Volume Title: Distribution of Economic Well-Being

Volume Author/Editor: F. Thomas Juster, ed.

Volume Publisher: NBER

Volume ISBN: 0-884-10478-8

Volume URL: http://www.nber.org/books/just77-1

Publication Date: 1977

Chapter Title: The Distribution of Earnings and Human Wealth in a Life-Cycle Context

Chapter Author: Lee A. Lillard

Chapter URL: http://www.nber.org/chapters/c4381

Chapter pages in book: (p. 557 - 620)

14

LEE A. LILLARD
National Bureau of Economic Research

The Distribution of Earnings and Human Wealth in a Life-Cycle Context

INTRODUCTION AND SUMMARY

Economists have long been interested in individual earnings differences and in the dispersion of earnings within populations. Recent development of explicit theoretical and empirical earnings functions from life-cycle human capital investment models increases the potential to explain existing earnings distributions and to predict changes in them. Life-cycle models suggest that current earnings are not a good index of well-being if choices about intertemporal transfers are available. Under certain conditions, the present value of earnings net of investments in human capital, human wealth, is an index of economic well-being. The purpose of this paper is to outline a set of conditions under which human wealth is an index of well-being in a life cycle, prefatory to empirical estimates of earnings and human wealth distributions for the 1960 Census population. Some tentative remarks on the interpretation of economic well-being in a life-cycle context when these conditions are not met are included. The basic conditions which allow human wealth to index well-being include the existence of a loan market for consumption

NOTE: This research was sponsored by National Science Foundation grant GS-31334 and U.S. Department of Labor grant L73-135 to the NBER. I have benefited from the comments of T. D. Wallace and Finis Welch. I wish to thank Christy Wilson for drawing the original figures.

expenditures, a fixed leisure-work time pattern, and no consumption of education or investment. If these are relaxed, appropriate adjustments to human wealth must be made.

The basic earnings equation used to predict earnings and human wealth is estimated on the NBER-Thorndike sample described later. Earnings are a function of age, schooling, and ability. This earnings function is used to predict earnings and human wealth distributions for the 1960 Census population, based on the joint distribution of age, schooling, and ability, as seen in age and schooling data from the 1960 Census of Population, and ability data within schooling classes from the NBER-Thorndike sample.

The purpose of this exercise is essentially to point out how earnings functions, which have been studied quite extensively, can be made more useful—that is, by predicting human wealth and by generating earnings distribution. Researchers often state: "If the distribution of such and such an independent variable has been this, then . . . ." These statements can be considered more formally, as I am trying to illustrate here. Even if the Thorndike sample is not like the 1960 Census (differences are noted later), the earnings function estimated from it can reproduce the general characteristics of the 1960 Census observed earnings distribution. In a previous paper, I have presented in more detail the statistical distribution theory necessary to go from the joint density of a population with respect to those characteristics which determine earnings through the estimated earnings function to a predicted earnings distribution. Even without any restrictive assumptions such as log-normality, the predicted distributions are positively skewed and the moments for subpopulations, such as schooling and age groups, behave similarly in actual and predicted distributions. The many caveats are pointed out in the paper.

Predicted earnings distributions are derived for the overall population, for schooling classes, for age groups, and for ability classes. Both the actual distribution and the distribution of earnings corrected for variation not explained by age, schooling, and ability are presented for each, along with selected summary statistics and Lorenz Curves. The predicted distributions reproduce the characteristics of the actual distributions for the 1960 Census population quite well, except for differences which can be explained between the 1960 Census population and the NBER-Thorndike sample.

Recognizing the degree of "fit" between predicted and observed 1960 Census earnings distributions and the reason for it, we then proceed to predict the distributions of mean human wealth, based on the same equations. This section attempts to estimate "What would be the distribution of the expected value of human wealth for employed men in the 1960 Census if they were like the NBER-Thorndike sample?"

Detailed mean human wealth distributions and selected statistics are presented, assuming a retirement age of sixty-six for several rates of discount. The sensitivity of the selected statistics, especially the mean, to discount rate and retirement age assumptions are then considered. Finally, some rough estimates of the variance of human wealth, rather than the variance of the mean, are constructed.

A lower bound on the variance of human wealth is defined as the variance in the present value of predicted earnings plus an error component which is completely transitory and independent from period to period. An upper bound is defined as the variance of the present value of predicted earnings plus a completely persistent error component which is constant over the life cycle but varies randomly over individuals, independently of the level of ability and schooling. Intermediate cases can be considered as combinations of these when the transitory and persistent variations are independent.

We study the effect of schooling level and of ability level on the distributions and on measures of inequality. These estimates are especially sensitive to discount rate assumptions. The effect of increased schooling level, for example, is to increase mean human wealth at discount rates below some level and to decrease mean human wealth at discount rates above that level. If this rate is below what we believe to be an appropriate discount rate—say, the rate appropriate to consumption loans or the real rate of return on physical assets—then the discrepancy could be accounted for by, for example, the consumption value of schooling or education discussed earlier. In this case, then, the human wealth measure is not a good index of economic well-being and the distribution of mean human wealth is not a good measure of the distribution of mean economic well-being. We may gain some insight into the partial effect of other attributes such as ability if they do not affect the consumption value of schooling. Ability increases the mean human wealth almost uniformly. Some inferences are made about the effect of retirement age on mean human wealth, but these results are tenuous, due to the limited upper age range in the sample.

MEASURING ECONOMIC WELL-BEING IN A LIFE-CYCLE CONTEXT

The life-cycle model is developed by assuming an individual maximizes lifetime utility, represented by an intertemporal utility function[1] within his opportunity set. Three components of the opportunity set are distinguished: endowment, market opportunities, and productive opportunities. All of these are relevant to an index of economic well-being.

Human capital investment models[2] assume that the individual has a homogeneous—across individuals and units within an individual—initial endowment of human capital, E_o, which can be rented in the labor market at the constant rate R per unit of time. This stock of human capital is subject to a given constant exogenous rate of deterioration δ, but the opportunity is available to use purchased inputs D, at price, P, and own human capital K to produce new human capital, according to the production function $Q(K, D)$. The net change in the stock of human capital at any point in time or age is then represented by $\dot{E}_a = Q(K_a, D_a) - \delta E_a$. These conditions relate to endowment and productive possibilities. Other endowments might include an initial endowment of nonhuman capital, an exogenous time stream of receipts or debts, and an exogenous time stream of educational inputs.[3]

Utility-maximizing behavior is clearly influenced by the existence or availability of market opportunities for intertemporal transfer of funds. When such funds are available, clearly earnings in a given time period cannot be considered an index of well-being.

There are many possible sets of assumptions. Consider market opportunities as they affect consumption, investment in human capital, and interperiod transfers of nonhuman wealth. The possibility of borrowing and loaning funds, endowed or earned, expands the permissible set of time paths of investment and consumption decisions. For example, there may be no market opportunities for borrowing or lending at all, in which case the individual must finance current investment in human and/or nonhuman capital and consumption out of current market earnings and exogenous receipts.

It is illustrative to introduce the concept of perfectly separable market opportunities—that is, funds borrowed for one purpose, consumption, investment in human capital, or investment in nonhuman capital, cannot be used for any other purpose. This is primarily introduced to capture the notion that investment in human capital accesses a different funds market, because (1) human capital is embodied in the individual, thus not subject to confiscation, which would imply a higher borrowing rate; and (2) there exist government-subsidized loan programs available only for educational investment at a lower rate. The nature of a perfectly separable funds market for financing direct educational expenditures, PD_a, will then affect only productive possibilities. Many additional constraints may also be imposed on the model, such as compulsory school attendance, and various school subsidy formulas.

Define:

$Y_a^* = R \cdot E_a$	Earning capacity at age a
$Y_a = R(E_a - K_a)$	Gross earnings at age a

$NY_a = R(E_a - K_a) - PD_a$	Earnings at age a net of direct educational investment
$I_a = RK_a + PD_a$	Total investment in human capital at age a
$N =$	Age at which working life and life cycle end, exogenous
$HW = \int_{t=0}^{N} e^{-rt} NY_t \, d_t$	Human wealth, present value of net earnings discounted at a rate dependent upon market opportunities

This development of the human capital model has ignored one sense of the time concept and has implicitly assumed that human capital is embodied in the individual, so that time and human capital enter the human capital production function in the same way. That is, $Q(K_a, D_a) = Q(SE_a, D_a)$, where S is the fraction of total time allocated to the production of new human capital. An equivalent model can be developed in terms of the use of time. A fuller discussion of the time interpretation is attempted in Ben-Porath (1967), Ghez (1972), Heckman (1974), and Lillard (1973).

The relevant index of lifetime economic well-being is lifetime utility. Consider a pedagogical construction under which human wealth defined as the present value of earnings net of educational investment is a relevant measure of economic well-being and the effect of failure to satisfy those conditions.

Human Wealth as an Index of Lifetime Well-Being

Human wealth is an index of economic well-being when the individual behaves in such a way as to maximize the present value of net earnings and there are no exogenous endowments of initial wealth or time stream of receipts or debts. The individual then maximizes his lifetime utility by arranging intertemporal consumption in an optimal manner, subject to the wealth constraint represented by human wealth. When exogenous endowments are present but do not affect the criteria of maximizing human wealth, their present value (positive for a time stream of receipts and negative for a time stream of debts) should be added to the wealth constraint and correspondingly to the index of economic well-being.

Under what conditions then will an individual behave in such a way as to maximize human wealth. We have already assumed the individual has perfect knowledge of himself and the world and faces no uncertainties. There is a fixed constant amount of time in each period to be allocated to either the labor market to produce earnings or to human capital production.[4] The utility function of the individual does not include as arguments either the stock of human capital or the use of time allocated to either the labor market or human capital production. This condition excludes the possibility that either investment or work is a more desirable activity, that obtaining education or going to school could be a consumption activity, and that the individual might derive utility directly from being more educated or highly trained. The individual has available a source of unlimited borrowing and lending at a constant rate of interest, r, for the purpose of consumption. This source of funds may or may not be available to finance educational expenditures as long as the loan markets are perfectly separable as defined earlier. If the unlimited funds are available for human capital investment, then the funds markets need not be separable and the model corresponds to the Ben-Porath (1967) specification. However, the loan market for human capital may contain any sort of imperfection as long as it is separable. This loan market may include low interest loans from parents or government agencies, high interest loans due to the embodied nature of human capital, or in the extreme no loan market for human capital investment expenditures at all. Under these conditions, clearly the relevant rate of discount of net earnings is the interest rate, r, on loans for consumption purposes.

The particular life cycle of earnings model specified by these conditions, assuming no loan market for direct educational expenditures and a Cobb-Douglas production function,[5] is capable of being fully solved analytically, which illustrates the simultaneity of schooling and earnings while providing an exact functional form for earnings and human wealth. This solution is exposited fully in Lillard (1973) and only summary results are presented here.

The solution implies that in the early period the individual specializes in the production of new human capital, full-time schooling, using all of his earning capacity for investment.[6] The period of specialization is

$$0 \leqq a \leqq a^*$$

where a^* denotes the age at which the individual stops specializing and begins investing only a fraction of his earning capacity. Specialization ends when earning capacity ceases to be an effective constraint on investment. One implication of assuming no loan market for educational expenditures, and the only qualitative difference from the Ben-Porath perfect loan market case is the prediction of positive labor force participation during the period of specialization. The individual supplies a constant fraction[7] of his human capital to the market to finance expenditures for direct educational expenditures, i.e., $R(E_a - K_a) = PD_a$.

Specialization with no loan market means investing exactly all of earnings capacity in the form of forgone earnings and purchased inputs. Specialization with the same perfect loan market available means using all of human capital in production and borrowing to finance purchased inputs. There are many intermediate assumptions, including availability of special loan markets, scholarships, and so on,[8] which may be available only during the period of specialization or formal full-time schooling. The effect of these conditions is summarized in the stock of human capital, earning capacity, at a^*. This earning capacity at a^* depends upon initial earning capacity, RE_o. It is important to note that the solution for earnings after the period of specialization takes earning capacity at a^* as a datum, both earnings and a^* are endogenous state variables and any exogenous change which affects earnings will also affect the length of time in specialization, and both must be considered jointly.

The length of the period of specialization is endogenous to the model. The optimum age to stop specializing in production and begin positive net earning is that point where the investment paths of the two regions cross. That is, the individual will invest according to the rule K_a and D_a for nonspecialization, except when he is constrained by his earning capacity, during which period he will invest all of his earning capacity. The solution for a^* as a function of the parameters and initial endowment of human capital, but not age, is an implicit simultaneous structural relationship which must be satisfied for each solution. The implicit solution for a^* must be considered simultaneously with earnings function to make any inferences. The expression allows inferences about the direction of effect of each characteristic on the length of the period of specialization.

For the particular solution reported in Lillard (1973), the length of the specialization period varies directly with N, R, and β, and inversely with E_o, P, and r. The effect of all other characteristics is ambiguous.[9]

For the rest of the life cycle, after the period of specialization ends, $a^* \leqq a \leqq N$, the individual invests some fraction of his earning capacity in producing more human capital. Neither forgone earnings nor direct educational expenditures, and thus investment in human capital, is a function of the initial stock of human capital E_o.[10] Gross investment declines with age after the period of specialization, reaching zero at retirement age N.[11] Earning capacity, observed earnings, and net earnings at any age after a^* depend upon the stock of human capital and the investments at that age. All of these results for the specific solution are presented in greater detail in Lillard (1973).

Given these assumptions so that human wealth is an index of well-being, what then does human wealth depend upon? As we have noted, it depends upon access to borrowing funds to finance human capital investment. Clearly access to such a loan market expands investment possibilities and enhances human wealth. Also, individuals may differ in the efficiency with which human capital is produced, the production parameters β, β_1, and β_2 in the specific model above. More efficiency in producing new human capital clearly increases human wealth. An empirical counterpart to β is introduced later.

An increase in the retirement age N, or a decline in the rate of interest, will clearly increase human wealth. A decline in the rate at which human capital deteriorates, δ, will clearly increase human wealth. Individuals may differ in some or all of these parameters. For empirical purposes, we shall assume that they differ only in ability representing efficiency of production, and schooling representing a^*. The effect of increased schooling on human wealth is less clear, since it represents the effect of all other differences between individuals, and these differences must satisfy the implicit simultaneous schooling relationship.

When Human Capital Is Not an Index of Well-Being

The life-cycle model makes it clear that when individual inter-temporal choice is available, individual period earnings are a myopic measure of well-being. Under certain conditions, when inter-temporal consumption choices are perfectly free, human wealth is a measure of lifetime well-being, and individual-period earnings observations and the age-earnings profile itself merely illustrate the optimal timing of a separable process. When these very stringent conditions are not met, the problem of indexing well-being falls ultimately back to considerations of the inter-temporal utility function. Human wealth and the lifetime pattern of earnings become variables of choice. Constructing an index based on observable values becomes extremely complex. The relevant models of life-cycle behavior have not yet been fully developed or analyzed. The problem is not solved here but relaxation of certain conditions one at a time may lend some additional insight into the problem. Let us begin with relatively simple deviations with the clearest implications.

The first potential problem is that schooling or education or the level of investment in human capital may enter the utility function directly. Alternatively, utility may be a function of the stock of human capital held by the individual—say, as a status measure, or by affecting the efficiency of consumption (see Michael [1972]). In these cases an investment in human capital yields returns not measured in the present value of net earnings. Human wealth will understate return to education. Human wealth may decline with increased schooling, discounted at the consumption borrowing rate, while total inter-temporal utility rises.

Secondly, consider the effect of allowing leisure time, as well as investment and work time, to be a subject for choice. This is the most widely considered generalization and has not yet been satisfactorily treated. Heckman (1973) and Stafford and Stephan (1973) attempt to model this generalized problem. Few specific conclusions have been obtained. Smith (1973) analyzes the problem of labor-leisure choice, assuming wages are exogenous. In the more general model, the individual must make inter-temporal choices about consumption of leisure and goods. The leisure-investment-work choice makes earnings endogenous and the result of previous decisions. We cannot say whether the present value of net earnings overstates or understates economic well-being. This depends on the individual's relative valuation of goods and leisure and their timing over the life cycle. What is needed is a measure of "full wealth." The usual suggestion is to value leisure time at the market wage and consider the net worth of total time. This approach seems to be inappropriate if the wage is endogenous. The individual will "choose" a low investment and wage pattern "because" he values his leisure time more. The full wealth at market-wage correction works in the wrong direction. What is needed in this situation is an index of initial endowments—say, of human capital—and the constraints the individual faces. A larger initial endowment makes an individual unambiguously better off, even if he chooses a lower value of human wealth than an individual beginning with less. This does not get us very far empirically but is meant as food for thought.

Another obvious omitted concept is nonhuman wealth, which must be included in any wealth calculations. The existence of initial nonhuman wealth clearly affects the access of the individual to funds for financing educational investments.

The effect of risk and uncertainty on investment in human capital is considered briefly by Levhari and Weiss (1973) and Razin (1973). Again the problem is exceedingly difficult and clear implications are few.

These tenuous statements are meant only as caveats in the interpretation of the empirical estimates which follow.

1960 CENSUS: PREDICTED EARNINGS DISTRIBUTIONS AND THE DISTRIBUTION OF HUMAN WEALTH

The previous sections considered the appropriateness of certain measures of economic well-being. This section considers the distribution of well-being if it is measured by either earnings or human wealth. Both the overall distributions and distributions within schooling and ability classes and age classes, where appropriate, will be considered. The format is to consider an earnings equation estimated using the NBER-Thorndike sample data, then to predict aggregate earnings distributions for the 1960 Census. The estimated age-earnings equations are a function of schooling and ability levels. This section may be characterized as answering the query, "What would be the distribution of earnings of the men in the NBER-Thorndike sample if they had the distribution of age and schooling present in the 1960 Census?" or "What would be the distribution of earnings of employed men in 1960 if they were like the men in the NBER-Thorndike sample?" As will be pointed out later, several caveats are in order in using one group to predict the other. Predicted and actual 1960 distributions are compared when possible.

Recognizing the degree of "fit" between predicted and observed 1960 Census earnings distributions and the reason for it, we then proceed to predict the distributions of mean human wealth based on the same equations. This section attempts to estimate "what would be the distribution of the expected value of human wealth" either "of the men in the NBER-Thorndike sample if they had the schooling distribution present in the 1960 Census" or "of employed men in the 1960 Census if they were like the NBER-Thorndike sample." Detailed mean human wealth distributions and selected statistics are presented, assuming a retirement age of sixty-six for several rates of discount. The sensitivity of the selected statistics, especially the mean, to discount rate and retirement age assumptions are then considered. Finally, some rough estimates of the variance of human wealth, rather than the variance of the mean, are constructed.

A lower bound on the variance of human wealth is defined as the variance in the present value of predicted earnings plus an error component which is completely transitory and independent from period to period. An upper bound is defined as the variance of the present value of predicted earnings plus a completely persistent error component which is constant over the life cycle but varies randomly over individuals, independently of the level of ability and schooling. Intermediate cases can be considered as combinations of these when the transitory and persistent variations are independent.

