middle Grades Symposium, Philadelphia, PA.

Barber, J. G., & Delfabbro, P. (2000). Predictors of adolescent adjustment: Parent-peer relationships and parent-child conflict. *Child and Adolescent Social Work Journal*, *17*(4), 275-288.

Barnes, G.M., Hoffman, J.H., Welte, J.W., Farrell, M.P., & Dintcheff, B.A. (2007). Adolescents' time use: Effects on substance use, delinquency and sexual activity. *Journal of Youth and Adolescence, 36*, 697–710.

Brand, S., Felner, R., Shim, M., Seitsinger, A., & Dumas, T. (2003). Middle school improvement and reform: Development and validation of a school-level assessment of climate, cultural

pluralism, and school safety. *Journal of Educational Psychology*, *95*, 570-588.

Brooks-Gunn, J., Klebanov, P. K., & Duncan, G. J. (1996). Ethnic differences in children's intelligence test scores: Role of economic deprivation, home environment, and maternal characteristics. *Child development*, *67*(2), 396-408.

Camara, W. J., & Schmidt, A. E. (1999). Group Differences in Standardized Testing and Social Stratification. Report No. 99-5. *College Entrance Examination Board*.

Conchas, G. Q., & Rodriguez, L. F. (2007). *Small schools and urban youth: Using the power of school culture to engage students*. Corwin Press.

Crosnoe, R. (2005). Double disadvantage or signs of resilience? The elementary school contexts of children from Mexican immigrant families. *American Educational Research Journal, 42*(2), 269–303.

Crosnoe, R., Johnson, M. K., & Elder, G. H. (2004). School size and the interpersonal side of education: An examination of race/ethnicity and organizational context. *Social Science Quarterly, 85*, 1257–1274.

Darling, N. (2005). Participation in extracurricular activities and adolescent adjustment: Cross-sectional and longitudinal findings. *Journal of Youth and Adolescence, 34*, 493–505.

Darling-Hammond, L. (1999). *Teacher quality and student achievement: A review of state policy evidence*. Seattle, WA: Center for the Study of Teaching and Policy, University of Washington.

Darling-Hammond, L. (2000). New standards and old inequalities: School reform and the education of African American students. *Journal of Negro Education*, 263-287.

Darling-Hammond, L., Ancess, J., & Ort, S.W. (2002). Reinventing high school: Outcomes of the coalition campus schools project. *American Educational Research Journal, 39*, 639–

673.

Deming , D., & Dynarski , S. (2008). The lengthening of childhood. *Journal of Economic Perspectives, 22*, 71-92.

Dobbie, W., & Fryer Jr, R. G. (2009). *Are high quality schools enough to close the achievement gap? Evidence from a social experiment in Harlem* (No. w15473). National Bureau of Economic Research.

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

Durlak, J. A., & Weissberg, R. P. (2007). The impact of after-school programs that promote personal and social skills: Collaborative for Academic, Social, and Emotional Learning (CASEL).

Fashola, O., & Cooper, R. (1999). Developing the academic talents of African American students during the non-school hours: Four exemplary programs. *Journal of Negro Education*, 130-137.

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

Foley, E., Mishook, J., Thompson, J., Kubiak, M., & Supovitz, J. (2008). Beyond Test Scores. Annenberg Institute for School Reform.

Ford, D.Y. (2012). Culturally different students in special education: Looking backward to move forward. *Exceptional Children, 78*(4), 391-405.

Frank, R. G., & Meara, E. (2009). *The effect of maternal depression and substance abuse on child human capital development* (No. w15314). National Bureau of Economic Research.

Fry, R. (2005). *The high schools Hispanics attend: Size and other key characteristics*.

Washington: Pew Center.

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

Goldhaber, D., Choi, H., & Cramer, L. (2007). A descriptive analysis of the distribution of NBPTS-Certified teachers in North Carolina. *Economics of Education Review*, *26*(2), 160–72.

Goldhaber, D. (2008). Teachers matter, but effective teacher quality policies are elusive. *Handbook of research in education finance and policy*, 146-65.

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of educational research*, *66*(3), 361-396.

Griffith, J. (2000). School climate as group evaluation and group consensus: Student and parent perceptions of the elementary school environment. *The Elementary School Journal*, *101*(1): 35-61.

Grossman, J. B., Lind, C., Hayes, C., McMaken, J., & Gersick, A. (2009). *The cost of quality out-of-school-time programs*. Public/Private Ventures.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2009). New evidence about brown v. board of education: The complex effects of school racial composition on achievement. *Journal of Labor Economics*, *27*(3), 349-383.

Hoffman, J.P. (2006). Extracurricular activities, athletic participation, and adolescent alcohol use: Gender-differentiated and school-contextual effects. *Journal of Health and Social Behavior, 47,* 275–290.

Jenkins, D. M., & McEwin, C. K. (1992). Which school for the fifth grader? Programs and practices in three grade organizations. *Middle School Journal*, *23*(4), 8-13.

