DOCUMENT RESUME

ED 116 102 CG 010 272

AUTHOR	Hill, C. Russell
TITLE	Dropping Out of High School: The Effects of Family, Ability, School Quality and Local Employment Conditions.
SPONS AGENCY	Department of Health, Education, and Welfare, Washington, D.C. Office of the Assistant Secretary for Planning and Evaluation.
PUB DATE	[75]
NOTE	45p.
EDRS PRICE	MF-$0.76 HC-$1.95 Plus Postage
DESCRIPTORS	*Aptitude; Critical Path Method; *Dropout Identification; Dropout Research; Family (Sociological Unit); Intelligence; *Labor Market; Multiple Regression Analysis; *Racial Differences; *School Environment; Secondary Education; *Socioeconomic Influences
IDENTIFIERS	National Longitudinal Surveying of Young Men

ABSTRACT

The author, after reviewing briefly the research literature on dropouts, mentions the following shortcomings of such studies: (1) lack of ability to control for differences in the students' abilities; (2) incomplete and inadequate measures of socioeconomic background; (3) inability to control for qualitative differences in the schools attended; and (4) inability to control for the effects of local market conditions on the decision to leave high school. The author used the National Longitudinal Survey of Young Men (NLSYM) data to control for these four factors and find the contribution of each to dropout rates (D). A theoretical model incorporating the four factors and others is explored, and seen as accounting for the reasons given for dropping out. Main findings from this study are: (1) IQ has the most important direct effect on D for whites, while knowledge of the labor market is the most important for non-whites; (2) the quality of high school attended does not have a direct affect on D for either racial group; and (3) parental education level is the only socioeconomic index affecting D rates for nonwhites, while other indexes affect D of whites. Based on these results, some policy implications for dropouts are discussed. (SE)

ED116102

DROPPING OUT OF HIGH SCHOOL: THE EFFECTS OF FAMILY, ABILITY, SCHOOL QUALITY AND LOCAL EMPLOYMENT CONDITIONS

by

C. Russell Hill*

* The research reported here was supported by the Office of the Assistant Secretary for Planning and Evaluation, Department of Health, Education, and Welfare during the author's tenure as a Brookings Institution Economic Policy Fellow. The analysis and conclusions expressed here are the author's own and do not necessarily reflect the positions or policy of either DHEW or the Brookings Institution.

Dropping Out of High School:
The Effects of Family, Ability, School Quality and Local Employment Conditions

I. INTRODUCTION

The major economic decision faced by young people enrolled in high school is whether to continue with the formal educational process. The primary alternatives to schooling are full-time participation in the labor force, service in the Armed Forces, marriage and work within a household. During the high school years these alternatives are chosen in varying degrees such that there is essentially a steady decline in the proportion of a schooling cohort that remains enrolled in the formal educational process. Recent data (U.S. Department of HEW, 1974, p. 14), for example, indicate that of those individuals entering the fifth grade in 1964, 97 percent reached the ninth grade, 87 percent entered the eleventh grade and only 75 percent graduated from high school in 1972. These data indicate that decisions are continually being made during these formative years as to whether to continue in school. Economists have found it both convenient and revealing to think of these decisions as investment decisions made by the student, his family and society. This paper will adopt that framework in looking at several of the factors that influence the decision of white and nonwhite male youth to continue or drop out of high school.

Of the several, sequential educational decisions made by youth, the decision to drop out of high school has seemed to arouse the greatest public and private concern. While there is now some evidence and concern regarding an overinvestment in college training (Freeman, 1975), the fact that young people drop out of high school usually raises the spectre of increased crime, drug usage, unemployment and a general alienation of youth from the adult community. The public result of this concern has been a variety of dropout prevention programs which have been supported under Title VIII of the Elementary and Secondary Education Act together with a provision of the Vocational Education Act which directs Federal monies to areas of high need, including local areas with a high concentration of school dropouts. A variety of instructional methods have been supported in the dropout prevention programs with work-training components, reading and math laboratories, and social adjustment laboratories being relatively common approaches to the problem.[1/] A recent example of an attempt to reform the educational process in order to make it more relevant to high school age youth is the development of "career education" programs of instruction (Marland, 1975). One of the purposes of this innovation is to incorporate instruction on the world of work into the primary and secondary education curriculum so as to both encourage youth to stay in school or, if they leave before graduation, increase their ability to cope with the problems of entering the labor force.[2/]

Knowledge of the causes and consequences of dropping out of high school was, of course, an important component of the design and evaluation

---

1/ See Weisbord (1965) for a pessimistic evaluation of the cost effectiveness of one of these programs.

2/ The goals and purposes of career education have been set forth in recent Federal legislation, viz, Education Amendments of 1974, (P.L. 93-380), Sec. 406(a) and (b).

of dropout prevention programs. Several excellent studies have been completed on this problem, among which could be cited the work of B. Duncan (1965), Conlisk (1969), Masters (1969), Bachman, et. al. (1971), Lerman (1972), Levin (1972) and Edwards (1975). Unfortunately, of those studies which looked specifically at the socioeconomic determinants of dropping out of high school, each used a data source that was deficient in at least one of the following respects: (1) inability to control for differences in the students' ability, (2) incomplete and inadequate measures of socioeconomic background, (3) inability to control for qualitative differences in the schools attended, and (4) inability to control for the effects of local labor market conditions on the decision to leave high school. This present study uses a data source that seems uniquely suited for a study of high school dropouts--the National Longitudinal Survey of Young Men (NLSYM).[3] These data not only contain a joint distribution of the several variables mentioned above which were missing in prior studies, but also have the advantage of being longitudinal rather than cross-sectional.

Section II of this paper presents a review of the theory of educational attainment treating formal schooling as a process of human capital acquisition. Section III presents the recursive model estimated here together with a discussion of some of the problems of measurement and estimation encountered. Section IV presents and discusses the empirical results and Section V contains a discussion of the policy implications implied by the analysis.

---

[3] For a description of the data, see Parnes, et. al. (1970).

II. THE PROCESS OF EDUCATIONAL ATTAINMENT

A simple model which explains interpersonal differences in investments in formal schooling has been developed by Becker (1967) with more formal extensions of this model having been presented by Ben-Porath (1967) and Wallace and Ihnen (1975). Becker's formulation is particularly useful for it leads to the specification of a simple recursive model whose parameters may be estimated using the NLSYM.

Basically, the model views the individual as maximizing the present value of his net earnings over the life cycle by investing in formal schooling up to the point at which the marginal rate of return from the investment equals the marginal financing costs. The returns from the investment in schooling are the product of two factors: the expectation of returns from a particular level of schooling achievement and the probability that the particular individual will in fact succeed in attaining this level.[4] The first factor is largely determined by the exogenous forces of the labor market where individual differences arise because of imperfect knowledge of the labor market. The second is largely a function of individual capacities (ability), the schooling environment, and the extent to which the individual believes the schooling environment and curriculum will actually lead to an increase in his stock of human capital. The private costs of investing in formal schooling are largely time costs, particularly at the high school level

---

4/ This useful distinction is taken from Griliches (1973).

which is largely supported by public revenues. The ability of a family and the student to finance these opportunity costs incurred by the student member depend upon some measure of family wealth with a student's foregone earnings making up a potentially large portion of many families' potential income (Solmon, 1970).5/

A convenient way of picturing the process described above is through the use of Figure I which shows the demand and supply curves for investments in schooling. The demand function shows an inverse relation between marginal returns and investments and is largely a function of ability as well as the schooling and other environmental influences which affect returns discussed above. The supply curve illustrates the marginal cost of financing an additional unit of capital and is dependent upon opportunity factors which could presumably be measured by variables reflecting family wealth and/or socioeconomic background.6/ The intersection of the two schedules, Do and So, defines an equilibrium where the optimal investment in formal schooling is Io.7/ The distribution of earnings and investments would both be more unequal and skewed the greater the elasticities of supply and demand and the more