The primary conclusions are that aggregate earnings distributions can be reproduced reasonably well even with the crude calculations made here, and that it is possible to generate estimates of human wealth distributions. In doing so, we can study the effect of schooling level and of ability level on the distributions and on measures of inequality. These estimates are especially sensitive to discount rate assumptions. The effect of increased schooling level, for example, is to increase mean human wealth at discount rates below some level and to decrease mean human wealth at discount rates above that level. The cutoff rate is in the neighborhood of 5.5 percent. If 5.5 percent is below what we believe to be the appropriate discount rate—say, the rate appropriate to consumption loans or the real rate of return on physical assets—then the discrepancy could be accounted for by, for example, the consumption value of schooling or education discussed earlier. In this case, then, the human wealth measure is not a good index of economic well-being and the distribution of mean human wealth is not a good measure of the distribution of mean economic well-being. We may gain some insight into the partial effect of other attributes, such as ability, if they do not affect the consumption value of schooling. Ability increases the mean human wealth almost uniformly. Some inferences are made about the effect of retirement age on mean human wealth, but these results are tenuous, due to the limited upper age range in the sample.

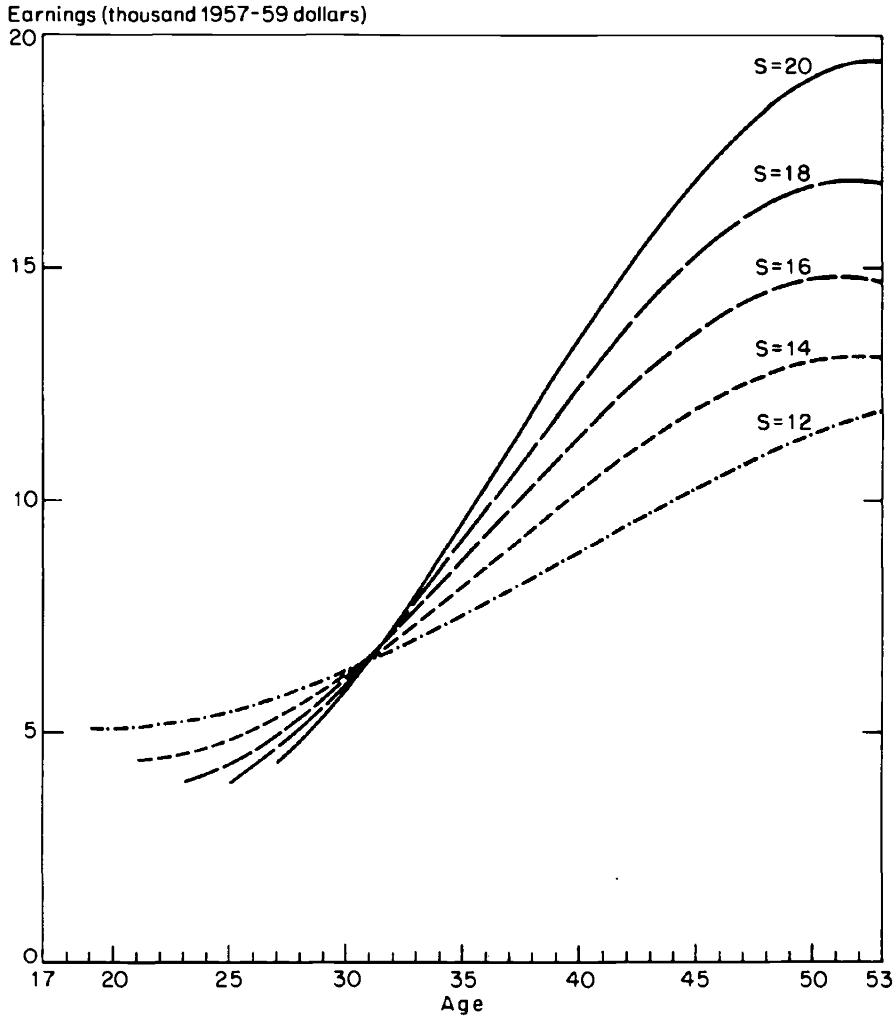
A Specific Earnings Function and Estimates

It is well founded theoretically and empirically that earnings depend upon schooling, ability, and age or experience.[12] The earnings function estimated and used here results from a life cycle of earnings model which is discussed elsewhere in detail, along with the empirical estimates.[13] The estimated earnings function is cubic in age, quadratic in schooling, and cubic in ability, including all interactions. This is the "best equation," in the sense that the age, schooling, and ability polynomials were determined by error variance criteria.[14] The estimated earnings function is[15]

$$
\begin{aligned}
Y(A, S, B) = {} & 21108.50 - 3921.20A + 877.25S + 148.02SA + 206.09A^2 \\
& - 794.20S^2 + 6.87SA^2 + 116.42S^2A - 7.82S^2A^2 - 45197.00B \\
& + 11015.00BA + 4721.40BS - 1820.80BSA - 594.93BA^2 \\
& + 1065.00BS^2 + 83.51BSA^2 - 122.05BS^2A + 8.56BS^2A^2 \\
& + 28134.00B - 6738.40B^2A - 5035.20B^2S + 1435.20B^2SA \\
& + 371.38B^2A^2 - 240.65B^2S^2 - 72.59B^2SA^2 + 5.86B^2S^2A \\
& + 0.99B^2S^2A^2 - 2.99A^3 - 0.31A^3S + 0.15A^3S^2 + 9.09BA^3 \\
& - 1.04BA^3S - 0.17BA^3S^2 - 5.74B^2A^3 + 1.04B^2A^3S \\
& + 0.03B^2A^3S^2
\end{aligned}
$$

where A = age, S = years, and B = ability index. The resulting age-earnings profiles are presented in Figures 1A, 1B, and 1C for various ability and schooling levels. Both schooling and ability raise earnings at

FIGURE 1A Cubic Estimated Age-Earnings Profiles Based on the NBER-Thorndike Sample for Several Schooling Levels at the Average Ability Level



NOTE: All earnings are in 1957–59 dollars.

FIGURE 1B Cubic Estimated Age-Earnings Profiles Based on the NBER-Thorndike Sample for Average Ability and One Standard Deviation (.25) Above and Below for High School Graduates (S = 12)

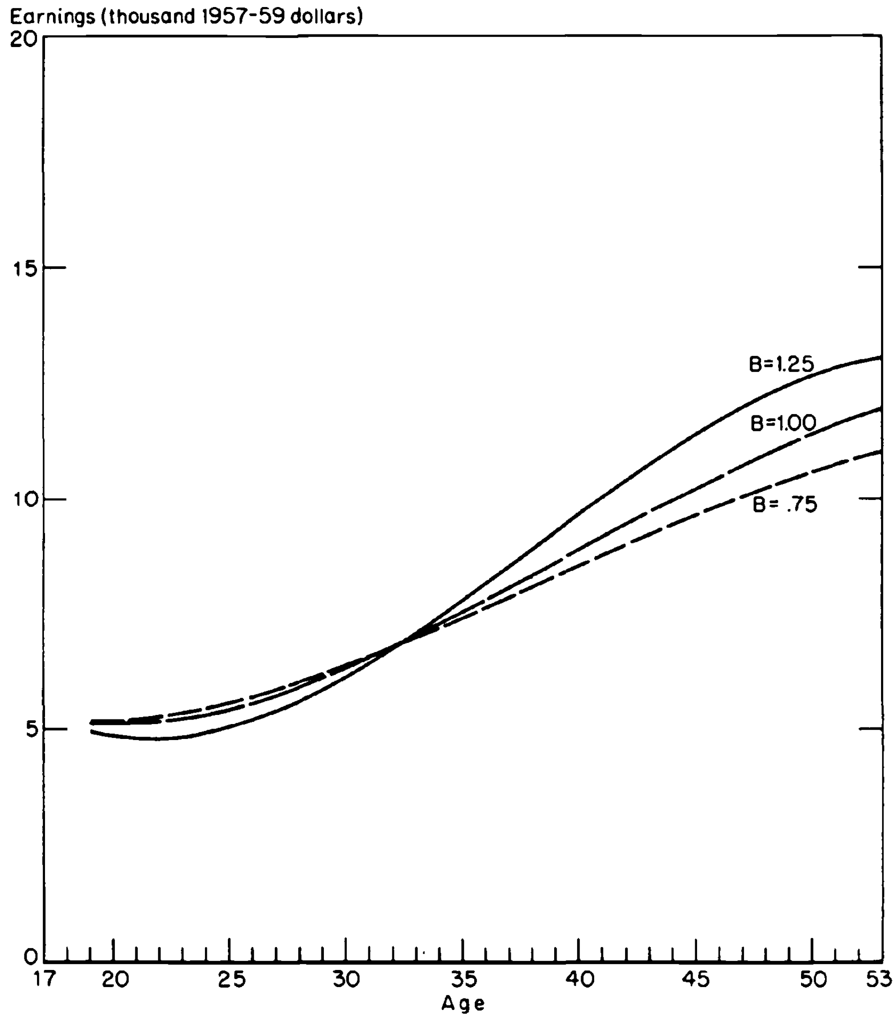
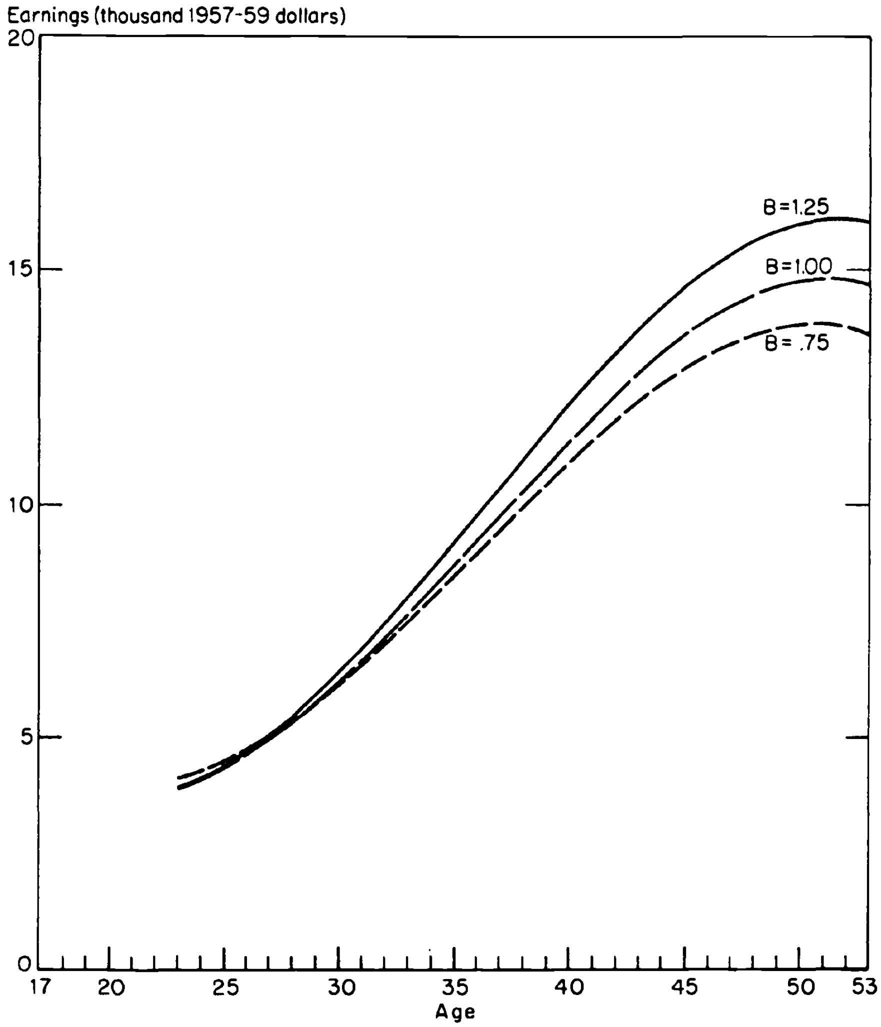


FIGURE 1C Cubic Estimated Age-Earnings Profiles Based on the NBER-Thorndike Sample for Average Ability and One Standard Deviation (.25) Above and Below for College Graduates (S = 16)



every age in the life cycle after some initial period.[16] Earnings estimated beyond age fifty-six are a pure prediction, in the sense that there are no individuals in the sample beyond that age. The resulting estimates of human wealth, defined as the present value of predicted earnings, are presented in Figures 2A and 2B for discount rates of 3, 5, and 7 percent.[17]

Consider the characteristics of the NBER-Thorndike sample which may make it different from the general population described in the 1960 Census. The NBER-Thorndike sample is based on a group of males volunteering for Air Force pilot, navigator, and bombardier programs in the last half of 1943. These volunteers were given initial screening tests and a set of seventeen tests to measure various abilities[18] in 1943. Thorndike and Hagen sent a questionnaire to a sample of 17,000 of these men in 1955, which included a question on 1955 earnings. In 1969, the NBER sent to a subset of these men a subsequent questionnaire, which included additional questions on earnings in later years and questions on schooling and initial job earnings.

The data include five separate approximately equally spaced points[19] on the age-income profile as well as the year of initial job, year of last full-time schooling, years of schooling, and seventeen separate measures of ability. The age-income points are approximately initial job, 1955, 1960, 1964, and 1968. The individuals in the Thorndike sample differ from the U.S. male population as a whole in several ways.[20] First, the sample includes a high-ability group. All of the men completed high school or high school equivalency examinations and passed the initial screening for the Air Force flight program. Their general health was better than the general population[21] in 1969. They were more homogeneous in height and weight due to military qualifications. They seem to have a high degree of self-confidence, self-reliance, and risk preference. They tend to be entrepreneurs. An unusual 20 percent work longer hours. Some of these factors may, however, be related to the high ability. The observed age range is nineteen to fifty-seven years but with less than 1 percent outside the range nineteen to fifty-five. The cubic earnings equation is quite a poor prediction above this range, since predicted earnings drop rapidly to large negative values; therefore, earnings are assumed constant at their peak level after the peak occurs.[22]

Earnings Distributions from the Estimated Earnings Function

The distribution of earnings derives from the distribution of the population with respect to age, ability, and schooling. Our predictions use 1960 U.S. Census of Population data on the distribution of the United States

FIGURE 2A Present Value of Predicted Observed Earnings from the Estimated Age-Earnings Profiles Based on the NBER-Thorndike Sample as a Function of Schooling (N = 66); Discounted at 3 Percent, 5 Percent, and 7 Percent

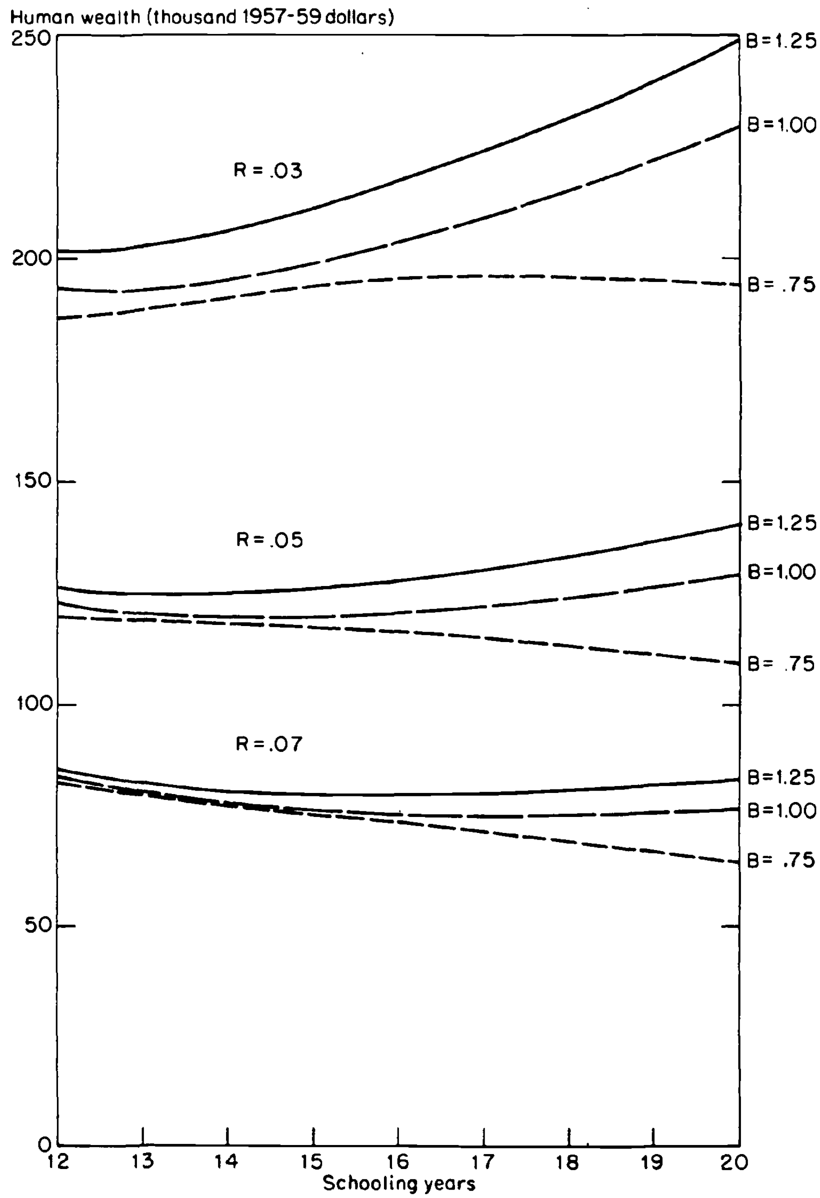
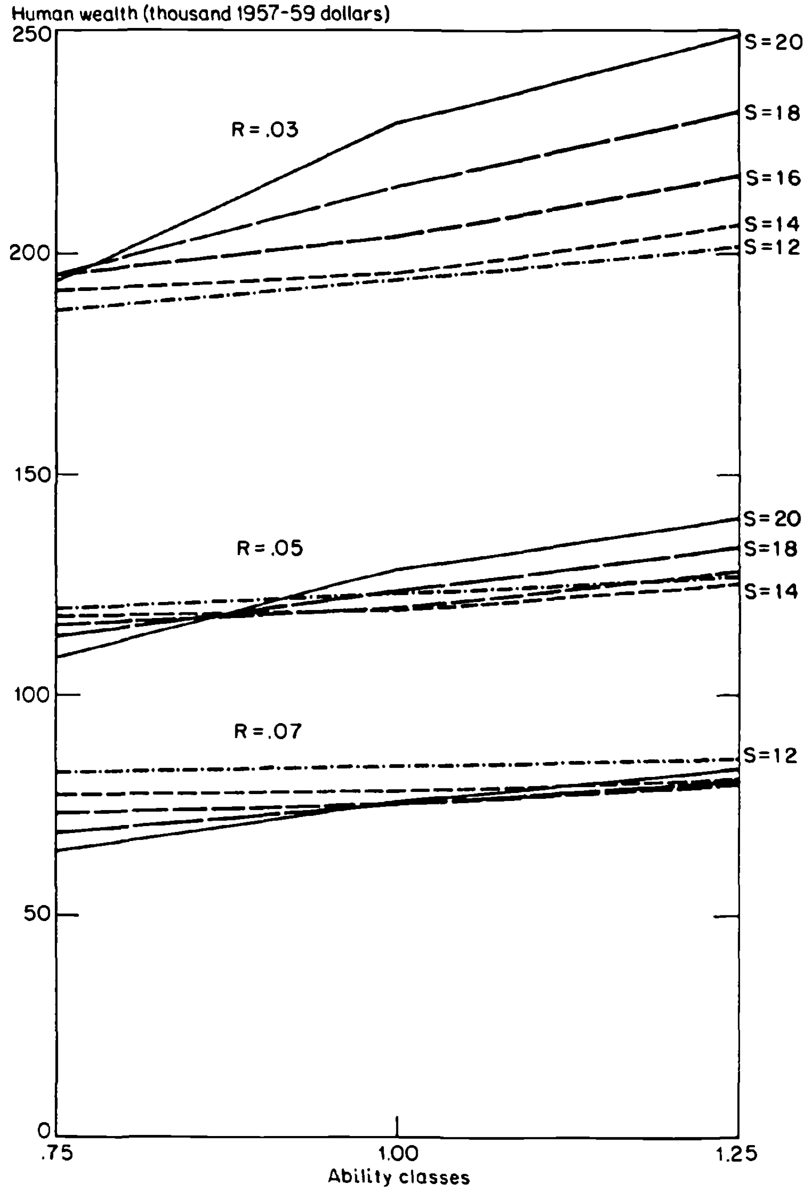


FIGURE 2B Present Value of Predicted Earnings as a Function of Ability from the Estimated Age-Earnings Profiles Based on the NBER-Thorndike Sample (N = 66); Discounted at 3 Percent, 5 Percent, and 7 Percent



population of males eighteen years old and over by labor force status, years of school completed, and age to predict earnings distributions based on the estimated earnings function.[23] A general framework for translating the joint density of age and characteristics which determine earnings through the earnings function into earnings or human wealth density is presented in Lillard (1973a).

