DOCUMENT RESUME

ED 333 840 HE 024 699

AUTHOR Altonji, Joseph G.
TITLE Controlling for Personal Characteristics, School and Community Characteristics, and High School Curriculum when Estimating the Return to Education. Technical Paper Series.
INSTITUTION National Center on Education and Employment, New York, NY.; Northwestern Univ., Evanston, Ill. Center for Urban Affairs and Policy Research.; Rand Corp., Santa Monica, Calif.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.; Spencer Foundation, Chicago, Ill.
REPORT NO NCEE-TR-14
PUB DATE Jul 90
CONTRACT OERI-G008690008
NOTE 29p.
PUB TYPE Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS College Bound Students; Community Characteristics; Data Analysis; *Educational Benefits; *Error of Measurement; Evaluation Problems; Family Characteristics; Higher Education; High School Graduates; High Schools; Individual Characteristics; Institutional Characteristics; *Outcomes of Education; *Research Methodology; *Research Problems; Secondary School Curriculum; *Wages

ABSTRACT

This paper examines whether failure to control for family background, aptitude, high school quality, high school curriculum, and community characteristics leads to bias in estimates of the economic return resulting from postsecondary education. High school specific fixed effects were used to control for all observed and unobserved characteristics common to students from a given high school and an instrumental variables procedure was used to deal with measurement error in reported education. The estimation was conducted using panel data on several persons from each of a large number of high schools sampled in the National Longitudinal Survey of the High School Class of 1972. Controlling for family background and aptitude and achievement measures led to a substantial reduction in estimates of the rate of return to education. Measurement error was not responsible for the reduction. However, the use of fixed effects to control for high school and community characteristics had only a modest effect on estimates of the rate of return to education, and controlling for high school curriculum made little difference. It is noted that this is an important finding because few data sets permit researchers to control for these factors when estimating the effect of education on wages. Included is an appendix of means and standard deviations. Contains 12 references. (Author)

National Center on Education and Employment

NCEE

CONTROLLING FOR PERSONAL CHARACTERISTICS, SCHOOL AND COMMUNITY CHARACTERISTICS, AND HIGH SCHOOL CURRICULUM WHEN ESTIMATING THE RETURN TO EDUCATION

Joseph G. Altonji
Center for Urban Affairs and Policy Research and
Department of Economics, Northwestern University

Technical Paper No. 14

CONTROLLING FOR PERSONAL CHARACTERISTICS, SCHOOL AND COMMUNITY CHARACTERISTICS, AND HIGH SCHOOL CURRICULUM WHEN ESTIMATING THE RETURN TO EDUCATION

Joseph G. Altonji
Center for Urban Affairs and Policy Research and Department of Economics, Northwestern University

Technical Paper No. 14

July 1990

This publication is based on work sponsored by the National Center on Education and Employment under grant number G008690008 from the Office of Educational Research and Improvement, U.S. Department of Education, by the Spencer Foundation, and by the Center for Urban Affairs and Policy Research at Northwestern University.

PREFACE

The National Center on Education and Employment is funded by the Office of Educational Research and Improvement (OERI) of the U.S. Department of Education. The Center is based at Teachers College, Columbia University, in New York City. The RAND Corporation of Santa Monica, California, and Washington, D.C., is a partner with Teachers College in this enterprise.

This publication appears in the Center's Technical Paper Series. This series is designed to communicate the interim technical results of a large-scale Center research project or the final results of a small project. The paper has been reviewed by one reader external to the project and the Center, and was approved for publication by Center leadership.

For information about ordering additional copies of this document, write or call:

National Center on Education and Employment
Box 174
Teachers College, Columbia University
New York, New York 10027

(212) 678-3091

This publication is based on work sponsored wholly or in part by the Office of Educational Research and Improvement, U.S. Department of Education, under grant number G008690008. Its contents do not necessarily reflect the views of OERI, the Department, or any other agency of the U.S. Government.

Acknowledgments

The author gratefully acknowledges the support of the National Center on Education and Employment, the Spencer Foundation, and the Center for Urban Affairs and Policy Research. Avner Grief, Christian Stadlinger, and James Spletzer provided excellent research assistance, and Thomas Dunn, Alan Gustman, Bruce Meyer, James Spletzer, and Barbara Wolfe made helpful comments on an earlier draft. The author is solely responsible for any errors or omissions.

**Controlling for Personal Characteristics,
School and Community Characteristics, and High School Curriculum
When Estimating the Return to Education**

ABSTRACT

This paper studies whether failure to control for family background, aptitude, high school quality, high school curriculum, and community characteristics leads to bias in estimates of the return to post secondary education. High school specific fixed effects are used to control for all observed and unobserved characteristics common to students from a given high school and an instrumental variables procedure is used to deal with measurement error in reported education. The estimation is conducted using panel data on several persons from each of a large number of high schools sampled in the National Longitudinal Survey of the High School Class of 1972.

Controlling for family background and aptitude and achievement measures leads to a substantial reduction in estimates of the rate of return to education. Measurement error is not responsible for the reduction. However, the use of fixed effects to control for high school and community characteristics has only a modest effect on estimates of the rate of return to education, and controlling for high school curriculum makes little difference. This is an important and reassuring finding because few data sets permit one to control for these factors when estimating the effect of education on wages.

Joseph G. Altonji
Department of Economics
Northwestern University
Evanston, IL 60208
708-491-8218

INTRODUCTION

Hundreds of studies have shown that more-educated workers receive higher wages and earnings than less-educated workers.[1] The typical estimate is that college graduates earn about 30 percent more than high school graduates. This earnings gap, though varying over time, has always been substantial.[2] But the broad agreement that more-educated workers earn more than less-educated workers is accompanied by disagreement on the extent to which the difference in earnings is due to the difference in education. Part of the earnings advantage of educated workers may arise from differences in ability, achievement in primary and secondary school, quality of curriculum in primary and secondary school, or school and community characteristics that happen to affect both wages and years of education.[3]

There is in fact an extensive literature on the effects of family background and ability on estimates of the return to education. Most studies have found that controlling for family background and ability reduces estimates of the rate of return to education. (See Griliches (1979) and Siebert (1985) for surveys.) However, some studies have paid attention to the fact that mismeasurement of education becomes a more serious problem when one controls for ability or family background; these studies find somewhat smaller levels of bias and in some cases obtain higher estimates of the return to education.