Jia, Y., Way, N., Ling, G., Yoshikawa, H., Chen, X., Hughes, D., ... & Lu, Z. (2009). The influence of student perceptions of school climate on socioemotional and academic adjustment: A comparison of Chinese and American adolescents. *Child Development*, *80*(5), 1514-1530.

Judge, S., & Watson, S. M. (2011). Longitudinal outcomes for mathematics achievement for students with learning disabilities. *Journal of Educational Research, 104*(3), 147-157.

Klingner, J. K. (2004). The science of professional development. *Journal of Learning Disabilities, 37,* 248–255.

Klugman, J. (2012). How resource inequalities among high schools reproduce class advantages in college destinations. *Research in Higher Education, 53*(8), 803-830.

Kuziemko, I. (2006). Using shocks to school enrollment to estimate the effect of school size on student achievement. *Economics of Education Review, 25*, 63–75.

Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2006). Out-of-school time programs: A meta-analysis of effects for at-risk students. *Review of Educational Research, 76*, 275–313.

Lee, J. (2012). Educational equity and adequacy for disadvantaged minority students: School and teacher resource gaps toward national mathematics proficiency standard. *Journal of Education Research, 105* (1)*,* 64-75.

Leithwood, K., & Jantzi, D. (2009). A review of empirical evidence about school size effects: A policy perspective. *Review of Educational Research*, *79*(1), 464-490.

Metropolitan Center for Urban Education. (2009). *A close look at the dropout crisis: Examining Black and Latino males in New York City*. New York, NY.

Moos, R.H. (1979). *Evaluating educational environments: Procedures, measures, findings, and policy implications.* San Francisco: Jossey-Bass.

Obradovic, J., Long, J.D., Cutuli, J. J., Chan, C., Hinz, E., Heistad, D., & Masten, A.S. (2009). Academic achievement of homeless and highly mobile children in an urban school district: Longitudinal evidence on risk, growth, and resilience. *Development and Psychopathology, 21*, 493–518.

Orfield, G. (Ed.). (2004). *Dropouts in America: Confronting the graduation rate crisis*. Cambridge: Harvard University Press.

Orfield, G., & Lee, C. (2006). *Racial Transformation and the Changing Nature of Segregation*. Cambridge, MA: The Civil Rights Project at Harvard University.

Oswald, D. P., Best, A. M., & Coutinho, M. J. (2006). Individual, Family, and School Factors Associated with the Identification of Female and Male Students for Special Education. *International Journal of Special Education*, *21*(3), 120-137.

Parthenon Group (2005). *NYC Secondary Reform Selected Analysis*. New York.

Pierce, K. M., Bolt, D. M., & Vandell, D. L. (2010). Specific features of after-school program quality: Associations with children's functioning in middle childhood. *American journal of community psychology*, *45*(3), 381-393.

Riccio, J., Dechausay, N., Greenberg, D., Miller, C., Rucks, Z., & Verma, N. (2010). Toward reduced poverty across generations: Early findings from New York City's conditional cash transfer program. *MDRC, March*.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, 247-252.

Roderick, M., Nagaoka, J., & Allensworth, E. (2005). Is the glass half full or mostly empty? Ending social promotion in Chicago. *Yearbook of the National Society for the Study of Education*, *104*(2), 223-259.

Rothstein, R. (2004). *Class and Schools: Using Social, Economic, and Educational Reform to Close the Achievement Gap*. Washington: Economic Policy Institute.

Rumberger, R.W. (2003). The causes and consequences of student mobility. *Journal of Negro Education, 72*(1), 6-21.

Rumberger, R. W., & Palardy, G. J. (2005). Test scores, dropout rates and transfer rates at alternative indicators of high school performance. *American Educational Research Journal, 42*, 3–42.

Sass, T. R., Hannaway, J., Xu, Z., Figlio, D., & Feng, L. (2012). Value added of teachers in high-poverty schools and lower poverty schools. *Journal of Urban Economics*.

Scafidi, B., Sjoquist, D.L., & Stinebrickner, T.R. (2007). Race, poverty, and teacher mobility. *Economics of Education Review, 26*(2), 145–159.

Schott Foundation. (2012). *The urgency of now*. Cambridge, MA: Author. http://blackboysreport.org/.

Skiba, R. J. (2002). Special education and school discipline: A precarious balance. *Behavioral Disorders*, *27*(2), 81-97.

Spera, C., Wentzel, K. R., & Matto, H. C. (2009). Parental aspirations for their children's educational attainment: Relations to ethnicity, parental education, children's academic performance, and parental perceptions of school climate. *Journal of youth and adolescence*, *38*(8), 1140-1152.

Stata, C. (2011). Stata statistical software: Release 12. *Special Edition*.