Since the earnings function predicts earnings only after the end of full-time schooling, the distribution of the population by age and schooling is taken only for persons employed and in the civilian labor force. The joint and marginal distributions of age and schooling are presented in Table 1. Since all persons in the NBER-Thorndike sample

TABLE 1 Joint and Marginal Distributions of Age and Schooling for Employed Males Eighteen to Sixty-Four Years of Age with at Least a High School Education, from the 1960 Census of Population

	Years of Schooling				
	12	13–15	16	17+	Age Marginal
18–19	.0247	—	—	—	.0247
20–21	.0280	.0146	—	—	.0426
22–24	.0455	.0182	.0080	—	.0716
25–29	.0803	.0313	.0215	.0148	.1480
30–34	.0793	.0306	.0241	.0194	.1534
35–44	.1670	.0560	.0361	.0330	.2920
45–54	.0980	.0405	.0218	.0216	.1819
55–64	.0399	.0227	.0123	.0108	.0858
Schooling marginal	.5628	.2139	.1237	.0996	1.0

have at least a high school education, predictions are restricted to that population. That is, the distribution of yearly earnings is predicted for persons who are between the ages of eighteen and sixty-four, have at least a high school education and are employed.[24] The distribution of the population with respect to ability is assumed to be the same as the NBER-Thorndike sample on which the earning function was estimated, since no ability data are reported in the 1960 Census of Population. Statistics for the distribution of ability by schooling class used is presented in Table 2.[25] For calculation of predicted yearly income, it is assumed that all individuals in an age or schooling class are at the midpoint of that class.[26]

TABLE 2 Selected Statistics for the Distribution of the Ability Index Overall and by Schooling Level from the NBER-Thorndike Sample for Schooling Interval Midpoints

	Mean	Standard Deviation
Overall	1.00	.25
By schooling:		
12 years	0.910	.219
14 years	0.971	.229
16 years	1.063	.255
18 years	1.071	.261

Yearly earnings are calculated for each age, schooling, ability combination corresponding to midpoints of class intervals. Each calculated yearly income assumes the relative frequency of the corresponding age, schooling, ability combination. The relative frequency of any (A, B, S) combination is calculated as the joint relative frequency of the age, schooling combination reported by the Census of Population times the relative frequency of the ability level within that schooling class.[27] These relative frequencies are then summed into relative frequencies of yearly earnings for intervals of a thousand dollars.[28]

The resulting predicted overall distribution of earnings and the predicted distribution for various subpopulations effectively represent distributions of mean earnings allowing no variation around the predicted value. However, only about 28 percent of the variation in earnings is explained by variation in age, schooling, and ability.

Consider the problem of correcting the distribution of earnings for variation not accounted for by variation in age, schooling, and ability. The error variance of the estimating equation is $\hat{\sigma}^2 = 36{,}593{,}472$ (standard error = 6,049.25). It is assumed that the errors are identically and independently[29] distributed, with mean zero and standard deviation 6,049.25. The obvious first-order approximation is simply to correct the standard deviations of the various distributions by using, for example

$$Y = \sqrt{\mathrm{Var}_{A,S,B}[Y(A,S,B)]} = \sqrt{\mathrm{Var}_{A,S,B}[Y(A,S,B)] + \hat{\sigma}^2}$$

This correction is unsatisfactory, however, because of the possibility of assigning a positive frequency to negative earnings, and it is desirable to see the effect on statistics other than the variance. Another simple approximate procedure based on the truncated normal is used to construct the distributions themselves, then selected statistics are calculated from these distributions.[30] This procedure is not entirely satisfactory either, since the truncation increases the mean and decreases the dispersion, but it allows a crude approximation. The probability density for any individual age, schooling, ability combination is calculated as before, but the density is allocated to earnings intervals according to the above-normal distribution centered on the midpoint of the interval in which the predicted value falls. This is an admittedly crude but simple correction. Better corrections can no doubt be obtained through more complex calculations. The interval in which the predicted earnings value falls receives an incremental relative frequency of .0662 times the relative frequency of that age, schooling, ability combination. Intervals adjoining the central interval receive an incremental relative frequency of .0643 times the relative frequency of (A, S, B) each, and so forth, until all relative frequency of the error is exhausted.

Finally, the actual distribution of earnings for employed males sixteen to sixty-four yours old with at least a high school education is calculated from more general distributions reported in the 1960 Census of Population.

All three overall earnings distributions and the corresponding Lorenz curves are presented in Figures 3A and 3B. Selected statistics and relative frequency tables are included in the tables of individual type distribution subsections.

The major caveats may be summarized as follows. The NBER-Thorndike sample and the population of employed males in 1960 differ in several ways, the most important of which is the high level of ability present in the NBER-Thorndike sample. Even though ability distributions by schooling class are used, the distribution of ability especially in lower schooling classes will overstate ability relative to the actual distribution in the 1960 population. The 1960 population is heavily concentrated at lower levels of schooling, especially high school, which is at the lower end of the range of observation for the Thorndike sample and thus subject to less confidence in estimation. Interval midpoints with respect to schooling are used for schooling classes 13–15 (14) and 17+ (18). More precise information about the distribution within these intervals would sharpen the prediction.

Predictions beyond age fifty-six are made assuming earnings constant after peak earnings. This is necessary due to the data limitations in the NBER-Thorndike sample. The age distribution used from the 1960 Census assumes individuals are at the midpoint of age intervals that increase in length from two years at early ages to ten at late ages. Approximately 10 percent of the 1960 population falls in the least reliable age interval, 55–64.

The unequal intervals also cause problems in comparing predicted and actual earnings distributions. Predicted distributions can be made for any interval groups and are made for equal $1,000 intervals here. The Census of Population earnings distributions are unequal beyond $7,000. Statistics are computed using interval midpoints and will vary with

FIGURE 3A Predicted Mean Earnings, Predicted Corrected Earnings, and Actual Income Distributions for Employed Males between the Ages of 18 and 64 with at Least a High School Education

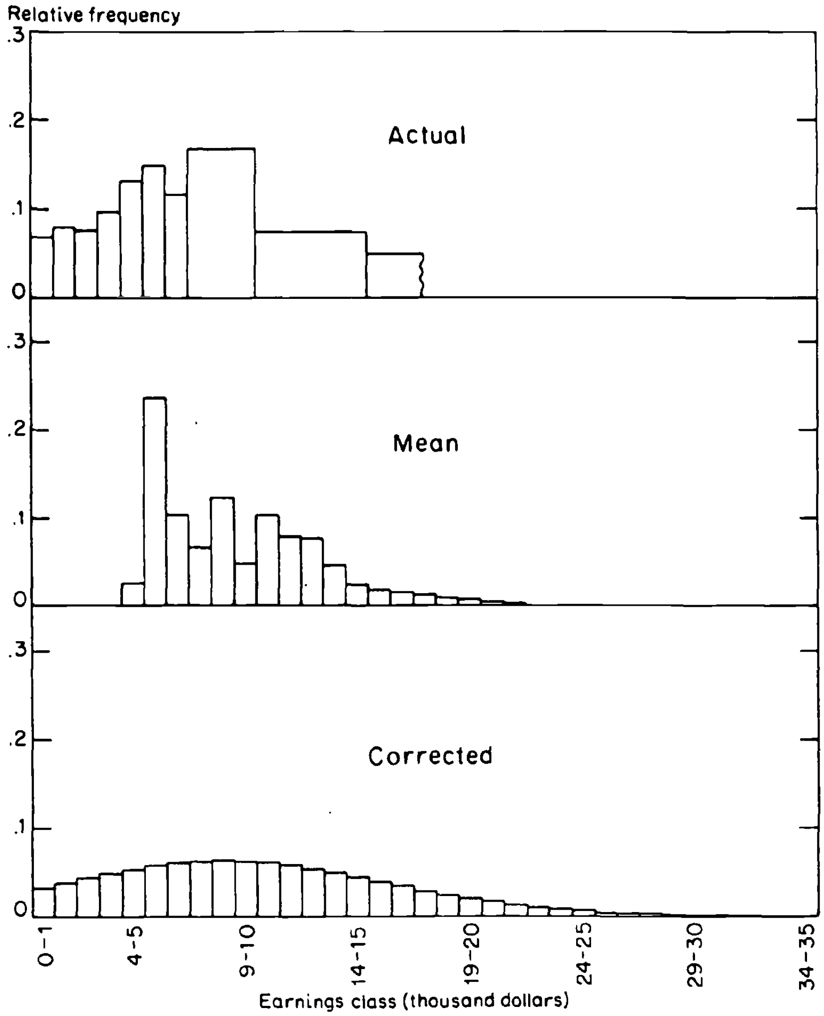
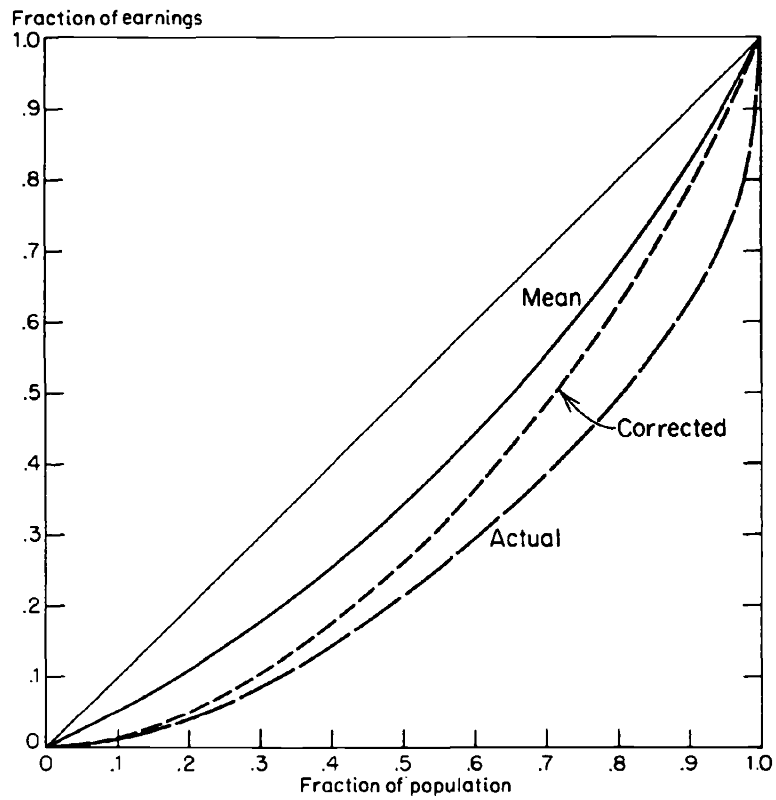


FIGURE 3B Lorenz Curves for Predicted Mean Earnings, Predicted Corrected Earnings, and Actual Income Distributions for Employed Males between the Ages of 18 and 64 with at Least a High School Education



different groupings. Interval midpoints predicted by the Pareto method were used for the interval fifteen thousand dollars and over in the Census of Population, while equal 1,000 intervals up to 90,000 are used for predicted distributions.

Several important differences remain. The 1960 Census figures are for total income, whereas the predicted figures are for earnings in the labor market. There may be important differences in weeks worked during the year, and hours worked during the week, between the sample and the population. There are indications that the men in the NBER-Thorndike sample tend to work longer hours and to spend less time unemployed.

Another very important difference is that the 1960 Census figures include employed students, whereas these persons are excluded in estimating the earnings function. This contributes to the large relative frequency of very low income at early ages in the actual Census distribution. For example, 53 percent of eighteen and nineteen year olds earned less than $1,000. These are likely to be employed students.

Predicted Mean Earnings Distributions

These earnings distributions are derived by transforming probability density from three-dimensional (age, schooling, ability) space through the estimated earnings function into the earnings dimension. Since age, schooling, and ability are not the only characteristics of an individual which determine earnings, these may be termed expected or mean earnings distributions. They are the distribution of the expected value of earnings.

Selected statistics relating to the earnings distributions are presented in Table 3. The relative frequency distributions for selected subgroups are in Figures 4A, 4B, 4C, and 4D.

Predicted Earnings Distributions Corrected for Unexplained Variation

These earnings distributions are mean earnings distributions corrected for variation in earnings not explained by age, schooling, and ability. Instead of transforming density from (age, schooling, ability) space into a single earnings point, it is spread over the positive real line in a manner proportional to the normal probability density, with its center at the predicted mean value and standard deviation equal to the estimated standard error of the regression.

Selected statistics are presented in Table 4. Relative frequency distributions for selected subgroups are presented in Figures 5A, 5B, 5C, and 5D.

Actual Earnings Distributions

These earnings distributions are those actually observed in the 1960 Census. Again they include total income and include employed students. Selected statistics are presented in Table 5. Relative frequencies for selected subgroups are presented in Figures 6A, 6B, and 6C.

TABLE 3 Selected Statistics for Predicted Mean Earnings Distributions for Employed Males Eighteen to Sixty-Six with at Least a High School Education, Based on an Earnings Function from the NBER-TH Sample

	Mean ($)	Median ($)	Standard Deviation ($)	Coefficient of Variation	Skewness	Gini Coefficient
Overall	9,182	8,545	3,530	.38	0.86	.21
By schooling:						
12 years	8,188	8,214	2,412	.29	0.50	.16
14 years	9,478	10,246	3,344	.35	0.11	.20
16 years	10,607	11,188	4,281	.40	0.20	.23
18 years	12,391	12,799	5,185	.42	0.03	.24
By age:						
19 years	5,679	5,582	506	.09	3.95	.03
21 years	5,556	5,518	503	.09	2.57	.03
23 years	5,457	5,455	585	.11	1.52	.04
27 years	5,528	5,504	484	.09	2.35	.03
32 years	7,011	6,921	622	.09	1.24	.04
39 years	9,992	9,617	1,632	.16	0.99	.09
49 years	13,283	12,479	2,681	.20	1.08	.11
59 years	13,833	12,917	2,599	.19	1.51	.10
By ability:						
<.75	8,359	8,313	2,537	.30	0.34	.17
.75–1.00	8,732	8,310	3,062	.35	0.65	.19
1.00–1.25	9,474	9,223	3,738	.39	0.62	.22
>1.25	11,241	10,739	4,892	.44	0.51	.25

NOTE: Skewness is measured by the square root of $E(X - \bar{X})^3/S^3$. Coefficient of variation is $S/\bar{X}$. The ability index is distributed with mean 1.0 and standard deviation .25.

FIGURE 4A **Distribution of Predicted Mean Earnings for Men 30–34 Years of Age**

FIGURE 4B **Distribution of Predicted Mean Earnings for Men 45–54 Years of Age**

FIGURE 4C **Distribution of Predicted Mean Earnings for College Graduates**

FIGURE 4D **Distribution of Predicted Mean Earnings for Men within One Standard Deviation Above Mean Ability (1.00–1.25)**

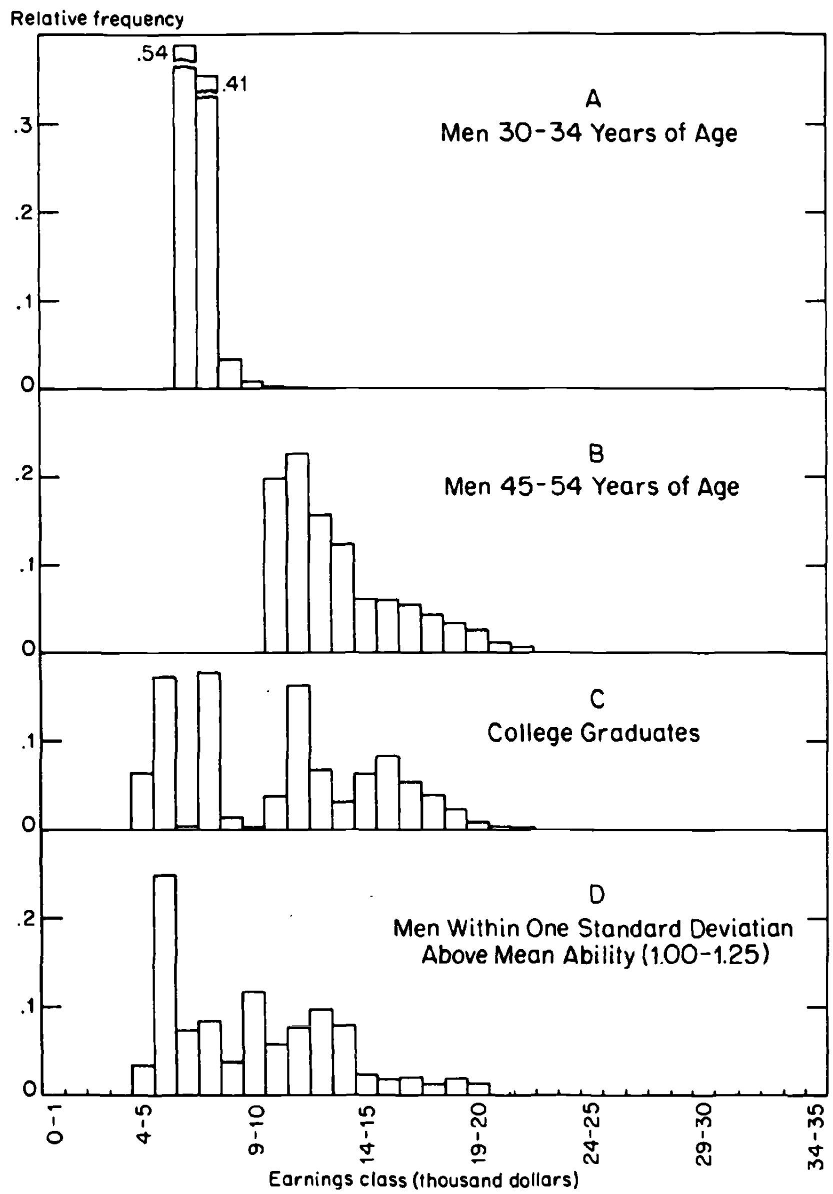


TABLE 4 Selected Statistics for Predicted Mean Earnings Distributions Corrected for Unexplained Variation