This paper examines whether failure to control for school quality, secondary school curriculum, and community characteristics leads to bias in estimates of the return to postsecondary education. Given the interest that college admissions committees show in this information, one might suspect that these variables are correlated with postsecondary education. Failure to control for them will lead to bias if they affect the wage

[1] Siebert (1985) and Willis (1986) provide recent surveys of the link between education and earnings.

[2] Recent research by Murphy and Welch (1988) suggests that the college differential has increased substantially during the 1980s.

[3] There are two main channels through which the spurious correlation between education and wages might arise. First, family background, primary and secondary school quality, and ability might affect the wage level independently of postsecondary schooling. Second, family background, ability, and primary and secondary school characteristics may affect the production function of human capital. Students who are more able, from better family backgrounds, or from better schools may choose more postsecondary education than the less advantaged because they receive a larger payoff to a year in college. In this case, a regression estimate of the difference in earnings of a high school graduate and a college graduate will exceed the gain in earnings that a typical high school graduate would receive if he or she had chosen college. See Siebert (1985), Willis (1986), and Griliches (1977) for discussions of these issues.

level or the return to education. However, little is known about the question, in part because most of the data sets that have been used to study the returns to education such as the Current Population Survey, the Decennial Census, the Panel Study of Income Dynamics, and the National Longitudinal Surveys of Labor Market Experience contain relatively little information about school curriculum and the community.[4] Furthermore, it is hard to envision a data set that would contain measures of all of the relevant school and community characteristics. The education production function literature has not been terribly successful in explaining differences in achievement across schools with school and community variables,[5] even though there are substantial differences across schools in the parental and school characteristics that we do observe. (See Appendix Table 1.) One naturally suspects that there are unobserved differences among high schools and communities that are important determinants of both education and wages.

Data from the National Longitudinal Survey of the High School Class of 1972 (NLSHS72) and from a matching postsecondary transcript survey (PETS) provide an opportunity to make some progress on this issue. The most important feature of the NLSHS72 for present purposes is that it contains several students from a large number of high schools (897 high schools in the subsample used in this analysis). This permits the use of fixed effect estimation procedures that control for all observed and unobserved characteristics common to students from the same high school. One may also use fixed effects to control for characteristics common to students in the same program (i.e., academic or nonacademic track) within a given high school. In addition, the data set contains information on parental background, high school curriculum, and test scores.

Of course, the more one controls for the observed and unobserved factors that lead to differences in education, the potentially more serious the problem of measurement error in reported education becomes, as Griliches (1977, 1979) has emphasized. I deal with the measurement error issue by using the information on education from the postsecondary transcript survey to form instrumental variables (IV) for the responses provided by the sample members.

The main conclusions of the study are fourfold. First, controlling for family background leads to a substantial reduction in estimates of the rate of return to education. For the combined sample of academic and nonacademic track high school students, the ordinary least squares (OLS) estimates imply that the return to a year of postsecondary academic education falls from 8.2 percent when one does not control for family

---

[4] Akin and Garfinkel (1977), Morgan and Sirageldin (1973), and Johnson and Stafford (1973) examine links between school quality proxies and labor market outcomes.

[5] See Hanushek (1986) for a recent survey.

background to 6.5 percent when one does. The IV results are similar, suggesting that measurement error is not responsible for the reduction. Similar reductions are found among the samples of students who are in high school academic programs and among those who are in nonacademic programs. Second, estimates of the rate of return to postsecondary academic education for academic and nonacademic track high school students are remarkably similar. This is true despite the fact that students from academic programs earn substantially more than those from nonacademic programs, even after controlling for observed family background characteristics and achievement and aptitude measures. Third, controlling for high school curriculum does not have much effect on the education coefficients. Fourth, the use of fixed effects to control for high school characteristics has only a modest effect on estimates of the rate of return to education. For the combined sample, controlling for these factors reduces estimates of the percentage increase in earnings from a year of college by about .5 (e.g., from 6.0 to 5.5). This suggests that the failure to control for differences in high school variables does not lead to serious biases in studies of education and wages. This is an important and reassuring finding, because few data sets would permit one to control for these factors.

The paper proceeds as follows: Section 1 presents the wage equation that underlies most of the econometric analysis below. It also discusses the potential biases in estimates of the effect of education on wages that may arise from failure to control for characteristics that correlate with educational attainment and that affect either the level of the wage holding the educational level constant or the relationship between wages and education. Section 2 discusses the data used in the study. Section 3 presents estimates of the return to education. The main conclusions are summarized in Section 4.

Econometric Framework

This section discusses sources of bias in estimates of the effect of education on wage rates, and presents the econometric methodology used in the paper.

Assume that the log real wage W_{iht} is determined by

$$
\begin{aligned}
(1.1) \qquad W_{iht} = {} & X_{rih}B_0 + X_{gh}B_2 + X_{aih}B_3 + C_{ih}B_4 \\
& + S_{ih}\rho + EY_{iht}B_5 + Z_hG + \omega_{iht} + S_{ih}\rho_i + S_{ih}\rho_h \quad ,
\end{aligned}
$$

where

W_{iht} = the log of the real average hourly wage rate of person i from high school h in year t.

X_{rih} = variables indicating whether the individual is female, black, and/or Hispanic.

X_{fi} = a set of family background characteristics.

X_{gh} = variables for the geographic location of the high school.

X_{ah} = a vector of aptitude and achievement measures.

C_{ih} = a vector of measures of the high school curriculum taken by person i from high school h.

S_{ih} = a measure of postsecondary education, such as years of schooling.

EX_{iht} = the labor market experience of person i in year t.

Z_h = a vector of observed high school and community characteristics.

The variable ϖ_{iht} is the composite error component

$$(1.2) \qquad \omega_{iht} = v_i + v_h + m_h + e_{iht}$$
$$v_i = v_i + v_h$$

where

v'_i = an index of student- and family-specific factors that affect W independently of the high school and community environment

v_h = the mean in the high school of v'_i.

v_i = the difference between v'_i and v_h for student i.

m_h = an index of high school and community factors that affect w.

f_{iht} = a transitory error component that is assumed to be uncorrelated with all explanatory variables in the wage equation and with the other error components.

The component v_i is uncorrelated with v_h and m_h by construction.

There are two additional error components in the wage equation. The rate of return to education $\rho + \rho_i + \rho_h$ varies across individuals and depends on an individual-specific component ρ_i and a high school component ρ_h, where ρ_i and ρ_h are uncorrelated by construction and have means of 0. The unobserved term $S_{ih}(\rho_i + \rho_h)$ is treated as part of the wage equation error in estimation. Below I allow ρ to depend on whether a student is in an academic or nonacademic program by estimating separate equations for these groups.