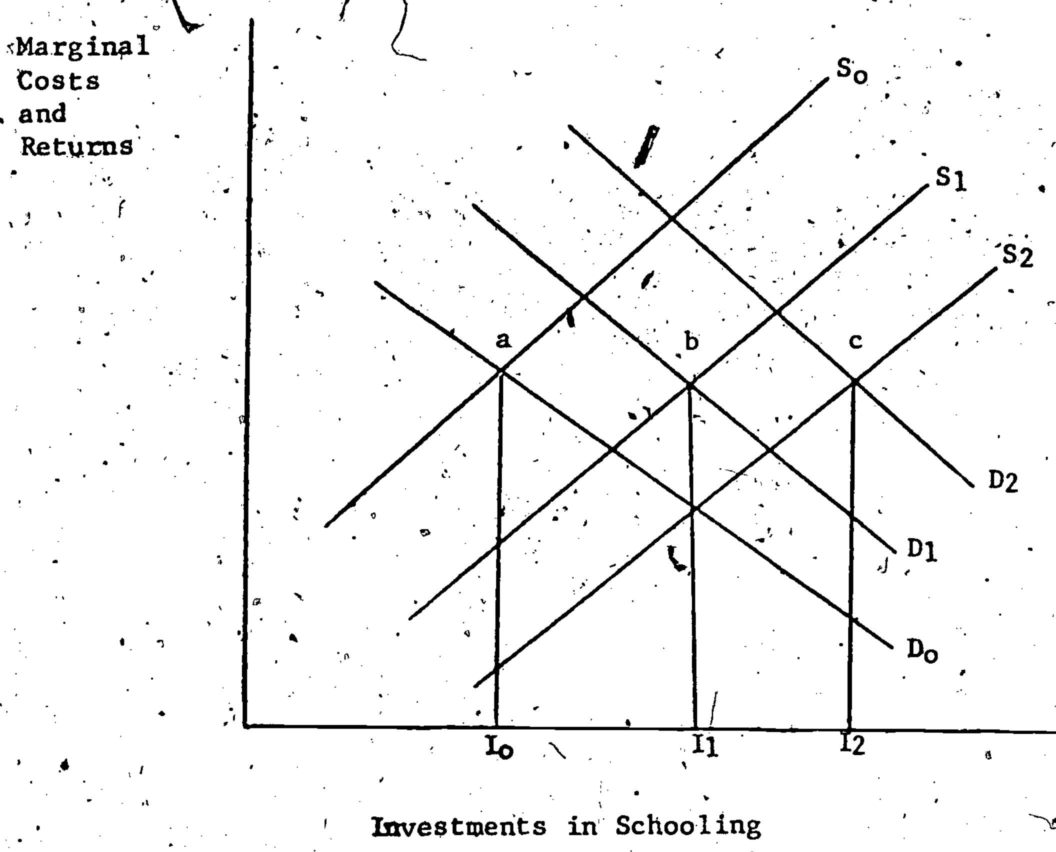
5/ If opportunities and abilities to finance investments in formal schooling differ among families, then the objective of maximizing the present value of net earnings in the formal model becomes distinct from maximizing the present value of utility. While important in the more formal models, the distinction is not critical for our purposes (Wallace and Ihnen, 1975, pp. 138-139).

6/ The factors affecting the slope of the demand curve (increased time costs of additional schooling, imperfect substitution between own time and other inputs, etc.) and the supply curve (segmented capital market) are discussed in detail in Becker (1967), Mincer (1970) and Ben-Porath (1967).

7/ The conditions under which the equilibrium is uniquely defined are discussed in Becker (1967, pp. 10-11).

Figure I

The Determinants of Optimal Educational Attainment



unequal and skewed are their distributions. In addition, however, the distribution of earnings and investments depend upon the correlation between different curves. If, for example, more favorable demand conditions (as reflected, for example, in the curve D_2 relative to D_1) are associated with more favorable supply conditions (S_2 relative to S_1) equilibrium positives such as a, b, and c would result. Consequently, if ability and social background are positively correlated, those with the greatest opportunities will have the greatest capacity to profit from schooling and, consequently, they may or may not have higher marginal returns but they will invest more.[8]

The probability that an individual will drop out of high school before completion may usefully be examined within the framework sketched here. This probability is conditional on what Becker has termed "capacities" and "opportunity" factors. The capacities include both inherited ability as well as those capacities determined by school quality and the curriculum in which the individual is enrolled. However, saying that some capacities are inherited is simply a restatement of the importance of the correlation between supply and demand elements. Individual ability, as measured by standardized IQ tests, is known to be subject to several of the environmental influences which are subsumed in the opportunity factors (Hill and Stafford, 1974). Consequently, a picture emerges of a recursive framework which affects the conditional probability of a

8/ The effects of the correlation of ability and opportunity factors on the dispersion of estimated rates of return is discussed in Wachtel (1975).

young man dropping out of high school: environmental home factors, directly affect the opportunities to invest in additional schooling and are reflected in shifts in the marginal finance cost (supply) curve. Second, as envisioned in Becker's formulation, the environmental factors <u>indirectly</u> affect the potential returns from the continuance of formal schooling by affecting the ability of the individual to profit from the investment. The next section of this paper will formalize this structure of the interdependence of opportunities and capacities, introduce intervening variables which will attempt to reflect both the importance of the individual's impression of labor market conditions and actual labor market conditions in his investment decision and culminate in a recursive model which will evaluate the relative importance of the direct and indirect factors affecting the probability that a male youth will drop out of high school.

III. THE MODEL AND DATA

Presently, the NLSYM provides data for the years 1966 through 1969 for a cohort of young men between the ages of 14 and 24. The sample chosen for study here consists of those individuals between the ages of 14 and 17 who were enrolled in high school in 1966 while the time at which the probability of dropping-out will be measured is two years later when the youth are between the ages of 16 and 19. At this time all of the individuals in the sample had some high school experience and to the extent that measures of school quality and/or curriculum choice have an effect on dropping-out, it should be most pronounced for this age group of young men. The dependent variable of prime interest for this study is defined, then, as follows: Dropout (D) = 1 if an individual between the ages of 16 and 19 has completed less than 12 years of formal schooling and is not enrolled in school in 1968. The complement of this set consists of both those individuals between 16 and 19 who (1) are still enrolled in high school and (2) those who have graduated and may or may not be enrolled in a formal postsecondary course sequence. If our purpose here was to undertake an earnings and employment comparison, it would be necessary to carefully distinguish among the behavior of these distinct groups. However, as our purpose here is only to analyze the factors which lead a young man to leave high school before completion, the simple dichotomous variable (D) seems appropriate.

The formal model of dropping-out to be developed here has two main goals: (1) to as closely as possible reflect the investment-theoretic model of Becker (1967), paying particular attention to the simultaneous effects of abilities and opportunities on the level of schooling completed and, (2) to incorporate variables which will approximate the effects of two programs of current Federal educational policy interest (career education and vocational education) into the model.

The probability that an individual will leave high school before graduation is conditional upon several factors which may be grouped as follows: (1) the student's socioeconomic background (SEB), (2) the student's ability to profit from the investment in schooling which may be measured, albeit imperfectly, by the student's IQ, (3) variables which reflect the quality (Q) of the high school attended and the curriculum (C) in which he is enrolled and, finally, (4) variables reflecting knowledge of the labor market and local employment conditions. Formally, the model may be structured as a causal system of four equations incorporating the factors mentioned above so as to measure their direct and indirect effects on D.

