

SCHOOL EFFECTS AND LABOR MARKET OUTCOMES FOR YOUNG ADULTS IN
THE 1980s AND 1990s

by

Amelie Constant

The University of Pennsylvania

and The Institute for the Study of Labor (IZA)

aconstan@pop.upenn.edu

and

Spyros Konstantopoulos

The University of Chicago

skonstan@midway.uchicago.edu

Abstract

This study examines high school effects on the labor market success of young adults, above and beyond individual and family characteristics. We employ data from two longitudinal, nationally probability samples: the National Longitudinal Study and the High School and Beyond study and the 5th and 4th follow-up studies respectively. We employ multi-level models and adjust for selection into the labor force. Our findings reveal that there are differential, enduring school effects on future earnings for individuals attending different schools net of the effects of important individual and family background characteristics.

Keywords: School effects; Multi-Level Models; Labor Market Performance; School Heterogeneity

Address Correspondence to: Amelie Constant, Population Studies Center, The University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104, USA.

This study was in part supported by a grant funded jointly by the NSF and ESR, Number REC-9987943.

1. INTRODUCTION

The quest of how to designate resources in order to effectively further the goals of educational systems is one of paramount importance for education researchers and practitioners as well as policy makers. Studying school-related assets that promote the educational thriving and prosperity of students in all educational levels and later in life becomes a necessity for every education system.

A major goal of educational systems is to provide high-quality educational experiences and adequate educational preparation for all groups that compose the national population. High-quality education in turn, can provide educational opportunities that improve the educational performance and attainment of all students. Solid scholastic training results in higher educational attainment, and leads to higher earnings via increases in labor market productivity. As the human capital theory posits, education enhances labor market productivity and, therefore, earnings (Becker, 1964). Schooling is pivotal to human capital formation, and lays the foundations of social capital development. Investments in human capital have been hypothesized to yield sizable economic and social rates of return.

The school effects literature typically involves the identification of certain school characteristics and examines their efficacy in predicting academic achievement or labor market performance via education production functions (Hanushek, 1986; Hedges, Laine, and Greenwald, 1994; Card, and Krueger, 1992a, 1992b). Another way of identifying schools effects is by examining the variation of an outcome of interest (i.e., academic achievement or earnings) across the distribution of school units (Hedges, and Konstantopoulos, 2002). In other words, a different approach treats the outcome of

interest as a random variable that varies across school units. This can be accomplished by computing, for example the average achievement or earnings by school and then defining its distribution across all school units. The variance of this distribution indicates how much average achievement or earnings differs from school to school and hence, suggests how important the impact of schools is on these outcome variables. This method decomposes the variation in the outcome variable (i.e., earnings) into two parts: the within and the between school variation. This variance decomposition is useful since it provides insight to questions such as, are there important differences in earnings across school units? In addition, typically the variation within schools is different from the variation between schools. By using mixed or nested models one can decompose the variation of earnings in within and between school components and measure the school heterogeneity in earnings. The between school component of variation is an index of the heterogeneity of earnings across school units. Significant between school variation suggests that there are differential school effects (not uniform) on earnings and hence, that the high schools individuals attend make a difference later in the labor force performance. It also indicates that there are important differences in earnings across schools.

Since individuals are nested within schools, school effects models are appropriately described by multi-level models (Goldstein, 1995, Raudenbush and Bryk, 2002). Consider the case, where students are nested within schools. This involves two levels of hierarchy: a within school level and a between school level. At the first level there are a series of within school regressions and at the second level there is a between school regression. The first level examines the within school or student

variation within schools and the second level determines the between school variation. Notice that the school specific intercepts from the first level are unadjusted or adjusted, depending on whether explanatory variables were included in the regressions or not, school specific indexes. These school specific indexes are treated as random variables at the second level. The school specific residuals in the second level model are called random effects and represent the school effects. The variance of these residuals indicates the impact of schools on future earnings. Hence, the significant role schools play in producing future success in the labor market amounts to whether there are significant differences (variation) among schools with respect to future earnings. This is well captured by the variance component of the school effects (Hedges, and Konstantopoulos, 2002).

The vast literature on the economic returns to schooling has unequivocally found that earnings increase with educational attainment while placing the emphasis on higher education. Studying high school effects on labor market outcomes, such as earnings, is essential since not all high school students attend college, and an even smaller percentage of them finish college. By default, secondary education plays a paramount role in forging indispensable skills into future workers who can contribute to a nation's productivity and a faster growth of a nation's economy. This has substantive policy implications; what resources can be manipulated by policymakers to increase school quality and affect beneficially economically disadvantaged groups; what level of spending should be invested in secondary education; and given the scarce resources what would be the optimal allocation of funds between schooling and welfare.