Suppose one is interested in ρ, which is the labor market value of an extra year of education for a randomly selected person. Ordinary least squares will lead to biased estimates if the unobserved factors that influence W also influence S. S is likely to be positively related to variables that increase the productivity of higher education, lower the direct costs to the student or lower the discount rate, or raise the nonpecuniary benefits of education. Consequently, one would expect family background, ability and achievement, course of study in high school, and other high school and community factors to affect not only wage rates but also postsecondary schooling, and the evidence for the NLSHS72 is that they do (see Altonji (1988)). In fact, many studies of the return to education have few controls for ability, family background, curriculum in high school, and other characteristics of the high school and community. Even when one uses as rich a data set as the NLSHS72, the fact that observed measures of family background and parental attitudes, student achievement and ability, average background and ability of students in the high school, and high school characteristics are assumed to influence education and wage rates suggests that the unobserved determinants of education are correlated with [illegible] and $(\rho_i + \rho_h)$. This is because the observed measures are likely to be incomplete or unreliable.

In the empirical work below I systematically add to the wage equation controls for family background, curriculum in high school, aptitude and achievement, and observed high school characteristics and examine the sensitivity of estimates of the return to education to the choice of control variables. Since the NLSHS72 has more than one student in the sample from each high school, I am also able to experiment with the use of a separate intercept, or "fixed effect," for each high school as a control for the high school error component $v_h + m_h$.[6]

It is important to keep in mind that the use of high school fixed effects does not solve all of the potential sources of endogeneity of the schooling variable. Specifically, it does not eliminate potential bias from the correlation between S_{hi} and the individual error component v_i, or between S_{hi} and the component ρ_i of the rate of return to education. Even after one controls for observed measures of family background and for aptitude and achievement, unobserved ability differences among students from the same high school may affect both S and W. Furthermore, the quality of instruction and peer group experiences of students probably varies substantially even within a given track within the same high school, and so the fixed effect analysis does not

[6] The estimation methods discussed here are standard approaches to estimation of models with an error components structure. A convenient reference is Hsiao (1986), Chapter 3. The standard errors for the OLS and instrumental variables regressions with and without high school fixed effects allow for arbitrary high school specific forms of heteroscedasticity, serial correlation, and correlation across students from the same high school.

control for all high school characteristics that influence particular students.[7] However, this study goes further than previous studies by including high school curriculum controls as well as family background and ability measures.

One should also keep in mind that including controls for high school, family background, and test scores may exacerbate downward bias in the education coefficient arising from measurement error in education. I address the measurement error issue by using the independent information in the Post Secondary Transcript Survey that accompanied the NLSHS72 to create instrumental variables for the education measures. I use the transcript information as instrumental variables rather than as direct measures of education, because the PETS survey did not obtain transcripts for all students who claimed to have attended postsecondary schools, in some cases because of lack of cooperation by the schools. Consequently, the PETS measure of postsecondary education will also differ from actual schooling. I assume that the measurement errors in the PETS data are independent of the information on years of schooling and degree attainment provided by the student.[8] When high school constants are included in the wage equation, I use deviations of the transcript variables from high school means as the instrumental variables.

Data

The NLSHS72 is a Department of Education survey of individuals who were high school seniors during the spring of 1972. Thus, high school dropouts are excluded. The individuals were resurveyed in 1973, 1974, 1976, and 1979. A subsample was also resurveyed in 1986. I restrict the sample to the 16,683 individuals from the schools that participated in the base-year survey. The sample is reduced to 15,680 by eliminating observations with missing high school test information and to 12,980 by eliminating individuals who did not respond to all of the first four follow-ups. Information from the 1986 follow-up was then added for persons who were in the earlier sample of 12,980.

The cross section-time series of observations for each individual used in the wage analysis was created using information on earnings divided by hours for 1977, 1978, and 1979, and information on the wage at the

---

[7] There is information on tracking in the NLSHS72, and in future work it would be interesting to use a fixed effect to control for observed and unobserved characteristics that are common to students from the same track in high school.

[8] Students were asked during each follow-up survey to identify any schools that they were attending or had attended. Correlated measurement errors could arise if a student attended college but said otherwise. In this case the student would not provide the name of the postsecondary school attended, and no transcript would be found. I assume that persons do not hide the fact that they attended college if they attended college for a significant period of time.

beginning and end of each job held between 1980 and 1986, up to a maximum of the four most recent jobs. An observation for 1977 is included if (1) the individual was a full-time student neither in October 1976 nor in October 1977, (2) the number of hours worked in 1977 was greater than 1,040, and (3) the log of the 1977 real wage was between $.50 and $75 in 1967 dollars. Observations for 1978 and 1979 were included if they met the corresponding three criteria for 1978 and 1979, respectively. Data for beginning and ending job dates (1980-1986) were included if (1) the number of hours worked in the appropriate year was greater than 1,040, and (2) the log of the real wage was between $.50 and $75 in 1967 dollars. Restriction of the sample to cases with nonmissing data on the variables used in the wage analysis reduced the sample size to 38,595 observations on 9,239 individuals from 897 high schools.[9]

The coding of a few of the variables requires discussion. The variable YRSACD79 is the number of years of postsecondary academic education completed by 1979. The variable YRSVOC79 is the number of years of postsecondary vocational education by 1979. To allow for nonlinearities in the return to education, a set of mutually exclusive education dummy variables were created. VOC79 is 1 if an individual never attended college but did attend a postsecondary vocational school, and is 0 otherwise. SOC1479 is 1 if a person has less than two years of college (regardless of whether that person also attended vocational school). SOC1579 is 1 if a person attended college for two or more years but did not receive a four-year degree. COLL79 is 1 if a person received a four-year degree but did not receive an advanced degree. ADV79 is 1 if a person received a graduate degree.

The instrumental variables constructed from the PETS survey include a number of transcripts found for each student and nine dummy variables for whether students had the following transcript combinations: (1) at least one transcript, (2) a transcript from a nonacademic institution, (3) a transcript from a two-year public academic institution, (4) a transcript from a four-year public academic institution, (5) a transcript from a private academic institution, (6) a licence or certificate but no academic degree, (7) an associate degree but no bachelor's or advanced degree, (8) a college degree but no advanced degree, and (9) an advanced degree.[10]

---

[9] The subsample of students in academic programs contains 18,653 person-year observations from 858 high schools. The corresponding figures for the nonacademic (general and vocational tracks) subsample are 19,942 and 864.