	Mean ($)	Median ($)	Standard Deviation ($)	Coefficient of Variation	Skewness	Gini Coefficient
Overall	10,201	9,652	5,971	.59	.53	.33
By schooling:						
12 years	9,343	8,883	5,433	.58	.48	.33
14 years	10,423	9,989	5,902	.57	.42	.32
16 years	11,435	10,959	6,447	.56	.42	.32
18 years	13,041	12,663	7,102	.54	.34	.31
By age:						
19 years	7,535	7,016	4,703	.62	.57	.35
21 years	7,460	6,935	4,678	.63	.58	.35
23 years	7,403	6,871	4,663	.63	.59	.35
27 years	7,442	6,917	4,672	.63	.58	.35
32 years	8,394	7,955	4,959	.59	.49	.33
39 years	10,658	10,373	5,574	.52	.33	.30
49 years	13,538	13,336	6,257	.46	.24	.26
59 years	14,032	13,859	6,283	.45	.24	.25
By ability:						
< .75	9,443	9,125	5,527	.59	.46	.33
.75–1.00	9,813	9,311	5,716	.58	.50	.33
1.00–1.25	10,457	9,912	6,081	.58	.50	.33
> 1.25	12,015	11,420	6,864	.57	.48	.33

FIGURE 5A **Predicted Earnings Distribution Corrected for Unexplained Variation for Men 30–34 Years of Age**

FIGURE 5B **Predicted Earnings Distribution Corrected for Unexplained Variation for Men 45–54 Years of Age**

FIGURE 5C **Predicted Earnings Distribution Corrected for Unexplained Variation for College Graduates**

FIGURE 5D **Predicted Earnings Distribution Corrected for Unexplained Variation for Men within One Standard Deviation Above Mean Ability (1.00–1.25)**

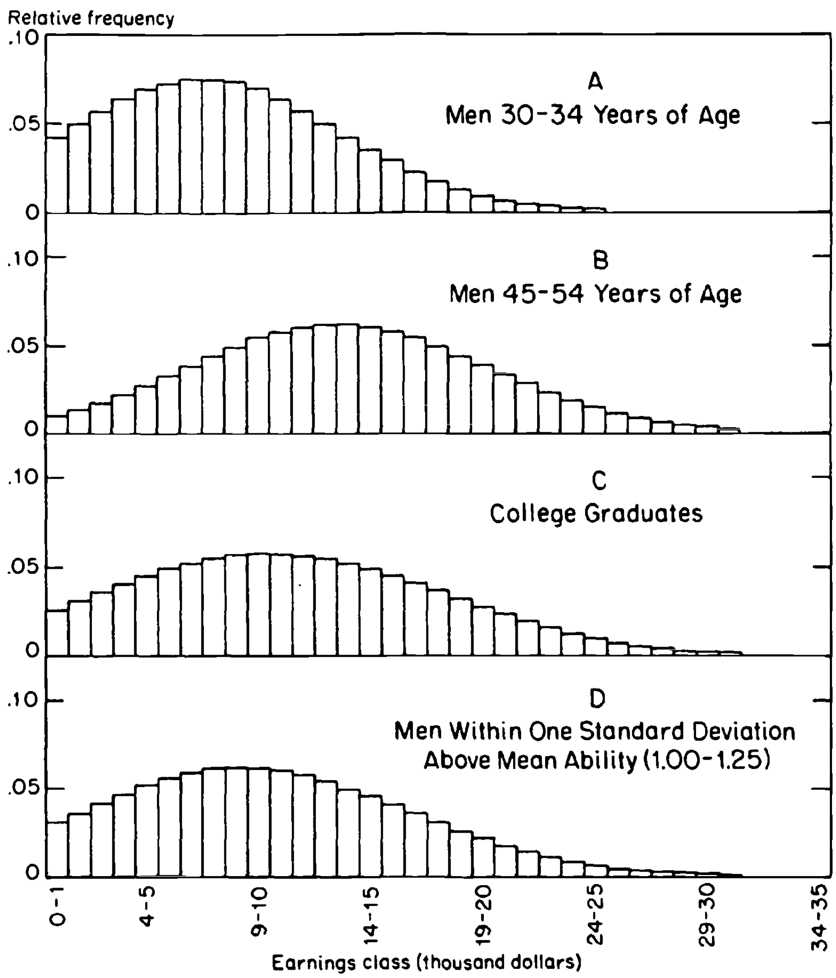


TABLE 5 Selected Statistics for the Actual Distribution of Earnings for Employed Males between the Ages of Eighteen and Sixty-Four with at Least a High School Diploma

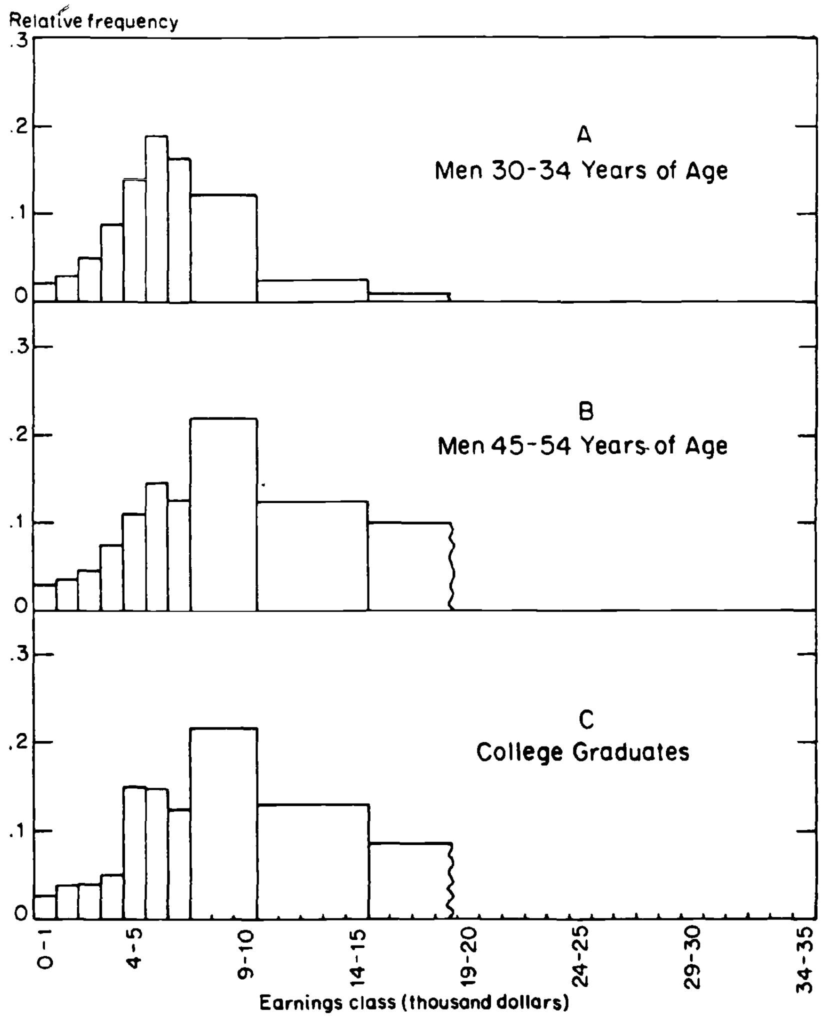
	Mean ($)	Median ($)	Standard Deviation ($)	Coefficient of Variation	Skewness	Gini Coefficient
Overall	6,429	5,358	5,356	0.83	2.24	.39
By schooling:						
12 years	5,346	4,933	3,696	0.69	1.42	.34
13–15 years	6,181	5,212	5,464	0.88	1.47	.42
16 years	8,756	6,424	7,650	0.87	1.48	.39
17+ years	12,334	7,258	12,475	1.01	1.26	.47
By age:						
18–19	1,399	993	1,315	0.94	3.36	.43
20–21	2,313	1,889	1,774	0.77	2.16	.39
22–24	4,023	4,109	2,299	0.57	1.03	.31
25–29	5,092	4,884	2,726	0.54	1.62	.27
30–34	6,674	5,940	3,839	0.58	2.01	.28
35–44	7,900	6,471	5,482	0.69	2.09	.33
45–54	9,050	6,529	8,006	0.88	2.01	.40
55–64	9,704	6,164	10,424	1.07	1.99	.48

NOTE: The interval means for the open-ended interval, $15,000 and above, are calculated separately for each class by the Lorenz procedure.

FIGURE 6A **Actual Distribution of Total Income Reported in the 1960 Census of Population for Men 30–34 Years of Age**

FIGURE 6B **Actual Distribution of Total Income Reported in the 1960 Census of Population for Men 45–54 Years of Age**

FIGURE 6C **Actual Distribution of Total Income Reported in the 1960 Census of Population for College Graduates**



Comparison of Mean Earnings, Corrected Earnings, and Actual Earnings Distributions

It should be remembered that any comparisons between predicted and actual distributions are subject to the qualifications implied by earlier comments. Another important factor in comparing actual and predicted statistics is the unequal 1960 Census income intervals, especially the open-ended interval "greater than $15,000." Better comparisons could be obtained from more detailed intervals, since the selected statistical estimates are quite sensitive to the interval midpoint chosen for the "greater than $15,000" interval.

Both the mean and corrected earnings distributions display the general characteristic of the actual distribution but tend to "overstate" earnings. All of the distributions display positive skewness, and have center and dispersion positively related to age and schooling. The predicted distributions also indicate increased center and dispersion with increased ability. The distributions corrected for unexplained variation tend to "overcorrect" in the sense that the resulting distributions are more smooth than the actual distribution.

The mean earnings distributions obviously have less dispersion than either the corrected or actual distributions, and the corrected distributions tend to overpredict mean earnings relative to the actual distribution, especially at young ages. The procedure used for "correcting" the mean earnings distribution to account for error variation seems to be inadequate. Evidence cited later with respect to human wealth will indicate that the error is not purely transitory, but has a persistent element that is related to age. That is, there are unobserved variables which may be uncorrelated with schooling and ability but which are not uncorrelated with age. An individual's profile may lie wholly above or wholly below the estimated profile and this is not captured in the correction to earnings distribution. Further evidence indicates that the distribution of this persistent component of earnings is itself positively skewed, which would further enhance the positive skewness of earnings as evidenced by the underprediction of positive skewness in the predicted, as opposed to actual, earnings distributions. These problems could be partially alleviated by a more complete accounting of the variation in earnings than is present in this earnings function. It should be remembered, however, that a source of the discrepancy in skewness is the large number of employed students at very low income level. The students' problem also partially explains the overprediction of the mean at young ages. For example, note the $1,000–$2,000 mean income of eighteen through twenty-one year olds. Fully employed males should have mean earnings greater than this even at young ages. The inclusion of these students will also pull down the mean of the overall actual distribution and the mean of the lower schooling groups. It should be noted, also, that the correction procedure by truncating the normal distribution at zero earnings and using conditional densities causes the corrected means to be too large.

Consider the properties of these distributions in more detail. With respect to central tendency, both the mean and median are overstated by the predicted distribution. Even so, the mean and median move in the right direction between age and schooling classes. The mean increases within higher schooling classes for both predicted and actual distributions. Mean and median earnings rise continuously with age in the actual distributions but decline very slightly before rising continuously after age twenty-four in both predicted distributions. The dip in mean earnings is clearly evident in the age-earnings profiles in Figure 1A. In the actual distribution, this property would be hidden by the inclusion of employed young students with very low earnings. Both mean and median earnings are predicted to rise sharply as the ability level of a subgroup rises. Again, the high ability level of the NBER-Thorndike sample itself is a source of the overstatement of earnings. It should be noted that the overall mean of the population is a weighted average[31] of individual subgroup means, whether grouped by age, schooling, or ability.

Dispersion is overstated in the corrected predicted distribution when measured by the standard deviation, but understated when measured by the coefficient of variation. The standard deviation increases continuously with schooling. As age increases, it dips slightly before age twenty-four in the predicted distributions, then rises continuously as it does throughout in the actual distribution. It is interesting to note here that the variance of overall earnings is the sum of the average of the variances of the subgroups and the variance of average earnings of subgroups.[32]

Another characteristic of earnings distributions widely discussed in the literature is concentration represented by the Lorenz curve and its summary statistic, the Gini coefficient.[33] The Gini coefficient is roughly the same between the corrected predicted and actual distributions, except that the predicted distributions always understate inequality at the extremes of age and schooling and overstate it in the middle range. This is partially caused by the large unequal income intervals in the actual distributions. Since the Lorenz curve is approximated by joining chords, the Gini is always understated but the understatement is much larger for the actual distributions.

The predicted distributions tend to indicate less skewness than the actual distribution, but this statistic is very sensitive to the unequal broad earnings classes in the actual distribution, and the results are not directly comparable. This statistic, as mentioned, is especially sensitive to the normality assumption used for the correction.

Predicted Human Wealth Distributions from the Estimated Earnings Function

The purpose of this section is to predict the distribution of human wealth overall, by schooling class, and by ability class for several interest rates and retirement ages. Human wealth is defined here as the present value of earnings, net of educational or human capital investments over the individual's lifetime. The earnings function and corresponding age-earnings profiles estimated from the NBER-Thorndike sample correspond to earnings somewhere between net and gross values, depending upon what fraction of investment is obtained on-the-job. The empirical measure of mean human wealth is then the integral of the discounted estimated earnings function with respect to age from the end of formal schooling to the retirement age.

Since the estimated earnings function corresponds to mean earnings, the estimated human wealth corresponds accordingly to the mean present value of observed earnings. Since the mean error for any age is zero and the estimation error is assumed to be uncorrelated with age, schooling, or ability, the expected discounted sum of errors over the life cycle is also zero. That is

$$PV(S, B) = \hat{P}V(S, B) + \int_{a=S}^{N} e^{-ra}u(a, S, B)\, da$$

where

$$\hat{P}V(S, B) = \int_{a=S}^{N} e^{-ra}\hat{Y}(a, S, B)\, da$$

so that

$$E_u[PV(S, B)] = \hat{P}V(S, B)$$

The predicted distributions presented in this section are the distributions of $\hat{P}V(S, B)$ and thus correspond to mean human wealth distributions. This should be carefully noted in observing the small measures of dispersion and inequality. Corrections for other sources of variation are considered later. The means should be unbiased estimates but the variation should be interpreted as variation in the mean, which obviously has much less dispersion. Thus, overall variation is due to differences in expected human wealth due to schooling and ability. Variation within a subgroup—say, schooling—is due to differences in expected human wealth due to the other factor, ability.

Everyone in the population is assumed to have the same discount rate and the same working life, but individuals differ in schooling and ability. Density is transformed from two-dimensional (schooling, ability)—space into human wealth—space through the integral function. The same schooling and ability distribution and midpoints are used as before.

Detailed selected statistics for mean human wealth are presented in Tables 6 and 7 for the overall population and for schooling and ability subgroups for discount rates three through seven and retirement age sixty-six and a retirement age that is a function of schooling level. The expected retirement ages as a function of schooling level, $N(S)$, are taken from Mincer (1973) and are reproduced in Table 8.

The relative frequency distributions for discount rates 3, 5, and 7 percent and retirement age sixty-six are presented in Figures 7A though 7G.

The most striking result is that there is much less inequality in mean human wealth than in mean earnings. Both the coefficient of variation and the Gini coefficient drop drastically. To the extent that perfect capital markets for consumption are available to everyone, the human wealth variation is a more appropriate index of the variation in expected economic well-being.

The clearest result of a more detailed study of the effect of schooling, ability, retirement age, and the discount rate is that an increased retirement age unambiguously raises mean human wealth, see Figure 8, and an increased discount rate unambiguously lowers it, see Figure 9. It is interesting to note that a 1 percent change in the rate of discount, within the range 3 to 7, has a much larger effect on mean human wealth than an increase of four years in retirement age from sixty-six to seventy. As expected, retirement age has an increasingly smaller effect at higher discount rates but the rate of discount has an increasingly greater effect for later retirement ages.

The effect of the discount rate on variation in human wealth is more ambiguous and is intimately related to the effect of schooling on human wealth. It is important to note that due to the year of forgone earnings and the short initial period of lower earnings associated with more schooling, increased schooling does not unambiguously increase predicted human wealth or mean human wealth averaged over ability levels. Predicted human wealth increases with increased schooling only if the discount rate is below the internal rate of return.[34] Figures 10 and 11 clearly illustrate this result for mean human wealth at various schooling levels. The reversal occurs at approximately 5.75 percent, except that high school graduates pass those with some college at approximately 4.5 percent. The effect of schooling declines as the discount rate increases up to the crossover, then has a negative effect on mean human wealth. Thus, an increased rate of discount decreases variation up to about 6 percent, at which point it causes the variation within ability groups, due to schooling, to increase. That is, at high discount rates, schooling differences cause variation, but because of their increasing negative effect on human wealth.