The first equation of the model will attempt to explain a student's IQ on the basis of his SEB and the quality of high school attended (Q). The form of this equation is due in large part to Griliches (1970) and variants of this equation have been recently estimated by Griliches and Mason (1972) and Ribich and Murphy (1975). In addition to the well-known problems of measuring "ability" by an IQ test (see, e.g., Griliches 1970), the student's home background, as we are able to measure it, only imperfectly captures the genetic and environmental influences which affect IQ. In this paper SEB is measured by the following variables: father's educational attainment (EF), mother's educational attainment (EM), number of the student's siblings (S) and the three year (1966-68) average income ($\overline{Y}$) of the student's family.[9/] While this list of variables is an improvement over those available in previous studies of the dropout decision, it by no means provides a totally informative insight into the process by which home background affects ability, educational attainment and, subsequently, lifetime earnings.[10/] Hill and Stafford (1974) have provided evidence, however, that parental educational attainment is positively related and family size is negatively related to the amount of time devoted by the parents to the physical care of the children in the preschool years. And there is some evidence from child development studies that this parental time positively affects the cognitive and affective abilities of the child. Family income ($\overline{Y}$) and family size are also direct measures of Becker's opportunity factor as they reflect the family's ability to finance investments in formal schooling for each of its children. These SEB variables may have a direct effect on D and, to the extent that the hypothesized positive correlation between supply and demand elements is present, they will have an indirect effect on D through IQ. Finally, the effect of family income on the dropout decision deals with questions of equality of educational opportunity which have been the focus of previous studies of this problem (see, e.g., Masters, 1969).

The quality of high school attended (Q) by the individual may also have an independent effect on his IQ for we know that intelligence is

---

9/ The exact definition and measurement scale of these and the other variables in the model is contained in Appendix Table A.1.

10/ Individual SEB variables rather than a socioeconomic status index (also included in the NLSYM) were used in the analysis because use of the index hides a considerable amount of interesting detail (Hauser, 1972).

not fixed and independent of schooling or other environmental influences. While school quality has recently become widely used in earnings regressions (Johnson and Stafford, 1973), it is typically measured by expenditures per-pupil in the district or State in which the student is enrolled. Consequently, there may be an aggregation bias in these estimates of the effect of school quality due to lack of data on a school-specific quality measure. The quality measure used here, however, is an ordinal index of the characteristics of the particular high school attended by the young men in our sample. The four elements upon which the construction of the index is based are (1) per-pupil availability of library facilities, (2) pupils per full-time teachers, (3) full-time equivalent counselors per 100 pupils and (4) annual salary of a beginning teacher with a bachelor's degree and no experience, adjusted for geographic price level differences.

The second equation in the causal model explains an individual's knowledge of the labor market (K) on the basis of his SEB, Q and the predetermined IQ. K is based on the score of a "Knowledge of the World of Work" test administered in 1966 to the young men in our sample. The student was tested on his knowledge of the kind of work performed by a variety of occupations (e.g., machinist, statistical clerk, economist, etc.) and the level of formal education usually attained by members of these occupations. Also included in the test was a series of questions regarding earnings comparisons between a series of occupations. For example, one such question is as follows: "Who do you think earns more in a year, a man who is a truck driver or a grocery store clerk?" The answers to these questions provides some evidence of a student's career awareness and the extent of his labor market information. This knowledge should ultimately have an impact on the probability of an individual dropping out of high school. Indeed, the effect of K on D will give some indication of the effectiveness of "career education" programs of instruction for it is precisely the development of career awareness that is one of the goals of the current "career education" movement in the education community.

The third equation of the model explains the young men's curriculum choice (C) using again the explanatory variables SEB, Q and IQ. C is a dichotomous variable taking the value of one if the individual is or was enrolled in a vocational or commerical (as opposed to a college preparatory or general) curriculum. Our purpose here is to attempt to assess the success of vocational educational programs as a dropout prevention device; a frequently cited justification for continued Federal funding of this program of instruction (Bell, 1975). While both K and C are determined subsequent to an individual's IQ, no assumption will be made as to the causal priority of these two variables. K is measured in 1966 when the individuals in the sample were between the ages of 14 and 17. Consequently, some of these young men had made a choice regarding their curriculum and some, no doubt, had not when the world of work test was administered.

Finally, the last equation in the model explains D on the basis of the individual's SEB, Q, and the three predetermined variables, IQ, K and C. In addition to these explanatory variables, a variable measuring an index of demand for teenage male labor (DI) in 1968 is included.[11]

11/Variables measuring local unemployment rates were also included in some of the regressions for D. These results will be discussed below.

A description of this index (which is contained in the orginal data source) is contained in Appendix A.1. Unfortunately, the expected sign of the estimated coefficient of DI in explaining the conditional probability of dropping out of high school is uncertain. On the one hand, DI may reflect increased opportunity costs of attending school and, consequently, have a positive partial effect on D. There is also an income effect imbedded in the coefficient, however. An increase in DI, for example, may be seen as an indication of an increased high school-dropout earnings differential over time and, consequently, as an increase in the expected return from completing high school leading to a reduction in D, ceteris paribis.[1]

Formally, the four equation causal model estimated here may be described as follows:

(1) $IQ = a_o + a_1\bar{Y} + a_2EF + a_3EM + a_4S + a_5Q + e_1$

(2) $K = b_o + b_1\bar{Y} + b_2EF + b_3EM + b_4S + b_5Q + b_6IQ + e_2$

(3) $C = c_o + c_1\bar{Y} + c_2EF + c_3EM + c_4S + c_5Q + c_6IQ + e_3$

(4) $D = d_o + d_1\bar{Y} + d_2EF + d_3EM + d_4S + d_5Q + d_6IQ + d_7K + d_8C + d_9DI + e_4$

where the disturbances, e_j, are taken as mutually uncorrelated in the probability limit (except for the correlation of e_2 with e_3) and uncorrelated with the regressors in their own and preceding equations. Given these assumptions, the coefficients of the model may be estimated by applying ordinary least squares to each equation, using the sample of young men described above as observations. Two problems of estimation remain, however. First, both equation (3) and (4) are linear probability models and therefore suffer from heteroskedastic error terms which, while not leading to biased estimates of the coefficients, will not provide efficient least squares estimators. Second, the NLSYM unfortunately suffers from the problem that several of the variables of interest having mission values for some of the observations. The methods used to address these problems will be discussed below. To the extent, have missing values for some of the observations. The methods used obtained, the total effect of, for example, $\overline{Y}$ on D is equal to $d_1 + d_6a_1 + d_7b_1 + d_8c_1 + d_7b_6a_1 + d_8c_6a_1$ where d_1 is the direct effect of $\overline{Y}$ on D and the other terms capture the indirect effect of family income as it affects the dropout decision through the predetermined variables in the model (IQ, K, C) and these, in turn, affect D. The recursive model specified in equations (1) - (4) seems a useful way to formalize Becker's (1967) investment-theoretic model of schooling decisions described in Section II. The direct effect of a change in an "opportunity" factor such as $\overline{Y}$ on D reflects an independent shift in the supply schedule while the indirect effects will enable us to determine the importance of the hypothesized correlation between demand and supply elements. On the other hand, the independent effect of changes in "ability" factors on D can be ascertained by, for example, an estimate of the direct effect of IQ on D which is measured by the coefficient d_6.

The econometric problems of estimating linear probability models such as equations (3) and (4) are by now well known. Nevertheless, this specification was employed for its linearity facilitates the computation of direct and indirect effects. The problem of heteroskedasticity may, however, be alleviated by following the procedure outlined by Goldberger

(1964) where, for example, each of the variables in equation (4) is weighted by $/\hat{D}(1-\hat{D})/^{-\frac{1}{2}}$ where $\hat{D}$ is the least squares fitted value of D. This procedure is followed here so that both equation (3) and (4) are estimated by generalized least squares (GLS). Unfortunately, there is no guarantee that $\hat{D}$ will lie between zero and one for all observations. Smith and Cicchetti (1972) have done Monte Carlo studies of alternative methods of handling inadmissable weights from the first stage ordinary least squares regression in a GLS analysis. However, for large samples none of their ad hoc procedures seems to be superior to simply deleting from the second stage those observations that have a predicted value of the dependent variable outside the admissable range. This is the procedure that is followed here for only a very small part of the total sample is lost by this technique. Finally, a problem of estimation which is generally not mentioned is the fact that both e_3 and e_4 are not normally distributed. Consequently, the classical tests of significance do not apply for these two equations. Asymptotic variances of the estimated coefficients of equations (3) and (4) can be estimated, however, and while the classical significance tests are not applicable we will frequently make comparisons of the size of the estimated coefficients and their standard errors as if a "t" test were being undertaken.