[10] The PETS survey contains at least one transcript for 83 percent of the sample members who reported some postsecondary education by 1979, for 74.8 percent of those who reported vocational education or some college but no degree, and for 96.16 percent of those who reported a college or advanced degree. Transcript evidence of a college or advanced degree was found for 82.29 percent of the sample members who reported a college or advanced degree. Transcript evidence of a college or advanced degree was found for 3.16 percent of sample who did not report a college or advanced degree by 1979. Also, transcript evidence of an advanced degree was found for 8.13 percent of the persons who reported college as their highest degree in 1979, which may in part be due to completion

The control variables for region and city size, family background, aptitude and achievement measures, high school curriculum (semester hours in each of eight subjects), and high school characteristics are listed in the footnotes to the tables. Descriptive statistics and variable definitions are provided in Appendix Table 1.[11] To save space, only the education coefficients are shown in the text tables below.

Estimates of Returns to Education

Table 1 presents least squares estimates of the effects of YRSACD79 and YRSVOC79. Columns 1-4 do not include dummy variables for each high school, while Columns 5-7 do. All equations contain controls for race, Hispanic background, sex, experience, experience squared, time, and time squared. In each column the letters R, F, and H indicate whether controls for region and city size (R), family background and achievement and aptitude measures (F), and high school curriculum and high school characteristics (H) are included. The high school dummy variables absorb the effect of any variables that are constant within the high school, thus, region and city size and fixed high school characteristics are implicitly controlled for in Columns 5-7. Table 2 provides a comparable set of results, with YRSACD79 and YRSVOC79 replaced by the dummy variables VOC79, SOC1479, SOC1579, COLL79, and ADV79. The coefficients on the education variables are relative to a high school graduate, so the coefficient on COLL79 is the difference in the log wages of a college graduate and a high school graduate.

The Returns to Academic Education

The coefficient on YRSACD79 is .0817 when only the basic controls are included (Column 1) and .0790 when one controls for region and city size. This is typical of estimates from other data sets. The coefficient falls to .0653 when family background and ability and aptitude measures are added, a decline of .0137. This reduction is consistent with findings of other studies that have used detailed controls for family

---

of their advanced degrees after 1979.

[11] Appendix Table 1 presents the means and standard deviations of the variables used in the wage equations for the full sample, the academic sample, and the nonacademic sample. The table also reports the standard deviation of each variable within a high school, and the fraction of the sample variance that is across high schools. The standard deviations and the variance decomposition in the table refer to the cross section-time series sample, to which individuals contribute different numbers of observations. Consequently, they provide only a rough indication of the relative importance of variation within the high school and variation across high schools in wages, education, and background characteristics. See Altonji (1988) for a more thorough treatment of this issue. However, the results indicate that there is substantial variation across high schools in background characteristics, aptitude and achievement measures, and curriculum. Note also that there are substantial differences in the means for the academic and nonacademic samples.

background and ability or made use of sibling pairs.[12] Adding controls for the student's courses and a set of high school characteristics lowers the YRSACD79 coefficient by only .0009 to .0644.

A major aim of the present paper is to determine whether the failure to control for unobserved high school and community characteristics leads to a substantial upward bias in estimates of the returns to education. The answer seems to be that there is only a small upward bias without these controls. For example, when one adds high school dummy variables to the specification in Column 2, which does not contain controls for family background, aptitude and achievement, or courses taken, the coefficient on YRSACD79 falls from .0790 to .0749 (see Column 5). When one controls for background and achievement, the comparable coefficients without and with high school dummy variables are .0653 and .0605. When one controls for curriculum and observed high school characteristics, adding the high school dummy variables reduces the coefficient on YRSACD79 from .0644 to .0598. Thus, failure to control for high school differences leads to an upward bias of .005 in the education coefficient, which (multiplying by 100) is an upward bias of .5 in the rate of return to education.

Similar results are obtained for students from academic and nonacademic programs. The coefficients for the two subgroups are remarkably similar. They are also a bit below the coefficients for the combined sample. This reflects the fact that both the wage level and the postsecondary schooling level are positively correlated with enrollment in an academic program, even after controlling for background, aptitude and achievement, and semester hours by subject area. Controlling for background and ability has a relatively large effect on the education coefficient, while curriculum has only a small effect (compare Columns 2 and 3 and Columns 3 and 4). For the academic subgroups, including a separate dummy for each high school reduces the estimated rate of return by about .6 (compare Columns 2 and 5 or Columns 3 and 6 and multiply by 100). For the nonacademic subgroup, the reduction in the rate of return is about .2.

Table 2 reports OLS estimates of the effects of academic education when the dummy variables VOC79, SOC1479, SOC1579, COLL79, and ADV79 are used to parameterize the model. The results are qualitatively consistent with those based upon the linear specification in Table 1.

Instrumental Variables Estimates

For the combined sample, the use of IV in Table 3 has no effect (to four digits) on the return to YRSACD79, when one does not control for family background and test scores, and leads to a slight reduction (relative to OLS) in estimates of the return to academic education, when one does. Evidently, measurement error

[12] See Griliches (1979) and Olneck (1979) for discussions of alternative estimates of the return to education based on sibling data.

in the education variable decreases from about .079 with only regional controls to .058 when family background and test scores are added. There is only a small drop in the IV estimate (from .058 to .056) when high school fixed effects are added to the equation with family background and test scores. This confirms the earlier OLS finding that failure to control for high school, community, or both sets of variables leads to only a small bias in estimates of the return to education.

The use of IV does not significantly change the conclusions based on OLS for the academic and nonacademic subgroups. The IV estimates of models that use five dummy variables for education outcomes indicate that controlling for high school makes almost no difference for academic education, if anything, it leads to an increase in the estimated return to vocational educational education (see Table 4).