This study examines high school effects on the labor market performance of young adults using two rich longitudinal datasets. The primary objective of our study is to identify and probe the importance of these effects on the subsequent earnings of young adults in the 1980s and the 1990s. To examine school effects on future earnings we control for important individual and family characteristics employing multi-level models in an empirical Bayes framework (see Morris, 1983, Raudenbush and Bryk, 2002). The advantage of multi-level models is that they control for the nesting of the data and allow for random effects at higher levels. We operationalize school effects in terms of significant between school variation in average earnings adjusting for important individual characteristics such as race, social class, and ability. Our dependent variable is annual labor earnings in 1985 and 1991 respectively. We make use of two rich comparable longitudinal datasets and compare the labor market performance of young adults 10 and 14 years after high school graduation. Our analyses are conducted separately for the women and the men samples. Finally, we properly adapt our statistical model to a labor market setting by employing at the first stage of our analysis a selection equation (probit model). At the second stage of our analyses we include the correction term for nonrandom selection of workers obtained from the probit model as a covariate in our multi level model. In addition, we fine-tuned our multi level model to adjust for the heteroskedastic error term and obtain the correct standard errors of estimates. . *Our analysis is We properly adapt this model to a labor market setting and fine-tune it to adjust for nonrandom selection of wage earners.*

When examining school effects on future outcomes such as earnings it is essential to control for individual characteristics and family background in high school.

One such important covariate is achievement test scores in high school. This variable summarizes the individual's educational experiences and family background up to a specific point in time. Other important covariates include social class and race/ethnicity of the individual.

The principal finding of our study indicates that there are indeed differential, enduring school effects on future earnings for individuals attending different schools net of the effects of student characteristics and family background. Other important secondary findings include that - as documented by other studies - family background plays a key role in determining education and earnings; and that academic achievement strongly influences economic status.

The remainder of this paper is structured as follows. First, we review previous notable research in this area. Second, we describe the datasets we based our analysis upon. Third, we discuss our methodological approach. Fourth, we present and discuss the results of our analysis. In the last section we draw our concluding remarks.

2. PREVIOUS RESEARCH ON SCHOOL EFFECTS

Nowadays the performance of educational systems is closely related to the efficiency of labor markets as well as to the countries' competitive position in the global economy. The role of education has become critical in raising and accelerating the nation's productivity growth. Therefore, the study of school effects on labor market performance is quite important. Coleman and his colleagues (1966) were the first to study the association between school inputs - such as teacher's education, and class size - and student achievement. The Coleman report postulated that when the

socioeconomic background of the students was held fixed school inputs had very little measurable effect on student achievement. In addition, the study also reported negligible between school variation in academic achievement. Hence, one of the key conclusions of the Coleman report was that schools do not make a notable difference with respect to student achievement.

The Coleman report spawned a series of studies that assessed the effects of school resources on academic achievement but with disparate conclusions, nonetheless. While Hanushek's (1986) influential study found little evidence of the relationship between school characteristics and student achievement, Hedges et al., (1994) notable study reports positive effects of school resources on student achievement (Hedges & Stock, 1983). Card and Krueger (1992a), the first to establish the link between school effects and earnings later in life, report that improvements in school quality characteristics, such as pupil/teacher ratios, teacher salary, and length of academic year had positive effects on black and white men's earnings and helped narrow the black-white gap between 1960 and 1980, thus offering cautious support to the thesis that, on average, resources do matter. In more recent work, Card and Krueger (1996) conclude that there is a positive association between school expenditures and student annual earnings later in life. In their qualitative review of the literature investigating the relationship between school resources and school quality and student's earnings Card and Krueger (1998) propose that the payoff to additional schooling is higher for students attending higher quality schools. Based on the High School and Beyond Data, Crawford et al. (1997) also found that "which high school a

student goes to makes a difference in annual earnings, but that most of the specific characteristics that account for the differences are not identifiable” (p. 255).

In sharp contrast, other studies failed to find an association between school inputs and earnings (Betts, 1996) or found that school inputs such as pupil/teacher ratios and term length have little effect, if any, on student wages (Grogger, 1996). Heckman et al., (1996) replicating the Card and Krueger results found that their positive effects are not robust to several assumptions in their model. In sum, even though there is some indication that school characteristics might affect future earnings, the findings remain uncertain and inconclusive (Card and Krueger, 1998). Much of the disparity lies in the level of aggregation at which the different analyses are conducted. Nonetheless, all previous studies in school effects ignored the issue of the variability of the outcome variables across school units.

