DOCUMENT RESUME

ED 286 268 EA 019 656

AUTHOR	Stern, David; And Others
TITLE	Labor Market Experience of Tennagers with and without High School Diplomas.
PUB DATE	86
NOTE	43p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Comparative Analysis; *Dropout Programs; *Dropouts; *Educational Status Comparison; Education Work Relationship; Employer Attitudes; *Employment Potential; *High School Graduates; High Schools; Labor Market; Outcomes of Education; Predictor Variables; Salary Wage Differentials; Statistical Analysis; *Student Characteristics; Student Needs

ABSTRACT

The rationale for high school dropout reduction programs rests mainly on evidence that graduates are more successful than dropouts in the labor market. It is not evident, however, whether this difference is attributable to the diploma itself or to underlying characteristics that affect both graduation and labor market success. This paper estimates the effect of the diploma on labor market success, over and above prior differences between graduates and dropouts, using a statistical procedure controlled for observable characteristics (ethnicity, gender, and academic achievement). The research also attempted to discover the role of unmeasured characteristics that may be associated with high school completion and labor market success. After selecting data from the 1980 sophomore cohort in the "High School and Beyond" survey, comparisons were made between a group who left school without a diploma and a group who graduated but did not attend postsecondary school. Results indicate that differences in employment and earnings between teenagers with and without diplomas are primarily not attributable to differences in measured prior characteristics. Dropouts' labor market difficulties are therefore not merely symptomatic of the prior characteristics measured by this research. The question remains whether the problems are due to discrimination by employers against dropouts or to other unmeasured variables. Society's gain from dropout reduction programs depends on whether employers' reasons for favoring graduates are arbitrary or valid. Conversely, if dropouts have unobserved characteristics causing them to be less productive than graduates, then society's gain from program efforts depends on whether would-be dropouts acquire characteristics during schooling that make them more productive. Seventeen references and 13 tables of data are appended. (CJH)

LABOR MARKET EXPERIENCE OF TEENAGERS WITH AND WITHOUT HIGH SCHOOL DIPLOMAS

David Stern, University of California, Berkeley
Il-Woo Paik, University of California, Berkeley
James S. Catterall, University of California, Los Angeles
Yoshi-fumi Nakata, University of California, Berkeley

Effect of High School Diplomas

LABOR MARKET EXPERIENCE OF TEENAGERS WITH AND WITHOUT HIGH SCHOOL DIPLOMAS

Abstract

The rationale for programs to reduce the high school dropout rate rests mainly on evidence that high school graduates are more successful than dropouts in the labor market. However, both high school completion and labor market success are to some extent results of prior characteristics. This paper estimates the effect of a high school diploma itself on success in the labor market, over and above the effects of prior characteristics, of which some are measured and some are not. Results indicate that differences in unemployment and earnings between teenagers with and without high school diplomas are mainly not attributable to differences in measured prior characteristics.

Objective

Reducing the high school dropout rate is a persistent aim of public policy. Evidence to support this policy consists mainly of findings that high school graduates are more successful in the labor marekt than high school dropouts (e.g., see Levin, 1972; Catterall, 1985). However, it is not self-evident whether the difference between the labor market experiences of dropouts and graduates is attributable to the high school diploma itself, or whether it is the result of underlying characteristics which affect both high school completion and success in the labor market. If these prior characteristics were entirely responsible for determining both high school completion and success in the labor market, then programs designed to reduce the high school dropout rate would not produce any gain in total economic output or societal well-being unless the programs were able to alter these underlying characteristics. Our objective in this paper is to measure the effect of the high school diploma, per se, over and above any such prior differences between graduates and dropouts. We attempt to control not only for the effects of characteristics which are ordinarily observable by both employers and researchers -- including variables such as race or ethnicity, gender, socioeconomic background, and cognitive or academic achievement -- but also for unmeasured characteristics which may be associated with both high school completion and success in the labor market.

Methods for Estimating the Effect of a High School Diploma, Per Se

If a high school graduate were suddenly stripped of the diploma, or if a high school dropout were suddenly given one, what would be the effect on the individual's experience in the labor market? This question cannot be answered with certainty, simply because there is no practical way to bestow or take away high school diplomas experimentally. Therefore, researchers are limited to addressing the following question instead: What is the difference in labor market experience between individuals who are alike in all respects except whether or not they have the high school diploma? In order to identify individuals who are alike in all respects, it is necessary to apply some kind of statistical control. Techniques for applying such statistical control have been evolving in recent years, as we will now explain.

The diagram in Figure 1 is useful for understanding the various statistical techniques. This is a convential "path" diagram, in which arrows represent possible causal influences. The aim is to estimate the magnitude of these influences, which could be positive, negative, or zero. Here we are interested especially in estimating the influence of the high school diploma on labor market success.

[Figure 1 about here]

Path diagrams like this one have been used extensively by sociologists and economists studying the "status attainment" process (e.g., Sewell and Hauser, 1975; Jencks and others, 1972). However, the estimation procedure used in most status attainment research has relied on the implicit or explicit assumption that the set of relationships is fully "recursive" (e.g., see Kiker and Condon, 1981). This means two things (e.g., see Hanushek and Jackson, 1977). First, if variable A is a possible cause of variable B, then B must not be considered a possible cause of A. This assumption is reflected in the one-directional flow of arrows in Figure 1. Since the sequence of variables in the causal flow is the same as their chronological order, this assumption is quite justifiable here.