Temple, J. A., & Reynolds, A. J. (2000). School mobility and achievement: Longitudinal findings from an urban cohort. *Journal of School Psychology*, *37*(4), 355-377.

Van Maele, D., & Van Houtte, M. (2010). The quality of school life: Teacher-student trust

relationships and the organizational school context. *Social Indicators Research, 100*, 85–100.

Voight, A., Shinn, M., & Nation, M. (2012). The longitudinal effects of residential mobility on the academic achievement of urban elementary and middle school students. *Educational Researcher, 41*(9), 385-392.

Willson , V. L., & Hughes , J. N. (2006). Retention of Hispanic/Latino students in first grade: Child, parent, teacher, school, and peer predictors. *Journal of School Psychology, 44*, 31-49.

Winsler, A., Hutchison, L.A., De Feyter, J.J., Manfra, L., Bleiker, C., Hartman, S.C., & Levitt, J. (2012). Child, family, and childcare predictors of delayed school entry and kindergarten retention among linguistically and ethnically diverse children. *Developmental Psychology, 48*(5)*,* 1299-1314.

Table 1. Percent Distributions of Selected Student
Characteristics of All Students

Total Sample size	7272
Family Background Characteristics	
Foreign Born	9.1
Free/Reduced Lunch	89.2
Student Level Characteristics	
Retained in the 5th Grade	0.6
Classified Special Education	19.8
Transferred Schools between 5th- 8th Grades	
Never	11.3
Once	73.7
More than once	15.0

Table 2. Percent Distributions and Means of School Level
Characteristics in 4th (1998/99) and 8th (2002/03) grade years

	4th Grade (1998/99)	8th Grade (2002/03)
Student Enrollment Characteristics		
Enrollment	2116 (1202)	1102 (475)
School Race/Ethnicity Makeup		
Black	62.8	59.3
Latino	27	27.9
White	5.5	6.7
Asian	4.2	5.7
Qualify for Subsidized Lunch	81.6	74.8
Math Scale Score	639.7	695.2
Resources		
Percent of Budget for Afterschool Programming	1.2	1.3
Teachers with fewer than 3 years experience	16.6	23.0
Climate		
Student Attendance Rate	84.3	83.7
Teacher Retention[a]	75.9	73.8
School Type		
Enrolled in small schools (<440)	4.2	15.0

Note: Standard deviations provided in parentheses.
[a]Teachers who return to the same school after 2 years.

Table 3. Predicting Math Exam Scores from MultiLevel Models (N=7,039)

	Model 1	Model 2	Model 3
Fixed Effects			
Intercept (initial status)	629.1**	0.05**	0.47**
	(0.43)	(0.01)	(0.07)
Year (rate of change)	15.0*	0.001	-0.02
	(0.09)	(0.002)	(0.01)
Family Background Variables			
Foreign Born			0.08
			(0.06)
Foreign Born x Year			0.01
			(0.02)
Ever free or reduced lunch status			-0.11*
			(0.05)
Lunch Status x Year			0.01
			(0.01)
Student Level Variables			
Special Education Status			-0.63**
			(0.05)
Special Ed. Status x Year			0.03*
			(0.01)
Retention in the 5th grade			-.75*
			(.31)
Retention x year			0.07
			(.09)
Transferred Schools between 5th-8th grades			-0.06**
			(0.01)
School Level: Enrollment Characteristics			
Students with free lunch status (in percent)			-0.0005*
			(0.0002)
Cohort Average Math Score			0.25**
			(0.008)
School Level: Resources			
Budget for Afterschool Programming (in percent)			0.01
			(0.008)
Teachers teaching outside area of certification (in percent)			0.0005
			(0.0006)

Teachers with fewer than 3 years of experience (in percent)			-0.0008
			(0.0006)
School Level: Climate			
Teacher Retention (in percent)			-0.0010**
			(0.0003)
Attendance Rate			-0.0004
			(0.0004)
School Level: School Type			
Small School			-.006
			(0.03)
Variance components			
Level 1			
Within-person	0.24	0.24	0.20
Level 2	(0.0020)	(.002)	(.004)
In initial status	0.72	0.72	0.63
	(0.0100)	(.01)	(.036)
In rate of change	0.01	0.01	0.01
	(0.0007)	(.0007)	(.003)
Covariance	-0.02	-0.02	-0.04
	(.002)	(.002)	(.01)
Model Fit			
AIC		70776.67	34532.55
BIC		70827.35	34710.82

Notes: **p<.01, *p<.05

[a] These models use unstandardized continuous math scores.

Figure captions

Figure 1. Math Proficiency Limits and Sample Mean Math Scale Scores in 1999 and 2003

Figure 2. Proficient and Below Proficient Limits in Math Scores and the Sample Mean Math Scale Scores in 1999 and 2003

Figure 3. Math proficiency levels of sample in 1999 and 2003

Figure 4. Math Exam Scores for Four Sample Students and the Average Trajectory in Math Scores for the Full Sample