3. DATABASES

We employ two rich, longitudinal, and representative samples of individuals who were high school students in 1972 and in 1980 and wage earners in 1986 and 1992 respectively. Specifically, we draw upon the National Longitudinal Study (NLS) of the High School Class of 1972 and the fifth follow-up in 1986, as well as on the High School and Beyond (HSB) surveys of 1980, 1982, and the fourth follow-up in 1992. 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

(Riccobono et al., 1981). These students were followed for 14 years after high school, and, thus, they were resurveyed in 1974, 1975, 1977, 1980, and 1986. We employ data collected during the base-year survey in 1972 and the fifth follow-up in 1986. Our final sample includes all individuals who were present in 1972 and in 1986 and were not in the military.

In the spring of 1980, two cohorts of tenth and twelfth grade students enrolled in public and private schools in the US were surveyed for the HS&B-80. This study targeted students who were high school sophomores and seniors in 1980 enrolled in public or private American schools. Both cohorts (the sophomore and senior class of 1980) were followed for six years and, thus, were resurveyed every two years through 1986. Only the sophomore class was surveyed again for the fourth time in 1992. Our final sample for the HS&B-80 includes individuals who were present in the base year, the first, and the fourth follow-up samples, and were not in the military in 1992. Hence, our sample consists of individuals who were high school seniors in 1982.

Both datasets are unique in providing valuable information on the educational attainment, student status, school resources, employment outcomes, and family formation of high school seniors in 1972 and 1982 respectively. Both NLS and HSB are longitudinal studies, which are part of the program of the National Center for Education Statistics (NCES) - established to study the educational, vocational, and personal development of young people. Their longitudinal feature allows us to follow these students 10 to 14 years after their high school graduation and examine their labor market performance during their prime time in the labor market. In addition, because we are also looking at students who have been in the labor market for a decade or more we

avoid any biases from school to work transitions. Typically, individuals tend to be more settled and change jobs less often after the first decade of employment.

A desirable property of these data sets is that they provide the unique opportunity of linking labor market information with high school data, and hence, identifying the high schools individuals attended. In addition, information on individuals' family socioeconomic status (SES) and achievement in high school was also provided. Throughout this study we conduct our analyses separately for women and men. Our samples consist of individuals who were high school seniors in 1972 and 1982 and participated in the respective last follow up studies. Notice that our samples include individuals with different levels of education since some of them joined the labor market immediately after high school while others pursued further education in community colleges or universities.

4. METHODOLOGICAL APPROACH

Most educational data have a multi-level structure, where students are nested within classrooms, classrooms are nested within schools, and schools are nested within districts and so forth. This clustering of the data needs to be taken into account when one conducts statistical analysis. One such analysis involves the use of multilevel models (Goldstein, 1995; Raudenbush, and Bryk, 2002). Such models conceptually permit analysis and pooling of school-specific regressions in a way that takes into account the clustering of the data by school. In addition, these models allow for random effects at the higher levels of the hierarchy. Classical regression models fail to take into

account the clustering nature of the data, which has the potential of underestimating the standard errors of the regression coefficients and may ultimately result in higher probabilities of Type I errors. Further, typical regression models do not allow for the estimation of between school variation, which is usually different than the within school variation. In other words, mixed models methods allow for the decomposition of the variation in earnings in within and between school variation, and present a more general case than the typical regression models. In fact, with multilevel models we can estimate the differential learning opportunities provided to students within schools and the role that these opportunities play in determining future success.

We employ a two-level model where student characteristics are introduced in the elementary level and school characteristics at the second or higher level. The elementary level conceptually involves school specific regressions, and the second level involves a between school regression. The first level model is an *earnings model* that describes the earnings of individuals within a school as a function of the particular school (school intercept) and important covariates adjusting for individual characteristics such as race/ethnicity, social class, achievement, etc. In essence this model includes a specific school index (the constant term) and other school-specific parameters that describe the association between individuals' characteristics and earnings for each school.

The school specific index is treated as a random variable at the second level of the model and hence, varies across school units. The school specific residuals or random effects at the second level model describe the degree to which the average earnings of individuals in a specific school differ from those of the other schools controlling for individuals' characteristics. These random effects are the school effects. The variance of

random effects in the between school model describes differences between school specific effects on earnings. The computation of the variance component of the school effects is the main objective of the study. Once the variance of the school effects is computed, we test for heterogeneity between school effects by determining the statistical significance of this variance. A significant variance indicates residualized differential school effects on future earnings. Notice that since the school effects represent the difference or deviation between a specific school's average earnings and the average earnings of the average school, their arithmetic mean is zero. Then, the most appropriate way to describe the distribution of the school effects is by estimating their dispersion or variance. In other words the objective of our analyses is the estimation of the variance or error components of the random effects (i.e., school effects). The importance of the school effects on earnings is represented by the significant variation of the random effects across schools.