The second assumption of a fully recursive model is that unmeasured influences on behavior at different points in the causal sequence are not correlated with each other. For instance, suppose that high school completion not only depends on family background, intelligence, performance in school, and other measured variables, but also depends on variables or events which are not measured, such as whether a student has an uncle who offers a full-time job and thereby induces the student to leave high school before graduating. That same uncle, whose existence is not known to the researcher, may continue to influence the student's success in the labor market after the student leaves high school. If so, the same unobserved uncle is clearly influencing outcomes at different points in the causal sequence. This is what a fully recursive model assumes does not happen. However, that assumption would be wrong in this hypothetical case.

Generally, it is hard to justify the assumption that unmeasured influences at different points in the causal sequence are not correlated with each other. We can easily imagine any number of idiosyncratic relationships or events which would not ordinarily be known to the researcher but would influence both high school completion and subsequent success in the labor market. These unobserved influences could be positive (the beneficent uncle) or negative (recurrent illness, stormy love life). Such influences could be observed if the researcher had enough resources, but resource limitations in practice mean there are always unobserved influences, and some of these may act at more than one stage of the causal sequence in the model. Therefore, Figure 1 shows the same unmeasured influences exerting possible effects all along the way. In other words, we assume that a fully recursive model is not appropriate.

This means we must discount the analysis by Bachman and others (1971) of data from the Youth in Transition longitudinal survey. That analysis attempted to estimate the effect of a high school diploma per se, but made no allowance for the fact that unmeasured influences on high school completion could be correlated with unmeasured influences on post-high school experience. The resulting estimate of the diploma effect could be too large, if many dropouts have unobserved problems which lead both to dropping out and to subsequent difficulties in the labor market. Alternatively, the effect of the diploma could be estimated to be smaller than it really is, if many students drop out because they have unobserved uncles or other influences which enable them to succeed in the job market without finishing high school.

In recent years, statistical procedures have been developed which allow us to estimate the bias due to unobserved influences in problems like this one. Heckman (1979) is often credited with devising these techniques. Willis and Rosen (1979) and Garen

(1985) have used these procedures to estimate economic returns to education. However, these applications have been based on models in which individuals are assumed to maximize future income, knowing in advance what income they are likely to obtain if they choose different levels of schooling. Such models are more plausible in analyzing choices about higher education than in understanding high school dropout behavior. Non-monetary factors reportedly play a big part in prompting high school students to drop out (e.g., see Wehlage, 1983). Furthermore, most high school dropouts, when interviewed within a few years after leaving school, declare that they have made a mistake (Combs and Cooley, 1968, pp. 352, 358; Bachman, 1971, p. 160; Jones and others, 1986, p. 8.4.9). This is not consistent with an economic model of "rational expectations." Therefore, our approach uses the new procedures to correct for bias due to unmeasured influences, but does not depend on the assumption that dropouts are maximizing their incomes.[7]

Specifically, we follow a procedure spelled out by Maddala (1983, pp. 223-225). For explanatory purposes here let y_i denote some measure of success in the labor market for person i, and x_i represent a variable that may predict y_i (in our actual model below, there is more than one predictor). Let

(1) $y_{1i} = \alpha_1 + \beta_1 x_i$

represent the predicted relationship between y_i and x_i for a high school graduate, where α_1 and β_1 are numerical constants for all individuals. This equation is meant to predict what the value of y_i would be if the person were a high school graduate, whether or not the person actually is a graduate in fact. Similarly, let

(2) $y_{2i} = \alpha_2 + \beta_2 x_i$

represent the relationship between y_i and x_i for anyone without a high school diploma. In other words, equation (2) predicts what the value of y_i would be for any individual if he or she were a dropout, whether or not he or she actually is a dropout in fact.

Since equations (1) and (2) are meant to apply to any individual regardless of his or her actual status, they can be used to measure the predicted effect of a high school diploma for any individual with measured characteristics x_i:

$$y_{1i} - y_{2i} = \alpha_1 - \alpha_2 + (\beta_1 - \beta_2)\ x_i.$$

Given the value of the predictor x_i, and estimates of the alphas and betas, this is how we can compute the difference a diploma makes for the labor market success of an individual with given x_i.

However, there is a problem: the labor market experience of any individual is actually observed when that person is a high school graduate or a dropout, but not both. Therefore, equation (1) can be estimated only with data for people who actually are high school graduates, and equation (2) only for actual dropouts. The reason this is a problem has to do with unmeasured influences. If graduates are different from dropouts in some unobserved way, and if that unobserved difference is also associated positively or negatively with success in the labor market, then ordinary least-squares estimates of alphas and betas in equation (1), based on actual graduates only, and in equation (2), based on actual dropouts only, will be different than if each equation could somehow be estimated for the whole population. For instance, if graduates have more of some unmeasured trait such as perseverance which makes them more likely not only to graduate but also to succeed in the labor market, then equation (1), estimated by ordinary least-squares regression from data on actual graduates only, will predict greater success for an individual with given x_i than it would if it could somehow be estimated with data for the whole population, including dropouts.

Futhermore, if there is overprediction from equation (1), it does not necessarily follow that there is underprediction from equation (2). For instance, while graduates might have more perseverance on mental tasks, dropouts might have more physical energy, which conceivably could make them more likely to drop out and also could give them an advantage in certain jobs that do not require a high school diploma. If so, then actual graduates would do better as graduates than would the average person in the whole population consisting of both graduates and dropouts, and actual dropouts would do better as dropouts than would the average person in the whole population consisting of both dropouts and graduates.

