Trends of School Effects on Student Achievement: Evidence from NLS:72, HSB:82, and NELS:92

Spyros Konstantopoulos

September 2005

Trends of School Effects on Student Achievement: Evidence from NLS:72, HSB:82, and NELS:92

Spyros Konstantopoulos
*Northwestern University
and IZA Bonn*

ABSTRACT

Trends of School Effects on Student Achievement: Evidence from NLS:72, HSB:82, and NELS:92[*]

The impact of schools on student achievement has been of great interest for the last four decades. This study examines trends of school effects on student achievement employing three national probability samples of high school seniors: NLS:72, HSB:82, and NELS:92. Hierarchical linear models are used to investigate school effects. The findings reveal that the substantial proportion of the variation in student achievement lies within schools not between schools. There is also considerable between school variation in achievement, which becomes larger over time. Schools are more diverse and more segregated in the 1990s than in the 1970s. In addition, school characteristics such as school region, school SES, and certain characteristics of the student body of the school, such as students' daily attendance, students in college preparatory classes, and high school graduates enrolled in colleges are important predictors of average student achievement. The school predictors explained consistently more than 50% of the variation in average student achievement across surveys. We also find considerable teacher heterogeneity in achievement within schools, which suggests important teacher effects on student achievement. Teacher heterogeneity in student achievement was larger than school heterogeneity, which may indicate that teacher effects have a relatively larger impact on mathematics and science student achievement than school effects.

Corresponding author:

Spyros Konstantopoulos
School of Education and Social Policy
Northwestern University
2120 Campus Drive
Evanston, IL 60208
USA
Email: spyros@northwestern.edu

A major goal of American education is to provide high quality educational experiences and adequate educational preparation for all of the groups that compose the national population. Many of the policies devised to meet this goal attempt to ensure that school materials and human resources are allocated equitably across schools. As a result, research about the impact of school characteristics on students' academic performance is of great interest.

The question of whether schools differ significantly in increasing students' academic achievement is essential in education. Hence, identifying school factors that make schools more effective is crucial. Coleman and his colleagues (1966) were the first who studied the association between school inputs and student achievement using national probability samples of elementary and secondary students. In their pioneering work Coleman et al. estimated education production functions in order to quantify the association between students' academic performance in standardized tests and school and family input measures. One of the key findings of the Coleman Report was that when the socioeconomic background of the students was held fixed, the differences among schools accounted "for only a small fraction of differences in pupil achievement" (Coleman et al., 1966, p. 21). In other words, variations in school characteristics were not closely associated with, and had hardly any effect on variations in student achievement.

The Coleman report generated a series of studies that were conducted to further assess the effects of school resources on academic achievement. It is noteworthy that for the last three decades, there have been disagreements among educational researchers, practitioners and policymakers about the relative impact-importance of school characteristics on students' academic achievement. The findings of numerous studies are rather mixed and inconclusive.

Some researchers have concluded that there is little or no evidence of a relationship between school factors and student achievement (Hanushek, 1986; 1989), while others report that the impact of school factors on test scores may be substantial (Greenwald, Hedges, and Laine, 1996).

The Present Study

This study examines the impact of schools on student achievement (mathematics, reading, and science) over time using national probability samples of high school seniors. Our objective is to determine whether schools "make a difference." There are at least two ways to gauge school effects. The first approach, which is typical in the school effects literature, is to identify the efficacy of certain school characteristics in predicting academic achievement via education production functions (Hanushek, 1986; Hedges, Laine, and Greenwald, 1994). The second approach of identifying schools effects is to compute the variation of academic achievement between schools. This approach involves the creation of the distribution of school level achievement by computing the average achievement for each school. The variance of this distribution indicates how much average achievement differs from school to school. A significant between-school variation in achievement is therefore an index of the impact of schools on student achievement. The advantage of this approach is that it does not need to identify and measure school characteristics. On the other hand, it does not single out specific school characteristics that make schools more or less effective. In this study, we employed both approaches.

Since individuals are nested within schools, school effects models are appropriately described by multi-level models (Raudenbush and Bryk, 2002). Consider the case where

students are nested within schools. This includes two levels of hierarchy: a within school level and a between school level. Conceptually the first level involves a series of within school regressions. The second level equation is a school level regression. The variance of the error term at the first level indicates the between student within school variation in achievement. The variance of the random school intercepts at the second level indicates the between school variation in achievement. This study employed two-level hierarchical linear models (HLM) to investigate school effects. Whenever teacher identifiers were available, we used three level HLM to examine teacher effects as well. Specifically, the three level model decomposes the total variation in achievement into between students within teacher within school, between teacher within school, and between school components. The between teacher variation in this case suggests teacher effects independently of school effects.

We investigated school effects on student academic achievement and determined how these effects changed over time from 1972 to 1992. We used data from three rich surveys spanning 20 years that queried nationally representative samples of high school seniors: the National Longitudinal Study of the High School Class of 1972 (NLS-72), High School and Beyond first follow-up from 1982 (HSB:82), and the 1992 second follow-up of the National Educational Longitudinal Study of the Eighth Grade Class of 1988 (NELS:92). A unique characteristic of NELS:92 sample was that the students were not only linked to schools, but to teachers as well (in mathematics and science). Hence, we were able to determine teacher effects as well and whether teachers or schools matter the most.

Definition of School Effects

We use the term school effects to indicate the associations between school structural features (e.g., school sector) and resources (e.g., pupil-teacher ratio) and student achievement, while controlling for important student background characteristics (e.g., student SES). The conceptual framework that guides the present study is based on the economic perspective of school effects research (Rumberger, & Palardy, 2005). This framework's empirical evidence has originated from education production function studies (Hanushek, 1986). We consider the associations between school factors and achievement as being strictly correlational not causal. Given the observational nature of our data and the type of school effects we examine, it would be difficult to infer causality (Raudenbush, & Wilms, 1995). In school effects research academic achievement is modeled as a function of school characteristics, controlling appropriately for student background. The school effects are estimated at the school level, where the adjusted for student background average school achievement is modeled as a function of school characteristics (Lee, 2000).

This study examines what Raudenbush and Wilms (1995) call Type A school effects. The Type A effects incorporate a variety of school characteristics that are not necessarily restricted to the practice of the school staff. For example, school SES and school composition are attributes of a school. In contrast, pupil-teacher ratio or college prep classes may be viewed as a treatment effect of a school (Raudenbush, & Wilms, 1995). Hence, Type A effects include different measures of school effects and school specific treatment effects are not easily detected.

Some of the school characteristics used indicated school context/composition or structure. For example, school region, urbanization, and sector may be categorized as

school structure, while school SES, minority concentration, daily attendance, dropout rates, and college attendance rates of high school graduates may be categorized as school context/composition. Other school factors indicated school resources (e.g., pupil teacher ratio), school organization/curriculum (e.g., college prep classes, advanced placement courses), and length of academic year. School resources and school organization/curriculum characteristics are typically more likely to be viewed as treatments effects of the school (or school effects that indicate causality). All of the school characteristics used in this study have been previously used in school effects research as well as in education production function studies as important correlates of school outputs such as student achievement (see Bryk et al., 1993; Card and Krueger, 1992; Coleman et al., 1966; D'Agostino, 2000; Lee, 2000).

Finally, we also defined school effects as the between school variation in achievement. By using HLM we were able to compute what proportion of the total variation in achievement is between schools. This between school variation provides a broad estimate of the importance of schools on student achievement. The use of school level random effects has been previously advocated by some researchers to represent school effects (see Constant, & Konstantopoulos, 2003; Raudenbush, & Wilms, 1995). The variance of these school specific random effects (typically intercepts) indicates school differences in average achievement, and shows that schools matter.