The second level model can simply be an unconditional model, which estimates the heterogeneity of the school specific intercepts. One can include school factors in the second level model to explain variation in the school effects, but their inclusion is not essential for the purposes of this study. Our analysis consists of two steps. The first step involves a probit model that estimates possible selection in the labor force for men and women separately. Since workers might differ from non-workers in unobservable ways we adjust our models for possible selection bias. Following Heckman (1979) we estimate the probability that an individual is working as:¹

$$\text{probit}(\pi) = \beta_0 + X\beta, \quad (1)$$

where β_0 is the intercept, X is the matrix of the predictors that identify the labor force participation equation, and includes race, family SES in high school, educational

attainment, student status, marital status, and nonlabor income, and β is the vector of the coefficients. The race variable is a dummy, which takes the value of 1 if the individual is not White or Asian and zero otherwise. Likewise family SES² is an indicator variable, which takes the value of 1 if the individual is in the top quartile of the SES distribution and zero otherwise. The educational attainment of individuals – extracted from the last follow up study - is captured by a dummy variable, which takes the value of one if the individual has a college degree or more education and zero otherwise. Student and marital status are also binary variables indicating whether an individual is a student or married/cohabitating in the last follow up study.

Nonlabor income includes income from interest, social security and veteran benefits, welfare, unemployment compensation, gifts, scholarships, etc. We hypothesize that the above variables are important predictors of labor force participation. For example, higher nonlabor income will decrease the probability to work by increasing the reservation wage, while a college degree will increase the probability to work. Marital status has also been shown to signal labor force attachment and higher productivity in the labor market affecting potential tenure on the job (e.g., Korenman and Neumark, 1991). We expect married individuals to have a higher probability to work. From the probit model we calculate the so-called inverse Mill's ratio or λ , which we include as an additional predictor in our two-level model (in the second stage) to adjust for possible non-random selection of workers.

The second step of our analysis involves a two-level model where individuals are nested within school units. This multi-level analysis was conducted separately for women and men following a modified Mincerian earnings regression.³ Assuming m

schools and n_j students within the j th school, the elementary within-school regression model for person i in school j adjusted for selection becomes:

$$Y_{ij} = \gamma_{0j} + Z_j \gamma_j + \varepsilon_{ij}, \quad (2)$$

where $i = 1, \dots, n_j, j = 1, \dots, m$, Y_{ij} is the log of annual earnings⁴ of individual i at school j , and γ_{0j} is the school specific intercept adjusted for individual characteristics. Z_j is the matrix of the predictors for school j including the selection correction term, which are hypothesized to influence labor market earnings, γ_j is the vector of coefficients for school j and shows the structural relations that occur within school j , and ε_{ij} is the student specific residual for school j . Notice that the regressors in the probit equation are carefully chosen to explain the proclivity to work, and avoid any identification problems with the within school regressions. Our elementary level model in equation (2) controls for individual and family characteristics captured by the dummy variables for race/ethnicity, family SES, and academic achievement of the individual in high school. Taking into account such characteristics is essential in determining the choice of school, and omission of them might lead to spurious associations between school effects and earnings (Card and Krueger, 1998). Further, we also control for selection bias, through the λ term. In an attempt to explain as much individual variation within each school as possible, we controlled for observed individual characteristics that are considered important predictors of earnings. Hence, we took into account the effects of race, family background, and ability or knowledge represented by achievement test scores (Heckman, 1995).

For example the Black-white as well as the Latino-white wage differential has been well documented in labor economics (Smith and Welch, 1986, 1989; O'Neill, 1990). Family SES is also an important predictor of earnings since educated, richer parents invest more on their children's education, which in turn results in higher rates of labor market performance. In the human capital model, achievement scores play an important role since they are hypothesized to generate higher human capital formation and affect educational and occupational attainment as well as labor market performance (Crawford et al., 1997; Miller, 1998). In fact, achievement scores have been shown to affect individuals' earnings even after controlling for social class (Murnane, Willett, and Levy, 1995). Notice, that achievement scores encompass to a certain extent school effects and in a way our model controls for more than it should. However, achievement scores consist a better predictor than high school GPA, which is heavily affected by teacher and school practices. By controlling for these important explanatory variables we contend that the dispersion of school specific effects (intercepts) is free of observed individual effects. If this assumption holds, then a significant between school variation of the school effects indicates the impact of schools on future earnings net of observed individual effects. Moreover, this specification controls in a way for the non-random assignment of students in schools.