As these examples illustrate, the prediction bias from ordinary least-squares estimation of equations (1) and (2) depends on whether unobserved influences on labor market success are correlated with unobserved influences on the probability of dropping out of high school. The statistical theory developed by Heckman and others tells us that, if these unobserved influences have a joint normal distribution, the amount of prediction bias in equations (1) and (2) can be estimated and, therefore, corrected. One technique for correcting the bias proceeds in two stages. In the first stage, a probit analysis is used to estimate the probability that each individual will graduate from high school, based on a set of observed predictors, z_i. The variables that predict dropping out should include some which are different from the x_i in equations (1) and (2). (In fact, in our analysis we orthogonalized the two sets of predictors by first regressing z_i on x_i and using the residuals as predictors of dropping out.) A function of the estimated probability of graduating is then added as a predictor to the right-hand sides of equations (1) and (2), which are then estimated by generalized least-squares regression. This is the second stage of the procedure. The estimated coefficient on this transformed probability is interpretable as an estimate of the correlation between unobserved influences on dropping out and unobserved influences on labor market success.

(3) $y_{1i} = \alpha_1 + \beta_1 x_i + \sigma_1 W_{1i}$ (graduates)

(4) $y_{2i} = \alpha_2 + \beta_2 x_i + \sigma_2 W_{2i}$ (dropouts)

Equations (3) and (4) are the corrected versions of equations (1) and (2), giving predicted labor market success y_i as a function of observed characteristics x_i (here, for explanatory purposes, there is still only one x_i; in our actual model below we use several), for graduates and dropouts, repectively. W_{1i} and W_{2i} represent functions of the probability that an individual graduates from high school, based on the first stage probit analysis. σ_1 and σ_2 are the estimated correlations between unobserved influences on high school graduation and labor market sucess. The alphas, betas, and sigmas are estimated by generalized least-squares regression.

Given these estimates, we can compute, for any individual with measured characteristics x_i, the difference in predicted labor market success resulting from presence or absence of a high school diploma. Moreover, we can measure how much of the average difference in labor market success between actual graduates and actual dropouts is attributable to each of the following three causes:

-- the average difference between dropouts and graduates in measured characteristics associated with labor market success (x_i);

-- the difference in how the labor market rewards graduates and dropouts who have identical measured characteristics (from comparing the alpha and beta in equation 3 with those in equation 4); and

-- the net effect of any unmeasured characteristics that are associated with both the probability of dropping out and success in the labor market after leaving high school.

These three magnitudes can be computed from the following identity:

$$(5) \quad \bar{y}_G - \bar{y}_D = 0.5\,(\beta_1 + \beta_2)\,(\bar{x}_G - \bar{x}_D) \\ + \alpha_1 - \alpha_2 + 0.5\,(\beta_1 - \beta_2)\,(\bar{x}_G + \bar{x}_D) \\ + \bar{y}_G - y_G - (\bar{y}_D - y_D),$$

where the three lines on the right-hand side correspond to the three components listed above, in the same order. In equation (5), the subscript G denotes graduates and D stands for dropouts; $\bar{y}_G$ and $\bar{y}_D$ are averages on a measure of labor market success; $\bar{x}_G$ and $\bar{x}_D$ are average values of a measured variable that predicts labor market success (in our actual model below there is more than one such variable); $y_G = \alpha_1 + \beta_1\bar{x}_G$ using coefficients from equation (3), and $y_D = \alpha_2 + \beta_2\bar{x}_D$ using coefficients from equation (4).

It is the second line on the right-hand side of equation (5) that we interpret as the effect of a diploma, per se. This is the difference between the intercepts in equations (3) and (4), plus the difference in the slope coefficients multiplied by the average of the mean observed characteristics of actual graduates and dropouts. The first line on the right-hand side of equation (5) is the average of the intercepts from equations (3) and(4) multiplied by the differences between the mean observed characteristics of actual graduates and dropouts. The third line is the difference between two differences. The first difference, $\bar{y}_G - y_G$, is the mean outcome for actual graduates minus the predicted outcome for anyone in the population -- including actual dropouts -- with the observed characteristics of the average actual graduate. It is, therefore, the effect of unobserved characteristics of actual graduates on the labor market outcome. Likewise, the second difference, $\bar{y}_D - y_D$, is the effect of unobserved characteristics of actual dropouts. The difference between these two differences is the net effect of unobserved characteristics on the difference in labor market outcome between actual graduates and actual dropouts.

Data and Findings

We used data on the 1980 sophomore cohort in the national High School and Beyond (HSB) survey (Jones and others, 1986). From this cohort we selected two groups. One group consisted of individuals who left high school without receiving a diploma, and who had not returned to school or obtained a diploma or equivalent at the time of the 1984 follow-up interview. The second group consisted of individuals who graduated from high school but attended no postsecondary school prior to the 1984 interview.

Tables 1 and 2 give means and standard deviations of all variables used in the analysis, for dropouts and graduates, respectively. Definitions of variables are as follows:

GRADUATE	=	1 if respondent had obtained a regular high school diploma by the time of the survey in spring 1984; 0 if the person had no diploma or equivalent.
UNEMP	=	amount of time respondent was without a job and looking for work, as a proportion of all time in the labor

		force (either employed or looking for work), from July 1982 to date of survey in spring 1984.
WAGE	=	total earnings from paid jobs, divided by total hours worked, from July 1982 to date of survey in spring 1984.
MALE	=	1 if male, 0 if female.
SES	=	index of socioeconomic status based on father's occupation, father's education, mother's education, family income, and material possessions in the household.
TEST	=	average of composite test score from 1980 and 1982.
NOMAN	=	1 if student reported no adult male in household in 1980; 0 otherwise.
HISPANIC	=	1 if Hispanic; 0 otherwise.
BLACK	=	1 if black; 0 otherwise.
WHITE	=	1 if non-Hispanic white; 0 otherwise.
SCHJOBHRS	=	number of hours per week worked on present or most recent job as of survey date in spring 1982.
CUNEMR82	=	1982 unemployment rate in the county where respondent's school was located.
CWAGE82	=	1982 mean wage in manufacturing in the county where respondent's school was located.