TABLE 6 Selected Statistics for the Predicted Distribution of Human Wealth for the Overall Population, by Schooling Class, and by Ability Class for Several Rates of Discount, Assuming Retirement Age Sixty-Six

	Discount Rate	Mean (Dollars)	Median (Dollars)	Standard Deviation (Dollars)	Coefficient of Variation	Skewness	Gini Coefficient
Overall	.03	198,240	193,759	15,264	.08	2.08	.04
	.04	154,150	150,940	9,762	.06	2.75	.03
	.05	121,943	120,145	6,236	.05	3.23	.02
	.06	97,965	97,846	4,651	.05	2.66	.02
	.07	79,948	80,847	4,479	.06	1.09	.03
By schooling:							
12 years	.03	190,969	189,523	7,262	.04	3.14	.02
13–15 years		197,873	194,327	8,932	.05	1.89	.02
16 years		211,515	206,880	15,669	.07	1.24	.04
17+ years		223,619	221,086	20,058	.09	0.74	.05
12 years	.04	150,781	149,527	5,501	.04	4.72	.02
13–15 years		153,052	150,484	6,152	.04	2.16	.02
16 years		160,407	156,880	11,277	.07	1.36	.04
17+ years		167,770	165,086	15,646	.09	1.14	.05
12 years	.05	121,389	120,418	3,850	.03	8.25	.01
13–15 years		120,215	118,660	4,525	.04	2.00	.02
16 years		123,504	120,880	8,480	.07	1.22	.04
17+ years		126,852	125,086	11,740	.09	1.07	.05
12 years	.06	99,121	98,727	3,078	.03	11.28	.01
13–15 years		96,294	95,422	2,960	.03	2.28	.01

16 years		96,145	94,457	6,226	.06	1.34	.03
17+ years		97,282	96,454	8,851	.09	0.97	.05
12 years	.07	82,418	82,023	2,562	.03	13.72	.01
13–15 years		77,941	77,352	1,961	.03	2.60	.01
16 years		75,866	74,464	4,662	.06	1.39	.03
17+ years		75,367	74,454	6,651	.09	0.98	.05
By ability:							
<.75	.03	186,977	184,817	7,085	.04	9.13	.01
.75–1.00		191,716	190,109	6,136	.03	1.49	.02
1.00–1.25		202,595	197,755	11,113	.05	1.45	.03
>1.25		227,941	224,668	19,376	.09	0.82	.05
<.75	.04	147,298	146,797	5,565	.04	14.35	.01
.75–1.00		149,638	149,141	2,685	.02	2.00	.01
1.00–1.25		156,840	155,143	6,063	.04	1.58	.02
>1.25		174,308	171,107	13,006	.07	1.70	.04
<.75	.05	118,572	118,841	5,184	.04	11.81	.01
.75–1.00		118,824	118,930	1,729	.01	−.026	.01
1.00–1.25		123,440	123,058	3,269	.03	1.24	.01
>1.25		134,963	132,833	8,120	.06	2.05	.03
>.75	.06	96,133	96,774	5,438	.06	7.26	.02
.75–1.00		96,214	96,943	2,977	.03	−1.00	.02
1.00–1.25		98,501	98,776	2,506	.03	−0.19	.01
>1.25		106,223	105,131	5,135	.05	2.16	.02
<.75	.07	79,640	80,835	5,744	.07	4.00	.03
.75–1.00		78,755	80,461	3,883	.05	−1.13	.03
1.00–1.25		80,159	82,073	3,631	.05	−0.39	.02
>1.25		84,605	84,128	3,922	.05	1.41	.02

TABLE 7 Selected Statistics for the Predicted Distribution of Human Wealth for the Overall Population, by Schooling Class, and by Ability Class for Several Rates of Discount, Assuming Retirement Age $N(S)$

	Discount Rate	Mean (Dollars)	Median (Dollars)	Standard Deviation (Dollars)	Coefficient of Variation	Skewness	Gini Coefficient
Overall	.03	202,736	196,074	18,243	.09	1.95	.04
	.04	156,950	153,159	11,451	.07	2.77	.03
	.05	123,653	121,659	6,925	.06	3.41	.02
	.06	98,887	98,590	4,803	.05	3.16	.02
	.07	80,423	81,106	4,345	.05	1.60	.03
By schooling:							
12 years	.03	193,662	192,990	7,587	.04	2.48	.02
13–15 years		200,873	197,066	9,342	.05	1.83	.02
16 years		218,248	214,880	16,304	.07	1.20	.04
17+ years		238,737	237,086	20,755	.09	0.49	.05
12 years	.04	149,053	147,672	5,757	.04	13.14	.01
13–15 years		152,212	151,192	3,996	.03	2.36	.01
16 years		159,845	156,777	8,313	.05	1.76	.03
17+ years		178,992	174,776	16,261	.09	1.75	.05
12 years	.05	122,403	121,705	4,080	.03	7.25	.01
13–15 years		121,573	119,928	4,223	.03	2.35	.02
16 years		125,920	123,430	8,553	.07	1.36	.04
17+ years		132,369	131,086	12,532	.09	1.26	.05
12 years	.06	99,705	99,202	3,171	.03	10.23	.01
13–15 years		96,699	95,623	2,991	.03	2.23	.01
16 years		97,682	95,605	6,412	.07	1.34	.03
17+ years		100,464	99,086	9,267	.09	1.18	.05
12 years	.07	82,658	82,393	2,626	.03	12.74	.01

13–15 years		78,010	77,386	2,067	.03	2.52	.01
16 years		77,032	75,605	4,695	.06	1.54	.03
17+ years		77,193	76,454	6,969	.09	1.10	.05
By ability:							
<.75	.03	189,966	186,817	8,337	.04	5.62	.02
.75–1.00		195,552	193,638	9,006	.05	2.00	.02
1.00–1.25		207,867	201,755	15,046	.07	1.55	.04
>1.25		234,869	230,440	23,534	.10	0.67	.06
<.75	.04	149,053	147,672	5,757	.04	13.14	.01
.75–1.00		152,212	151,192	3,996	.03	2.36	.01
1.00–1.25		159,845	156,777	8,313	.05	1.76	.03
>1.25		178,992	174,776	16,261	.09	1.75	.05
<.75	.05	119,297	119,037	4,914	.04	14.82	.01
.75–1.00		120,541	120,356	1,797	.01	0.79	.01
1.00–1.25		125,321	124,176	4,106	.03	1.70	.02
>1.25		137,616	134,300	10,018	.07	2.08	.04
<.75	.06	96,849	97,183	5,122	.05	8.75	.02
.75–1.00		96,881	97,913	2,457	.03	−0.76	.01
1.00–1.25		99,608	99,912	2,710	.03	−0.03	.01
>1.25		107,891	106,278	5,939	.06	2.57	.03
<.75	.07	79,873	80,835	5,434	.07	4.94	.03
.75–1.00		79,214	80,615	3,653	.05	−0.84	.02
1.00–1.25		80,553	82,073	3,202	.04	−0.30	.02
>1.25		85,822	85,776	4,273	.05	1.55	.03

TABLE 8 Estimated Average Retirement Age by Years of Schooling, from Mincer (1974)

Years of Schooling	Estimated Average Retirement Age
8 years	65
9–11 years	66
12 years	67
13–15 years	67
16 years	68
17+ years	70

The effect of increased ability is to increase unambiguously mean human wealth as illustrated in Figures 12 and 13. The magnitude of the effect of ability declines at higher discount rates, since the returns to higher ability come late in the life cycle.

The human wealth distributions are corrected for error variation by decomposing the error into purely random or transitory and persistent components. A lower bound on the variance of human wealth is defined as the variance of the present value of predicted earnings plus an error component which is completely transitory and independent from period to period. An upper bound is defined as the variance of the present value of predicted earnings plus a completely persistent error component which is constant over the life cycle but varies randomly over individuals independently of the level of ability and schooling. Intermediate cases can be considered as combinations of these when the transitory and persistent variations are indepedendent. The upper and lower bounds allow no comparisons of inequality in human wealth versus earnings, since the human wealth coefficient of variation lower bound lies below, and the human wealth coefficient of variation upper bound lies above, the earnings coefficient of variation. The answer lies in the "persistence" of the error over an individual's lifetime. The standard deviation of the persistent component is estimated and used to estimate standard deviation and coefficient of variation for human wealth. Corresponding estimates are also made by calculating the actual present value of the residuals for each individual.

FIGURE 7A Predicted Distribution of Mean Human Wealth for the Overall Population for Several Rates (R) of Discount

FIGURE 7A Predicted Distribution of Mean Human Wealth for the Overall Population for Several Rates (R) of Discount

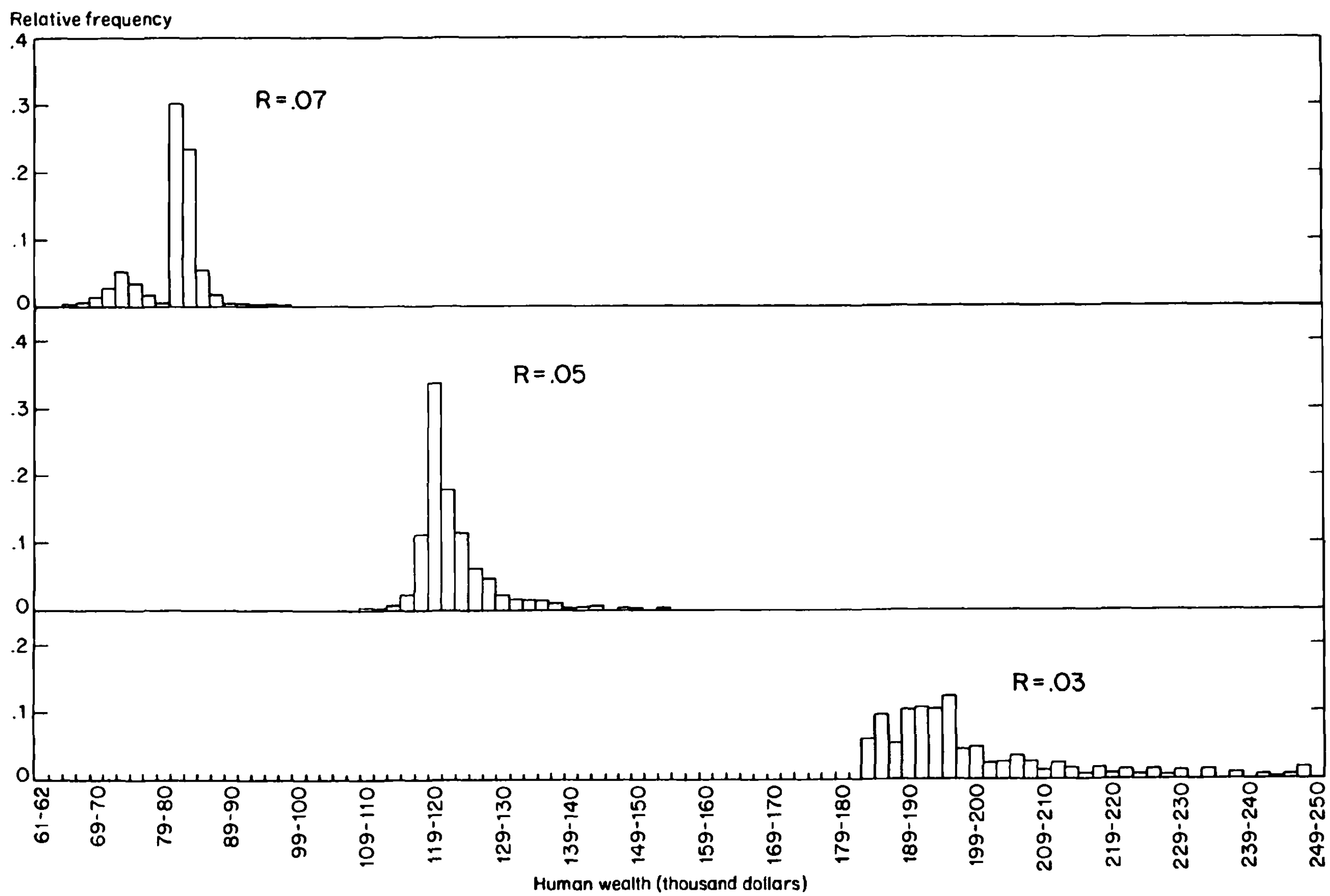


FIGURE 7B Predicted Distribution of Mean Human Wealth by Schooling Level, Discounted at 3 Percent

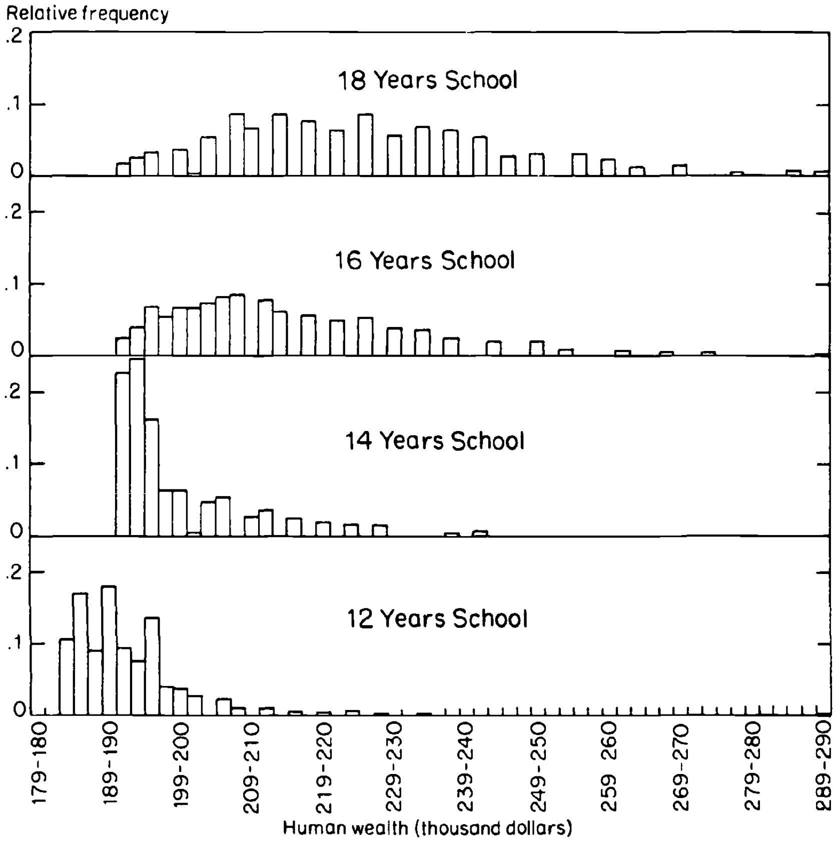


FIGURE 7C Predicted Distribution of Mean Human Wealth by Ability Group, Discounted at 3 Percent

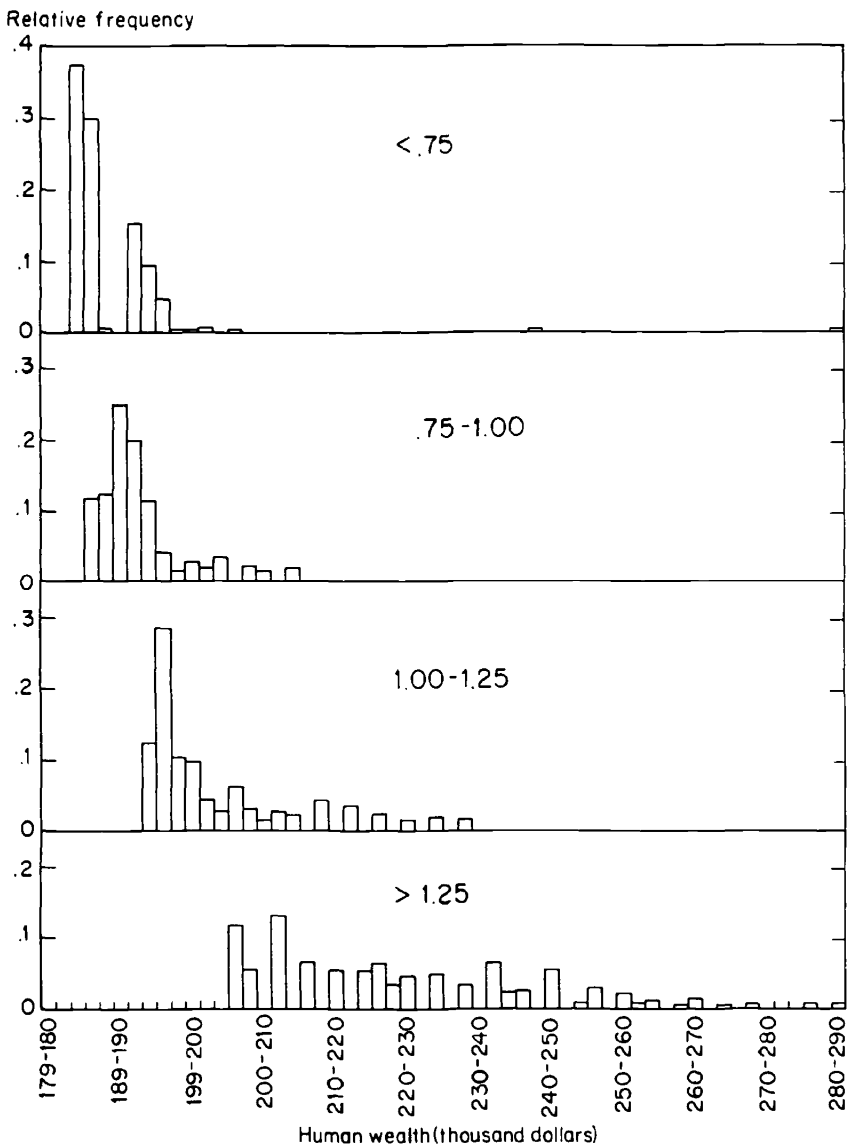


FIGURE 7D Predicted Distribution of Mean Human Wealth by Schooling Level, Discounted at 5 Percent

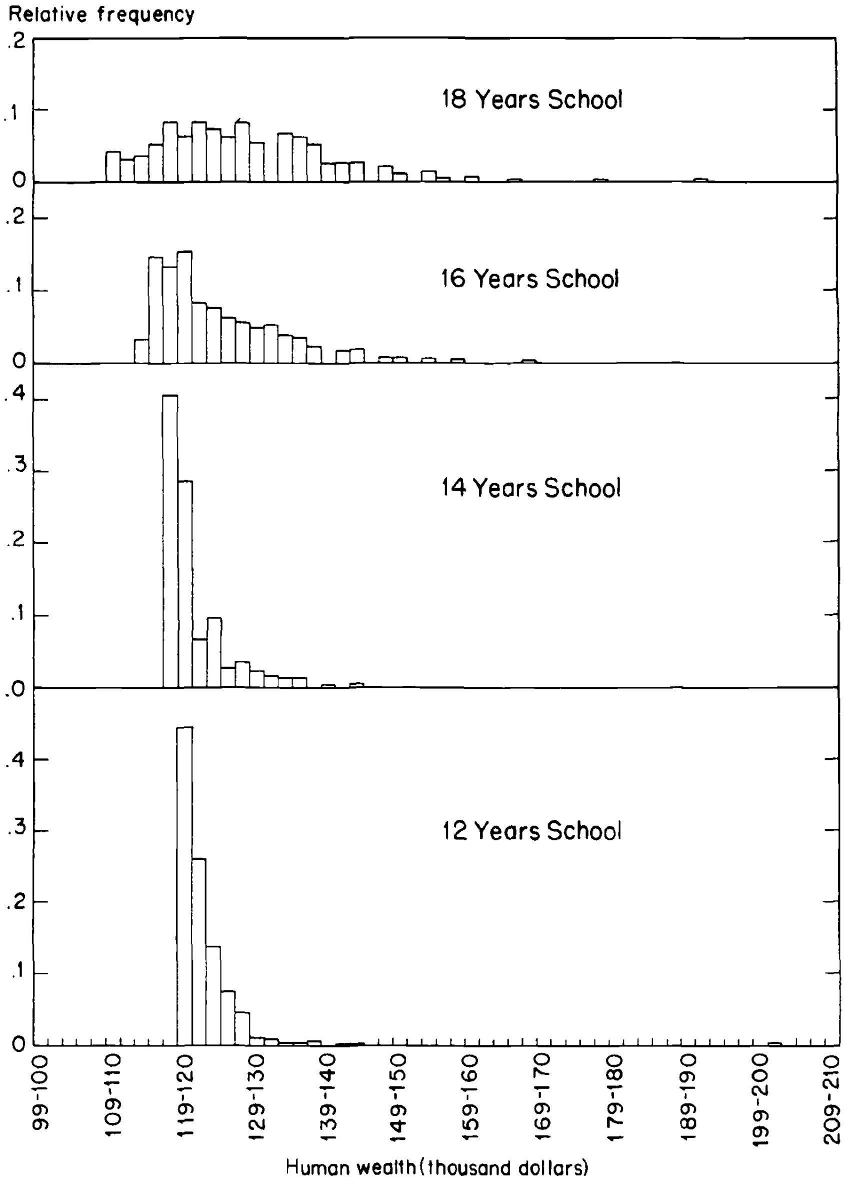


FIGURE 7E Predicted Distribution of Mean Human Wealth by Ability Group, Discounted at 5 Percent