The Returns to Vocational Education

Tables 1 and 3 report OLS and IV estimates of the effect of years of vocational education (YRSVOC79) on wages for the combined sample and the academic and nonacademic subgroups. The mean of YRSVOC79 is .5110 for the combined sample and .5031 and .5183 for the academic and nonacademic subsamples, respectively. For the combined sample, the OLS results for the linear specification indicate a much lower return for vocational education than for academic education, with a coefficient of .0145 in the absence of controls (Column 1) and .0154 when one controls for background, aptitude and achievement, high school curriculum, and the high school (Column 7). The estimates for the academic and nonacademic samples are in the same general range, but the small coefficients should be treated cautiously for two reasons. First, vocational education is a very heterogenous category, and if programs lasting just a few months are coded as lasting a year, a downward bias would result. Second, there is evidence that the value of vocational education is lower if one has also obtained academic postsecondary education. I have excluded an interaction term from the specifications in Tables 1 and 3 to simplify interpreting the results. However, the categorical specification in Tables 2 and 4 excludes persons who have both vocational and academic education from the vocational category. Therefore, the mean of VOC79 is much lower for the academic sample than the mean for the nonacademic sample, despite the fact that the mean of YRSVOC79 is similar for the two groups. For the categorical specification and the combined sample, the OLS estimates imply that the vocational education dummy raises wages by .048 percent to .065 percent, depending upon what one controls for. I suspect that there are differences between academic and nonacademic students in the content of postsecondary vocational education, and that these may underlie the larger coefficient on VOC79 for the academic sample.

The instrumental variables estimates for YRSVOC79 and VOC79 follow the same general pattern as the OLS estimates, but are imprecise, particularly for the academic sample. Some of the point estimates for YRSVOC79 are actually negative but not statistically significant. However, for the combined sample the coefficient on VOC79 is quite substantial (.118) when one controls for the high school, family background, curriculum, and test scores, although the standard error is .064. A possible explanation (other than sampling error) is that the returns to vocational programs that are sufficiently well established to lead to a transcript, a licence, certificate, or any combination of the three are larger than the returns to other programs.

The impact of controlling for high school and community characteristics and for family background and achievement measures on estimates of the return to vocational education is sensitive to whether one uses OLS or IV, to the choice of dependent variable, and to whether the student was in an academic or nonacademic program in high school. I will not discuss the detailed results in the tables. There are a number of possible explanations for this, but I do not have the evidence on relative quality of high schools and of students who go on to vocational education to sort them out.

Conclusion

The OLS and instrumental variables estimates with high school fixed effects indicate that only modest biases result from the failure of previous studies to control for differences in high schools and for differences in primary school and community characteristics common to students from the same high school. This is good news for researchers, because few data sets permit one to study clusters of students from the same high school. This paper also confirms previous studies that have found that estimates of the rate of return to education are reduced substantially when one controls for family background and aptitude and achievement measures. The reduction does not result from measurement error in reported education.

REFERENCES

Akin, John S., and Irwin Garfinkel, "School Expenditures and the Economic Returns to Schooling," *Journal of Human Resources*, 12 (winter 1977): 460-477.

Altonji, Joseph G., "The Effects of Family Background and School Characteristics on Education and Labor Market Outcomes," unpublished paper, Northwestern University (December 1988).

Griliches, Zvi, "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica* 45 (January 1977): 1-22.

Griliches, Zvi, "Sibling Models and Data in Economics: Beginnings of a Survey," *Journal of Political Economy*, 87 (October 1979): S37-S64.

Hanushek, Eric, "The Economics of Schooling," *Journal of Economic Literature*, 24 No. 3 (September 1986): 1141-1177.

Hsiao, Cheng, *Analysis of Panel Data*, New York: Cambridge University Press, 1986.

Johnson, George, and Frank Stafford, "Social Returns to Quantity and Quality of Schooling," *Journal of Human Resources*, 8 (winter 1973): 139-155.

Morgan, James, and Ismail Sirageldin, "A Note on the Quality Dimension in Education," *Journal of Political Economy* 76 (September 1968): 1069-1077.

Murphy, Kevin, and Finis Welch, "The Structure of Wages," unpublished paper, Unicon Corporation (September 1988).

Olneck, Michael R., "The Effects of Education," in C. Jencks et al. (editors), *Who Gets Ahead?* New York: Basic Books, 1979.

Siebert, W. Stanley, "Developments in the Economics of Human Capital," in *Labour Economics*, London and New York: Longman, 1985.

Willis, Robert J., "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions," in O. Ashenfelter and R. Layard (editors) *Handbook of Labor Economics*, Elsevier Science Publishers (1986).

Table 1
Estimates of the Return to Education
Dependent variable: log wage

	OLS[a]				OLS Dummies for each high school[g]		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Combined Sample[a]			
YRSACD79	.0817	.0790	.0653	.0644	.0749	.0605	.0598
	(.0028)	(.0028)	(.0035)	(.0036)	(.0029)	(.0036)	(.0036)
YRSVOC79	.0145	.0150	.0133	.0135	.0173	.0163	.0154
	(.0053)	(.0052)	(.0032)	(.0052)	(.0053)	(.0052)	(.0052)
Other controls[b,c,d,e,f]		R	R,F	R,F,H	R	R,F	R,F,H
				Students in Academic Program[a]			
YRSACD79	.0731	.0734	.0636	.0637	.0663	.0568	.0567
	(.0043)	(.0042)	(.0047)	(.0047)	(.0046)	(.0050)	(.0051)
YRSVOC79	.0133	.0152	.0165	.0166	.0183	.0201	.0180
	(.0076)	(.0074)	(.0075)	(.0075)	(.0081)	(.0081)	(.0080)
Other controls[b,c,d,e,f]		R	R,F	R,F,H	R	R,F	R,F,H
				Students in Nonacademic Programs[a]			
YRSACD79	.0669	.0670	.0572	.0563	.0651	.0547	.0550
	(.0046)	(.0046)	(.0073)	(.0057)	(.0053)	(.0065)	(.0065)
YRSVOC79	.0196	.0192	.0162	.0178	.0138	.0121	.0116
	(.0074)	(.0073)	(.0073)	(.0074)	(.0078)	(.0078)	(.0078)
Other controls[b,c,d,e,f]		R	R,F	R,F,H	R	R,F	R,F,H

[a] The combined sample contains 38,595 person-year observations on 9,239 persons from 897 high schools. The academic sample contains 18,653 person-year observations on persons from the academic programs in 858 high schools. The nonacademic sample contains 19,942 person-year observations on persons from the vocational or general program in 864 high schools.
[b] All equations include BLACK, HISP, CSEX, PEXP10 PEXP10SQ, YEAR, AND YEARSQ.
[c] R = NO.CENTRAL, SOUTH, WEST, SMALLTOWN, MED.CITY, BIGCITY, BIGSUBURB, HUGECITY, MED.SUBURB,BIGSUBURB, HUGESURB, COLL_PROX.
[d] F = FATHER_ED, MOTHER_ED, LOWSES ED_MONEY MOTHER_WORK BLUECOLF, ENGLISH, FATH_COLL, MOTH_COLL, DISC_PLANS, PAR_INTEREST, PAR_INFL, IMPTAVER,

COLLEGE_ABILITY, TEACHER-ASSESMENT, VOCABULARY, PICTURE.NUMB, READING, LETTER.GROUP, MATH, MOSAIC.COMP, HOMEWORK, AND dummy variables for whether data were missing for FATH_COLL, MOTH_COLL, OR BLUECOLF.