EXPGRAD	=	1 if respondent expected to graduate from high school, as indicated in spring 1980 survey; 0 otherwise.
DEFINJOB	=	1 if respondent had a definite job lined up for after high school, as of survey in spring 1980; 0 otherwise.
GRADES	=	self-reported average grades in high school (on percentage scale, 0 to 100), as of spring 1980.
ABSENT	=	self-reported frequency of absence from school (number of days from beginning of school year to Christmas) during first half of sophomore year, from 1980 survey.
LATE	=	self-reported frequency of lateness to school (number of times) during first half of sophomore year, from 1980 survey.
RESWAGEBEF	=	lowest hourly wage respondent would accept for a job while still in high school, from 1980 survey.
RESWAGEAFT	=	lowest hourly wage respondent would accept after graduating from high school, as indicated in 1980 survey.
EXTRACURR	=	number of extracurricular activities in which respondent participated during sophomore year, from 1980 survey.
OWNVALUGRADES	=	1 if respondent thinks well of a student who gets good grades; 0 otherwise.

FRNDVALUGRADES	=	1 if respondent's friends think well of a student who gets good grades; 0 otherwise.
SCHLSATIS	=	1 if true that "I am satisfied with the way my education is going" in 1980 survey; 0 if false.
DISCPROBS	=	1 if true that "I have had disciplinary problems in school during the last year" in 1980 survey; 0 if false.
INTERESTSCHL	=	1 if true that "I am interested in school" in 1980 survey; 0 if false.
SUSPORPROBATN	=	1 if true that "I have been suspended or put on probation in school"; 0 if false.
CUTCLASS	=	1 if true that "Every once in a while I cut a class"; 0 if false.
NOTSAFESCHL	=	1 if true that "I don't feel safe at this school"; 0 if false.
EXPSCHLASCH	=	level of schooling respondent expects to achieve, as of 1980 survey.

[Tables 1 and 2 about here]

Readers who are familiar with U.S. data on youth employment may wonder why the unemployment rates in Tables 1 and 2 are so low. One reason is that respondents for whom data are available on all variables in Tables 1 and 2 have relatively low unemployment rates. There were 1,244 dropouts and 3,239 graduates for whom we can compute post-high school unemployment and wages, but only 439 dropouts and 1,690 graduates with complete data in Tables 1 and 2. For the larger group of 1,244 dropouts, the unemployment rate was 19.7 percent, and for the whole group of 3,239 graduates it was 13.8 percent. These are closer to the unemployment rates usually reported for this age group.

Tables 3 and 4 show results of ordinary least-squares regressions for unemployment and hourly wages in the pooled sample of dropouts and graduates. The coefficient on GRADUATE is an estimate of the effect of a high school diploma. Controlling for the other predictors, graduates experience an unemployment rate 2.5 percentage points less than dropouts, and graduates earn $0.23 an hour more than dropouts when they are employed. Morgan (1981) obtained comparable results using the same statistical procedure with data from the National Longitudinal Survey of Youth Labor Market Experience. However, as explained above, these estimates are not valid if there are unmeasured variables which affect both the probability of dropping out and the degree of success in the labor market.

[Tables 3 and 4 abut here]

To correct for possible bias due to unmeasured influences, the first stage is estimation of a probit equation predicting graduation from high school, using data available in 1980, before any sample members dropped out. Results of the probit analysis are in Table 5. As mentioned above, all predictors in this equation consist of residuals from regressions where each predictor was itself regressed on the set of "x" variables used to predict labor market success. The point was to ensure that the predictors in the probit equation are not correlated with the "x" predictors, so that the "x" predictors do not enter the probit analysis by the back door. Significant predictors of high school graduation in Table 5 are EXPGRAD, GRADES, ABSENT, DISCPROBS, SUSPORPROBATN, and CUTCLASS. All have the signs we would expect.

[Table 5 about here]

Tables 6 through 9 show regression equations for predicting labor market outcomes. The ordinary least-squares results are estimates of equations (1) and (2) above, and the generalized least-squares results are estimates of equations (3) and (4). The selection term used in the generalized least-squares equation is a function of the predicted probability of graduating from high school, as estimated by the probit equation in Table 5 (see Maddala, 1983, p.

224). In Tables 6 through 9, the coefficient on the selection term is significant only in Table 8. This particular coefficient implies that individuals who, as high school graduates, would have higher than predicted unemployment (as predicted by the observed "x" characteristics), are also less likely to graduate. Likewise, the coefficient on the selection term in Table 6, though not quite statistically significant, implies that individuals who, as dropouts, would have higher than predicted unemployment are more likely to graduate. Both coefficients indicate some degree of rational self-selection.