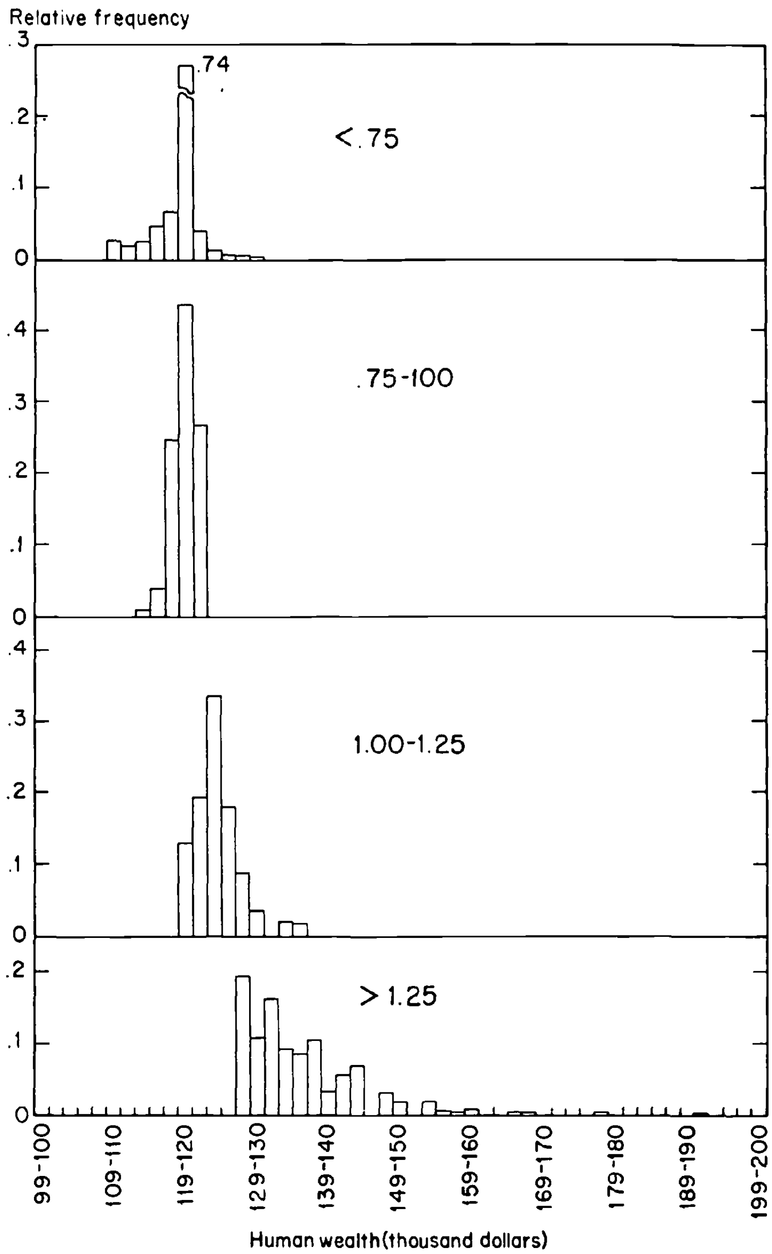


FIGURE 7F Predicted Distribution of Mean Human Wealth by Schooling Level, Discounted at 7 Percent

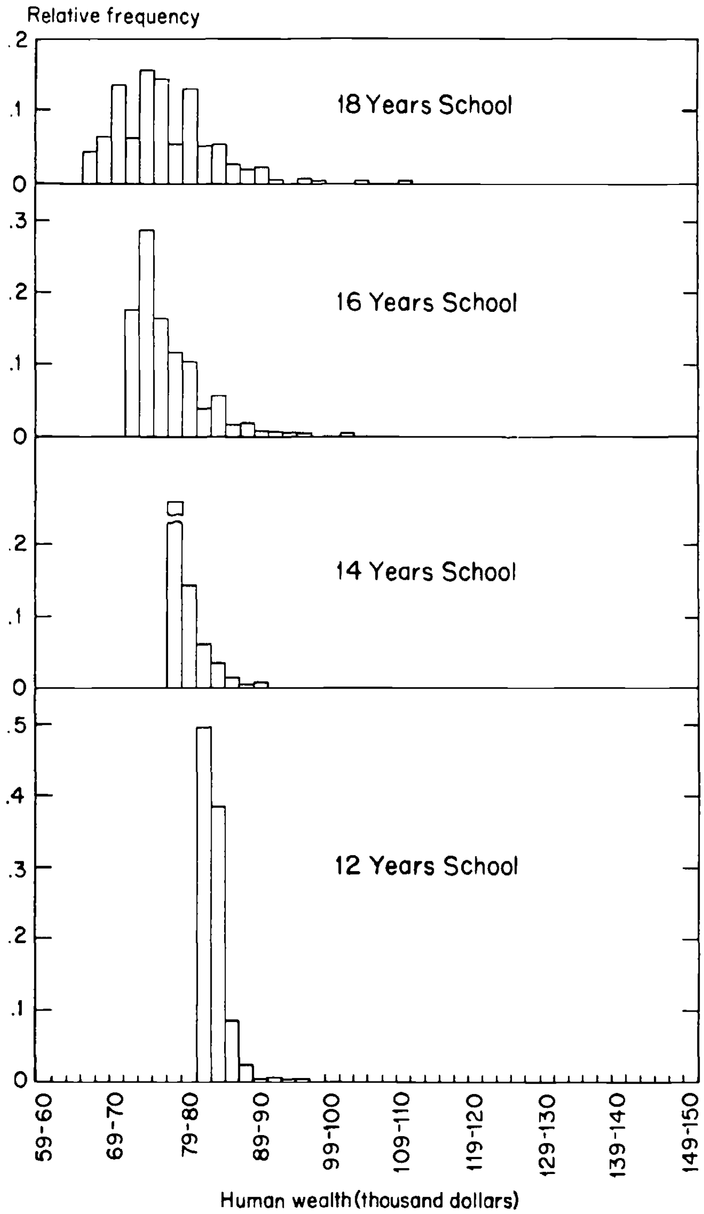


FIGURE 7G Predicted Distribution of Mean Human Wealth by Ability Group, Discounted at 7 Percent

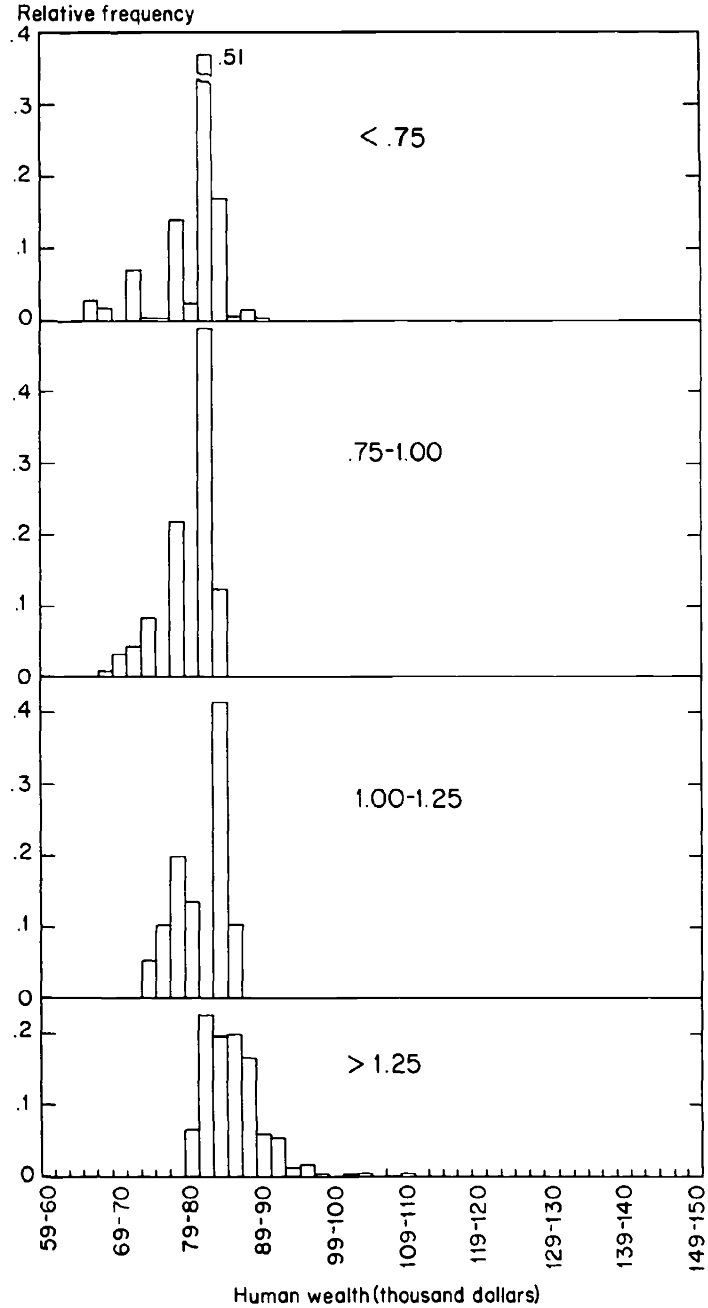


FIGURE 8 Overall Mean Human Wealth as a Function of Retirement Age for Several Rates of Discount

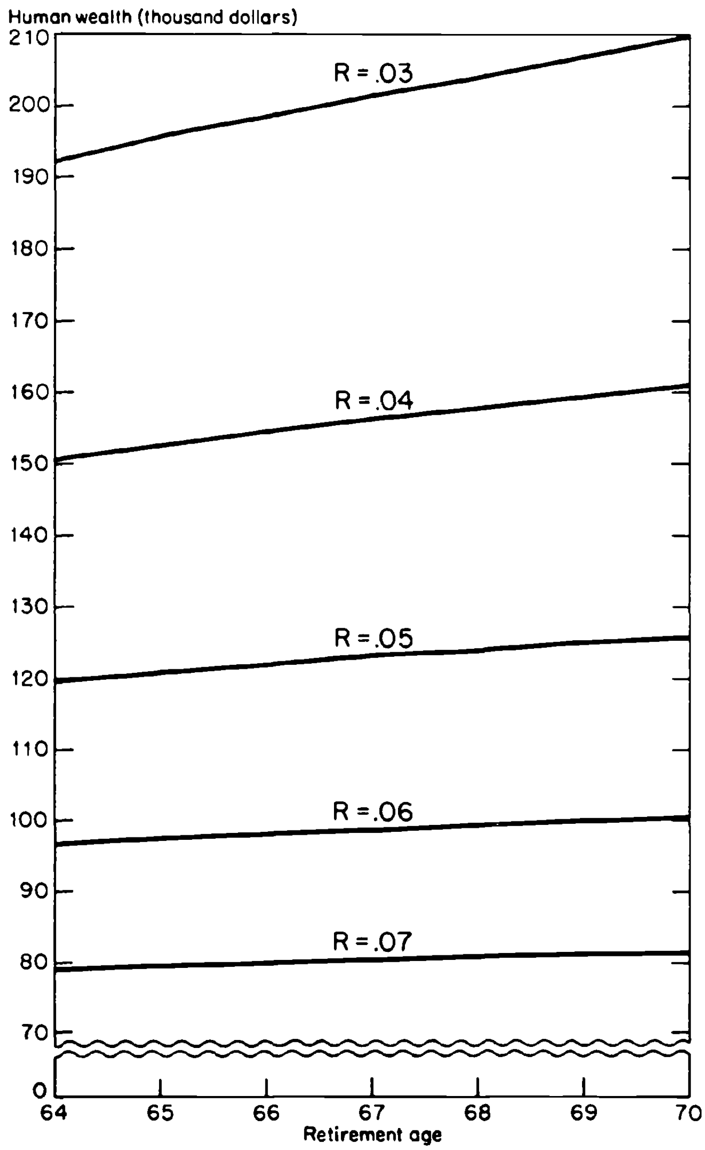


FIGURE 9 Overall Mean Human Wealth as a Function of the Discount Rate for Retirement Ages 66 and 70 and Retirement Age as a Function of Schooling Level, *N(S)*

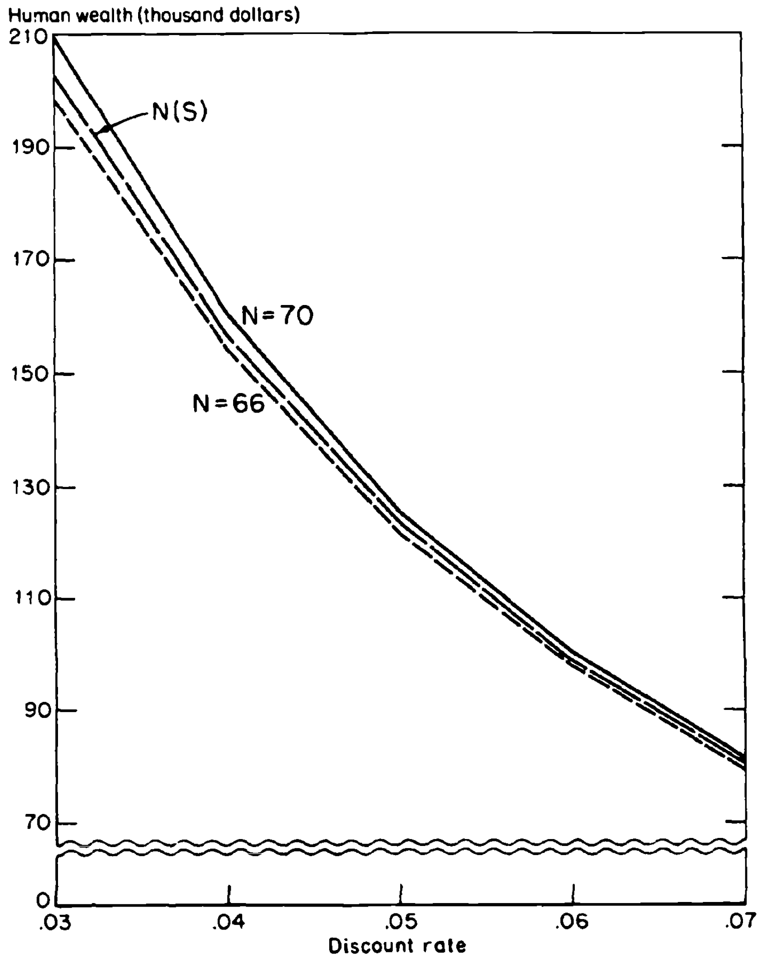


FIGURE 10 Mean Human Wealth by Schooling Level as a Function of the Rate of Discount for Retirement Age 66

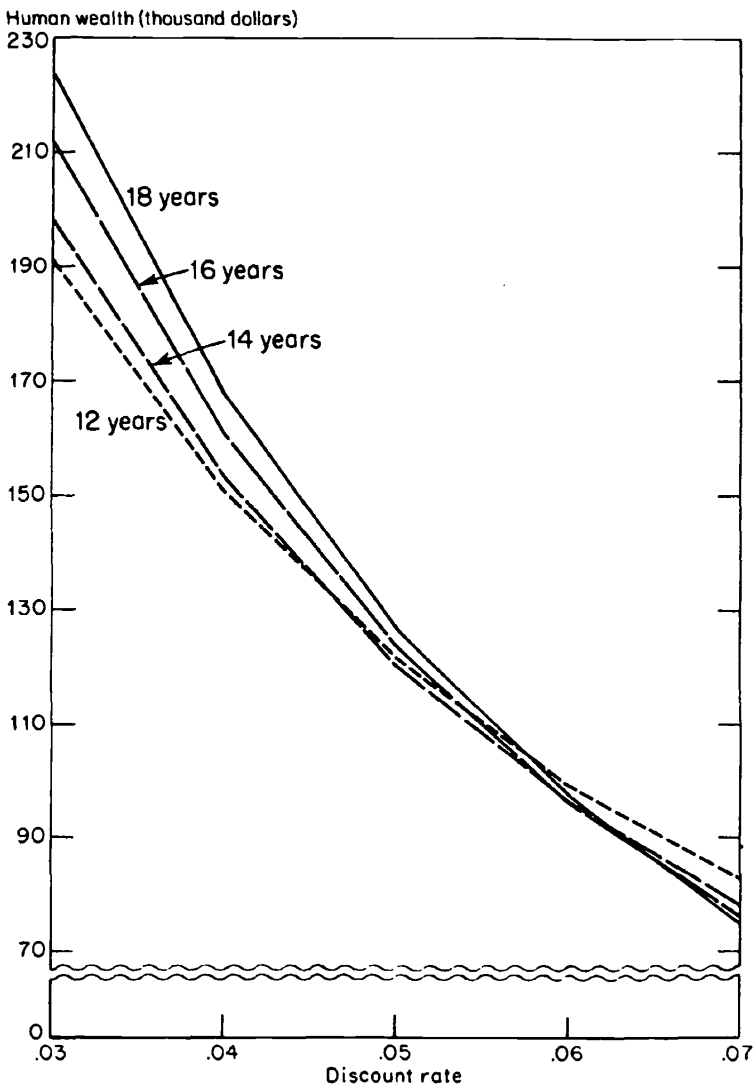


FIGURE 11 Mean Human Wealth as a Function of Schooling Level for Several Discount Rate and Retirement Combinations