Related Literature

In the very early stage of school effects research studies examined the association between school inputs and outputs such as student achievement (Coleman et al., 1966).

The main findings of these studies were the importance of family background characteristics, such as socio economic status of the family, in explaining variation in student achievement, and the relatively small impact of school characteristics on student achievement (see Mosteller & Moynihan, 1972). The Coleman report in particular encouraged a considerable body of research that examined the usefulness of school factors in predicting student achievement the last 30 years.

In the 1980s methodological advances in the school effects research helped to more accurately assess the importance of school factors in predicting student achievement. During this period multi-level statistical models were introduced and allowed the use of student characteristics and school factors at the appropriate level of analysis (Raudenbush, & Bryk, 1986). Specifically, the flexibility of multi level models allowed for the use of student characteristics at the student level and the school factors at the school level.

Student Background and Achievement

Previous research has demonstrated the relation between student characteristics and student outcomes such as academic achievement. There is little disagreement over the existence of a positive correlation between family background and student achievement (Jencks et al, 1979). For example, the relationship between test scores and family SES characteristics is well replicated in the social sciences (Neff, 1938; White, 1982; White, Reynolds, Thomas, & Gitzlaff, 1993). The strength of the relationship between SES variables and achievement varies from study to study in part because researchers operationally define socioeconomic status in different ways, and this can affects the magnitude or strength of the association (White, 1982). Traditional measures of

socioeconomic status include parental educational level, and family economic resources (see Coleman, 1969; Konstantopoulos, Modi, & Hedges, 2001). In addition, other factors such as parent's occupation, family size, family structure, quality of housing, and household possessions have been considered SES measures (White, 1982; White et al., 1993). The importance of gender and race effects on student achievement has also been demonstrated (Hedges, & Nowell, 1995; 1998). The student background variables used in this study were student gender, race, and family SES. Family SES is a composite measure, which was created by using information about parental educational attainment, occupation, and family income.

School Variables and Achievement

The social composition of students in a school has also been found to influence achievement. For example, school composition measured as percent of minority or disadvantaged students in the school is negatively associated with achievement and accounts for a substantial amount of variability in achievement (Bryk & Raudenbush, 1988). In particular, schools with higher proportions of minority and disadvantaged students have lower average achievement than other schools. Other school composition variables such as school SES are also significantly associated with student achievement (Lee & Bryk, 1989). Higher SES schools have typically higher average achievement than lower SES schools. In addition, the effect of another potential compositional variable such as the length of the school year on achievement has also been studied. Specifically, the length of school year has been shown to have positive effects on learning (D'Agostino, 2000), and to provide positive returns in education (Card & Krueger, 1992).

The usefulness of school structure has also been demonstrated. School structure variables such as school location or urbanization and school sector are significantly related to student achievement. For example, Coleman and Hoffer (1987) found that, on average, students' verbal and mathematics achievement growth in Catholic schools was higher than that in public schools. This sector effect holds even when student characteristics such as academic background, minority status, and SES were held constant (Bryk, Lee, and Holland, 1993; Raudenbush & Bryk, 1989).

There is a debate in the school effects literature about whether school resources are consistently important predictors of achievement. There is some evidence however, that class size has a significant effect on student achievement and student dropout rates (Nye, Hedges, & Konstantopoulos, 2000; Rumberger, & Thomas, 2000). For example, a recent study on allocation of education resources such as class size demonstrated a positive relationship between small classes and academic achievement (Nye, Hedges, & Konstantopoulos, 2000). In addition, pupil-teacher ratio, a proxy of class size, has been an important factor of successful preschool and school programs (Zigler and Styfco,1994). In this study we measured class size as the average pupil-teacher ratio in a school.

Method

<u>Data</u>

Data from three major surveys conducted the last 30 years were used in this study. All surveys tested nationally representative samples of high school students, that is, each survey used a stratified national probability sample of high school students. In all data sets we used the twelfth grade samples and thus we investigated the academic performance of

high school seniors who participated in each survey. All variables used were comparable across datasets. Sampling weights, which permitted inferences about specifically defined national populations (e.g., high school seniors) were provided.

The National Longitudinal Study of the High School Class of 1972 (NLS-72) is a national probability sample of high school seniors designed to represent all twelfth graders enrolled in public or private American high schools in the spring of 1972. Of the 16,860 seniors, a sample of 15,800 students who completed a 69 minute, six-part battery measuring both verbal and non-verbal skills was used in the analyses. We used the NLS reading and mathematics test scores in this study.

In the spring of 1980, two cohorts of tenth and twelfth grade students enrolled in public and private schools were surveyed for the High School and Beyond study (HS&B-80). The sophomores were resurveyed in 1982 when they were seniors (HS&B-82). To maintain comparability with the other samples we limited the 1982 sample to students still enrolled in school. We used data from the 1982 follow-up national probability sample of 26,216 seniors. Students completed a 68 minute test battery similar in format to the battery used in NLS-72, but with slightly different content. We used the HSB reading, mathematics, and science test scores in this study.

The National Educational Longitudinal Study of the Eighth Grade Class of 1988 (NELS-88) used a two stage national probability sample of 24,599 eighth graders enrolled in public and private schools in 1988. These students were followed for four years and were resurveyed in 1992, when they were high school seniors. Our sample consisted of 12,921 seniors of the second follow-up (1992). Students completed an 85-minute battery of four cognitive tests with a similar format as in HSB and NLS, with a slightly different

content. Nonetheless, in all three surveys there was some content comparability. We used the NELS reading, mathematics, and science test scores in this study.

Variables of Interest

The outcome variables we used in this study were mathematics, reading, and science test scores. We standardized all achievement measures to assure that all scores are in the same metric. This also allowed us to interpret the between school variances as the percentage of variation in student achievement accounted for by schools.

The set of explanatory variables included both student and school level characteristics. At the student level we included student gender, race/ethnicity, and a composite measure of student SES (a composite of parental education, occupation, and income). The school level variables included indexes of school structure such as school region, school urbanization, school sector, indexes of school composition such as school SES, minority concentration, daily attendance, dropout rates, college attendance rates of high school graduates, and length of school year (in weeks), indexes of school resources such as pupil-teacher ratio, and indexes of school organization/curriculum such as students in college preparation courses and advanced placement courses. The coding for some of the predictors is summarized in the appendix.

Analysis

Most educational data have hierarchical structure, where students are grouped-nested within organizational units such as schools. These kinds of data provide information that describes both students and schools. Nonetheless, until recently, classical statistical

methods, such as linear regression, were used extensively in school effects research. In multiple linear regression settings typically school and student level predictors are introduced simultaneously at the student level, and hence the analysis is conducted at the individual level. Such regression models fail to take into account the clustering nature of the data and its consequences. In addition, the typical regression models do not allow the estimation of the between school variation. In contrast, HLM take the clustering of students within schools into account, allow the use of student and school variables at different levels, and permit the computation of between school variances (see Raudenbush, and Bryk, 2002). Each of the levels in this structure is represented by its own submodel. Each submodel reveals associations between the set of explanatory variables and the outcome at that level.

The proposed analysis depended on the use of two-level HLM to explore the between-school variability and the effects of school characteristics on average student achievement. The first level (or student level) was specified by a linear regression additive model, where we control for student background. The second level (or school level) renders the associations between school characteristics and student achievement net of the effects of student background. In our specifications, all school specific intercepts were treated as random variables at the school level. The residual terms at the second level are random effects, and the variance components of these random effects represent the between school variation, which indicates the variability of the impact of schools on student achievement or school effects. Important student characteristics such as gender, race/ethnicity and socio-economic status (SES) were included in the student level model. At the school level, the school specific intercepts are regressed on a set of school

characteristics described in the previous section. In addition, the gender, race, and SES achievement gap were allowed to vary across schools. In all HLM analyses individual weights were employed at the first level to make projections to the national population of high school seniors.