The second between-school level regression model for the school specific intercepts for school j is depicted by:

$$\gamma_{0j} = \delta_{00} + \eta_{0j}, \quad (3)$$

where δ_{00} is the average earnings across school units, and η_{0j} is the school specific residual. Notice that η_{0j} is a school effect that describes the difference/deviation

between the average earnings in school j and that of the average school. The term η_{0j} is a unique random effect associated with school j , and follows a normal distribution with a mean of zero and a variance namely τ^2 . The variance of η_{0j} measures how much the average earnings for each school varies across schools, and indicates the impact of schools on earnings. The computation of this between school variance is the major object in this study. Finally, in the two-level model we adjusted for heteroskedasticity in the error term to obtain the correct standard errors for the estimates.

The remaining elementary level coefficients were treated as fixed variables⁵, namely:

$$\gamma_{kj} = \delta_k. \quad (4)$$

The estimates are computed using the EM algorithm within the empirical Bayes framework.

5. RESULTS

Estimation Results based on NLS

Table 1 summarizes the descriptive statistics for selected variables in the base year and the fifth follow-up of NLS for the total sample and by gender. Overall, our sample consists of 51% women and 49% men. Seventeen percent of our sample are in the minority category - blacks and Hispanics - and there are slightly more minority women than men, while the majority of our sample are in the White/Asian category. On average, women in 1986 while in their early thirties earn 40% less than men.⁶ When we break annual earnings down by race, we see that minority women earn nearly as much as White or Asian women, while minority men earn 20% less than White or Asian men.

Women in 1986 also have lower labor force participation than men (20%). These statistics also show that although a smaller percentage of women than men have finished college in 1986, still a quarter of them have acquired higher education. The majority of both women and men in our 1986 sample are married but a higher percentage of women than men are in the divorced category. With respect to the family socioeconomic status in 1972, 51% of both men and women are in the medium SAS category. For the tails of the SES distribution, we find more women at the lower end and more men at the higher end. Lastly, from the achievement scores in high school we see that men performed slightly better than women.

Table 1 Here

In Tables 2 and 3 we present the estimated results of our two-level model based on the NLS data for women and men respectively. The results of Table 2 show that, as expected, family SES is a positive and significant predictor of future earnings. Women from educated, well-to-do families seem to earn higher salaries later in life than other women. To our surprise the race coefficient is positive and significant indicating that minority women - Blacks or Hispanics - earn higher salaries annually than White or Asian women. This indicates that when adjusting for race, SES, ability, and selection minority women are performing better in the labor market than White or Asian women. Notice that even though race and SES are moderately correlated ($\Delta = 0.3$) within gender categories for both datasets, our results indicate independent effects of each variable in all of our specifications.

As hypothesized, higher ability depicted by higher achievement test scores in high school is also a positive and significant predictor of future earnings. Women with high levels of achievement in high school are more likely to have higher earnings later in life. The coefficient of the correction term, λ , is positive and significant indicating that there was a positive selection in our female sample of wage earners. Hence, its inclusion in our model was necessary to correct for selection bias. Most importantly for this study, the variance of the school effects was at the borderline of statistical significance (see row 2). Thus, for the female sample of NLS no definitive conclusions can be drawn regarding differential effects of schools on earnings. It seems that in this case the impact of schools is rather uniform.

 Table 2 Here

The results of our two-level model for the male sample are presented in Table 3. Overall, the results in Table 3 resemble those in Table 2. As in Table 2 all coefficients are significant at the 0.05 level. The family SES and achievement coefficients are as hypothesized by the theory, positive and significant. The sign of the race coefficient is also positive and significant in this specification, indicating that men in the minority group earn significantly lower salaries (12% less) than White or Asian men. The correction term, λ , is also positive and significant suggesting the male workers sort themselves into higher paying jobs. Most importantly, the variance of the school effects is quite significant indicating that for men the school they attended as high school students definitely affects their future earnings (see row 2).

Table 3 Here
-----Estimation Results Based on HSB

Descriptive statistics for selected variables from the HSB study are reported in Table 4. The gender and race distribution in 1992 is slightly different than in 1986. Specifically, the sample contains a higher percentage of women than men, and a higher percentage of individuals in the minority category than Whites or Asians. On average, women in 1992 - while in their late twenties - earn 25% less than men. While this shows that women in the 1990s are doing a lot better than in the 1980s the gender earnings gap is still considerable. In fact, these raw statistics on annual earnings by race and gender show that minority women earn somewhat less than White or Asian women, while minority men earn nearly as much as White or Asian men. With regards to their attachment to the labor force, we see that women in 1992 have still a lower percentage (14%) in labor force participation than men. The percentage of both men and women who have finished college or acquired a higher degree is at the same level of 25%. With regards to demographics, the majority of individuals are married, but a much higher percentage of women than men are married and a higher percentage of them are also in the divorced category. The family SES distribution in 1980 is somewhat different than in 1972. For example, the percentage of the individuals in the medium SES category is smaller, while the percentage of individuals in the low SES category has increased by 2.5%. As in Table 1, we find that more women than men come from the lower end of the SES. Lastly, from the achievement scores in high school we see again that men performed slightly better than women.