[Tables 6, 7, 8, 9 about here]

For comparison with results using HSB data, we also analyzed data from the National Longitudinal Survey of Youth Labor Market Experience (NLS). This survey began in 1979 with a nationally representative sample of 11,406 individuals age 14 to 22 (Center for Human Resource Research, 1983). Follow-up interviews have been held in each succeeding year. For our analysis, we had data from 1979 through 1982. We focused on 1979 seniors who graduated from high school but did not attend any postsecondary school, and on members of the same 12-month age cohort who had not received a high school diploma or equivalent as of 1982. These two samples thus parallel the samples of dropouts and graduates from the HSB data. Unlike the HSB survey, however, the NLS did not collect much information about experience in school. Therefore, we could not use the NLS data to estimate a probit equation to predict dropping out, as we did in Table 5 with the HSB data. Consequently, the NLS data permit us to estimate equations (1) and (2) above, but not equations (3) and (4), which require the first-stage probit results. Tables 10 and 11 show the estimates of equations (1) and (2) using NLS data. These can be compared directly with the ordinary least-squares results from HSB data in Tables 6 through 9.

[Tables 10 and 11 about here]

When we use the ordinary least-squares regression results in Tables 6-11 to compute how much of the difference between the labor market experience of graduates and dropouts is due to differences in measured characteristics and how much is due to differences in the way these characteristics are rewarded in the job market, we get the results in Table 12. Results with HSB and NLS data are qualitatively very similar, although the differences in labor market success between graduates and dropouts are bigger in the NLS data. The qualitative similarity is that most or all of the difference between graduates and dropouts is attributable to the coefficient effect -- differences in how measured characteristics are translated into labor market success -- rather than to differences in these measured characteristics (the "x" variables) themselves.

[Table 12 about here]

Table 12 partitions the observed difference between the labor market experience of graduates and dropouts into two components, which correspond to the first two lines on the right-hand side of equation (5). To obtain the full three-way partitioning of equation (5), we are limited to the HSB data. Table 13 shows the results. With respect to unemployment, the coefficient effect is still dominant, even more than in Table 12. However, with respect to wages, the signs in Table 13 are the opposite of those in Table 12, and the difference in labor market success between graduates and dropouts seems entirely attributable to differences in measured and unmeasured characteristics.

Discussion

Most of our results indicate that the higher unemployment rates and lower hourly earnings of teenage dropouts compared to high school graduates are not attributable to prior differences between dropouts and graduates in race or ethnicity, sex, family socioeconomic status, absence of father, test scores, work experience while in high school, or local labor market conditions. The only di ergent result is in the last line of Table 13, based on the Heckman-Maddala procedure for estimating the influence of unmeasured variables. However, although this procedure is the best yet developed to correct for unmeasured variables when direct experimentation is not possible, it has not been found to be very accurate by LaLonde (1986), who compared it to results obtained from a true experiment. We conclude, therefore, that the labor market difficulties of teenage dropouts are for the most part not merely symptomatic of the prior characterisitcs just mentioned, but whether they are due to discrimination by emloyers against dropouts or to unmeasured characteristics of dropouts remains an open question.

From the standpoint of an individual student deciding whether or not to finish high school, it does not matter whether employers discriminate against dropouts in a purely arbitrary fashion, or whether employers favor graduates because they believe correctly that lack of a diploma signals a lack of productive virtues such as reliability, perseverance, self-discipline, and obedience to authority. In either case, the individual student can improve his or her prospective employment and earnings by finishing high school. However, the gain to society does depend on whether the reasons why employers favor graduates are arbitrary or valid. If discrimination against dropouts is purely arbitrary, then inducing a would-be dropout to finish high school is equivalent to increasing the effective supply of labor, which leads to an unambiguous increase in economic output and well-being (provided that demand for labor is not perfectly inelastic). On the other hand, if dropouts on average really have unobserved characteristics that make them less productive than graduates, then the gain to society from inducing would-be dropouts to graduate depends on whether the would-be dropouts acquire more productive characteristics in the process of finishing high school.

REFERENCES

Bachman, J. G., Green, S., and Wirtanen, I. D. (1971). Youth In Transition (Vol. III). Ann Arbor, Michigan: Institute for Social Research, The University of Michigan.

Catterall, J. S. (1985). On the Social Costs of Dropping Out of High School. (Report 86-SEPI-3), Stanford, CA: Stanford Education Policy Institute, Stanford University.

Center for Human Resource Research (1983). *NLS Handbook.* Columbus, OH: Ohio State University.

Combs, J. and Cooley, W. W. (1968). Dropouts: In High School and After School. American Educational Research Journal, 5 (3), 343-363.

Garen, J. (1984). The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable. Econometrica, 52 (5), 1199-1218.

Hanushek, E. A. and Jackson, J. E. (1977). Statistical Methods for Social Scientists. New York: Academy Press.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. Econometrica, 47(1), 153-161.

Jencks, C. S., Smith, M., Ackland, H., B ne, M. J., Cohen, D., Gintis, H., Heyns, G., and Michelson, S. (1972). Inequality: A Reassessment of the Effect of Family and Schooling in America. New York: Basic Books.

Jones, C., Sebring, P., Crawford, I., Spencer, B., and Butz, M. (1984). High School and Beyond 1980 Senior Cohort. Data File User's Manual. Base year through Second Follow-up. Washington, D.C.: Office of Educational Research and Improvement, U.S. Department of Education.

Kiker, B. F. and Condon, C. M. (1981). The Influence of Socioeconomic Background on the Earnings of Young Man. Journal of Human Resources, 16 (1), 94-105.

Levin, H. M. (1972). The Costs to the Nation of Inadequate Education. (Study prepared for the Select Committee on Equal Educational Opportunity, U.S. Senate.) Washington, D.C.: U.S. Government Printing Office.