A three-level model was also employed to gain some insight on the role of teachers and schools in student achievement. Specifically, we initially ran the simplest possible three-level model (unconditional) where only the constant terms were included in the level specific equations. Such a model decomposes the variance into three parts: the within teacher between student variation, the within school between teacher variation (or teacher effects), and the between school variation (or school effects). Significant variation in student achievement at the teacher and school levels indicates important teacher and school effects or that teachers and schools matter. We also ran a three level model including level 1 predictors and computed the variation of the achievement gap between teachers and schools.

Centering Student Predictors

The major objective of the study is to estimate school level effects adjusting for student characteristics. In other words, our objective is to examine the association between school characteristics and average school achievement net of the effects of the student level covariates such as gender, race, and SES. In an HLM setting this means that the school specific intercepts (or school average achievement), which are treated as random at the school level should be adjusted for the effects of gender, race, and SES. As

Raudenbush and Bryk (2002) argue when the main interest is to "estimate the association between a level 2 predictor and the mean of Y, adjusting for one or more level-1 covariates" (page 142), then grand mean centering is more appropriate. Hence, we used grand mean centering for the student level predictors to examine school level effects net of the effects of student characteristics.

However, another objective of the study is to estimate school effects as between school variation in achievement. As Raundenbush and Bryk show, the choice of centering affects the estimation of the variance components of the student level coefficients (including the intercept). That is different types of centering provide different estimates of the variances of the random effects at the school level. Specifically, the use of grand mean centering may underestimate the variance of the school level random effects (see Raudenbush, & Bryk, 2002). In this case, the level-1 predictors may explain between school variation, and hence the estimates of the variance components are smaller than in group mean centering. We followed Raudenbush and Bryk's recommendation and used group mean centering in order to estimate the between school variance components. That is, we conducted all analyses twice: use of group mean centering to obtain "correct" estimates of the between school variances, and use of grand mean centering to estimate the association between school predictors and achievement controlling for student characteristics.

Model Building

Overall, three different two-level HLM were examined. The first model was an unconditional model. This model is used to describe how much of the variation in

achievement is between schools and how much is within schools. The second model introduced important student level predictors such as family SES (the effect of high levels of social class), gender (the effect of being female), and race (the effect of being minority). All student level coefficients (including the intercept) were treated as random at the school level. However, school predictors were not used in the second model. The third model added school characteristics as school level predictors. Hence, the school specific intercepts were regressed on the set of school predictors at the school level. We also ran two three-level models: one unconditional model, and one with all level one predictors.

Comparability of Measures Across Surveys

All datasets that were used in this study were acquired from three major studies (NLS, HSB, and NELS) that are part of the National Education Longitudinal Studies program instituted by the National Center of Education Statistics (NCES). One objective of this longitudinal program was to represent the educational experiences of our students in the 1970s, 1980s, and 1990s. NCES reports contend that cross-sectional time-lag comparisons for high school seniors in 1972, 1982, and 1992 are possible, and that these data can be regarded as a series of repeated cross-sections of high school seniors (see Green, Dugoni, & Ingels, 1995). The sample designs of all three studies are similar and the achievement tests while not identical, are rather similar (see Green et al., 1995). All achievement tests intended to capture similar domains of academic achievement (see Hedges, & Nowel, 1999). Some NCES reports argue that there were common items in NLS and HSB, and HSB and NELS for mathematics and reading, and hence some content

comparability of the achievement measures over time is warranted (see Rock et al., 1985; Green et al., 1995).

The use of equating methods that put mathematics and reading scores for high school seniors in 1972 and 1982 on a common scale have been previously demonstrated (see Rock et al., 1985). Rock et al. concluded that comparisons of test scores in NLS and HSB can reasonably indicate change along the same dimension over time. In this study we used linear equating methods (e.g., creating z scores) to put mathematics, reading, and science scores on a common scale (see Hedges, & Nowel, 1999). Of course the standardization creates comparable indexes of achievement across surveys under the assumption that the tests are linearly equatable. Previous research has documented that, even though typically IRT equating methods lead to greater stability of equating results, linear equating also performs acceptably when tests are comparable (Kolen, & Brennan, 1995; Petersen, Cook, & Stocking, 1983). In addition, linear equating is widely used by commercial test publishers and it is known to provide reasonably good results. Nonetheless, even though NLS, HSB, and NELS were designed to be as similar as possible, caution should "be exercised in comparing NLS-72, HS&B, and NELS:88 data." (Green et al., 1995, page 125).

In addition, the items used to construct the independent variables are very similar across all three datasets. We coded all independent variables similarly to achieve comparability for all predictors.

Results and Discussion

<u>NLS:72</u>

The results of model II for mathematics and reading are presented in the first and third columns of Table 1 respectively. On average, male students performed better than female students in math achievement by ¼ of a standard deviation, but female students outperformed male students in reading achievement by 1/17 of a standard deviation. Minority students had significantly lower achievement than white students in mathematics and reading achievement (about 2/3 of a standard deviation). As expected, there was a positive and significant relationship between high levels of family SES and student achievement, indicating that students from affluent families have higher achievement than other students net of gender and race effects. The social class gap was about ½ of a standard deviation. The average school mathematics and reading achievement varied significantly between schools. Similarly, the race and social class achievement gap varied significantly between schools. Overall, group and grand mean centering of the level 1 predictors produced similar estimates.

Insert Table 1 Here

In the third specification (or Model III) both student and school level predictors were introduced in the level specific linear equations, with school characteristics predicting

the school intercepts. The predictive efficacy of the school characteristics is summarized in columns two and four. On average, schools in the North East and North Central region of the country had higher achievement in mathematics than schools in the South. In reading, schools in the North East part of the country also outperformed schools in the South. Schools with high daily attendance and high proportions of high school graduates in college had higher mathematics and reading achievement than other schools. Affluent schools had higher mathematics and reading achievement than less affluent schools. The gender, race, and SES gap was somewhat smaller in Model III. Overall, group mean centering produced similar results.

HSB:82

The results for the second model are presented in the first, third, and fifth columns of Table 2. On average, white and high SES students performed better in mathematics, reading, and science than other students. The race gap was more than ½ of a standard deviation, and the social class gap somewhat smaller than ½ of a standard deviation. Male students performed better than female students in mathematics and science. The gender gap was insignificant in reading however. The average school mathematics, reading, and science achievement varied significantly across schools. The gender, race, and SES gap also varied significantly between schools. As in NLS, the results using group mean centering were comparable.

Insert Table 2 Here

The predictive efficacy of the school characteristics is summarized in columns two, four, and six. On average, high SES schools as well as schools in the North East, North Central, and West region of the country had higher mathematics, reading, and science achievement than other schools. Schools with high daily attendance, low dropout rates, and high proportions of high school graduates in colleges had also higher mathematics, reading, and science achievement than other schools. School sector (private school) had a positive effect on reading achievement, and high minority schools had a negative effect on science achievement. The gender, race, and SES gap was somewhat smaller in Model III. The results using group mean centering were similar.