Table 4 Here

The results of the two-level analyses for the HSB female sample are portrayed in Table 5. The results in Table 5 are comparable to those presented in Table 2. The main difference is that the race coefficient although still positive is not significant. The rest of the coefficients are positive and significant, as expected. Family SES and achievement are, as in NLS, important predictors of earnings. Similar to the results on NLS, the correction factor is positive and significant, justifying our Heckman selection procedure. Contrary to NLS however, the variance of the school effects varied significantly across school units, indicating a strong differential effect of schools on earnings. This result shows the important impact of schools on the earnings of women in 1992.

Table 5 Here

In Table 6 we present the results of the two-level analyses based on the HSB for the male sample. The results in Table 6 are comparable to those presented in Table 3. Minority men earn significantly less than White or Asian men in the 1990s but compared to the NLS results this gap has been considerably reduced to a half. All other coefficients are positive and significant. Both the family SES and the achievement indicators are, as in NLS, important predictors of future earnings for men. Similar to the results on NLS, the correction factor is positive and significant, justifying our Heckman selection procedure. It is noteworthy that, similar to the results in NLS, the variance of the school effects varied significantly across school units. This striking result indicates

that there is a strong differential effect of schools on the earnings of men in the early 1990s. This result shows once more the important impact of schools on earnings and is in congruence with that in the NLS sample for men and HSB sample for women. In sum, our analysis on two datasets a decade apart, demonstrates the importance of school effects on earnings three out of four times (75%).

Table 6 Here

6. CONCLUSION

In this paper we undertook a thorough analysis of school effects on future earnings of young women and men in the 1980s and 1990s, controlling appropriately for individual, and family characteristics, as well as for selection into the labor market. We aimed to assess the extent to which schooling holds enduring and long-lasting effects on young adults' labor market trajectory above and beyond individual and family characteristics. Using the NLS and the HSB, two rich data sets of representative samples of high school students, we estimated individual and school effects on the natural logarithm of annual earnings adjusting for employment selection. Overall, our results provided evidence that high schools have important differential effects on future earnings. In almost all cases for both men and women and for both datasets average earnings varied significantly across schools suggesting the important impact of school effects on future earnings. In other words, we found important differences in earnings across schools, and since we appropriately controlled for individual variation, these differences point to school effects on earnings. There was also a wide dispersion of

earnings for individuals who attended the same school indicating student and perhaps school specific effects (e.g., teachers or other resources or climate of school). Our findings are congruent with other recent findings on school effects and achievement (Hedges, and Konstantopoulos, 2002).

Unlike previous studies we examined school effects for men and women separately via the variability of average school-specific earnings across the sample of schools controlling for important student predictors such as race, family SES, and student achievement. Using multi-level methods within an empirical Bayes framework we were able to decompose the variability in earnings into within and between schools. In our study the most important component is the between school variability, which indicates the degree to which schools affect future earnings. We controlled for important covariates such as the race, social class, and ability of individuals within each school, and hence, the between school variability is net of these effects.

However, we are aware that the effects of other unobserved student characteristics such as motivation or ambition were not controlled in our model, but it is quite hard if not impossible to control for them. In addition, since students are not randomly assigned to schools, it is possible that our results might be compromised from neighborhood effects. Nonetheless, we contend that given the data limitations we captured a good deal of student characteristics. Still, the between school variation of average earnings was significantly different from zero in almost all cases indicating that the high schools individuals attend do matter later in life. The fact that this result was replicated in three out of four cases might also indicate that our results are robust in cohort effects. These differential effects of high schools are quite important since

students in our samples had finished high school for at least 10 years and many of them have attended college or graduate/professional schools. In addition, notice that by controlling for student achievement at the end of the senior year in high school in a way we over control for school effects and thus it is possible that our estimates are somewhat underestimating the effects of high schools.

The within school variability can be thought of as the error variance in the usual regression models and indicates differences in earnings among students who attended the same high school. A major part of this variability is definitely due to student characteristics. However, one could also argue that such differences can be attributed to differences in school-specific effects (i.e., teachers, resources, school climate, school leadership), assuming that schools are treatments and as such have different impact on student groups.

To conclude, our findings suggested that schools make a difference in the subsequent labor market performance of their graduates. The average earnings differed significantly across school units, a finding indicating that schools do matter. Further research on school effects using other rich datasets is needed. Specifically, longitudinal analyses using three level models that examine the differences in earnings over time would provide useful information regarding school effects on the growth trajectory of earnings. In addition, cross-sectional analyses employing three levels models to decompose the variability in earnings into three parts, within classroom, across classrooms (within schools), and across schools would provide information on how important the teacher effects are on top of school effects.