[e] H = BLACK1, BLACK1^2, STU/TEACHER, PRIVATE, GRADES, ATTEND_RATES, DROP_RATE, TEACHER_TURNOVER, TEACHER_MA, SCIENCE_COURSE, HS_SIZE, SCIENCE, FOR.LANG, SOCIAL.ST, ENGLISH, MATH, IND.ARTS, COMMERCIAL, FINE.ARTS. Summary statistics and definitions of all of the variables are given in Appendix Table 1.

[f] Variables that do not vary across high schools, such as the geographic variables and the high school variables in footnote c and high school variables in footnote e, are implicitly controlled for in the equations with high school dummies.

[g] "White" standard errors in parentheses account for arbitrary forms of heteroscedasticity and correlation across observations on students from a given high school.

Table 2
Estimates of the Return to Education
Dependent variable: log wage

	OLS[a]				OLS Dummies for each high school[b]		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Combined Sample[c]			
VOC79	.0634	.0615	.0479	.0479	.0608	.0528	.0521
	(.0137)	(.0134)	(.0138)	(.0134)	(.0138)	(.0137)	(.0137)
SOC1479	.0773	.0674	.0241	.0251	.0479	.0149	.0146
	(.0133)	(.0111)	(.0119)	(.01020)	(.0118)	(.0125)	(.0127)
SOC1579	.1054	.1083	.1191	.1160	.1531	.1017	.0984
	(.0120)	(.0117)	(.0133)	(.0134)	(.0125)	(.0141)	(.0143)
COLL79	.3437	.3365	.2571	.2543	.3130	.2390	.2364
	(.0126)	(.0126)	(.0155)	(.0158)	(.0133)	(.0159)	(.0162)
ADV79	.5057	.4906	.4101	.4040	.4658	.3787	.3736
	(.0291)	(.0289)	(.0310)	(.0313)	(.0299)	(.0317)	(.0319)
Other controls[d,e,f]		R	R,F	R,F,H	R	R,F	R,F,H
				Students in Academic Programs[c]			
VOC79	.1528	.1507	.1342	.1297	.1840	.1629	.1568
	(.0310)	(.0314)	(.0336)	(.0325)	(.0320)	(.0130)	(.0337)
SOC1479	.0876	.0701	.0462	.0477	.0711	.0402	.0440
	(.0244)	(.0245)	(.258)	(.257)	(.0260)	(.0274)	(.0274)
SOC1579	.1933	.1757	.1427	.1403	.1560	.1254	.118?
	(.0237)	(.0236)	(.0159)	(.0260)	(.0271)	(.0294)	(.0294)
COLL79	.3319	.3321	.2677	.2667	.2975	.2450	.2390
	(.0237)	(.0238)	(.0271)	(.0273)	(.0265)	(.0294)	(.0296)
ADV79	.4734	.4658	.3995	.3997	.4370	.3711	.3668
	(.0347)	(.0341)	(.0360)	(.0366)	(.0363)	(.0407)	(.0408)
Other controls[d,e,f]		R	R, F	R, F, H	R	R, F	R, F, H
				Students in Nonacademic Programs[c]			
VOC79	.0368	.0379	.0293	.0306	.0303	.0243	.0264
	(.0140)	(.0146)	(.0144)	(.0144)	(.0150)	(.0157)	(.0166)
SOC1479	.0681	.0405	.0294	.0310	.0414	.0210	.0235
	(.0135)	(.0133)	(.0130)	(.0130)	(.0146)	(.0155)	(.0156)
SOC1579	.1610	.1473	.1167	.1165	.1310	.1004	.0998
	(.0159)	(.0156)	(.0167)	(.0160)	(.0173)	(.0184)	(.0186)
COLL79	.3069	.3044	.3407	.3388	.2836	.2379	.2400
	(.0194)	(.0194)	(.0219)	(.0221)	(.0222)	(.0254)	(.0253)
ADV79	.4882	.4809	.4252	.4223	.3648	.3318	.3327

	(.1436)	(.1438)	(.1456)	(.1479)	(.1461)	(.1477)	(.1478)
Other controls[illegible]		R	R, F	R, F, H	R	R, F	R, F, H

FOOTNOTE: See Table 2

Table 3
Estimates of the Return to Education
Dependent variable: log wage

	Instrumental variables [a,b]				Instrumental Variables Dummies for each high school [a,b]		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Combined Sample[a]			
YRSACD79	.0817	.0783	.0582	.0572	.0770	.0560	.0551
	(.0031)	(.0031)	(.0037)	(.0037)	(.0033)	(.0039)	(.0040)
YRSVOC79	.0254	.0080	-.0127	-.0151	.0084	-.0104	-.0133
	(.0170)	(.0167)	(.0169)	(.0169)	(.0171)	(.0173)	(.0175)
Other controls[c,d,e,f]		R	R,F	R,F,H	R	R,F	R,F,H
				Students in Academic Program[a]			
YRSACD79	.0765	.0765	.0618	.0615	.0699	.0570	.0568
	(.0058)	(.0056)	(.0062)	(.0062)	(.0065)	(.0071)	(.0072)
YRSVOC79	.0550	.0379	.0338	.0327	.0283	.0336	.0314
	(.0340)	(.0331)	(.0333)	(.0329)	(.0379)	(.0379)	(.0378)
Other controls[c,d,e,f]		R	R,F	R,F,H	R	R,F	R,F,H
				Students in Nonacademic Programs[a]			
YRSACD79	.0724	.0722	.0589	.0578	.0758	.0636	.0641
	(.0054)	(.0055)	(.0061)	(.0062)	(.0064)	(.0073)	(.0073)
YRSVOC79	.0180	.0033	-.0074	-.0095	-.0041	-.0143	-.0173
	(.0212)	(.0213)	(.0212)	(.0213)	(.0225)	(.0225)	(.0226)
Other controls[c,d,e,f]		R	R,F	R,F,H	R	R,F	R,F,H

[a,c,d,e,f] See Table 1.