<u>NELS:92</u>

The results for Model II are presented in columns one, three, and five of Table 3. As in NLS:72 and HSB:82, on average, white and high SES students performed better in mathematics, reading, and science than other students. The race gap ranged from about 0.4 standard deviations in mathematics and reading to more than ½ of a standard deviation in science. The SES gap was consistently larger than ½ of a standard deviation. As in NLS:72 and HSB:82, male students performed better than female students in mathematics (1/12 of a standard deviation) and science (1/4 of a standard deviation), but contrary to HSB:82 female students achieved significantly higher than their male counterparts in reading (1/4 of a standard deviation). As in NLS:72 and HSB:82, the variance component estimates revealed that the average mathematics, reading, and science achievement varied

significantly across schools. In addition, the gender, race and SES gap varied significantly between schools. Again, the results from the group mean centering analyses were similar.

Insert Table 3 Here

The predictive efficacy of the school characteristics is summarized in columns two, four, and six. On average, affluent schools, as well as schools in the North East, North Central, and West region of the country had higher mathematics, reading, and science achievement than other schools. In addition, schools with high proportions of students in college preparatory courses had higher mathematics and reading achievement than other schools. Schools with low pupil/teacher ratios and high proportions of high school graduates in colleges had higher mathematics achievement than other schools. High minority schools have lower average science achievement than other schools. The gender, race, and SES gap was somewhat smaller in Model III. Group mean centering provided comparable estimates.

Between School Variation

NLS:72

The variance components estimates of the random school intercepts are reported in the right panel of Table 4. The unconditional model, which included only level 1 and level 2 intercepts suggested that the school specific mathematics and reading achievement varied significantly across school units. The between school variance for both mathematics and reading was nearly 10% of the total variation in achievement. Notice that because we standardized student achievement, these variance components estimates also reflect the intraclass correlation. The significant variation in average achievement among schools indicates that schools are heterogeneous in student achievement. The majority of variation in achievement is within, not between, schools in 1972 (about 90% of the total variation). Besides student effects, this type of variation may indicate the importance of resources in each school (including teachers). The school predictors explained 75% of the between school variation in average mathematics achievement, and approximately 60% of the between school variation in reading. Still, the between school variation was statistically significant. In addition, the race and social class achievement gap varied significantly across schools in mathematics and reading.

We employed likelihood ratio tests to examine whether the school predictors produced a significant reduction in the between school variation in achievement. All likelihood ratio tests were significant at the 0.001 level indicating the importance of school predictors.

Insert Table 4 Here

HSB:82

As in NLS:72, in the unconditional model the school mathematics, reading, and science achievement varied significantly among schools. The between school variation was somewhat less than 20% of the total variation in mathematics and science, and a little more than 10% in reading. The between school variation in mathematics is 35% larger in 1982. As in NLS:72 it appears that in 1982 the majority of the variation in achievement is within schools. The average mathematics, reading, and science achievement varied significantly among schools even when school characteristics were taken into account. Nonetheless, the school level predictors reduced the between school variation in student achievement by more about 75% in mathematics and reading and about 70% in science. In addition, the gender, race, and SES achievement gap varied significantly between schools for all test scores.

We employed likelihood ratio tests to examine whether the school predictors produced a significant reduction in the between school variation in achievement. All likelihood ratio tests were significant at the 0.001 level indicating the importance of school predictors.

NELS:92

As in NLS:72 and HSB:82, in the unconditional model the average school mathematics, reading, and science achievement varied significantly across schools. The between school variation in mathematics, reading, and science was approximately 20%. Consistently, over time, the majority of the variation in achievement is within schools, which may partly indicate the important effects of school resources (including teacher effects). The average mathematics, reading, and science achievement varied significantly among schools even when school characteristics were taken into account. Nonetheless, the school level predictors reduced the between school variation in student achievement by nearly 60% in reading and science, and 65% in mathematics. It is remarkable that across all surveys the school predictors explained consistently more than 50% of the between school variation in achievement. As in HSB, the gender, race, and SES achievement gap varied significantly between schools for all test scores.

We employed likelihood ratio tests to examine whether the school predictors produced a significant reduction in the between school variation in achievement. All likelihood ratio tests were significant at the 0.001 level indicating the importance of school predictors.

Overall the estimates of the variance components in the unconditional models of HSB and NELS are comparable to variance components estimates reported in previous studies. For example, in HSB mathematics, the between school variance estimate is 0.19, while Raudenbush and Bryk (2002) report an intraclass correlation of 0.18, and Lee and Bryk (1989) report an intraclass correlation of 0.19. Similarly, in NELS the between school variance estimate for reading is 0.19, while Lee and Croninger (1993) report an

intraclass correlation of 0.19. Lee and Smith (1996) provide an estimate of the intraclass

correlation for science gain scores of about 0.20 for NELS, while our variance component

estimate of science achievement status is 0.23. Finally, our variance components estimates

are comparable to those reported in a recent study that used all data sets that are included

in the present study (see Hedges, & Hedberg, 2004).

Analyses Using Data from All Surveys

We also conducted analyses using data from all three surveys. Specifically, since

all three surveys provide comparable data we decided to pool all data across surveys and

use hierarchical linear models to analyze them. Pooling data from comparable studies has

been used in previous work (see Wong, & Rosenbaum, 2004). Although the sample size

for each of the surveys is large, analyses using data from all three surveys should in

principle produce tests (for the coefficients of the school characteristics in particular) that

have higher statistical power. This indicates a higher probability of detecting school effects

assuming that these effects exist. The between school model remained the same as in the

cross-sectional analyses by survey. The within school model changed slightly, since we

included dummies to control for the effects of the year of the survey. We constructed two

dummies for the year of the survey for reading and mathematics: one for 1992 and one for

1982, with 1972 being the comparison group. Only one dummy was constructed for

science (e.g., 1992) since science data in 1972 were not available. In order to conduct these

analyses we assumed that the data from these different surveys are comparable (see Green

et al., 1995). We also assumed that the student and school characteristics used in our

models have the same effects across all surveys.

Results

The results from the pooled analyses are summarized in Table 5 (left panel). In mathematics, males outperformed females by 1/6 of a standard deviation. The race gap was even larger and hovered around ½ of a standard deviation (favoring whites). The social class gap was somewhat smaller than the race gap. High SES students outperformed their peers by about ½ of a standard deviation. Students in 1982 scored on average higher than students in 1972 in mathematics, but the gap was small (1/20 of a standard deviation). On average, schools in the North East, North Central, and West region of the country had higher achievement in mathematics than schools in the South. Private schools performed higher in mathematics on average than public schools. Schools which offer advanced placement courses also performed higher than other schools in mathematics. As expected, schools with high daily attendance and high proportions of high school graduates in college had higher mathematics achievement than other schools. Finally, school with lower proportions of dropouts as well as affluent schools had higher mathematics achievement than other schools. The between school variation in mathematics achievement (unconditional model) was 17% of the total variation. The school predictors explained nearly 70% of the between school variation in student achievement, and this variance reduction is statistically significant.

Insert Table 5 Here

In reading, females outperformed their male peers by 1/13 of a standard deviation. The race gap was the same as in mathematics and the social class gap was slightly smaller (about 4/10 of a standard deviation). The HSB82 effect was the same as in mathematics. The results for reading regarding the school characteristics were identical with those reported in mathematics, with the exception that proportion of dropouts was not statistically significant. The between school variation in reading achievement (unconditional model) was 13% of the total variation. The school predictors explained nearly 65% of the between school variation in student achievement, and this variance reduction is statistically significant.

In science, males outperformed their female peers by ¼ of a standard deviation. The race and social class gap was the same as in reading and mathematics. The results for school characteristics were identical to those reported for mathematics. In addition, in science, rural schools and schools with low proportions of minority students had higher achievement on average than other schools. The between school variation in science achievement (unconditional model) was 21% of the total variation. The school predictors explained nearly 65% of the between school variation in student achievement, and this variance reduction is statistically significant.