NOTES

¹ The general specification of the employment index function is based on the rational that individuals work if the offered wage is greater than the reservation wage. The selection issue arises from the fact that wages are missing for the nonworkers.

² The SES composite index is constructed by NLS and HSB using information on parental educational attainment, father's education, family income, and household items (e.g., dishwasher, color TV, cars, etc.). We define high SES as the top quartile of the SES distribution.

³ Because our sample is of the same age and education levels, age or experience are not useful determinants of earnings. This restriction of range in age or experience attenuates the correlation between these variables and earnings.

⁴ Annual earnings is the only information available in our data sets, while there was no information on the hours of work, or distinct categories of full time part time status.

⁵ This restriction was deemed appropriate for the purposes of this study. However, typically school specific slopes can also be treated as random variables. For example, we also ran models treating the school specific slopes as random and the estimate of the variance component of the school effects was comparable to that obtained in our primary analyses.

⁶ This number is not adjusted for hours of work.

ACKNOWLEDGMENTS

We wish to thank Larry V. Hedges and Klaus F. Zimmermann for their insightful comments. However, all opinions and inadvertent errors are our own responsibility. We also thank Dave Ribar and the participants at the George Washington University labor and social insurance workshop for their comments on an earlier version that was the precursor of this study. An earlier version of this paper was also presented at the EALE 2002 in Paris, France.

Table 1. Arithmetic Means and Proportions for Selected Variables by Gender, NIOS on Earnings 25

	<u>U.S. Population</u>	<u>Female</u>	<u>Male</u>
<u>Individual Characteristics</u>			
Female	50.9%	-	-
Male	49.1%	-	-
Minority	17.0%	18.0%	16.0%
White/Asian	83.0%	82.0%	84.0%
Annual Earnings in 1986	21,754	15,859	26,660
Minority Earnings	18,882	15,578	22,129
White/Asian Earnings	22,334	15,989	27,427
Labor Force Participation	79.1%	69.9%	88.9%
Low SES in 1972	24.8%	26.9%	22.5%
Medium SES	51.3%	51.3%	51.5%
High SES	23.9%	21.8%	26.0%
Student in Last Follow Up	8.5%	8.5%	8.6%
College Degree or More	26.0%	24.7%	27.4%
Acievement Scores	22.6	21.6	23.7
Single	17.1%	15.3%	18.8%
Married	67.7%	67.9%	67.6%
Divorced-Separated-Widowed	10.5%	12.3%	8.6%
Living in Common Law	3.8%	3.8%	3.9%

Table 2. Two Level Estimates and Standard Errors of Annual Log-Earnings: NLS Female Sample

	Estimate	SE	P-Value
Average Log-Earnings	9.397	0.016	0.000
Variance Across Schools	0.026	0.014	0.060
High SES	0.129	0.040	0.001
Minority	0.235	0.043	0.000
Achievement Test Scores	0.013	0.002	0.000
Selection Correction Factor	0.478	0.040	0.000

Table 3. Two Level Estimates and Standard Errors of Annual Log-Earnings: NLS Male Sample

	Estimate	SE	P-Value
Average Log-Earnings	10.039	0.011	0.000
Variance Across Schools	0.018	0.004	0.000
High SES	0.149	0.024	0.000
Minority	-0.121	0.033	0.000
Achievement Test Scores	0.008	0.001	0.000
Selection Correction Factor	0.285	0.033	0.000

Table 4. Arithmetic Means and Proportions for Selected Variables by Gender: HSB

	<u>U.S. Population</u>	<u>Female</u>	<u>Male</u>
<u>Individual Characteristics</u>			
Female	51.6%	-	-
Male	48.4%	-	-
Minority	24.9%	24.4%	25.4%
White/Asian	75.1%	75.6%	74.6%
Annual Earnings in 1992	22,449	19,100	25,489
Minority Earnings	21,479	17,794	24,594
White/Asian Earnings	22,749	19,485	25,778
Labor Force Participation	80.3%	73.5%	87.6%
Low SES in 1982	27.3%	29.1%	25.4%
Medium SES	48.9%	48.3%	49.5%
High SES	23.8%	22.6%	25.1%
Student in Last Follow Up	2.2%	2.2%	2.2%
College Degree or More	25.4%	25.2%	25.5%
Acievement Scores	22.3	21.2	23.4
Single	37.1%	30.5%	44.1%
Married	52.8%	58.0%	47.2%
Divorced-Separated-Widowed	9.3%	10.7%	7.8%
Living in Common Law	0.8%	0.8%	0.8%

Table 5. Two Level Estimates and Standard Errors of Annual Log-Earnings: HSB Female Sample