LaLonde, R. J. (1986). Evaluating the Econometric Evaluation of Training Program with Experimental Data. American Economic Review, 76 (4), 604-620.

Maddala, G. S. (1983). Limited-Dependent and Qualitative Variables in Econometrics. New York, NY: Cambridge University Press.

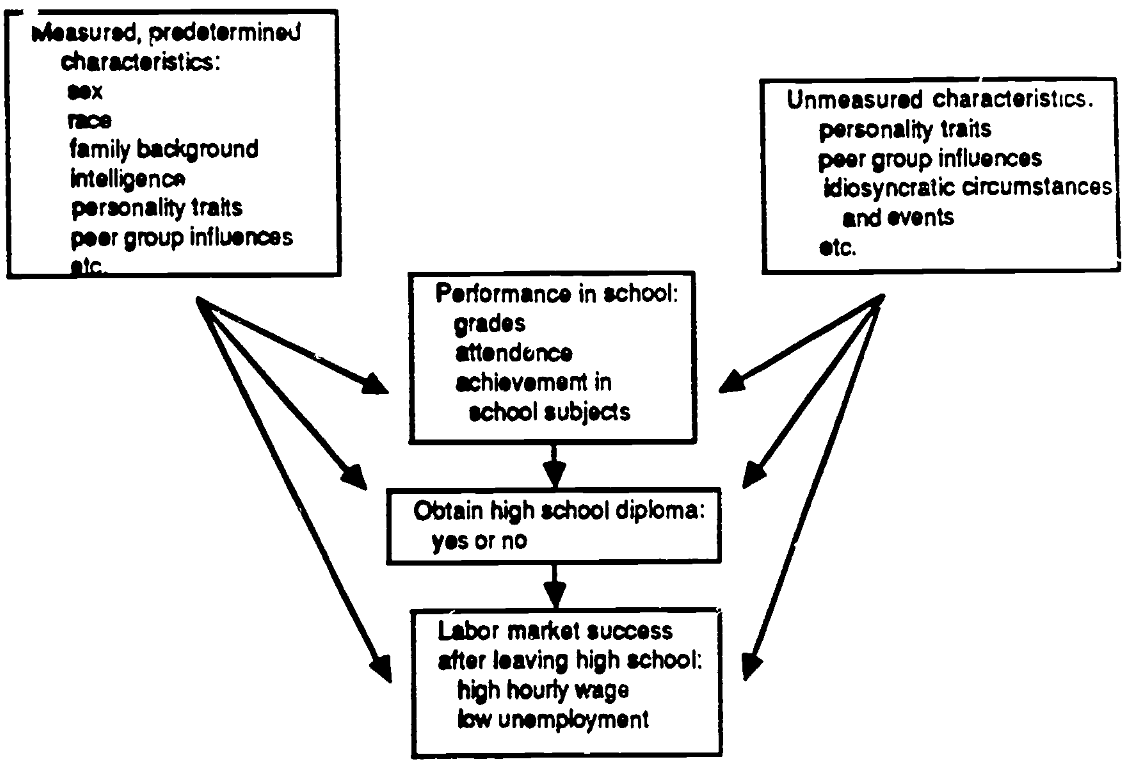
Morgan, W. R. (1981). The High School Dropout in an Overeducated Society. In Pathways to the Future, Worthington, Ohio: Ohio State University Center for Human Resource Research, 215-275.

Sewell, W. H. and Hauser, R. M. (1975). Education, Occupation, and Earnings. New York: Academic Press.

Wehlage, G. G. (1983). Effective Programs for the Marginal High School Student. Bloomington, Indiana: Phi Delta Kappa Educational Foundation.

Willis, R. J. and Rosen, S. (1979). Education and Self-Selection. Journal of Political Economy, 87(5), S7-S36.

FIGURE 1



Effect of High School Diplcmas

Table 1

Means and standard deviations of variables for dropouts (N=439), HSB data

Variable	Mean	Std Dev
GRADUATE	0.00	0.00
UNEMP	0.12	0.19
WAGE	3.74	2.06
MALE	0.56	0.50
SES	-0.48	0.65
TEST	43.69	6.24
NOMAN	0.28	0.45
HISPANIC	0.25	0.43
BLACK	0.10	0.30
WHITE	0.63	0.48
SCHJOBHRS	28.10	13.74
CUNEMR82	10.30	3.77
CWAGE82	7.82	1.59
EXPGRAD	0.72	0.45
DEFINJOB	0.31	0.46
GRADES	71.43	7.77
ABSENT	7.01	6.80
LATE	4.13	5.32
RESWAGEBEF	2.79	0.67
RESWAGEAFT	3.55	0.52
EXTRACURR	8.84	2.29
OWNVALUGRADES	0.31	0.46
FRNDVALUGRADES	0.27	0.44
SCHLSATIS	0.56	0.50
DISCPROBS	0.46	0.50
INTERESTSCHL	0.46	0.50
SUSPORPROBATN	0.32	0.47
CUTCLASS	0.58	0.49
NOTSAFESCHL	0.20	0.40
EXPSCHLACH	12.15	1.74