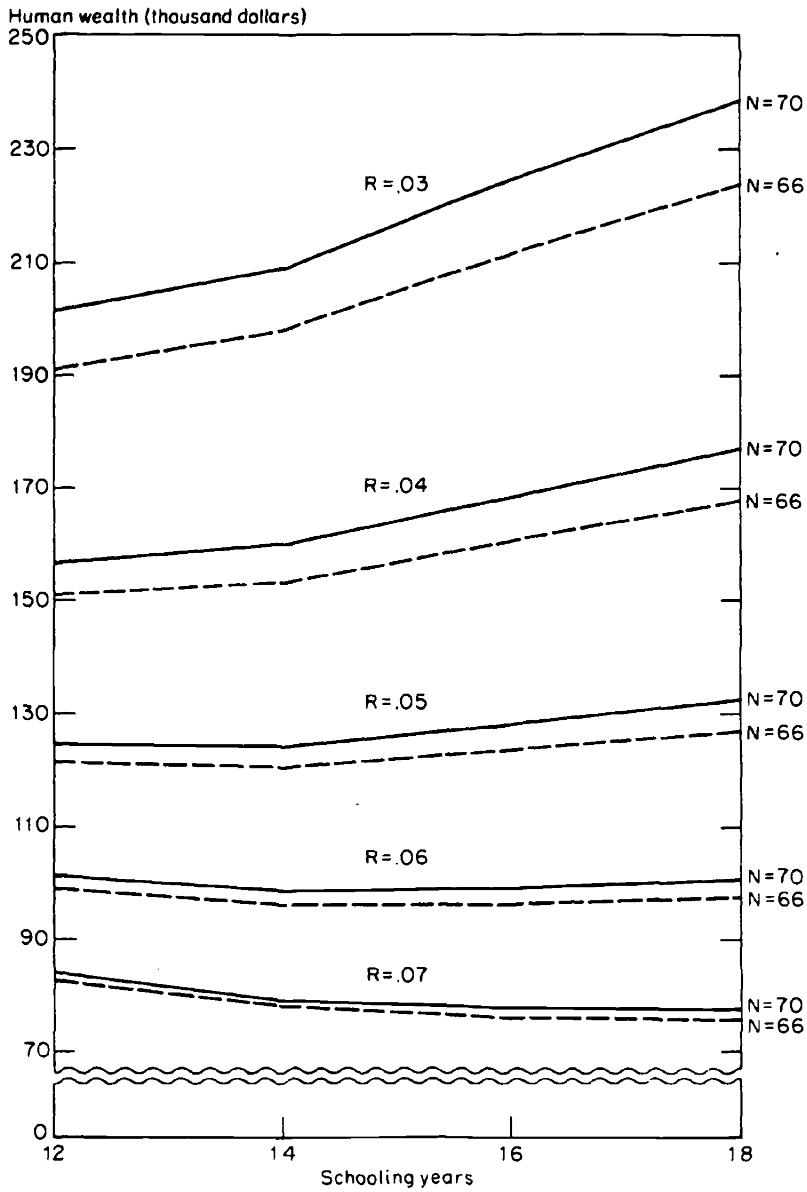


FIGURE 12 Mean Human Wealth by Ability Class as a Function of the Rate of Discount for Retirement Age 66

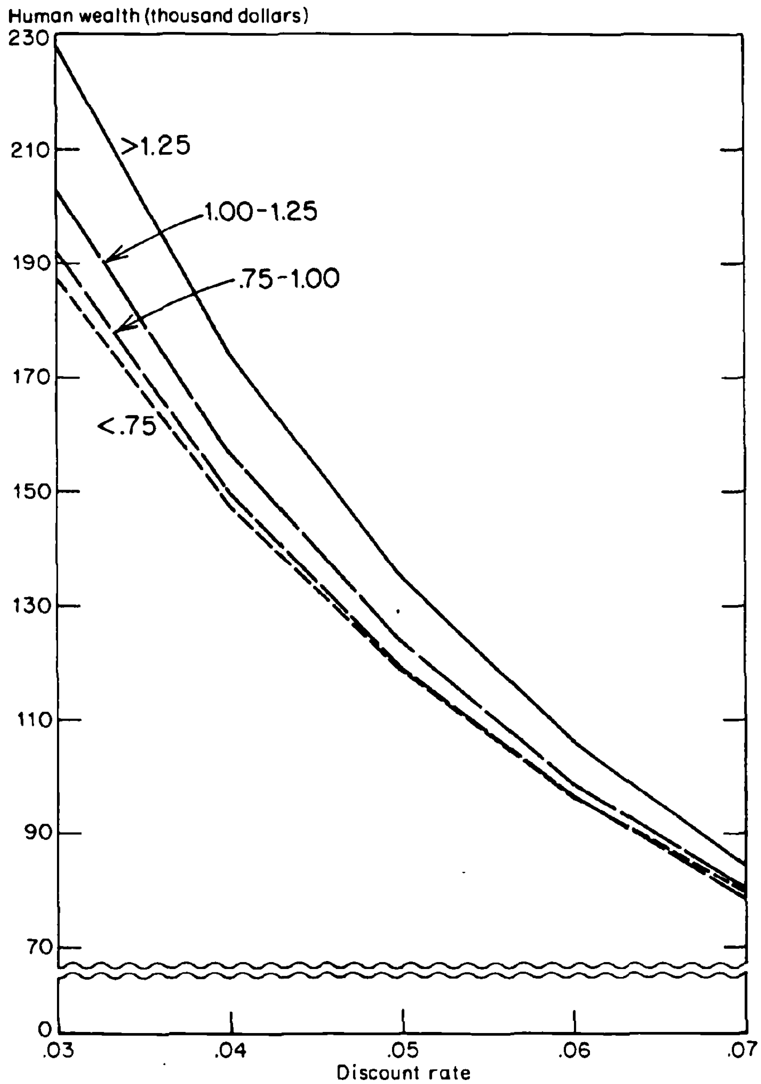
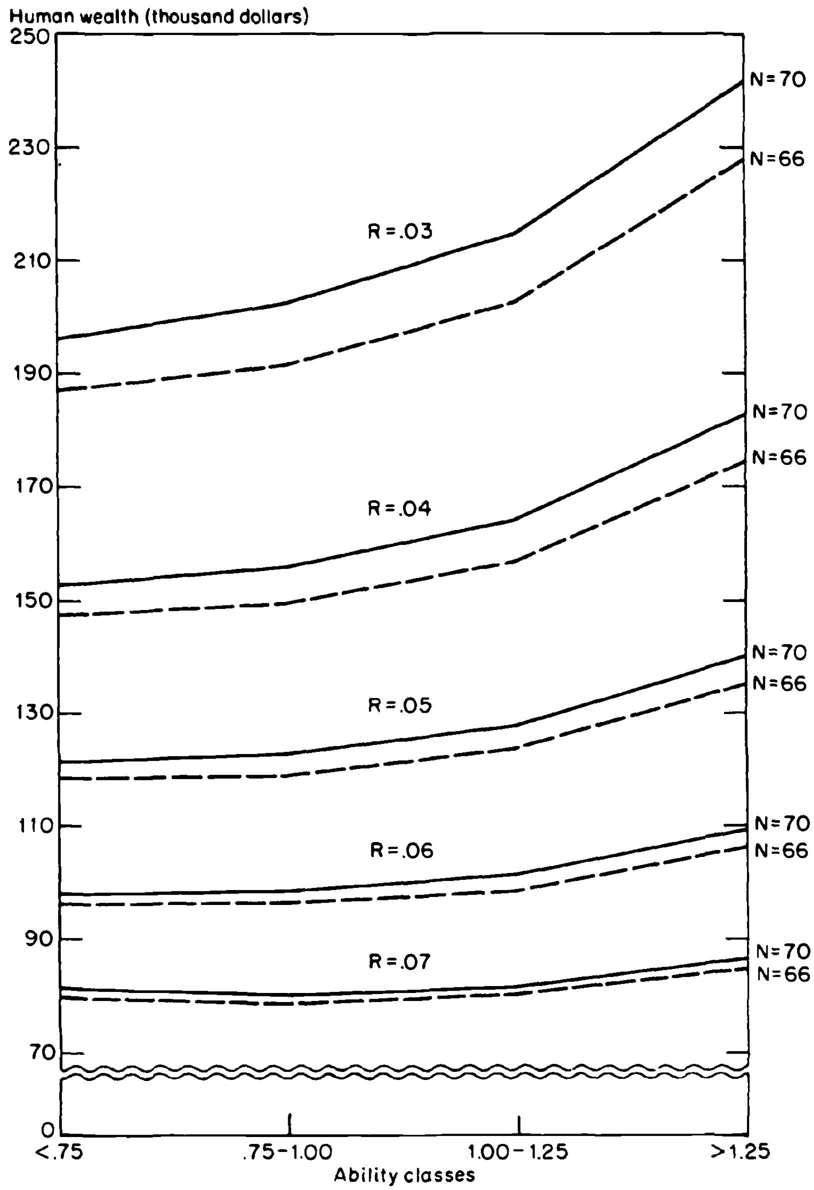


FIGURE 13 Mean Human Wealth as a Function of Ability Class for Several Discount Rate and Retirement Age Combinations



Consider the more general combination of these two variance components.

$$Y_i(a, S, B) = \hat{Y}(a, S_i, B_i) + \delta_i + \eta_{ia}$$

where i indicates individual.

The error components δ_i and η are assumed independent of each other and over i and are uncorrelated with a, S, and B; therefore $\delta \sim (0, \sigma_\delta^2)$ and $\eta_a \sim (0, \sigma_\eta^2 I)$ where I is of dimension equal to the number of age points specified. We still obtain

$$E_{\delta,\eta}[Y(a, S, B)] = \hat{Y}(a, S, B)$$

and

$$E_{\delta,\eta}[\hat{P}V(S, B)] = \int_{a=S}^{N} e^{-ra}\hat{Y}(a, S, B)\, da$$

Consider the variance for fixed values of schooling; that is, for both ability and schooling fixed or simply within a schooling class.

$$\begin{aligned}
\mathrm{Var}_{\delta\eta}[PV(S, B)] &= \mathrm{Var}_{\delta\eta}\left[\int_{a=S}^{N} \delta\, e^{-ra}\, da + \int_{a=S}^{N} \eta_a\, e^{-ra}\, da\right] \\
&= E_{\delta\eta}\left[\int_{a=S}^{N} \delta\, e^{-ra}\, da + \int_{a=S}^{N} \eta_a\, e^{-ra}\, da\right]^2 \\
&= E_{\delta\eta}\left[\int_{a=S}^{N} \delta\, e^{-ra}\, da\right]^2 + E_{\delta\eta}\left[\int_{a=S}^{N} \eta_a\, e^{-ra}\, da\right]^2 \\
&\quad + 2E_{\delta\eta}\left[\left(\int_{a=S}^{N} \delta\, e^{-ra}\, da\right)\left(\int_{a=S}^{N} \eta_a\, e^{-ra}\, da\right)\right] \\
&= \sigma_\delta^2(e^{-rS} - e^{-rN})^2/r^2 + E_\eta\left[\int_{a=S}^{N} \eta_a\, e^{-ra}\, da\right]^2
\end{aligned}$$

Since δ and η are independent

$$= \sigma_\delta^2(e^{-rS} - e^{-rN})^2/r^2 + \sigma_\eta^2(e^{-2rS} - e^{-2rS})/2r$$

since

$$E_\eta\left[\int_{a=S}^{N} \eta_a\, e^{-ra}\, da\right]^2 = \sigma_\eta^2 \int_{a=S}^{N} e^{-2ra}\, da$$

Similarly

$$\begin{aligned}
\mathrm{Var}_{B,\delta,\eta}[PV(S)] &= \mathrm{Var}_B[\hat{P}V(S, B)] + \sigma_\delta^2(e^{-rS} - e^{-rN})^2/r^2 \\
&\quad + \sigma_\eta^2[e^{-2rS} - e^{-2rN}]/2r
\end{aligned}$$

However, when schooling varies as within ability classes, or in the overall distribution, we must take an expected value with respect to the lower limit of the present value integral.

$$\begin{aligned}
\mathrm{Var}_{S,\delta,\eta}[PV(B)] &= \mathrm{Var}_S[\hat{P}V(S, B)] + \sigma_\delta^2[E_S(e^{-2rS}) - 2e^{-rN}E(e^{-rS}) + e^{-2rN}]/r^2 \\
&\quad + \sigma_\eta^2[E(e^{-2rS}) - e^{-2rN}]/2r
\end{aligned}$$

TABLE 9 Error Variance Lower and Upper Bound Correction Factors for Retirement Age Sixty-Six

	Discount Rate				
	.03	.04	.05	.06	.07
Overall, $[E_S(e^{-2rS})-e^{-2rN}]/2r$ and $E_S(e^{-rS}-e^{-rN})^2/r^2$					
	12.7594	9.3207	7.1006	5.5911	4.5180
	514.3049	341.5083	234.0726	165.1963	119.7322
By schooling, $(e^{-2rS}-e^{-2rN})/2r$ and $(e^{-rS}-e^{-rN})/r^2$					
12 years	13.9522	10.4229	8.1199	6.5346	5.3919
	573.8171	387.8743	270.7686	194.6646	143.7127
14 years	12.2807	8.8479	6.6358	5.1359	4.0736
	489.5750	321.1340	217.0589	150.8151	107.4442
16 years	10.7982	7.5058	5.4207	4.0356	3.0771
	416.3503	265.1138	173.5719	116.6002	80.1921
18 years	9.4833	6.3622	4.4259	3.1701	2.3241
	352.8105	218.1631	138.4082	89.9325	59.7336
By ability, $[E_{S\mid B}(e^{-2rS})-e^{-2rN}]/2r$ and $[E_{S\mid B}(e^{-2rS})-2E_{S\mid B}(e^{-rS})e^{-rN}+e^{-2rN}]/r^2$					
<.75	13.2224	9.7462	7.4921	5.9515	4.8501
	537.3420	359.3718	248.1428	176.4407	128.8384
.75–1.00	12.9530	9.4981	7.2633	5.7404	4.6552
	523.9226	348.9463	239.9150	169.8519	123.4912
1.00–1.25	12.5895	9.1648	6.9576	5.4597	4.3971
	505.8616	334.9714	228.9324	161.0955	116.4171
>1.25	11.8130	8.4522	6.3029	4.8578	3.8432
	467.2546	305.0703	205.4099	142.3217	101.2336

NOTE: Upper figure lower bound. Lower figure upper bound.

The lower bound obtains when $\sigma_\delta^2 = 0$, and the upper bound obtains when $\sigma_\eta^2 = 0$, for a given total variation $\sigma_\delta^2 + \sigma_\eta^2$ from the estimated earning function. The coefficients of the variance components are presented in Table 9 for discount rates 3 through 7 percent and retirement age sixty-six. The upper and lower bounds on variance of human wealth overall and within subgroups are presented in Table 10. The corresponding coefficients of variation are presented in Table 11.

Both the standard deviation and the coefficient of variation differ widely between the lower and upper bound. The inequality in mean human wealth is much less than either the lower bound or upper bound. This indicates that the error component is very important in determining human wealth inequality and indicates that the persistent component is

TABLE 10 Lower and Upper Bounds on the Standard Deviation of Human Wealth and the Standard Deviation of the Mean

	Discount Rate				
	.03	.04	.05	.06	.07
			Mean		
Overall	$ 15,264	$ 9,762	$ 6,236	$ 4,651	$ 4,479.09
By schooling:					
12 years	7,262	5,501	3,850	3,078	2,562.40
14 years	8,932	6,152	4,525	2,960	1,961.16
16 years	15,669	11,277	8,480	6,226	4,662.79
18 years	20,058	15,646	11,740	8,851	6,651.07
By ability:					
<.75	7,085	5,565	5,184	5,438	5,744.34
.75–1.00	6,136	2,685	1,729	2,977	3,883.49
1.00–1.25	11,113	6,063	3,269	2,506	3,631.12
>1.25	19,376	13,006	8,120	5,135	3,922.32
			Lower Bound		
Overall	26,456	20,890	17,284	15,041	13,616.00
By schooling:					
12 years	23,734	20,290	17,662	15,767	14,279.00
14 years	23,004	19,017	16,227	14,025	12,366.00
16 years	25,311	20,046	16,440	13,655	11,591.00
18 years	27,374	21,855	17,314	13,941	11,370.00
By ability:					
<.75	23,110	19,688	17,351	15,728	14,508.00
.75–1.00	22,620	18,836	16,394	14,796	13,617.00
1.00–1.25	24,170	19,291	16,288	14,355	13,194.00
>1.25	28,421	21,874	17,222	14,288	12,491.00
			Upper Bound		
Overall	138,033	112,215	92,760	77,889	66,344.00
By schooling:					
12 years	145,089	119,264	99,615	84,457	72,564.00
14 years	134,146	108,578	89,238	74,348	62,734.00
16 years	124,424	99,139	80,147	65,617	54,371.00
18 years	115,382	90,709	72,130	58,046	47,224.00
By ability:					
<.75	140,404	114,811	95,432	80,537	68,903.00
.75–1.00	138,599	113,033	93,714	78,894	67,335.00
1.00–1.25	136,509	110,881	91,587	76,820	65,370.00
>1.25	132,189	106,455	87,078	72,349	60,991.00

NOTE: The assumptions underlying these bounds are outlined in the text.

TABLE 11 Coefficient of Variation for Mean, Lower Bound and Upper Bound for Human Wealth Distributions

	Discount Rate				
	.03	.04	.05	.06	.07
			Mean		
Overall	.08	.06	.05	.05	.06
By schooling:					
12 years	.04	.04	.03	.03	.03
14 years	.05	.04	.04	.03	.03
16 years	.07	.07	.07	.06	.06
18 years	.09	.09	.09	.09	.09
By ability:					
<.75	.04	.04	.04	.06	.07
.75–1.00	.03	.02	.01	.03	.05
1.00–1.25	.05	.04	.03	.03	.05
>1.25	.09	.07	.06	.05	.05
			Lower Bound		
Overall	.13	.14	.14	.15	.17
By schooling:					
12 years	.12	.13	.15	.16	.17
14 years	.12	.12	.13	.15	.16
16 years	.12	.12	.13	.14	.15
18 years	.12	.13	.14	.14	.15
By ability:					
<.75	.12	.13	.15	.16	.18
.75–1.00	.12	.13	.14	.15	.17
1.00–1.25	.12	.12	.13	.15	.16
>1.25	.12	.13	.13	.13	.15
			Upper Bound		
Overall	.70	.73	.76	.80	.83
By schooling:					
12 years	.76	.79	.82	.85	.88
14 years	.68	.71	.74	.77	.80
16 years	.59	.62	.65	.68	.72
18 years	.52	.54	.57	.60	.63
By ability:					
<.75	.75	.78	.80	.84	.87
.75–1.00	.72	.76	.79	.82	.85
1.00–1.25	.67	.71	.74	.78	.82
>1.25	.58	.61	.65	.68	.72

very important in determining inequality in human wealth. We can note, however, that inequality in mean values before correcting for error variation is much less for human wealth than for earnings—in the overall values, the difference being 38 percent for earnings as opposed to about 5 or 6 percent for human wealth. When the correction for error variation is made, the lower and upper bound on the coefficient variation for human wealth brackets the coefficient of variation for either the predicted distribution of earnings or the actual coefficient of variation observed for earnings. It is necessary, then, to estimate the variance of the persistent component in revising our estimate of the standard deviation and coefficient of variation of human wealth.

The standard deviation of the persistent component of the error term is estimated in the following way. For each individual of the roughly 5,000 in the sample, the persistent component is measured as that value of a constant error, deviation from the predicted profile, such that the present value of deviations from it, the purely transitory part, is zero, i.e.,

$$\delta = \frac{\sum \mu(a)e^{-ra}}{\sum e^{-ra}}$$

The standard deviation of the error term is \$6,048 and the standard deviation of the persistent component, $\hat{\sigma}_\delta$, is roughly \$4,000, depending on the discount rate. The corresponding estimates of the standard deviation and coefficient of variation of human wealth based upon this estimate of the standard deviation of the persistent component are presented in Table 12 under the heading "Estimated for 1960 Census Groups Using $\hat{\sigma}_\delta$."

Corresponding estimates are made by calculating the actual present value of the residuals in the sample and inflating them to the equivalent of a working life of observations,[35] and taking the standard deviation. These are presented as the "directly estimated" values also shown in Table 12. These estimates correspond quite closely to those of the previous procedure and are larger probably because of the greater schooling present in the Thorndike sample than in the 1960 Census of Population.