	Estimate	SE	P-Value
Average Log-Earnings	9.649	0.015	0.000
Variance Across Schools	0.039	0.007	0.000
High SES	0.179	0.032	0.000
Minority	0.023	0.032	0.472
Achievement Test Scores	0.013	0.001	0.000
Correction Factor	0.253	0.029	0.000

Table 6. Two Level Estimates and Standard Errors of Annual Log-Earnings: HSB Male Sample

	Estimate	SE	P-Value
Average Log-Earnings	9.997	0.011	0.000
Variance Across Schools	0.012	0.004	0.002
High SES	0.081	0.023	0.001
Minority	-0.069	0.024	0.005
Achievement Test Scores	0.008	0.001	0.000
Correction Factor	0.158	0.028	0.000

REFERENCES

- Becker, Gary S., 1964, Human capital. (New York: Columbia University Press).
- Betts, Julian, 1996 Do School resources matter only for Older Workers? The Review of Economics and Statistics 78, 638-652.
- Card, David and Alan B. Krueger, 1992a. School Quality and Black-White Relative Earnings: A Direct Assessment. Quarterly Journal of Economics : 151-200.
- Card, David and Alan B. Krueger. 1992b. Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. Journal of Political Economy, 1-40.
- Card, David and Alan B. Krueger, 1996. The economic return of school quality. In Assessing educational practices: The contribution of economics eds William Becker and William Baumol, 161-181. (New York Russell Sage).
- Card, David and Alan B. Krueger, 1998. School Resources and Student Outcomes. Annals of the American Academy of Political and Social Sciences 559, 39-51.
- Coleman, James S. et al. 1966. Equality of Educational Opportunity. Washington, DC: U.S. Government Printing Office.
- Crawford, David L., Amy W. Johnson, and Anita A. Summers, 1997. Schools and Labor Market Outcomes, Economics of Education Review, 16, 255-269.
- Goldstein, H. (1995). Multilevel statistical models. London: Arnold Publishers.
- Grogger, Jeff. 1996. Does School Quality Explain the Recent Black/White Wage Trend? Journal of Labor Economics 14, 231-253.
- Hanushek, Eric A. 1986. The Economics of Schooling: Production and Efficiency in Public Schools. Journal of Economic Literature 24, 1141-77.

- Heckman, James, 1979 Sample Selection Bias as a Specification Error. *Econometrica*, 47(1), 153-161.
- Heckman, James, J. (1995). Lessons from the Bell Curve. *Journal of Political Economy*, 103, 1091-1120.
- Heckman, James, et al., 1996 Human Capital Pricing Equations with an Application Estimating the Effect of Schooling Quality on Earnings. *The Review of Economics and Statistics* 78, 562-610.
- Hedges, Larry V. and W. Stock. 1983. The Effects of Class Size: An Examination of Rival Hypotheses. *American Educational Research Journal*, 20, 63-65.
- Hedges, Larry V., Richard Laine, and Rob Greenwald, 1994. Does Money Matter: A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes. *Educational Researcher*, 23(Apr.): 5-14.
- Hedges, Larry V., and Spyros Konstantopoulos (2002). How large an effect can we expect from school reform? Manuscript Submitted for Publication.
- Korenman, Sanders and David Neumark, 1991. Does Marriage Really Make Men More Productive? *Journal of Human Resources*, 26(2), 282-307.
- Miller, Shazia Rafiullah, 1998. Shortcut: High School Grades as a Signal of Human Capital. *Educational Evaluation and Policy Analysis*, 20(4), 299-311.
- Mincer, Jacob, 1974. *Schooling, Experience, and Earnings*, New York, Columbia University Press.
- Morris, Carl N. (1983). Parametric empirical Bayes inference: Theory and applications. *Journal of the American Statistical Association*, 78, 47-65.

- Murnane, Richard, J., John B. Willett, and Frank Levy. 1995. The growing importance of cognitive skills in wage determination. Review of Economics and Statistics, 77, 251-266.
- O'Neill, June, (1990) The Role of Human Capital in Earnings Differences between Black and White Men, *Journal of economic Perspectives*, 4, 25-46.
- Raudenbush, Steven W., and Anthony S. Bryk, 2002. Hierarchical linear models: Applications and data analysis methods. (Newbury Park: Sage).
- Riccobono, John, Louise B. Henderson, Graham J. Burkheimer, Carol Place, and Jay R. Levinsohn, 1981. National Longitudinal Study. Research Triangle Institute.
- Smith, James and Finis R. Welch, 1986. Closing the Gap: Forty years of Economic Progress for Blacks. (Santa Monica, CA: Rand Corp).
- Smith, James and Finis R. Welch. 1989. Black Economic Progress after Myrdal. *Journal of Economic Literature* 27(June): 519-64.
-