Unfortunately, the theory of estimation when several of the variables of interest have missing values for some observations is not well developed in a multivariate regression framework. However, in the case of a simple linear regression we do know (Kmenta, 1971, pp. 336-344) that if we ignore those observations with missing values there will only be a small loss in efficiency if the missing values of the independent variable have a small dispersion and, at the same time, the mean of the missing values of the independent variable is close to the mean of the available values of the variable. In our sample these conditions are unlikely to hold given that the missing values seem to be concentrated in that portion of the sample which is most likely to contain the dropouts, e.g., young men of low SEB. Consequently, the option of ignoring an entire observation if any one of its variables is missing does not seem appropriate. Not only would this lead to inefficient estimators but would also give biased estimates for the usable sample would consist primarily of relatively socially and economically advantaged young men. Instead, we have adopted the procedure that a missing value for a particular variable causes that observation to be eliminated from the calculations involving that variable only. This option has the advantage of utilizing as much of the data as possible in estimation. The risk of this procedure is that, under certain circumstances, the partial coefficients are based on a very different number of observations and perhaps on quite different subpopulations. Fortunately, each of the four equations estimated in the model described above used the same minimum number of observations in each regression so it is unlikely that each regression is examining the response of different subpopulations.

IV. Empirical Results

As discussed above, equations (1) and (2) were estimated by ordinary least squares while equations (3) and (4) were estimated by generalized least squares using as observations the white and nonwhite male youth between the ages of 16 and 19 contained in the NLSYM. The number of observations used in each regression was 1073 individuals where, given the missing data correction used here, this is the minimum number of cases any correlation is based on. In order to discover if there were significant differences in the response of the white and nonwhite subsample to the several independent variables, an analysis of covariance was employed (see Johnston, 1972).[12] For each equation the null hypothesis that the two subsamples had identical slope coefficients was rejected at the 0.01 level of significance.[13] As a result of this test, separate regressions are reported for the white and nonwhite youth in the sample where the minimum number of observations used for each regression was 849 and 224 for whites and nonwhites, respectively.[14] Finally, all cases whose predicted value of the dependent variable fell outside of the unit interval in the second stage of the generalized least squares estimation were dropped from the analysis. For the white sample, 3 and 16 cases were dropped from the second stage estimation of equation (3) and (4),

---

12/Ninety-six percent of the nonwhite sample are black.

13/The computed F Statistics were 12.16, 8.38, 6.91 and 15.56 for equations (1)-(4), respectively. Note again, however, that equations (3) and (4) do not have normally distributed error terms. Consequently, the F test is not strictly appropriate although it is indicative of a significant racial difference in the slope coefficients.

14/Zero order correlations, sample means and standard deviations for the variables in the two sub-samples are contained in Appendix Tables A.2 and A.3.

respectively, while the corresponding number of unusable observations for the nonwhite sample was 4 and 26.

There are several ways of presenting the results of estimating the parameters of the four equation recursive model described above. A technique familiar to sociologists which clearly shows how the white and nonwhite structures differ is path analysis (see Blalock, 1971). The path diagrams are shown in Figures 2 and 3 for the white and nonwhite sample, respectively. The single-headed arrows indicate significant causal paths from cause to effect where the coefficient assigned to each path is the estimated standardized regression coefficient, sometimes called the beta or path coefficient. For equation (1) and (2) significance implies a classical test of significance which satisfies the 0.95 level of confidence or better while for equations (3) and (4) we adopt the convention that the estimated coefficient must be at least twice its standard error before the coefficient is labeled "significant." The two-headed arrows denote correlations not analyzed in causal terms with the coefficient assigned to each such arrow being the zero-order correlation coefficient.

The unstandardized regression coefficients together with their estimated standard errors and associated summary statistics for each regression are found in Tables 1 and 2 for whites and nonwhites, respectively. It will be recalled that the estimated coefficients of the equation for D, for example, can be used to compute the probability of dropping out of high school for our sample of young men conditional upon the values taken by the various independent variables. However, given the recursive nature of our model, the estimated coefficients of equation

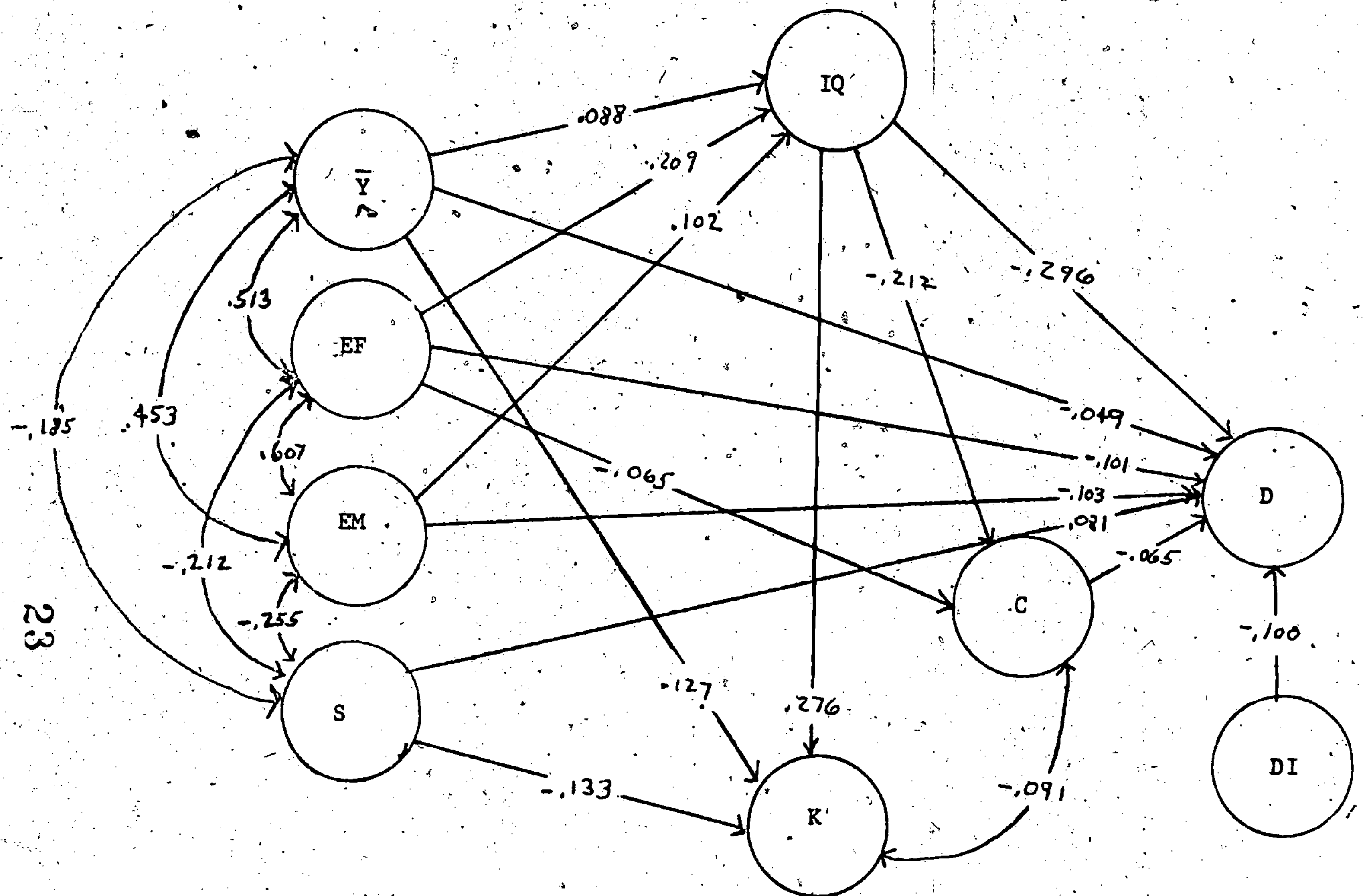
(4) show only the <u>direct</u> effects of the predictor variables on D. That is, any one of these coefficients indicates only the partial effect of that variable on the conditional probability of dropping out. However, as has been discussed above, a variable such as $\overline{Y}$ may affect D not only directly but also indirectly through its effect on IQ, K, or C and these variables subsequent effect on D. Consequently, the major interest of this paper will be the calculation of the <u>total</u> effect of a predictor variable on D using the chain-rule formula illustrated for the variable $\overline{Y}$ in Section III of the paper. The total, direct and indirect effects of the several independent variables on D are presented in Table 3 using the estimated coefficients of Tables 1 and 2. While, of course, complementary the information provided by the path analysis and the effects computed in Table 3 provide somewhat different views of the underlying structure of the model. Consequently, each technique will be discussed in turn.