We also conducted analyses where we centered the predictors in the within school regression around their mean and the results were overall comparable. In the latter analyses rural schools in mathematics, rural and low minority schools in reading, and suburban schools in science also outperformed other schools. Overall the results using group mean centering of the level 1 predictors overestimated the coefficients of the school

characteristics, which is expected, since the estimates are not adjusted for student characteristics (see Raudenbush, & Bryk, 2002). These results are overall consistent with the results obtained for each survey.

Overall, school effects measured as between school variation were more pronounced in science and mathematics than in reading. This may indicate that science and mathematics are mostly learned in school and thus may be more directly influenced by school resources (such as teachers), or that there is more variation in how (or how well or how much) science and mathematics are taught in schools.

Sensitivity Analyses

Dropouts

Since the samples include high school seniors, individuals who are not in school at grade 12 are excluded from the analyses. Hence it is possible that the 12th grade samples are selected and the estimates may be biased (positively or negatively) or different from their "true" population parameters. On the other hand, if the students who drop out are not systematically different that those who stayed in school, then one would expect the potential bias to be close to zero. Alternatively, this would suggest that the estimates would be similar for samples without dropouts and samples including dropouts. Fortunately, some of the samples provided such data. Specifically, HSB and NELS provide data including dropouts. Hence, we conducted sensitivity analyses using the samples that included dropouts in order to examine whether the estimates from the samples which excluded

dropouts are similar to those which included dropouts. Unfortunately, such samples were not available for NLS.

Results

Both in HSB and NELS and for all subject matters, the samples which included dropouts and had complete student and school data were very similar to the samples that excluded dropouts and had complete data. As a result, the estimates of the analyses that included dropouts were almost identical to those reported in the analyses that excluded dropouts. In other words, we were not able to detect selection bias, but this may be due to the fact that the complete data in the two samples were very similar.

Grade 10 Analyses

Another way to examine the effects of possible selection bias is to conduct analyses using samples of high school students in previous grades. Such samples of students are available for HSB and NELS, but not in NLS. Our assumption is that these samples should not have experienced the same selection effects because of dropping out as the 12th grade samples. In other words, selection effects should be smaller at earlier grades than at Grade 12. Hence, if the results of the grade 10 analyses are comparable to those from the grade 12 analyses, this would indicate that dropout effects at grade 12 are minimal or that the dropout effects are similar in grades 10 and 12. HSB provided samples of 10th grade students and NELS provided samples of 8th and 10th grade students. For comparability purposes we decided to use the 10th grade samples for HSB and NELS.

Results

The results of these analyses are reported in Table 6 (right panel). Overall the results obtained from analyses on the Grade 10 samples are comparable to those reported for 12th graders. Hence, it appears that either the selection effects were similar for 10 and 12 graders or that the selection effects for 12 graders is minimal. Even though we did not find strong evidence for selection bias in HSB and NELS, it is difficult to generalize these results to NLS.

Insert Table 6 Here

Achievement Gains Analyses

Since information on previous achievement was available for HSB and NELS we also conducted analyses using 10th grade achievement as a covariate at the first, within school, level. We decided to use grade 10 achievement in NELS to achieve comparability with HSB. These analyses examine school effects on achievement gains. Notice that the use of sophomore achievement as an adjustment of prior student ability is hardly the most appropriate pre school measure in a school effects study, since it can be contaminated with school effects and its inclusion as a covariate would most likely underestimate the cumulative nature of school effects. We report these results from the pooled data across

both surveys in the right panel of Table 5, since the results from the cross –sectional

analyses and those obtained from the pooled analyses were similar.

Results

As expected controlling for previous achievement at grade 10 changed dramatically

most of the level 1 and level 2 coefficients. The gender, race, and social class achievement

gaps were still significant, but were reduced by more than 50%. The coefficients of the

school predictors were insignificant except region, proportion of high school graduates

going to college and schools offering advanced placement courses in mathematics, and

school sector in reading. Hence, once previous achievement is controlled for, the

predictive power of student and school characteristics was decreased. However, it is

plausible that grade 10 achievement entails school effects and hence its presence in the

equation influences the school effects considerably. In sum, in the achievement gains

model the effects of school characteristics are most likely underestimated.

Trends in the Achievement Gap and the Between School Variation

We also used meta analytic methods to determine linear trends in the achievement

gap and the between school variation over time (see Konstantopoulos, & Hedges, 2004).

Specifically, we regressed the female, race, and social class regression coefficients and

their variance components estimates (since the achievement gaps were treated as random at

the between school model) on year of survey. We also regressed the variances of the

school specific random intercepts on year of survey (for models I and III). The regressions included weights so that more precise survey estimates would have a larger influence on the average estimate.

<u>Results</u>

The results from the trend analyses are summarized in Table 4. In mathematics, the female gap became smaller over time by about 2/3. Still, it was significant favoring males in 1992. The race gap decreased by somewhat less than 50%, but it was still 1/3 of a standard deviation in 1992. The SES gap became slightly larger over time. All trends were significant and positive. In reading, the gender gap favoring females increased over time by more than four times. The race gap, as in mathematics, decreased significantly over time by 40%, but still it was nearly 1/3 of standard deviation in 1992. The SES gap increased over time as in mathematics. In science the gender gap remained the same, the race gap decreased, and the SES gap increased.

The variance components estimates of the random school specific intercepts obtained from the unconditional model indicate that the between school variation in achievement increased over time. Specifically, in mathematics and in reading the between school variation in achievement increased by about 1.8 times. In science the increase was nearly 1.2 times. Hence, it appears that schools become more heterogeneous in achievement over time. Similar patterns were observed in the residual between school variation in achievement (controlling for school predictors). Specifically, the residual between school variation in achievement more than doubled in mathematics and in reading

over time. In science the residual between school variation was twice as large in 1992 than in 1972. In addition, the variation of the average school gender, race, and SES gap also increased over time in mathematics, reading, and science. Notice that all trend estimates are positive indicating that the between school variation increased over time. Overall these results indicate that schools have become more heterogeneous and hence more segregated with respect to student achievement. In other words, the distance between low achieving and high achieving schools has increased over time. Similarly, the distance between schools with smaller and larger average achievement gap also increased.

Three Level HLM

The NELS:92 data permitted analyses that examined teacher effects since the students were linked to teachers and schools. Specifically, students enrolled in math or science courses were linked to math or science teachers. Each student was assigned to one teacher, and therefore there is no dependency between groups. To determine teacher and school effects simultaneously we used a three level HLM. The results of the three-level HLM analyses for mathematics and science are presented in Table 7. The unconditional model in mathematics indicated that 34% of the total variation in achievement is between teachers within schools and 18% of the variation is between schools. In science 23% of the variation in achievement was between teachers within schools and 18% between schools. It is striking that the teachers' heterogeneity is nearly twice as large as the schools heterogeneity in mathematics. In science, the between teacher variation is nearly 25% larger than the between school variation. However, the between teacher variation was consistently larger than the between school variation, which indicates that naturally

occurring teacher effects are larger than naturally occurring school effects. This finding is consistent with findings reported in previous work using elementary school data from a large scale experiment (see Nye, Konstantopoulos, & Hedges, 2004). In addition, it is noteworthy that the estimates of the between school variation in the three level model were comparable to those obtained from the unconditional two level model. This may suggest that only a small part of the between school variation is due to teacher effects. In contrast, our data indicate that a considerable part of the within school variation is due to teacher variation or teacher effects.