[b] For columns 1-4 the instruments consist of dummies for whether the individual had a postsecondary transcript, a transcript from a vocational school, a transcript from a 2-year public college, a 4-year public college, a private college; dummies for whether the individual's highest degree was a license or certificate, an associate degree, a bachelor's degree, or an advanced degree; and a count of the number of transcripts for the individual. The IV estimates with high school dummies are computed from deviations from high school means (for the combined sample, the academic sample, and the nonacademic sample, as appropriate) for all variables, including the instruments. CSEX, BLACK, and HISP were also included. The education coefficients are numerically identical to what one would obtain if all of the other variables that appear in a particular equation were used as instruments for that equation.

BEST COPY AVAILABLE

Table 4
Estimates of the Return to Education
Dependent variable: log wage

	Instrumental variables [a,b]				Instrumental Variables Dummies for each high school [a,b]		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Combined Sample[c]						
VOC79	.0404	.0611	.0673	.0625	.01196	.1202	.1179
	(.0617)	(.0613)	(.0614)	(.0606)	(.0641)	(.0648)	(.0642)
SOC1479	.0640	.0362	-.0171	-.0177	.0363	-.0136	-.0146
	(.0293)	(.0287)	(.0289)	(.0290)	(.0296)	(.0296)	(.0297)
SOC1579	.1968	.1828	.1112	.1032	.1571	.1143	.1073
	(.0275)	(.0271)	(.0278)	(.0280)	(.0280)	(.0288)	(.0292)
COLL79	.3134	.3304	.3066	.3029	.3049	.3097	.3063
	(.0209)	(.0208)	(.0216)	(.0217)	(.0215)	(.0227)	(.0227)
ADV79	.6364	.6160	.4904	.4779	.6177	.4883	.4837
	(.0643)	(.0643)	(.0646)	(.0642)	(.0659)	(.0665)	(.0663)
Other controls[d,e,f,g]		R	R,F	R,F,H	R	R,F	R,F,H
	Students in Academic Programs[c]						
VOC79	.0058	.0672	.0065	.0411	.3581	.3225	.3179
	(.1590)	(.1544)	(.1580)	(.1560)	(.1725)	(.1727)	(.1721)
SOC1479	.0093	-.0297	-.0731	0.0064	-.0283	-.0659	-.0779
	(.0743)	(.0716)	(.0713)	(.0714)	(.0761)	(.0753)	(.0754)
SOC1579	.1931	.1669	.1143	.1060	.1668	.1187	.1090
	(.0667)	(.0662)	(.0677)	(.0676)	(.0625)	(.0645)	(.0644)
COLL79	.2722	.2534	.1773	.1726	.2607	.1893	.1815
	(.0647)	(.0636)	(.0652)	(.0651)	(.0695)	(.0610)	(.0609)
ADV79	.5433	.5333	.4163	.3983	.5443	.4201	.4198
	(.0811)	(.0797)	(.0816)	(.0818)	(.0880)	(.0921)	(.0818)
Other controls[d,e,f,g]		R	R, F	R, F, H	R	R, F	R, F, H
	Students in Nonacademic Programs[c]						
VOC79	.0273	.0666	.0616	.0677	.0261	.0287	.0276
	(.0644)	(.0641)	(.0634)	(.0630)	(.0698)	(.0704)	(.0701)
SOC1479	.0923	.0609	.0331	.0312	.0690	.0409	.0414
	(.0304)	(.0298)	(.0303)	(.0304)	(.0332)	(.0342)	(.0342)
SOC1579	.1401	.1494	.1102	.1027	.1373	.1006	.0966
	(.0345)	(.0345)	(.0346)	(.0355)	(.0375)	(.0384)	(.0366)
COLL79	.2877	.2909	.2324	.2283	.2921	.2377	.2383
	(.0299)	(.0300)	(.0316)	(.0313)	(.0340)	(.0364)	(.0362)
ADV79	.7301	.7546	.6784	.7012	.8309	.7805	.8112
	(.3276)	(.3425)	(.3420)	(.3485)	(.3740)	(.3731)	(.3749)
Other controls[d,e,f,g]		R	R, F	R, F, H	R	R, F	R, F, H

Footnotes: See Table 1 and 3.

Appendix

Means and Standard Deviations

Explanatory Variables	Combined Sample (1) Mean	Combined Sample (2) SD	Combined Sample (3) SD in HS	Combined Sample (4) Fraction of variance across high schools	Academic (5) Mean	Academic (6) SD	Nonacademic (7) Mean	Nonacademic (8) SD
Wages								
LOGWAGE log of real average hourly wage, 1967 dollars	.9196	.4635	.4402	.0980	.9916	.4746	.8523	.4425
Education								
YRSAC79 years of postsecondary education	1.910	1.766	1.594	.1853	2.835	1.660	1.044	1.362
YRSVOC79 years of postsecondary education	.5110	.7613	.7154	.1160	.5031	.7908	.5183	.7325
VOC79 1 if some vocational, no college by 79	.0851	--	--	--	.0381	--	.1290	--
SOC1479 1 if less than two years college by 79	.1803	--	--	--	.1413	--	.2168	--
SOC1579 1 if more than two years college, no degree	.1738	--	--	--	.1693	--	.1497	--
COLL79 1 if college degree, no advanced degree by 79	.3022	--	--	--	.0534	--	.0053	--
ADV79 1 if advanced degree by 79	.0285	--	--	--	.0534	--	.0053	--