(a) <u>Path Analysis</u>

The beta (path) coefficients shown in Figures 2 and 3 indicate the relative "importance" of the independent variables in explaining the variation in each of the dependent variables (Goldberger, 1964, pp. 197-198). That is, when the variables are standardized, the unit of measurement of each variable is comparable, being measured in standard deviation units. Consequently, a beta coefficient indicates the effect of a one standard deviation change in a given dependent variable on the standarized dependent variable. It should be noted, however, that we are discussing "importance" only in a narrow statistical sense and this may or may not

Figure 2

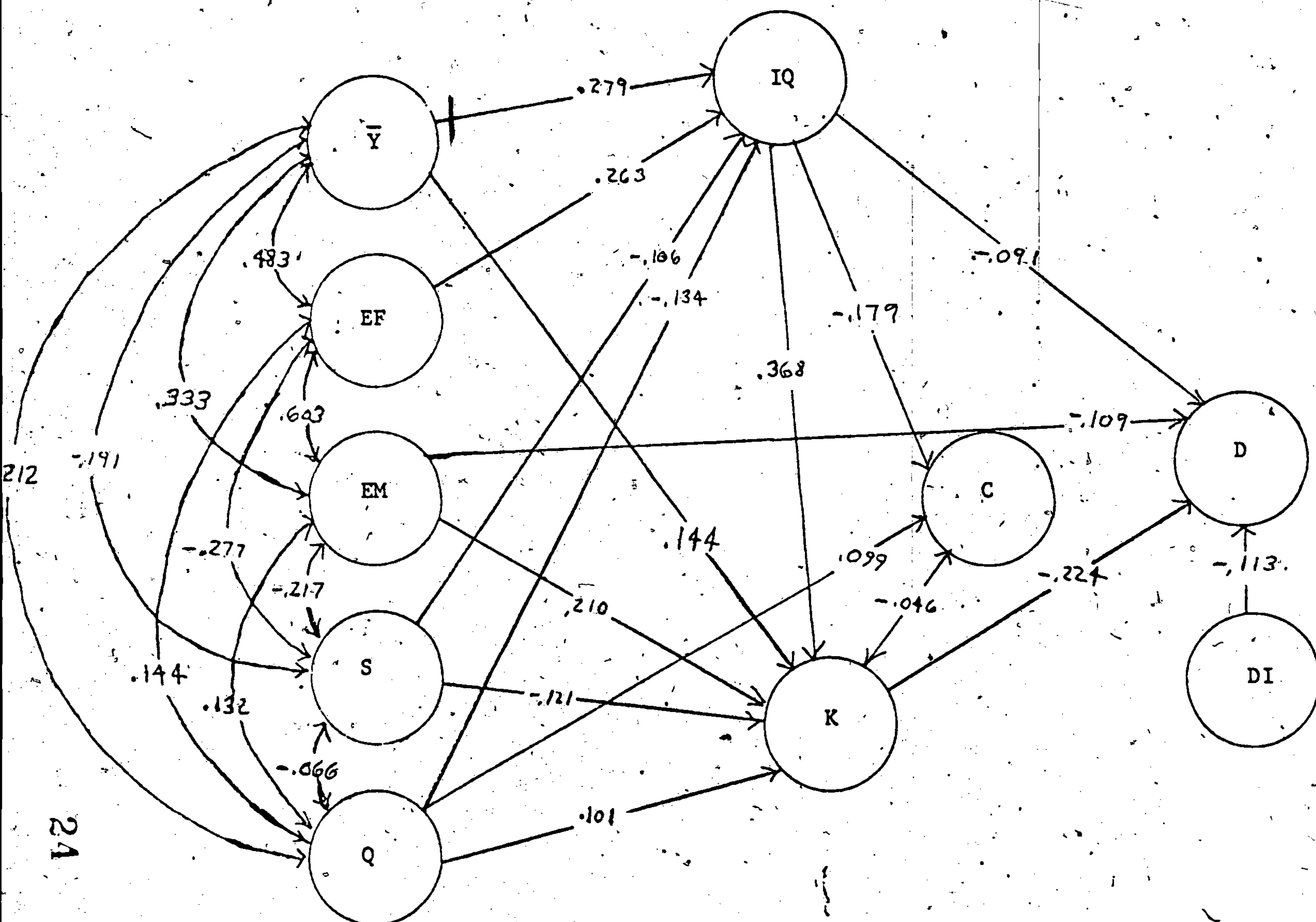
Standardized Regression Coefficients, White Males (N = 849)*



* All paths shown are Significant at the 0.05 Level or Better.

Figure 3

Standardized Regression Coefficients, Nonwhite Males (N = 224)*



* All paths shown are significant at the 0.05 level or better.

have implications for public policy. The question of whether a variable is important in terms of its effect on D is one which statistical inference cannot, by itself, answer. Ultimately, one's perception of importance depends on the policy actions in mind and the degree to which the relevant explantory variables are manipulable through public policy. Our discussion in this section will concentrate on the relative importance of the direct effects of the several independent variables on D. The discussion of the total effects of these variables on D is contained in the following section.

Of the four variables used here which reflect an individual's SES, only mother's educational attainment has a significant direct effect on the probability that a nonwhite male youth will drop out of high school. This is in contrast to the results shown for white youth where all four variables have a significant direct effect on D. Even for whites, however, the educational attainment of the parents is the most important background variable affecting D, and again the education of the youth's mother is slightly more important than that of his father's in affecting the drop out decision. As expected, for both whites and nonwhites the higher the educational attainment of the parents, the lower is the conditional probability of dropping out. As discussed above, however, we really do not yet know exactly what process within the household this relationship between EF, EM and D reflects. For a sample of high school graduates Sewell and Hauser (1972) have demonstrated that the parent's education is associated with parental encouragement and this positively affects the son's subsequent education independent of several other socioeconomic factors.

Hill and Stafford (1974) have found that the parent's education positively affects the educational aspirations they have for their children (as measured by whether they expect their children to attend college) and this expectation is positively associated with increased parental time devoted to the children's physical care in the preschool years. This physical care is, in turn, associated with the development of cognitive and affective abilities. Interestingly, this latter result of Hill and Stafford also indicates that the mother's educational attainment is relatively more important than the father's in affecting educational expectations and that this effect is largely independent of family income-- both of which are consistent with the results presented here. Whatever the process at work, the results of this present paper together with the earlier work of several others (see expecially Leibowitz, 1974, and Hill and Stafford, 1974) indicates the independent importance of the mother's socioeconomic characteristics on her children's educational attainment. This result is noteworthy for the great majority of previous studies of the determinants of educational attainment have used only variables which reflect the father's characteristics.

If equal educational opportunity is defined in this context as a situation in which each individual's probability of not completing high school is unaffected by his SEB and reflects, instead, only differences in ability, the relative unimportance of the direct effects of these background variables on nonwhite youth's D is heartening. However, as pointed out above, we are looking here at only the direct effect of these variables on D. To the extent that the SEB variables affect IQ, for example, and

IQ affects D, the total effect of SEB on D may be substantial.