In addition, the female, race, and SES achievement gap varied significantly between teachers within schools, but not between schools both in mathematics and in science (except for the minority gap). This indicates that teachers may play a more significant role in the mathematics and science achievement gap than schools do. Nonetheless, the statistical significance of these estimates should be treated with caution since the tests were computed using only part of the data.

Insert Table 7 Here

Conclusion

The present study included three major national surveys conducted in the early 1970s, 1980s and 1990s that provided information about student achievement, student background, and school characteristics. We examined the between school variation in

achievement as well as the importance of school characteristics in predicting student achievement and explaining variation in achievement over time.

Our analyses provided conclusive evidence about the importance of school factors in predicting student academic achievement over time (net of the effects of student background). Across all surveys school region, school SES, and certain characteristics of the student body in the school had a considerable effect on student achievement. Students attending schools in the South had lower average achievement than students attending schools in other regions. In contrast, students in high SES schools had higher average achievement than students in lower SES schools. In addition, schools with high levels of student attendance, high proportions of graduates in colleges, and low dropout rates had higher average achievement than other schools. The school predictors explained consistently more than 50% of the between school variation in achievement across test scores and surveys and this variance reduction was significant.

We also examined school effects as between school variation in achievement. The variance decomposition suggested that most of the variation in student achievement is within schools. The within school dispersion in achievement was nearly five times larger than the between school variation across test scores and data sets. Nonetheless, the between school variation in achievement was significant and indicated the importance of schools. Over time, the between school variation in achievement or school effects increased significantly. School in the 1990s were more heterogeneous in student achievement than in the 1970s. This also indicates that schools in the 1990s are more segregated since there is a larger gap between low and high achieving schools. Overall, the

predictive efficacy of the school predictors and the significant between school variation provided convincing evidence that schools matter.

The results from the three-level HLM provided additional interesting findings. Specifically, these results suggested that teachers matter as well. The between teacher within school variation was consistently larger than the between school variation in mathematics and science. It seems that differences in achievement between teachers in the same school are much larger than differences in achievement between schools. This may suggest that there are large differences in how (or how well) teachers teach mathematics and science within schools. It appears that the teachers students are assigned to may be more important than the schools they attend. Policy initiatives will have to take into consideration this teacher heterogeneity as well as the school heterogeneity. It is also noteworthy, that the between teacher variation or teacher effects was a considerable part of the within school variation, not the between school variation. Hence, an important part of achievement differences within schools is due to teachers.

The gender gap in mathematics favoring male students decreased over time significantly. In reading however, the gender gap favored females in 1972, was reversed in 1982 favoring males, and was reversed again in 1992 favoring females. The trend was not significant. The gender difference in science achievement favored male students consistently. The race achievement gap favoring whites decreased significantly over time for mathematics and reading. The race gap in science also decreased. In contrast, the positive effect of family SES increased over time. The trend was significant for mathematics and reading.

The gender, race, and SES gap varied significantly across schools. This indicates that schools treat the gender, race, and SES gap differently. This further indicates the importance of schools. The results from the three level model however, showed that differences in the achievement gap are mainly due to differences in teachers within schools. This finding supports the significance of teachers in treating the achievement gap. The between school variation in the achievement gap became larger over time suggesting that in the 1990s there was a larger gap between schools with smaller and larger achievement gaps.

Across all datasets and specifications the between school variation in mathematics and science is much larger than that in reading. This may be because mathematics and science is mostly learned in school or that there is more variation in how mathematics and science are taught in schools. Reading, on the other hand, is more likely to be learned (in part) outside of school and thus the influence of schools on reading may be smaller, or there is less variation in how (or how well or how much) reading is taught in school. This finding is also consistent with previous work using elementary school data (see Nye et al, 2004)

In sum, our findings indicated important school and teacher effects and that school factors are important predictors of student achievement net of the effects of student background. Nonetheless, future studies should further examine how the allocation of school resources affects students in different ages using representative samples of students.

References

Bryk, A. S., & Raudenbush, S. W. (1988). Toward a more appropriate conceptualization of research on school effects: A three-level hierarchical linear model. American Journal of Education, 97, 65-108.

Bryk, A.S., Lee, V.E., & Holland, P.B. (1993). Catholic schools and the common good. Cambridge, MA: Harvard University Press.

Card, D., & Krueger, A. (1992). Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. Journal of Political Economy, 100, 1-40.

Coleman, J. S. (1969). Equality and achievement in education. Boulder, CO: Westview Press.

Coleman, J. S., & Hoffer, T. B. (1987). Public and private schools: The impact of communities. New York: Basic Books.

Coleman, J. S., Campbel, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1996). Equality of Educational Opportunity. Washington DC: U.S. Government Printing Office.

Constant, A., & Konstantopoulos, S. (2003). School effects and labor market outcomes for young adults in the 1980s and 1990s. Applied Economics Quarterly, 49, 5-22.

D'Agostino, J. V. (2000). Instructional and school effects on students' longitudinal reading and mathematics achievements. School Effectiveness and School Improvement, 11, 197-235.

Green, P., Dugoni, B. L., & Ingels, S. J. (1995). <u>Trends among high school seniors 1972-1992</u>. Washington, D.C.: National Center for Education Statistics.

Greenwald, R., Hedges, L. V., & Lane, R. D. (1996). The effects of school resources on student achievement, <u>Review of Educational Research, 66</u>, 361-396.

Hanushek, E. A. (1986). The Economics of Schooling: Production and Efficiency in Public Schools. <u>Journal of Economic Literature 24</u>, 1141-77.

Hanushek, E. A. (1989). The impact of differential expenditures on school performance. <u>Educational Researcher, 18,</u> 45-51.

Hedges, L. V., & Hedberg, E. (2004). <u>The variance structure of academic achievement in America</u>. Paper presented at the annual meeting of the American Educational Research Association, San Diego, April 2004.

Hedges L. V., & Nowel, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. <u>Science, 269</u>, 41-45.

Hedges, L. V., & Nowell, A. (1999). Changes in the Black-White gap in achievement test scores. <u>Sociology of Education, 72,</u> 111-135.

Hedges, L. V., Laine, R., & Greenwald, R. (1994). Does Money Matter: A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes. <u>Educational Resercher, 23</u>, 5-14.

Jencks, C. S., Bartlett, S., Corcoran, M., Crouse, J., Eaglesfield, D., Jackson, G., McClelland, K., Mueser, P., Olneck, M., Schwartz, J., Ward, S., & Williams, J. (1979). <u>Who gets ahead? The determinants of economic success in America</u>. New York: Basic Books.

Kolen, M. J., & Brennan, R. L. (1995). Test equating: Methods and practices. New York: Springer.

Konstantopoulos, S., & Hedges, L. V. (2004). Meta-analysis. In D. Kaplan (Ed.), The SAGE handbook of quantitative methodology for the social sciences. Thousand Oaks, CA: Sage.

Konstantopoulos, S., Modi, M., & Hedges, L. V. (2001). Who are America's gifted? American Journal of Education, 109, 344-382.

Lee, V. E. (2000). Using hierarchical linear modeling to study social contexts: The case of school effects. Educational Psychologist, 35, 125-141.

Lee, V. E., & Bryk, A. S. (1989). A multilevel model of the social distribution of high school achievement. Sociology of Education, 62, 172-192.

Lee, V. E., & Smith, J. B. (1996). Collective responsibility for learning and its effects on gains in achievement for early secondary school students. American Journal of Education, 104, 103–145.

Mosteller, F, & Moynihan, D., P. (1972). A pathbreaking report. In F. Mosteller and D., P. Moynihan (Eds.), On Equality of Educational Opportunity (pp. 3-66). New York: Random House.