Table 2

Means and standard deviations of variables for graduates (N=1690), HSB data

Variable	Mean	Std Dev
GRADUATE	1.00	0.00
UNEMP	0.10	0.18
WAGE	3.93	1.76
MALE	0.51	0.50
SES	-0.31	0.62
TEST	48.67	7.30
NOMAN	0.17	0.38
HISPANIC	0.24	0.42
BLACK	0.07	0.26
WHITE	0.66	0.47
SCHJOBHRS	17.53	12.96
CUNEMR82	10.40	3.73
CWAGE82	7.84	1.46
EXPGRAD	0.95	0.23
DEFINJOB	0.20	0.40
GRADES	77.65	7.59
ABSENT	2.79	3.95
LATE	2.27	3.96
RESWAGEBEF	2.56	0.73
RESWAGEAFT	3.45	0.59
EXTRACURR	8.52	2.19
RFEELGD	3.35	0.48
FFEELGD	0.29	0.45
SCHLSATIS	0.36	0.48
DISCPROBS	0.21	0.41
INTERESTSCHL	0.26	0.44
SUSPORPROBATN	0.13	0.33
CUTCLASS	0.27	0.44
NOTSAFESCHL	0.12	0.33
EXPSCHLACH	12.90	1.91

Table 3

Ordinary least-squares regression for unemployment,
in pooled sample of dropouts and graduates, HSB data

Variable	Coefficient	t-statistic
MALE	-0.0137	-1.768
TEST	0.0004	0.752
SES	-0.0303	-4.635
NOMAN	0.0098	0.973
SCHJOBHRS	-0.0013	-4.406
CUNEMR82	0.0044	4.136
CWAGE82	0.0059	2.206
HISPANIC	-0.0324	-1.391
BLACK	0.0235	0.903
WHITE	-0.0270	-1.202
GRADUATE	-0.0252	-2.450
INTERCEPT	0.0555	1.327
R^2	0.051	
N	2129	

Table 4

Ordinary least-squares regression for hourly wage, in pooled sample of dropouts and graduates, HSB data

Variable	Coefficient	t-statistic
MALE	0.4639	5.826
TEST	-0.0000	-0.011
SES	0.2891	4.306
NOMAN	-0.0281	-0.272
SCHJOBHRS	0.0067	2.202
CUNEMR82	0.0124	1.126
CWAGE82	-0.0012	-0.044
HISPANIC	0.4703	1.966
BLACK	0.2170	0.812
WHITE	0.2268	0.993
GRADUATE	0.2316	2.196
INTERCEPT	3.0448	7.094
R^2	0.037	
N	2129	

Table 5

Probit analysis predicting graduation from high school, HSB data

Variable	Coefficient	t-statistic
EXPGRAD	.4667	4.349
DEFINJOB	-.0484	-.592
GRADES	.0199	3.777
ABSENT	-.0404	-5.753
LATE	.0054	.682
RESWAGEBEF	-.0501	-.911
RESWAGEAFT	-0.169	-2.55
EXTRACURR	-.0192	-1.256
OWNVALUGRADES	-.1982	-1.828
FRNDVALUGRADES	-.0044	-.040
SCHLSATIS	-.0615	-.855
DISCPROBS	-.1513	-1.901
INTERESTSCHL	-.0939	-1.249
SUSPORPROBATN	-.1742	-1.925
CUTCLASS	-.2889	-3.853
NOTSAFESCHL	-.0247	-.261
EXPSCHLACH	.0348	1.822
INTERCEPT	5.8982	177.165
CHI-SQUARE	2222.401	
DEGREES OF FREEDOM	2111	

Table 6

Ordinary and generalized least-squares regressions for predicting unemployment among high school dropouts, HSB data

Variable	Ordinary least-squares		Generalized least-squares	
	Coefficient	t-statistic	Coefficient	t-statistic
MALE	0.0176	0.933	0.0179	0.950
TEST	0.0035	2.268	0.0033	2.101
SES	-0.0296	-1.933	-0.0293	-1.916
NOMAN	0.0293	1.379	0.0317	1.483
SCHJOBHRS	-0.0013	-1.812	-0.0012	-1.665
CUNEMR82	0.0023	0.937	0.0025	0.998
CWAGE82	0.0031	0.523	0.0033	0.553
HISPANIC	-0.0145	-0.206	-0.0122	-0.174
BLACK	0.0277	0.373	0.0353	0.476
WHITE	0.0022	0.032	0.0023	0.033
INTERCEPT	-0.0835	-0.756	-0.0083	-0.387
SELECTION TERM	------	------	-0.0313	-1.457
R^2	.0481		------	
N	439		439	

Table 7

Ordinary and generalized least-squares regressions for predicting hourly wages among high school dropouts, HSB data

	Ordinary least-squares		Generalized least-squares	
Variable	Coefficient	t-statistic	Coefficient	t-statistic
MALE	0.3731	1.854	0.3754	1.865
TEST	-0.0228	-1.371	-0.0249	-1.493
SES	0.4121	2.526	0.4193	2.568
NOMAN	-0.2443	-1.077	-0.2225	-0.977
SCHJOBHRS	0.0038	0.513	0.0045	0.608
CUNEMR82	0.0613	2.311	0.0629	2.370
CWAGE82	-0.0249	-0.396	-0.0245	-0.389
HISPANIC	-0.0538	-0.072	-0.0404	-0.054
BLACK	0.0869	0.110	0.1479	0.187
WHITE	-0.3828	-0.522	-0.3898	-0.532
INTERCEPT	4.5019	3.822	9.7718	4.016
SELECTION TERM	-------	-------	-0.1766	-0.768
R^2	.0559		-------	
N	439		439	

Table 8

Ordinary and generalized least-squares regressions for predicting unemployment among high school graduates, HSB data