As measured by the (absolute) value of the beta coefficients, IQ has the most important direct effect on D for whites while K is the most important for nonwhites. In Becker's (1967) framework, both IQ and K reflect "ability" factors which affect the marginal returns from a given investment in schooling. For both IQ and K, higher scores are associated with a lower probability of dropping out which is as expected for the direct effects of these variables reflect an upward shift in the demand curve for educational investments given a fixed supply schedule. While both IQ and K reflect the scores of tests administered to the sample of young men, they presumably measure different types of "ability" or knowledge. IQ is thought to reflect the cognitive development of individuals as measured by their ability to read, write, calculate and articulate. Knowledge of the world of work, on the other hand, would seem to reflect such affective development as a spirit of inquiry into the nature of the labor market and/or motivation for the acquisition of more applied information. The positive effect of IQ on educational attainment is well known, most recently having been analyzed by Sewell and Hauser (1975). A high score on K reflects, in part, knowledge of the degree to which educational credentials are required for labor market success and this is, consequently, an inducement to complete high school for nonwhites.

The actual state of the local labor market, as measured by DI, reduces the probability of dropping out for both racial groups. Although DI is a very incomplete and imperfect measure of labor market conditions, the results shown here indicate the dominance of the expected future income effect over the substitution effect in affecting schooling decisions. In some unreported regressions, the local unemployment rate in 1968 was first used as a regressor and then the 1966-68 average rate was used instead of DI in the D equation. In both cases and for both races, the estimated coefficients were negative but had estimated standard errors which were equal to or exceeded the coefficients. The direct effect of labor market conditions on the drop out decision is, therefore, somewhat mixed. For a more complete discussion of the problems of interpretation encountered here, the reader should consult the paper of Lerman (1972).

The quality of high school attended (Q) by the individuals in our sample does not have a direct effect on D for either racial group. For nonwhites, Q does affect D indirectly (through both IQ and K) and this will be discussed in the next section. One could, of course, argue that our measure of Q really is an inadequate index of school "quality" and that school characteristics do make a difference if properly measured. There is no totally satisfactory rebuttle to that criticism except to note again that Q measures the characteristics of the particular high school attended by the youth in the sample rather than a State or school district average. Consequently, while not completely satisfactory, we would argue that the measure of quality used here is the best now available. The high school curriculum (C) in which the individuals are enrolled directly affects D only for white youth. For these young men, enrollment in a vocational or commercial curriculum does lower the probability of dropping out; thereby verifying the claims of vocational education administrators for at least part of our sample. It should also be noted here that the vocational curriculum attracts those white students who are relatively dropout prone. This is demonstrated by the negative path coefficients from both EF and IQ to C.

(b) Total, Direct and Indirect Effects

Table 3 contains the estimated total, direct and indirect effect of the several independent variables on the variable of interest, D. These are computed from the significant coefficients shown in Tables 1 and 2 so can be interpreted as the effect of changes in these predictor variables on the probability of dropping out of high school. Given the structure of the model, K, C, and DI have only a direct effect on D which can be read directly from Tables 1 and 2. For the two policy variables, for example, each additional point scored on the K test reduces the probability of dropping out for nonwhites by 1.2 percent while for white youth being enrolled in a vocational curriculum lowers the probability of leaving high school before graduation by five percent, ceteris paribus.

As noted in the path diagrams, the direct effects of the SEB variables on D are quite strong for white youth while largely absent for nonwhite youth. However, when the indirect effects of these variables are computed it is clear that family background characteristics make a difference for nonwhite youth also. Looking at $\overline{Y}$, for example, the direct effect for whites accounts for 67 percent of the total effect of this variable ($-.008 \div -.0119$) in reducing the chance of dropping out. For nonwhites, however, the total effect of $\overline{Y}$ on D is entirely an indirect one as $\overline{Y}$

TABLE 1

ESTIMATED REGRESSION COEFFICIENTS (STANDARD ERRORS) FOR WHITE MALES (N = 849)

Dependent	Predetermined											
	$\bar{Y}$	EF	EM	S	Q	IQ	K	C	DI	Constant	R^2	F
IQ[a]	.693 (.301)	.864 (.177)	.520 (.213)	-.270 (.229)	.371 (.235)					84.03	.129	25.02
K[a]	.489 (.143)	.127 (.085)	.071 (.101)	-.447 (.109)	.079 (.112)	.135 (.016)				13.02	.185	31.94
C[b]	-.001 (.005)	-.006 (.003)	.0002 (.003)	.0004 (.004)	.003 (.004)	-.005 (.001)				.712	.062	19.76
D[b]	-.008 (.004)	-.009 (.002)	-.011 (.003)	.011 (.003)	.003 (.003)	-.006 (.0005)	-.001 (.001)	-.053 (.016)	-.002 (.0005)	1.072	.186	54.82

a/ Estimated by ordinary least squares.

b/ Estimated by Generalized least squares.

TABLE 2

ESTIMATED REGRESSION COEFFICIENTS (STANDARD ERRORS) FOR NONWHITE MALES (N = 224)

Dependent	Predetermined											
	$\bar{Y}$	EF	EM	S	Q	IQ	K	C	DI	Constant	R^2	F
IQ[a]	2.029 (.508)	1.086 (.339)	-.458 (.397)	-.557 (.330)	-.888 (.407)					83.17	.208	11.55
K[a]	.515 (.239)	-.091 (.157)	.541 (.181)	-.313 (.151)	.327 (.187)	.180 (.031)				2.79	.325	17.48
C[b]	.008 (.009)	-.006 (.006)	.003 (.007)	-.002 (.006)	.016 (.007)	-.004 (.001)				.479	.047	3.82
D[b]	.002 (.009)	.001 (.006)	-.015 (.007)	.002 (.006)	.002 (.007)	-.002 (.001)	-.012 (.003)	.023 (.042)	-.004 (.002)	.967	.087	6.32

a/ Estimated by ordinary least squares.

b/ Estimated by generalized least squares.

TABLE 3

MARGINAL CONTRIBUTION OF PREDICTOR VARIABLES[a] TO THE CONDITIONAL PROBABILITY OF DROPPING OUT OF HIGH SCHOOL FOR WHITE AND NONWHITE MALES[b]

Predictor Variable	Total Effect	Direct Effect	Indirect Effects				
			Through IQ	Through C	Through K	Through C via IQ	Through K via IQ
$\bar{Y}$	-.0119 (-.0140)	-.008 (0)	-.004 (-.004)	0 (0)	0 (-.006)	.0001 (0)	0 (-.004)
EF	-.0134 (-.004)	-.009 (0)	-.005 (-.002)	.0003 (0)	0 (0)	.0003 (0)	0 (-.002)
EM	-.014 (-.021)	-.011 (-.015)	-.003 (0)	0 (0)	0 (-.006)	0 (0)	0 (0)
S	.011 (.006)	.011 (0)	0 (.001)	0 (0)	0 (.004)	0 (0)	0 (.001)
Q	0 (0)	0 (0)	0 (.002)	0 (0)	0 (-.004)	0 (0)	0 (.002)
IQ	-.0057 (-.004)	-.006 (-.002)	0 (0)	.0003 (0)	0 (-.002)	0 (0)	0 (0)

a/ Only those effects which are significant at the 0.05 level or better are shown. For equations (3) and (4), the estimated coefficient must be at least twice its estimated standard error.

b/ The effects enclosed in parentheses ( ) are for nonwhite males.

increases IQ which decreases D, $\overline{Y}$ increases K which decreases D, and finally, $\overline{Y}$ increases IQ which increases K which decreases D. The circuitous route that $\overline{Y}$ follows in reducing the probability of dropping out for nonwhites may be one reason that Masters (1969), using a model which only captured direct effects, found an ambiguous and frequently insignificant effect of family income on the probability of dropping out for nonwhite youth. Clearly, for nonwhite youth the correlation between Backer's "opportunity" factors (as measured here by the SEB variables) and the "capacity" factors (measured here by IQ, K and C) is an important determinant of the total investment made in formal schooling.