Neff, W. S. (1938). Socioeconomic status and intelligence. Psychological Bulletin, 35, 727-757.

Nye, B. A., Hedges, L. V., & Konstantopoulos, S. (2000). The effects of small classes on academic achievement: The results of the Tennessee class size experiment. American Educational Research Journal. 37, 123-151.

Nye, B. A., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects. Educational evaluation and Policy Analysis, 26, 237-257.

Petersen, N. S., Cook, L. L., & Stocking, M. L. (1983). IRT versus conventional equating Methods: A comparative study of scale stability. Journal of Educational Statistics, 8, 137-156.

Raudenbush, S. W., & Bryk, A. S. (1986). A hierarchical model for studying school effects. Sociology of Education, 59, 1-17.

Raudenbush, S. W., & Bryk, A. S. (1989). Methodological advances in analyzing the effects of schools and classrooms on student learning. Review of Research in Education, 15, 423-475.

Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical Linear Models. Thousand Oaks, CA: Sage Publications.

Raudenbush, S. W., & Wilms, J. D. (1995). The estimation of school effects. Journal of Educational and Behavioral Statistics, 20, 307-335.

Rock, D.A., Hilton, T.L., Pollack, J.M., Ekstrom, R.B., Goertz, M.E. (1985). Psychometric analysis of the NLS-72 and the High School and Beyond Test Batteries. Washington, D.C.: National Center for Education Statistics.

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

Rumberger, R. W., & Thomas, S. L. (2000). The distribution of dropout and turnover rates among urban and suburban high schools. Sociology of Education, 73, 39-67.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. Psychological Bulletin, 91, 461-481.

White, S. W., Reynolds, P. D., Thomas, M. M., & Gitzlaff, N J. (1993). Socioeconomic status and achievement revisited. Urban Education, 28, 328-343.

Wong, M., & Rosenbaum, J. E. (2004). Changes in the educational attainment process over the past three decades: How have things changed? Paper presented at the annual meeting of the American Sociological Association, San Francisco, CA. August, 2004

Zigler, E., & Styfco, S. J. (1994). Is the Perry Preschool Project better than Head Start? Early Childhood Research Quarterly, 9, 269-287.

Appendix
Coding of Independent Categorical Variables

Student Characteristics

Gender
Female   Equals 1 if individual is female, 0 otherwise

Race/Ethnicity
Minority   Equals 1 if individual is Black, Hispanic, American Indian
                                Asian, or Other Race, and 0 otherwise

Family SES
High SES   Equals 1 if student lives in a high SES family (top quartile), 0 otherwise

---

School Characteristics

School Region
North East   Equals 1 if school is in the North East region of the country, 0 otherwise
North Central   Equals 1 if school is in the North Central region of the country, 0 otherwise
West   Equals 1 if school is in the West region of the country, 0 otherwise

School Urbanization
Rural   Equals 1 if student attends school in a rural community, 0 otherwise
Suburban   Equals 1 if student attends school in a suburban community, 0 otherwise

School Sector
Private School   Equals 1 if high school is private, 0 otherwise

High Minority School   Equals 1 if school is in the top quartile of percent of minority students, 0 otherwise

Advanced Placement Courses   Equals 1 if school offers advanced placement courses, 0 otherwise

Table 1.
Two-Level HLM Fixed and Random Effects Estimates: NLS:72 Mathematics and Reading: Grade 12

	Mathematics Achievement		Reading Achievement	
	Model II	Model III	Model II	Model III
Female	-0.233*	-0.236*	0.059*	0.056*
Minority	-0.669*	-0.618*	-0.661*	-0.632*
SES	0.500*	0.435*	0.451*	0.370*
North East		0.172*		0.138*
North Central		0.117*		0.038
West		-0.070*		0.001
Rural School		0.013		0.013
Suburban School		-0.049		-0.003
Private School		0.043		0.122
Pupil-Teacher Ratio		-0.006		-0.003
Advanced Placement Courses		0.050		0.032
Students in College Prep Classes		0.0004		-0.0002
Length of School Year		-0.006		0.002
Percent of High School Graduates in College		0.005*		0.004*
Students Daily Attendance		0.008*		0.006*
High Minority School		-0.039		-0.007
Dropout Rates		0.001		0.003
School SES		0.268*		0.370*

* $p < 0.05$

Table 2.

Two-Level HLM Fixed and Random Effects Estimates: HSB:82 Mathematics, Reading, and Science: Grade 12

	Mathematics Achievement		Reading Achievement		Science Achievement	
	Model II	Model III	Model II	Model III	Model II	Model III
Female	-0.156*	-0.155*	-0.031	-0.029	-0.268*	-0.263*
Minority	-0.573*	-0.525*	-0.589*	-0.543*	-0.690*	-0.641*
SES	0.531*	0.473*	0.460*	0.406*	0.432*	0.386*
North East		0.226*		0.133*		0.179*
North Central		0.203*		0.097*		0.193*
West		0.139*		0.093*		0.217*
Rural School		0.016		0.006		0.048
Suburban School		0.001		-0.042		0.008
Private School		0.081		0.121*		-0.036
Pupil-Teacher Ratio		-0.0002		0.001		-0.002
Advanced Placement Courses		0.045		-0.014		-0.014
Students in College Prep Classes		0.001		0.001		0.0002
Length of School Year		-0.001		-0.003		-0.012
Percent of High School Graduates in College		0.003*		0.002*		0.002*
Students Daily Attendance		0.008*		0.007*		0.010*
High Minority School		-0.009		-0.045		-0.139*
Dropout Rates		-0.005*		-0.004*		-0.005*
School SES		0.342*		0.208*		0.170*

$* \ p < 0.05$

Table 3.

Two-Level HLM Fixed and Random Effects Estimates: NELS:92 Mathematics, Reading, and Science: Grade 12

	Mathematics Achievement		Reading Achievement		Science Achievement	
	Model II	Model III	Model II	Model III	Model II	Model III
Female	-0.086*	-0.082*	0.246*	0.246*	-0.270*	-0.267*
Minority	-0.393*	-0.346*	-0.418*	-0.383*	-0.550*	-0.488*
SES	0.620*	0.527*	0.536*	0.446*	0.511*	0.440*
North East		0.178*		0.144*		0.180*
North Central		0.140*		0.106*		0.125*
West		0.255*		0.214*		0.246*
Rural School		0.015		-0.073		0.070
Suburban School		-0.014		-0.082		0.021
Private School		0.012		-0.011		-0.052
Pupil-Teacher Ratio		-0.007*		-0.001		-0.004
Advanced Placement Courses		0.049		0.030		0.053
Students in College Prep Classes		0.002*		0.002*		0.001
Length of School Year		-0.001		-0.004		-0.002
Percent of High School Graduates in College		0.003*		0.0007		0.001
Students Daily Attendance		0.004		0.005		0.005
High Minority School		-0.023		-0.043		-0.095*
Dropout Rates		-0.002		-0.0001		0.0005
School SES		0.485*		0.442*		0.442*

* $p < 0.05$

Table 4.
Trends in Level 1 Estimates and Variance Components of Random Effects: NLS:72 to NELS:92

	Mathematics							
	Coefficient (Model III)			VC (Model III)			VC of School Intercept	
Survey	Female	Minority	SES	Female	Minority	SES	Unconditional Model	Model III
NLS:72	-0.236*	-0.618*	0.435*	0.013	0.106*	0.021*	0.125*	0.031*
HSB:82	-0.155*	-0.525*	0.473*	0.039*	0.030*	0.088*	0.191*	0.049*
NELS:92	-0.082*	-0.346*	0.527*	0.132*	0.158*	0.108*	0.220*	0.074*
Trend	0.008*	0.014*	0.005*	0.006	0.004	0.004	0.005	0.002