Variable	Ordinary least-squares Coefficient	Ordinary least-squares t-statistic	Generalized least-squares Coefficient	Generalized least-squares t-statistics
MALE	-0.9219	-2.588	-0.0217	-2.549
TEST	-0.0001	-0.234	0.0001	0.170
SES	-0.0302	-4.184	-0.0293	-4.092
NOMAN	0.0021	0.182	0.0065	0.564
SCHJOBHRS	-0.0013	-3.955	-0.0011	-3.409
CUNEMR82	0.0049	4.113	0.0052	4.375
CWAGE82	0.0070	2.290	0.0084	2.940
HISPANIC	-0.0283	-1.157	-0.0173	-0.744
BLACK	0.0319	1.146	0.0510	1.921
WHITE	-0.0287	-1.221	-0.0173	-0.768
INTERCEPT	0.0503	1.119	-0.0042	-0.587
SELECTION TERM	------	-------	-0.1027	-2.865
R^2	0.0597		-------	
N	1690		1690	

Table 9

Ordinary and generalized least-squares regressions for predicting hourly wages among high school graduates, HSB data

	Ordinary least-squares		Generalized least-squares	
Variable	Coefficient	t-statistic	Coefficient	t-statistic
MALE	0.4807	5.598	0.4824	5.615
TEST	0.0043	0.704	0.0037	0.594
SES	0.2476	3.387	0.2420	3.308
NOMAN	0.0591	0.510	0.0544	0.468
SCHJOBHRS	0.0071	2.143	0.0077	2.312
CUNEMR82	-0.0011	-0.093	-0.0016	-0.132
CWAGE82	0.0091	0.297	0.0099	0.322
HISPANIC	0.5167	2.083	0.5213	2.100
BLACK	0.1339	0.475	0.1357	0.481
WHITE	0.3213	1.350	0.3273	1.374
INTERCEPT	3.0129	6.610	5.1706	6.300
SELECTION TERM	-----	-----	-0.1635	-0.672
R^2	0.0384		-----	
N	1690		1690	

Table 10

Ordinary least-squares regressions for predicitng unemployment and hourly wages among high school dropouts, NLS data

Variable	Predicting unemployment Coefficient	Predicting unemployment t-statistic	Predicting hourly wage Coefficient	Predicting hourly wage t-statistic
NOMAN	0.0290	0.548	-0.0692	-0.228
ED	0.0026	0.366	-0.0377	-0.922
SCHJOBHRS	0.0008	0.500	-0.1036	-1.462
MALE	0.0428	1.005	1.0978	4.484
BLACK	0.1948	3.285	-0.4403	-1.292
HISP	-0.0111	-0.185	0.5240	1.518
ASVAB SCORES:				
QUANT	0.0160	0.268	-0.4013	-1.607
VERBAL	0.0103	0.296	-0.1895	-0.949
CODING	0.0010	0.043	-0.0635	-0.478
AUTO	-0.0239	-1.039	-0.0359	-0.272
MECH	-0.0070	-0.315	0.0971	0.763
ELEC	-0.0202	-0.783	-0.0924	-0.624
INTERCEPT	0.1396	1.289	3.6377	5.851
R^2	0.1177		0.2152	
N	170		170	

Table 11

Ordinary least-squares regressions for predicting unemployment and hourly wages among high school graduates, NLS data

Variable	Predicting unemployment Coefficient	Predicting unemployment t-statistic	Predicting hourly wage Coefficient	Predicting hourly wage t-statistic
NOMAN	0.0424	1.113	-0.5710	-1.559
ED	-0.0107	-1.845	0.0944	1.701
SCHJOBHRS	-0.0010	-0.909	0.01223	1.122
MALE	0.0832	2.993	0.5079	1.899
BLACK	0.1111	3.211	0.0014	0.004
HISP	-0.0450	-0.964	0.2934	0.653
ASVAB SCORES:				
QUANT	-0.0116	-0.410	-0.3695	-1.358
VERBAL	0.0205	0.911	-0.4529	-2.098
CODING	0.0063	0.380	0.2690	1.686
AUTO	-0.0224	-1.310	0.1703	1.038
MECH	0.0087	0.524	0.1161	0.729
ELEC	0.0073	0.410	0.0266	0.156
INTERCEPT	0.2225	2.795	2.4155	3.156
R^2	0.1113		0.1063	
N	264		264	

Table 12

Differences between high school graduates and dropouts in unemployment and hourly wages, attributed to differences in measured characteristics and regression coefficients, with dropouts and graduates treated as separate populations, HSB and NLS data

	GRADUATES MEAN	DROPOUTS MEAN	DIFFERENCE IN MEANS	DIFFERENCE ATTRIBUTED TO: CHARACTS.	COEFFICIENTS
HSB data (1982-84 outcomes)					
Unemployment rate	0.101	0.116	-0.015	0.015	-0.303
Hourly wage	3.930	3.744	0.186	-0.061	0.247
NLS data (1980-82 outcomes)					
Unemployment rate	0.163	0.265	-0.102	0.001	-0.102
Hourly wage	3.829	2.875	0.954	0.154	0.800

Table 13

Differences between high school graduates and dropouts in unemployment and hourly wages, attributed to differences in measured characteristics, unmeasured characteristics, and differences in regression coefficients, with dropouts and graduates treated as members of the same population, HSB data

	GRADUATES MEAN	DROPOUTS MEAN	DIFFERENCE IN MEANS	DIFFERENCE ATTRIBUTABLE TO: MEASURED CHARACTS.	COEFFICIENTS	UNMEASURED CHARACTS.
UNEMPLOYMENT RATE	0.103	0.117	-0.014	0.012	-0.090	0.064
HOURLY WAGE	3.930	3.745	0.185	0.581	-0.662	0.266