The effects of parent's educational attainment on D have been discussed in detail above. Consequently, it will only be noted here that for both white and nonwhite youth the effect of EM on D is primarily a direct one. EF affects nonwhite youth's D only indirectly while for white youth the effect is again largely a direct effect. The greater the number of siblings of the youth in our sample, the greater is the likelihood that they will leave high school before graduation. Again, however, the method by which S affects D is quite different between the two races. For white youth, the total and direct effect are equivalent implying that the number of the respondent's brothers and sisters primarily reflects a parental "ability-to-pay" factor which leads to an independent shift upward in the marginal financing cost schedule and a consequent increase in D. For nonwhites, the effect of S on D is entirely an indirect one as S reduces both IQ and K, holding the several other variables constant. That the effect of S on D works through IQ only for nonwhites is consistent with some evidence presented by Belmont and Morolla (1973). They present evidence that the effect of family size on IQ differs across socioeconomic status groups with those youth from "high status" families suffering the smallest debilitating effects on their IQ as S increases.

As noted above, school quality has no independent effect on D for white youth. As Tables 1 through 3 demonstrate, Q also has no indirect effect for whites. For nonwhites, however, we obtain a result that is not easily explained. For this subsample Q significantly reduces IQ and consequently increases the probability of dropping out both because of IQ's direct effect on D and the effect of IQ on K. This anamolous result, it must be admitted, is not well understood by the author. It is some small comfort to note that Ribich and Murphy (1975) found that Q (measured by expenditures per pupil in their study) negatively affected IQ for their sample of individuals from the Project Talent data. It is also true that the total effect of Q on D is zero for nonwhites as well as whites because the indirect positive effect of Q on K and K's subsequent effect on D offsets the dropout inducing effects of Q. Nevertheless, an explanation of the negative relationship between Q and IQ is not forthcoming from the data. A hypothesis that was investigated suggested that the effect of Q on IQ for nonwhites was dependent on the percentage of blacks enrolled in the high school. That is, to the extent that high quality schools are predominately white, the pressure of well-endowed white classmates and the lack of a significant number of individuals from a racial peer group may lower nonwhite achievement.[15] To test this hypothesis, an interaction term between percentage black

[15]There is some evidence for this effect of peer group composition on achievement contained in Armor (1972) and Winkler (1975).

enrollment and Q was added to equation (1) for the nonwhite sample. While the estimated coefficient of the interaction term was positive (as implied by the hypothesis), it was extremely small and never significant by conventional standards.[16]

IQ's negative effect on white youth's D is predominately a direct effect. As noted before, IQ does have a small positive effect on D for IQ negatively affects C and C, in turn, reduces the probability of dropping out. For nonwhites the total negative effect of IQ is equally divided between a direct effect and an indirect effect through K. It is worth noting once again that for the nonwhite sample a high score on the K test is more influential in reducing D than is IQ. While an additional point on the IQ scale has the total effect of reducing the probability of dropping out by 0.4 percent, the total effect of an additional point on the K test reduces the probability by 1.2 percent.

[16] The same specification of equation (1) was tried for the sample of white youth with the estimated coefficient of the interaction term being zero.

V. Some Policy Implications

The model of educational attainment sketched out in Part II suggested that the probability of leaving high school before graduation could be viewed as reflecting the interaction of capacity and opportunity factors which determined the optimal investment in human capital. This solution was an optimal or income-maximizing one for the individual subject to the constraints of the model, particularly those embodied in the opportunity factors. But if this is an "optimal" solution for the individual, is it also optimal for society? If the view is widely shared that dropping out is socially undesirable, are there policies by which we can encourage a greater investment in formal schooling?

The matter of dropping out becomes a social problem when the consequences of leaving high school early begin to impinge seriously on others; or when dropping out results from incomplete information--e.g., on income, employment or other prospects--which was available but not known to the student or when dropping out reflects inequality of educational opportunity.[17/] The external diseconomies from dropping out most frequently mentioned and measured are the increased unemployment of the dropout and the resulting loss of taxable earnings or national output.[18/] Of course, there may be other social costs, some of which were mentioned in the introduction to this paper. But the extent of the relationship between dropping out and

---

17/ These issues are discussed in more detail by Weisbrod (1965).

18/ Weisbrod, ibid, pp. 139-149 and Levin (1972).

increased crime and drug usage, for example, is unknown. Nevertheless, if the social costs of leaving high school before graduation are considered severe enough to warrant public concern, the model estimated here has several implications for public policy.

First, to the extent that a relatively low IQ has an important direct effect on increasing D, dropping out may indeed be optimal for both the individual and society. For the white youth in our sample IQ had the largest direct effect (as measured by the beta weight) of the several independent variables used to explain D. The policy implications one can infer from the indirect effect of the SEB variables on D through IQ are unclear. Family income could, of course, be increased through some type of income maintenance mechanism but exactly how this would translate itself into increased cognitive and affective abilities is not well understood.[19]

The direct effects of family background on D do indicate the extent of inequality of opportunity in that these direct effects represent shifts of the marginal financing curve due to family background for a given marginal return schedule (see Becker, 1967, pp. 27-29). Surprisingly, these effects are relatively unimportant for nonwhite youth but are influential in affecting a white youth's dropout decision. Here, again, some type of income support mechanism would seem to be the primary short-term policy instrument through which to reduce D, although adult education programs which would affect EF and EM and family planning services to reduce S may have long-run effects.

---

[19] For a discussion of the relationship between a guaranteed income and cognitive development, see Hill (1975).

Knowledge of the world of work as measured here by K has the largest direct effect in reducing D for nonwhites. To the extent that this type of information is provided in "career education" programs of instruction, this result suggests that a restructuring of the secondary school curriculum in order to increase career awareness and labor market information may, indeed, be useful as a dropout prevention device. Vocational Education, which has long been touted as a means to reduce dropouts does reduce D for white youth while at the same time attracting those youth who are dropout-prone as judged by their family background and IQ.

After controlling for the other variables in the model, the total effect of school quality (Q) on D was zero for both races. That is not to say, however, that schools don't make a difference. Indeed, the impact of the C and K variables suggest otherwise. What is implied is that whatever effects the various components of our quality measure have on student achievement or other student outcomes, a community's investment in the physical facilities of its school has little or no effect on the probability of dropping out once family background and ability are controlled for.

Finally, whatever the public policies which could be utilized to reduce the probability of dropping out of high school and their cost-effectiveness, leaving school is not an irrevocable decision. Presumably, some of those who drop out will return to school at a later date if they find it advantageous to do so. An indication of this phenomenon is contained in the answer to a question posed in the 1968 wave of the NLSYM: "Do you expect to return to high school?" For the young men in our sample with D = 1 and who answered this question (N = 97), a series of cross tabulations were developed which enabled us to test the independence of the expectation to return to school and several other classification variables.[20] Only the cross tabulation of "expect to return" with race yielded a significant X^2 of 4.20 with one degree of freedom. Hence the null hypothesis that expecting to return to high school after dropping out is independent of race is rejected at the 0.05 level of significance. For this cross tabulation the observed frequency of nonwhites responding "yes" to the "expect to return" question is considerably larger than the expected frequency. While the reason for this racial difference is subject only to conjecture at this point, it seems reasonable to suppose that for nonwhites the importance of a high school credential is more important for subsequent labor market success than for whites. The validity of this hypothesis could be demonstrated by following white and nonwhite dropouts into the labor force and then observing their career progression as compared to high school graduates. This analysis will be the subject of a subsequent study.

20/ Those variables include race; curriculum last attended; age; highest grade completed; race by curriculum; race by highest grade completed.

REFERENCES

Armor, D. J., "The Evidence on Busing", The Public Interest 28 (Summer, 1972): 90-126.

Bachman, J. G., Youth In Transition, Vol. II, Ann Arbor: Institute for Social Research, 1971.

Becker, G. S., Human Capital and the Personal Distribution of Income, Ann Arbor: University of Michigan Department of Economics, 1967.

Bell, T. H., U.S. Commissioner of Education. Statement before the House Subcommittee on Elementary, Secondary and Vocational Education, February 19, 1975.

Belmont, L. and F. Morolla, "Birth Order, Family Size and Intelligence", Science 182 (December, 1973): 1096-1100.

Ben-Porath, Y., "The Production of Human Capital and the Life Cycle of Earnings", Journal of Political Economy 75 (August, 1967): 352-365.

Blalock, H., Causal Models in the Social Sciences, Chicago: Aldine Press, 1971.