	Reading							
	Coefficient (Model III)			VC (Model III)			VC of School Intercept	
Survey	Female	Minority	SES	Female	Minority	SES	Unconditional Model	Model III
NLS:72	0.056*	-0.632*	0.370*	0.008	0.137*	0.039*	0.105*	0.029*
HSB:82	-0.029	-0.543*	0.406*	0.028*	0.019	0.059*	0.133*	0.035*
NELS:92	0.246*	-0.383*	0.446*	0.139*	0.207*	0.092*	0.192*	0.081*
Trend	0.009	0.012*	0.004	0.007	0.004	0.003	0.004	0.003

	Science							
	Coefficient (Model III)			VC (Model III)			VC of School Intercept	
Survey	Female	Minority	SES	Female	Minority	SES	Unconditional Model	Model III
HSB:82	-0.263*	-0.641*	0.386*	0.046*	0.105*	0.051*	0.188*	0.055*
NELS:92	-0.267*	-0.488*	0.440*	0.145*	0.163*	0.122*	0.234*	0.099*

* $p < 0.05$
Note: VC: Variance Component; Unconditional Model: No Predictors Included;
Model III: Student and School Level Predictors Included

Table 5.
Pooled Estimates of Grade 12 Samples for HSB and NELS

| | Pooled Estimates | | | | | |
| | Achievement Status | | | Achievement Gains | | |
Variable	Mathematics	Reading	Science	Mathematics	Reading	Science
Female	-0.160*	0.075*	-0.263*	-0.067*	0.042*	-0.106*
Minority	-0.509*	-0.520*	-0.573*	-0.076*	-0.125*	-0.195*
SES	0.485*	0.413*	0.411*	0.122*	0.115*	0.121*
North East	0.193*	0.142*	0.174*	0.052*	0.062*	0.073*
North Central	0.162*	0.085*	0.163*	0.008	0.033*	0.049*
West	0.090*	0.100*	0.228*	0.007	0.066*	0.103*
Rural School	-0.002	-0.008	0.065*	-0.011	0.009	0.024
Suburban School	-0.024	-0.032	0.025	-0.003	0.009	0.033
Private School	0.088*	0.124*	-0.017	0.028	0.066*	0.019
Pupil-Teacher Ratio	-0.003	-0.0004	-0.002	-0.001	0.001	0.0003
Advanced Placement Courses	0.056*	0.020	0.004	0.032*	0.016	0.023
Students in College Prep Classes	0.001	0.0005	0.0004	0.0001	0.0006	0.00005
Length of School Year	-0.001	-0.002	-0.002	0.00006	-0.002	0.001
Percent of High School Graduates in College	0.003*	0.002*	0.001*	0.001*	-0.0003	0.0004
Students Daily Attendance	0.006*	0.005*	0.007*	0.001	0.002	0.002
High Minority School	-0.009	-0.020	-0.099*	0.040*	-0.008	-0.017
Dropout Rates	-0.002*	-0.001	-0.004*	-0.0006	0.0002	-0.0007
School SES	0.373*	0.318*	0.330*	0.072*	0.096	0.020

* $p < 0.05$

Table 6.
Estimates of Grade 10 and 12 samples for HSB and NELS

	HSB:82					
	Grade 12			Grade 10		
Variable	Mathematics	Reading	Science	Mathematics	Reading	Science
Female	-0.155*	-0.029	-0.263*	-0.098*	-0.011	-0.233*
Minority	-0.525*	-0.543*	-0.641*	-0.452*	-0.425*	-0.543*
SES	0.473*	0.406*	0.386*	0.406*	0.415*	0.353*
North East	0.226*	0.133*	0.179*	0.218*	0.134*	0.133*
North Central	0.203*	0.097*	0.193*	0.235*	0.136*	0.185*
West	0.139*	0.093*	0.217*	0.162*	0.127*	0.170*
Rural School	0.016	0.006	0.048	-0.011	-0.032	0.040
Suburban School	0.001	-0.042	0.008	-0.064	-0.084*	-0.032
Private School	0.081	0.121*	-0.036	-0.005	-0.005	-0.141*
Pupil-Teacher Ratio	-0.0002	0.001	-0.002	-0.001	-0.0003	-0.005*
Advanced Placement Courses	0.045	-0.014	-0.014	0.026	-0.011	-0.021
Students in College Prep Classes	0.001	0.001	0.0002	0.001	0.0007	0.0004
Length of School Year	-0.001	-0.003	-0.012	-0.008	0.005	-0.001
Percent of High School Graduates in College	0.003*	0.002*	0.002*	0.003*	0.002*	0.003*
Students Daily Attendance	0.008*	0.007*	0.010*	0.005	0.008*	0.010*
High Minority School	-0.009	-0.045	-0.139*	-0.102*	-0.087*	-0.214*
Dropout Rates	-0.005*	-0.004*	-0.005*	-0.006*	-0.007*	-0.006*
School SES	0.342*	0.208*	0.170*	0.550*	0.361*	0.379*

	NELS:92					
	Grade 12			Grade 10		
Variable	Mathematics	Reading	Science	Mathematics	Reading	Science
Female	-0.082*	0.246*	-0.267*	-0.014	0.183*	-0.266*
Minority	-0.346*	-0.383*	-0.488*	-0.358*	-0.362*	-0.439*
SES	0.527*	0.446*	0.440*	0.450*	0.436*	0.416*
North East	0.178*	0.144*	0.180*	0.140*	0.145*	0.159*
North Central	0.140*	0.106*	0.125*	0.152*	0.061	0.149*
West	0.255*	0.214*	0.246*	0.200*	0.181*	0.207*
Rural School	0.015	-0.073	0.070	0.004	-0.042	0.001
Suburban School	-0.014	-0.082	0.021	-0.011	-0.076*	-0.053
Private School	0.012	-0.011	-0.052	-0.022	0.066	-0.070
Pupil-Teacher Ratio	-0.007*	-0.001	-0.004	-0.003	-0.005	-0.008*
Advanced Placement Courses	0.049	0.030	0.053	-0.014	-0.002	-0.012
Students in College Prep Classes	0.002*	0.002*	0.001	0.002*	0.001	0.0007
Length of School Year	-0.001	-0.004	-0.002	0.002	-0.0007	-0.001
Percent of High School Graduates in College	0.003*	0.0007	0.001	0.003*	0.004*	0.003*
Students Daily Attendance	0.004	0.005	0.005	0.004	0.007*	0.008*
High Minority School	-0.023	-0.043	-0.095*	-0.072	-0.071	-0.155*
Dropout Rates	-0.002	-0.0001	0.0005	-0.021	-0.011	-0.014
School SES	0.485*	0.442*	0.442*	0.422*	0.276*	0.480*

* $p < 0.05$

Table 7.

Three-Level HLM Variance Components Estimates: NELS:92 Mathematics and Science (Grade 12)

	Mathematics Achievement		Science Achievement	
	Unconditional Model	Model II	Unconditional Model	Model II
Intercept				
Level 2 Variance Component	0.342*	0.375*	0.230*	0.270*
Level 3 Variance Component	0.185*	0.185*	0.182*	0.184*
Female				
Level 2 Variance Component		0.021*		0.058*
Level 3 Variance Component		0.040		0.026
Minority				
Level 2 Variance Component		0.101*		0.102*
Level 3 Variance Component		0.059		0.028*
SES				
Level 2 Variance Component		0.055*		0.064*
Level 3 Variance Component		0.007		0.012

* $p < 0.05$