Appendix

Means and Standard Deviations

Explanatory Variables	Combined Sample (1) Mean	(2) SD	(3) SD in HS	(4) Fraction of variance across high schools	Academic (5) Mean	(6) SD	Nonacademic (7) Mean	(8) SD
BLACK	.0892	--	--	--	.0663	--	.1106	--
HISP	.0366	--	--	--	.0213	--	.0509	--
CSEX	.4916	--	--	--	.4780	--	.5043	--
Family background								
FATHER_ED father's ed.	12.75	2.545	2.171	.2723	13.49	2.632	12.06	2.250
MOTHER_ED mother's ed.	12.43	2.098	1.854	.2191	12.96	2.158	11.93	1.910
LOWSES 1 if low SES	.2340	--	--	--	.1343	.3410	.3273	.4692
ED_MONEY 1 if worry over money interfered with HS education	.2891	--	--	--	.2261	.4183	.3480	.4764
MOTHER_Work 1 if mother worked while in elementary school	.4021	--	--	--	.3804	.4855	.4223	.4939
BLUECOLF 1 if father blue collar	.3216	--	--	--	.2925	.4549	.3488	.4766
ENGLIS 1 if English spoken at home	.9207	--	--	--	.9183	.2739	.9230	.2666
FATH_COLL 1 if father wants college or grad school	.5787	--	--	--	.7814	.4133	.3890	.4875
MOTH_COLL 1 if mother wants college or grad school	.6140	--	--	--	.8183	.3856	.4229	.4940
DISC_PLANS 1 if student often discussed plans with parents	.7902	--	--	--	.8510	.3561	.7332	.4423
PAR_INTEREST 1 if uninterested parents interfered with high school	.2019	--	--	--	.1272	.3332	.2718	.4449
PAR_INFL 1 if parents influenced post-HS plans a great deal	.4398	--	--	--	.5037	.5000	.3799	.4854
Geographic variables								
SMALLTOWN	.2953	--	--	--	.3027	.4574	.2883	.4501
MED.CITY	.0832	--	--	--	.0863	.2797	.0802	.2699
MED.SUBURB	.0487	--	--	--	.0596	.2358	.0386	.1911
BIGCITY	.1020	--	--	--	.0855	.2925	.1080	.3084
BIGSUBURB	.1046	--	--	--	.1171	.3203	.0930	.2904
HUGECITY	.0785	--	--	--	.0890	.2834	.0696	.2505

Appendix

Means and Standard Deviations

Explanatory Variables	Combined Sample				Academic		Nonacademic	
	(1) Mean	(2) SD	(3) SD in HS	(4) Fraction of variance across high schools	(5) Mean	(6) SD	(7) Mean	(8) SD
HUGESURB	.0950	--	--	--	.1127	--	.0784	--
NO.CENTRAL	.2018	--	--	--	.2750	--	.3072	--
SOUTH	.3176	--	--	--	.2737	--	.3585	--
WEST	.1678	--	--	--	.1431	--	.1881	--
COLL-PROX	1.785	.7471	--	--	1.701	.7022	1.863	.7787
High school program								
HSACAD student is in academic program	.4833	--	.4378	--	1.0	--	0.0	--
SCIENCE semester hours	19.21	10.11	8.228	.3377	23.54	9.655	15.16	8.752
FOR.LANG semester hours	11.39	11.43	9.203	.3517	16.98	11.02	6.152	9.089
SOCIAL.ST semester hours	26.27	7.643	5.465	.4887	26.56	7.635	25.99	7.641
ENGLISH semester hours	30.08	6.665	4.509	.5423	30.75	6.516	29.46	6.741
MATH semester hours	19.67	9.839	8.078	.3259	24.05	8.573	15.58	9.159
IND.ARTS semester hours	5.919	12.18	10.86	.2024	3.661	8.704	8.030	14.35
COMMERCIAL semester hours	13.32	15.77	13.87	.2264	8.223	10.32	17.90	18.34
FINE.ARTS semester hours	8.768	13.22	11.25	.2758	9.288	13.73	8.282	12.70
Aptitude and achievement measures								
IMPTAVER grades	15.64	7.586	3.469	.7909	14.44	7.437	16.76	7.553
COLLEGE_ABILITY 1=definitely; 5=definitely not	1.843	.9659	.8984	.1349	1.477	.6991	2.186	1.052
TEACHER_ASSESSMENT expectation of student (5=low, 1=high)	2.085	.8701	.9227	.1060	1.816	.7948	2.337	.8621
VOCABULARY	52.31	9.896	8.641	.2376	56.34	9.328	48.54	8.872
PICTURE.NUMB associative memory	51.57	9.680	8.938	.1474	53.75	9.138	49.54	9.456
READING	52.31	9.424	8.459	.1943	56.07	8.321	48.80	9.033
LETTER.GROUP inductive reasoning	52.34	8.878	8.035	.1809	55.33	7.082	49.54	9.456
MATH quantitative comparisons (basic competence in math)	52.51	9.539	8.486	.2086	57.06	7.939	48.25	8.924
MOSAIC.COMP perceptual speed and accuracy	51.46	9.187	7.407	.3500	53.11	8.681	49.92	9.377

Appendix

Means and Standard Deviations

Explanatory Variables	Combined Sample (1) Mean	(2) SD	(3) SD in HS	(4) Fraction of variance across high schools	Academic (5) Mean	(6) SD	Nonacademic (7) Mean	(8) SD
HOMEWORK homework hrs/week	4.467	3.278	3.018	.1523	5.315	3.442	3.874	32.900
High school characteristics for individual student								
ESSAYS essays used fairly often or frequently in classes	.6502	--	--	--	.6655	--	.6359	--
NO_COURSES 1 if lack of courses interfered somewhat or great deal with HS education	.5043	--	--	--	.4957	--	.5124	--
High school characteristics								
BLACK1 fraction black or Native American reported in school survey	11.54	19.24	--	--	10.57	17.97	12.46	20.32
BLACK1^2	503.5	14671.	--	--	434.5	1338.	568.1	1583.
STU/TEACHER student-teacher ratio	20.30	4.248	--	--	20.40	4.093	20.20	4.385
PRIVATE if school is private or parochial	.0751	--	--	--	.1128	--	.0398	--
GRADES 5 minus # of grades in high school	1.620	.8628	--	--	1.572	.8104	1.664	.8884
ATTEND_RATE daily attendance rate	91.42	3.685	--	--	91.47	3.864	91.38	3.509
DROP_RATE weighted average of dropout rates for males and females	6.359	6.398	--	--	5.835	6.260	6.849	6.486
TEACHER_TURNOVER percentage of full-time teachers leaving between 1970-71 and survey in spring 1972	8.204	8.306	--	--	7.747	7.608	8.631	8.888
TEACHER_MA percentage of full-time teachers with master's or Ph.D.	40.59	20.72	--	--	43.89	20.90	37.51	20.07
SCIENCE_COURSE sum of dummy variables for whether a course offered in general science, biology, chemistry, and physics	3.62	.7115	--	--	3.639	.6457	3.606	.7676
HS_SIZE high school senior class enrollment	351.2	235.9			372.2	237.2	331.6	233.0

National Center on Education and Employment

Teachers College,
Columbia University
Box 174
New York, New York 10027

The RAND Corporation
2100 M Street, N.W.
Washington, D.C. 20037