The estimated inequality in human wealth is slightly less than the inequality in the predicted earnings distributions corrected for unexplained variations, the coefficient of variation being 60 percent for earnings and 50 percent for human wealth. The actual distribution of earnings is even more unequally distributed, with the coefficient of variation of 83 percent. It should be noted that the coefficient of variation for the actual distribution of earnings is larger than even the upper bound of the coefficient of variation for the human wealth distributions.

These crude estimates seem to indicate that human wealth is more equally distributed over individuals than is earnings, but that the varia-

TABLE 12 Estimated Standard Deviation of Human Wealth and Coefficient of Variation

	Discount Rate .03	.04	.05	.06	.07
Standard Deviation of the Persistent Component					
σ_δ	\$ 4,102	\$ 3,943	\$ 3,799	\$ 3,671	\$ 3,559
Standard Deviation of Human Wealth					
Directly Estimated from Present Value of Sample Residuals					
Overall	98,760	78,292	62,640	50,612	41,319
Estimated for 1960 Census Groups Using $\hat{\sigma}_\delta$					
Overall	95,617	74,847	59,795	48,764	40,562
By schooling:					
12 years	99,938	79,254	64,061	52,772	44,233
14 years	92,539	72,234	57,455	46,483	38,245
16 years	86,414	66,391	51,941	41,279	33,339
18 years	80,800	61,410	47,266	36,933	29,270
By ability:					
<.75	96,730	76,318	61,443	50,456	42,208
.75–1.00	95,461	75,056	60,220	49,309	41,124
1.00–1.25	94,274	73,747	58,906	48,003	39,918
>1.25	92,056	71,351	56,312	45,358	37,283
Coefficient of Variation					
From Direct Estimate					
Overall	.49	.50	.51	.51	.51
From Estimate Using $\hat{\sigma}_\delta$					
Overall	.47	.48	.48	.49	.50
By schooling:					
12 years	.52	.53	.52	.53	.54
14 years	.46	.47	.47	.48	.49
16 years	.40	.42	.41	.42	.43
18 years	.34	.34	.36	.37	.38
By ability:					
<.75	.51	.51	.52	.52	.53
.75–1.00	.49	.49	.50	.51	.52
1.00–1.25	.45	.46	.47	.48	.50
>1.25	.39	.40	.41	.42	.43

NOTE: $\hat{\sigma}_\mu = \$6,048$

tion in these measures due to factors other than schooling and ability are quite important, and that much further analysis is necessary to really pinpoint sources of human wealth inequality. This analysis is meant to be suggestive of the procedure by which more precise implications can be obtained. This general type of analysis can be carried out using any earnings function describing age-earnings profiles, or alternatively, experience-earnings profiles, as a function of characteristics for which data is available on the joint distribution of those characteristics.

NOTES

1. The individual is also assumed to have perfect knowledge of himself and the world and faces no uncertainties.
2. Many aspects of the following discussion are considered in more detail in the growing literature on this subject, including Rosen (1972), Rosen (1973), Mincer (1974), Stafford and Cohen (1973), Stafford and Stephan (1972), Heckman (1973), Smith (1973), Weiss (1971), Razin (1971). The original works of Becker (1962, 1967) and Ben-Porath (1967) are obviously important.
3. The effect of educational doles on the length of the schooling period are considered by Wallace and Ihnen (1972).
4. Time spent in on-the-job training is considered in human capital production, as is investment time off-the-job, rather than in the labor market. The distinction of where investment occurs, on or off the job, has no implication for total investment, assuming a single production function, but does have empirical implications for the interpretation of earnings per unit time for time intervals within a period. They may represent net or gross earnings or even earning capacity. More detail on this issue is considered in Lillard (1973b).
5. $$Q(K_a, D_a) = K_a^{\beta_1} D_a^{\beta_2}$$

   such that $(\beta_1 + \beta_2)\varepsilon(0, 1)$ and $\beta_1 > 0$. It is also assumed in the equation presented here that $\delta\varepsilon(0, 1)$.
6. A general proof that if specialization occurs, it occurs in the initial periods is provided by Ishikawa (1973).
7. The constant fraction result is due to the Cobb-Douglas production function.
8. Becker (1967) provides a discussion of loan markets.
9. For a more detailed discussion of these implicit partials see Wallace and Ihnen (1972).
10. Even though the level of investment is not a function of initial earning capacity, the fraction of earning capacity invested, I_a/RE_a, will be, since earning capacity is.
11. This result obtains from the assumption of no bequest and no restriction on the objective function at N.
12. See for a review, Mincer (1970).
13. Lillard (1973b).
14. Additional polynomial terms were added until they failed to significantly reduce error variance.
15. $R^2 = .2759$. Age and schooling in this equation are years beyond age sixteen. No individual in the sample had less than a high school education. Caution should be taken for predicting below this schooling level, especially late in the life cycle. The

estimates are based on observation of 15,578 age-earnings points from 4,956 individuals. The upper age range of the sample is 54 years and the age-earnings profiles turn down sharply, because there are four men who are three to four years older than the rest of the sample, older when applying in 1943, who have unusually low earnings. All predictions of earnings are restricted as closely as possible to the age range observed.

16. Again, these results are discussed in detail in Lillard (1973b).
17. Due to the data limitations in age mentioned earlier for human wealth predictions, it was assumed that the earnings profiles are flat after the end of the sample range where the profiles peak. I prefer this to either the quadratic or linear profile estimates. For example, in the quadratic estimates, the profiles rise parabolically, since the convexity at early ages dominates the concavity at older ages, which is even more unrealistic.
18. The ability index used in this paper is the first principal component of a subset of the ability test scores corresponding approximately to I.Q.-type attributes. The effect of each individual ability measure and their interactions on earnings and schooling is also discussed in Lillard (1973).
19. Any observation which might cause special problems is omitted. These include those individuals disabled, unemployed, in the military, or who are pilots as their major occupation. Particular year observations for an individual are omitted if, for example, the year of initial job was questionable.
20. Many of these comments originated with F. T. Juster, who directed the data collection for the NBER.
21. The model response was excellent, with 57 percent; 38 percent were good; 3 percent, fair; and less than 1 percent each were poor or nonresponse.
22. When mean earnings predicted distributions are also derived with this assumption.
23. U.S. Census of Population: 1960 (Final Report PC(a)-5A) Subject Reports, School Enrollment: Personal and Family Characteristics of Persons Enrolled in School or College and of Persons Not Enrolled (U.S. Bureau of the Census, 1963, Table 4, page 54).
24. The age is extended to sixty-four because it corresponds to the closest Census of Population age classification, 35–64 years old. The distributions do include persons employed while going to school full time and are correspondingly incorrectly estimated.
25. Forty ability intervals were actually used in calculations.
26. Any assumption about how observations are distributed within reported class intervals is arbitrary. This assumption facilitates calculation of earnings but adds a source of error in the predicted distribution of earnings. The predicted relative frequencies are created in a discrete rather than a continuous manner.
27. Assuming this distribution of ability is a source of error in the predicted distribution to the extent that the distribution of ability of Air Force pilot and navigator school candidates in 1943 is different from the distribution of ability of employed males in 1960.
28. The equal intervals of $1,000 are used to allow the greatest perspective and skewness, since the discrete and widely spaced midpoints of the age and schooling intervals distort the continuity of the predicted distribution. The predicted distributions with unequal interval lengths for higher incomes used in Census of Population tabulations are presented later for comparisons with the actual distributions calculated from Census of Population data.
29. Each individual observations error is distributed independently of age, schooling, ability and the error in any other observation.

30. All interval probabilities are corrected according to the truncated normal, so that only positive earnings are counted and the total relative frequency of all positive earnings is unity.
31. The weights are obviously the relative frequency of the subgroups.
32. Both the average of variances and the variance of averages are calculated weighted by the relative frequency of the subgroups. Formally, $\mathrm{Var}(Y) = E[\mathrm{Var}(Y/\text{subgroup})] + \mathrm{Var}[E(Y/\text{subgroup})]$.
33. The Gini coefficient is the area between the diagonal and the Lorenz curve relative to the area of the triangle, one-half. A larger Gini coefficient implies more inequality. The extremes are zero when every individual gets an equal share of total income and one when one individual holds total income.

    An alternative interpretation of the Gini coefficient is the mean absolute difference between all possible pairs of values relative to their mean, i.e.

$$\frac{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}|X-Y|f(X)f(Y)\,dX\,dY}{2\bar{X}}$$

    X and Y represent all possible pairs of values, earnings, and the numerator is the coefficient of mean difference. The mean difference due to Gini (1912) is dependent on the dispersion of the values among themselves and not on deviations from the mean as in the case of the standard deviation and thus coefficient variation. The Lorenz curve and Gini coefficient are unambiguous measures of concentration only if the Lorenz curves do not cross. An infinite number of Lorenz curves may have the same Gini concentration coefficient if they cross. If two Lorenz curves cross once, say at the point (.7, .3) and have the same Gini coefficient, the population underlying the Lorenz curve which is beneath in the region bounded by (0, 0), (0, .3), (.7, 0), and (.7, .3) may be said to have income distributed more unequally among low income holders (lower 70 percent) than among high income holders, relative to the other population. This says nothing about location of high and low, only about the concentration of low relative to high income holders. This may be thought of as if populations have the same Gini coefficient, and thus, their Lorenz curves must cross, and the same variance and mean, the population with the largest positive skew having its Lorenz curve above the other in the lower earnings region.
34. More detailed comments on calculations of an internal rate of return for the NBER-Thorndike data based on both log equations and present value equalization are presented in Lillard (1975).
35. A maximum of 5 and an average of 3.2 age-earnings points are observed. These are then inflated by the factor $(N-S)$/Number of Points. These estimates are slightly different from the others in that the underlying schooling distribution is that of the Thorndike sample rather than the 1960 Census population.

REFERENCES

Becker, G., 1964. *Human Capital.* Columbia University Press, New York.

———. 1967. *Human Capital and the Personal Distribution of Income.* W. S. Woytinsky Lecture No. 1, Department of Economics, Institute of Public Administration, University of Michigan, Ann Arbor.

Becker, G., and Chiswick, B. R., 1966. "Education and the Distribution of Earnings." *American Economic Review* 56 (2): 358–369.

Ben-Porath, Y., 1967. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy* 75 (4): 352–365.

Ghez, G., and Becker, G. S., 1976. *The Allocation of Time and Goods Over the Life Cycle.* NBER, New York.
Haley, W. J., 1973. "Human Capital: The Choice Between Investment and Income." *American Economic Review* 63 (5):929–944.
Heckman, J. J., 1974. "Estimates of a Human Capital Production Function Embedded in a Life Cycle Model of Labor Supply." Processed. NBER, New York.
Ishikawa, T., 1973. "A Simple Jevonian Model of Educational Investment Revisited." Harvard Institute of Economic Research, Discussion Paper 289, April.
Johnson, T., 1970. "Returns from Investment in Human Capital." *American Economic Review* 60 (4): 546–560.
Levhari, D., and Weiss, Y., 1972. *The Effect of Risk on Investment in Human Capital.* Foerder Institute for Economic Research, Tel Aviv University, October.
Lillard, L. A., 1973b. "An Essay on Human Wealth." NBER Working Paper No. 4.
———. 1973a. "From Age-Earnings Profiles to the Distribution of Earnings and Human Wealth." NBER Working Paper No. 9, September.
———. 1975. "Ex Post Rates of Return to Schooling When There Are Interactions Among Variables in the Earnings Function." Processed, April.
Mincer, J., 1970. "The Distribution of Labor Incomes: A Survey with Special Reference to the Human Capital Approach." *Journal of Economic Literature* 8 (1): 1–26.
———. 1974. *Schooling, Experience, and Earnings.* NBER, New York.
Morgan, J., 1962. "The Anatomy of Income Distributions." *Review of Economics and Statistics* 44 (3): 270–283.
Razin, A., 1973. "On Investment in Human Capital Under Uncertainty." University of Minnesota, processed, March.
Rosen, S., 1973. "Income Generating Functions and Capital Accumulation." Harvard Institute of Economic Research, Discussion Paper 306, June.
———. 1972. "Measuring the Obsolescence of Knowledge." University of Rochester, Working Paper, June.
Schultz, T. W., 1971. *Investment in Human Capital: The Role of Education and of Research.* Free Press, New York.
Smith, James P., 1973. "Family Labor Supply over the Life Cycle." Processed. NBER, New York.
Stafford, F. P., and Stephan, P. E., 1973. "Labor, Leisure, and Training Over the Life Cycle." Working Paper, Georgia State University, November.
Stafford, F. P., and Cohen, M. S., 1973. "A Life Cycle Model of the Household's Time Allocation." Working Paper, University of Michigan, June.
Taubman, P., and Wales, T., 1972. *Mental Ability and Higher Educational Attainment in the Twentieth Century,* Carnegie Commision Technical Report.
U.S. Bureau of the Census, 1963. *Census of Population: 1960.* Final Report PC(2)-5A. U.S. Government Printing Office, Washington, D.C.
Welch, F., 1970. "The NBER Approach to Human Resource Problems." NBER monograph.

14 COMMENTS

Zvi Griliches
Harvard University

This is a long and difficult paper, the result of much research, only part of which is described in the paper at hand. I found it hard going and some of my comments may be based on a misunderstanding or misreading of the paper.

The paper does four things: (1) it presents a brief summary of the state of the optimal-investment-in-human-capital theory and outlines a special version of it to be used further on; (2) it reports on the results of estimating an earnings function (as a function of age, schooling, and ability) based on this theory and on the NBER-Thorndike data; (3) it adjusts these estimates to correspond to the 1960 Census data for the population as a whole; and (4) it computes estimates of the distribution of human wealth as a function of different discount rates and discusses the sources of its variance and alternative measures of inequality. Since I have difficulties with all four of the steps, I shall discuss them in turn.

We need a theory to interpret our data and we must make simplifications and "unrealistic" assumptions to be able to comprehend the world. However, assumptions must be chosen so that they do not eliminate the essence of the problem. Lillard assumes a perfect consumption loan market, implying that one can borrow enough to eat no matter how long one is in school or who one is. This essentially removes any short-term investment funds constraint from the model. (It is true that he must finance tuition and other direct costs from current earnings, but the bulk of investment costs are forgone earnings, not direct costs.) In this model, then, maximizing individuals will differ in their schooling attainment only because they differ in the productivity of learning (ability) and for no other reason. Moreover, no allowance is made in this (or other similar models) for the fact that schooling is subsidized, primarily by parents but also by various public bodies, implying a rather different calculus of the optimal amount of investment in human capital than is assumed by the model and by the interpretation of the estimates.

The version of the model actually computed takes formal schooling as predetermined and concentrates on the effects of postschooling on-the-job training on the age-earnings profiles of individuals. In this version, the family background variables do not appear explicitly, since they are assumed to work entirely via the previously achieved formal schooling levels. Also one should note that ability enters only to the extent that it represents individual differences in the productivity of *equal* amounts of time spent on on-the-job training. It does not reflect the total (reduced form) contribution of ability differences to observed differential in earnings, since it does not allow for the effects of ability on achieved schooling levels or on the supply of hours of work.

Another, more specific difficulty of this type of model is the assumption of the same production function of human capital both in school and in on-the-job training. First, the fact that specialization occurs would make one suspect that the technology of human capital accumulation may not be the same in these rather different pursuits. More importantly, since schooling is largely accomplished in the nonprofit public sector, it is not obvious that one can assume that the observations are "on" the production function or that behavior can be described in terms of it. Of course, one might want to interpret this "production function" just as a summary device describing the constraints facing an individual, but there is no market mechanism then which would force the technology of the public sector to be similar to, and as efficient as, that of on-the-job training, which presumably is occurring largely in the private sector of the economy.

When all is said and done, Lillard estimates earnings as a cubic function, with interactions, of age, schooling, and ability. The relationship of that to the particular theory outlined earlier is obscure. There are a number of practical difficulties with these estimates. Imposing a cubic age structure is a mixed blessing, producing rather strange age-earnings profiles. It is quite clear that in the "real" world earnings of highly educated people do not turn down at age 50, Figure 1A notwithstanding. Also, the prediction that at age 27 high school dropouts earn more than high school or college graduates is inconsistent with all other data known to me. Since the original age distribution of this sample is rather narrow, the estimated effect of age is confounded with that of time, a problem that is not really solved by deflating the earnings series by a general price index. Finally, the estimated ability effects, the primary reason for using the NBER-Thorndike data, are surprisingly small. Moreover, they seem to account for little of the observed dispersion in schooling, mean ability differing by little over one-half of a standard deviation between the lowest and highest schooling levels (see Table 2). This leaves most of the schooling dispersion to be thought of as exogenous, a reasonable but unfortunate conclusion in the context of an *investment* theory of human capital.

It is not clear to me, really, what the purpose is of adjusting the estimated distribution to the 1960 Census levels. The discrepancy between it and the NBER-Thorndike sample is so large, that any extrapolation to the whole population appears to be unwarranted. To recapitulate, the NBER-Thorndike sample covers only the union of the upper half of the schooling and the ability distributions. Extrapolating it to the other quadrants on the assumption that the equations are the same there seems rather farfetched. Moreover, both the schooling of this population and the correlation of ability with schooling are based on economic and institutional conditions as they were in the 1930s and early forties. Much of the 1960 population was educated and selected (by others and itself) in the late 1940s and the 1950s, a much different time period, with different conditions of access to schooling opportunities and economic resources (such as the G.I. bill). There are many peculiarities in the actual tables which may be due to the extrapolation procedure. For example, in Table 3, the mean of predicted earnings is below the median in the highest schooling class, a not very likely occurence. The situation improves in Table 4, when the residuals are put back into the tables, but the estimates still remain rather unrealistic. Comparing the overall predicted results in Table 4 to the actual figures in Table 5, one finds that Lillard's extrapolation procedure significantly underestimates the dispersion and skewness of the actual earnings distribution. If I understand Figures 5A through 5D, they tell us that because the original earnings function accounts for very little of the observed variation in earnings, adjusting the observed distributions for differences in the arguments of this function does not change our view of the inequality of earnings by much (which is a point already made by Jencks, among others).

This problem, the interpretation of the residual variance, also affects the major product of this paper—the estimated dispersion of human wealth. The reasons why wealth distributions would be less unequal than earnings distribution are: (1) the averaging out of the life-cycle inequality of income over time; and (2) the averaging out of the transitory components of earnings over time. Since in Lillard's model, the effects of schooling and ability persist over time, the dispersion due to these sources should not cancel out. The role of unexplained sources of variation is illustrated in my Table 1, culled from Lillard's paper.

TABLE 1 Overall Coefficients of Variation

		Human Wealth	
	Earnings	At .03 Discount Rate	At .07 Discount Rate
Predicted	.38	.09	.05
Adjusted for unexplained variation	.59		
Lower bound		.13	.17
Upper bound		.70	.83

Whether human wealth is really more equally distributed than earnings depends on our assumption about the persistence of the random unexplained components. If "luck" or unobserved characteristics persist, then it is not true that wealth dispersion is much smaller than earnings dispersion. This is clearly an important area for further research. Even within Lillard's own data, it would have been possible to estimate the fraction of the random variance that persists and get a narrower bound.

Finally, I am not sure that the computed wealth distributions have much operational content, since the human capital market is far from perfect and the average twenty-six year old cannot really borrow $70,000 plus solely on the basis of his human capital.