Conlisk, J., "Determinants of School Enrollment and Performance", Journal of Human Resources 4 (Spring, 1969): 140-157.

Duncan, B., Family Factors and School Dropouts 1920-1960, Ann Arbor: University of Michigan Press, 1965.

Edwards, L. N., "The Economics of Schooling Decisions: Teenage Enrollment Rates", Journal of Human Resources 10 (Spring, 1975): 155-173.

Freeman, R. B., "Overinvestment in College Training?", (mimeo), 1975.

Goldberger, A., Econometric Theory, New York: John Wiley, 1964.

Griliches, Z., "Notes on the Role of Education in Production Functions and Growth Accounting" in Education, Income and Human Capital, W. L. Hansen (ed.), New York: Columbia University Press, 1970.

Griliches, Z., "Economic Problems of Youth" in Youth: Transition to Adulthood, J. S. Coleman (ed.), Washington: GPO, 1973.

Griliches, Z., and W. Mason, "Education, Income and Ability", Journal of Political Economy 80 (May/June, 1972, Part 2): 74-104.

Hauser, R. M., "Disaggregating a Social-Psychological Model of Educational Attainment", Social Science Research 1 (June, 1972): 159-188.

Hill, C. R., "Guaranteeing Income and Encouraging Child Care", Review of Social Economy 33 (April, 1975): 15-25.

Hill, C. R. and F. P. Stafford, "Family Background and Lifetime Earnings", paper presented to an NBER Conference on Income and Wealth, 1974.

Johnson, G. E. and F. P. Stafford, "Social Return to Quantity and Quality of Schooling", Journal of Human Resources 8 (Spring, 1973): 139-155.

Johnston, J., Econometric Methods, New York: McGraw-Hill, 1972.

Kmenta, J., Elements of Econometrics, New York: The MacMillan Co., 1971.

Kohen, A. I., Determinants of Early Labor Market Success Among Young Men, Columbus, Ohio: Center for Human Resource Research, 1973.

Leibowitz, A., "Home Investments in Children", Journal of Political Economy 82 (March/April, 1974, Part 2): 111-131.

Lerman, R. I., "Some Determinants of Youth School Activity", Journal of Human Resources 7 (Summer, 1972): 366-379.

Levin, H. M., The Costs to the Nation of Inadequate Education, Washington: GPO, 1972.

Marland, S. P., Career Education: A Proposal for Reform, New York: McGraw-Hill, 1975.

Masters, S. H., "The Effect of Family Income on Children's Education: Some Findings on Inequality of Opportunity", Journal of Human Resources 4 (Spring, 1969): 158-175.

Mincer, J., "The Distribution of Labor Incomes: A Survey", Journal of Economic Literature 8 (March, 1970): 1-26.

Parnes, H. S., et. al., Career Threshholds, Washington: U.S. Department of Labor, 1970.

Ribich, T. I. and J. L. Murphy, "The Economic Returns to Increased Educational Spending", Journal of Human Resources 10 (Winter, 1975): 56-77.

Sewell, W. H. and R. M. Hauser, "Causes and Consequences of Higher Education: Models of the Status Attainment Process", American Journal of Agricultural Economics 54 (December, 1972): 851-861.

Sewell, W. H. and R. M. Hauser, Education, Occupation and Earnings: Achievement in the Early Career, New York: Academic Press, 1975.

Smith, V. K. and C. J. Cicchetti, "Estimation of Linear Probability Models with Dichotomous Dependent Variables", Resource for the Future, (mimeo), 1972.

Solmon, L. C., "A Note on Equality of Educational Opportunity", American Economic Review 60 (September, 1970): 768-771.

U.S. Department of HEW, Digest of Educational Statistics, Washington: GPO, 1974.

Wachtel, P., "The Returns to Investment in Higher Education: Another View" in Education, Income and Human Behavior, F. T. Juster (ed.), New York: McGraw-Hill, 1975.

Wallace, T. D. and L. A. Ihnen, "Full-Time Schooling in Life Cycle Models of Human Capital Accumulation", Journal of Political Economy 83 (February, 1975): 137-156.

Weisbrod, B. A., "Preventing High School Dropouts" in Measuring Benefits of Government Investments, R. Dorfman (ed.), Washington: Brookings Institution, 1965.

Winkler, D. R., "Educational Achievement and School Peer Group Composition", Journal of Human Resources 10 (Spring, 1975): 189-204.

APPENDIX TABLE A.1

Variable Definitions and Measurement Scales

Dropout (D)	= 1 if highest grade completed is less than 12 and respondent is not enrolled in school in 1968
Father's Education (EF)	Highest grade of formal schooling completed by father
Mother's Education (EM)	Highest grade of formal schooling completed by mother
Mean Family Income ($\bar{Y}$)	Arithmetic average of total family income of respondent's parental family for 1966, 1967 and 1968. Income is coded in the following intervals: 01: under $1000 02: $1000 - $1999 03: $2000 - $2999 04: $3000 - $3999 05: $4000 - $4999 06: $5000 - $5999 07: $6000 - $7499 08: $7500 - $9999 09: $10,000 - $14,999 10: $15,000 - $24,999 11: $25,000 and over
Number of Siblings (S)	Total number of siblings of respondent in 1966.
Ability (IQ)	Respondents' actual IQ score constructed by using pooled data from several different tests (see Kohen, 1973).
School Quality (Q)	Normalized school quality index of last high school attended with range from 1 (lowest) to 11 (highest) (see Kohen, 1973).
Knowledge of Work (K)	Total score of "Knowledge of World of Work" test administered in 1966 with possible range from 0 to 56.
Curriculum (C)	= 1 if current (last) high school curriculum in which respondent is (was) enrolled is (was) vocational or commercial.
Demand Index (DI)	Index of demand for teenage male labor for labor market of current residence in 1968. The index represents the sum of the percentage of total employment in the area represented by agriculture and retail trade with possible range from 0 to 99.

APPENDIX TABLE A.2

ZERO-ORDER CORRELATIONS, MEANS AND STANDARD DEVIATIONS OF VARIABLES FOR WHITE MALES

	D	EF	EM	DI	IQ	K	Q	S	$\overline{Y}$	C
D	1.00									
EF	-.263	1.00								
EM	-.258	.607	1.00							
DI	-.036	-.141	-.039	1.00						
IQ	-.332	.328	.283	-.087	1.00					
K	-.189	.265	.238	-.072	.355	1.00				
Q	-.023	.067	.084	.078	.081	.063	1.00			
S	.166	-.212	-.255	.075	-.126	-.212	-.004	1.00		
$\overline{Y}$	-.205	.513	.453	-.198	.254	.269	.090	-.185	1.00	
C	.035	-.139	-.092	-.027	-.216	-.091	.005	.035	-.085	1.00
MEAN	-.110	10.8	10.9	22.7	105.8	32.3	6.04	2.99	7.69	.144
SD	.313	3.49	2.82	11.8	14.4	7.04	1.97	2.10	1.83	.351

APPENDIX TABLE A.3

ZERO-ORDER CORRELATIONS, MEANS AND STANDARD DEVIATIONS OF VARIABLES FOR NONWHITE MALES

	D	EF	EM	DI	IQ	K	Q	S	$\bar{Y}$	C
D	1.00									
EF	-.125	1.00								
EM	-.157	.603	1.00							
DI	.028	-.301	-.244	1.00						
IQ	-.134	.355	.170	-.143	1.00					
K	-.270	.330	.333	-.368	.461	1.00				
Q	-.026	.144	.132	-.181	-.042	.146	1.00			
S	.092	-.277	-.217	.215	-.204	-.264	-.066	1.00		
$\bar{Y}$	-.084	.483	.333	-.390	.368	.372	.212	-.191	1.00	
C	.055	-.068	.003	-.073	-.161	-.046	.113	.030	-.020	1.00
MEAN	.247	7.76	8.89	25.1	89.9	25.8	5.40	4.91	4.86	.192
SD	.431	3.77	2.96	12.4	15.6	7.63	2.36	2.96	2.14	.394