NBER WORKING PAPER SERIES

From Age-Earnings Profiles to the Distribution of Earnings and Human Wealth

L. A. Lillard

Working Paper No. 9

CENTER FOR ECONOMIC ANALYSIS OF HUMAN BEHAVIOR AND SOCIAL INSTITUTIONS
National Bureau of Economic Research, Inc.
261 Madison Avenue, New York, N.Y. 10016

September, 1973

Preliminary; Not for Quotation

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (concluded)

September 1973

From Age-Earnings Profiles to the Distribution of Earnings and Human Wealth

by L. A. Lillard*

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 it. The purpose of this paper is to suggest how these earnings functions can be used more directly to derive predicted earnings and human wealth distributions for populations and sub-populations, with an empirical illustration. This is accomplished by an application of statistical distribution theory as a link between the earnings function and the earnings distribution. The appropriate statistical distribution theory is first presented in very general terms. The means of translating the joint distribution of a population's characteristics into earnings distribution through the earnings function is presented. When an explicit relationship

*The research reported here is an extension of a part of my doctoral dissertation which benefited from comments by T. D. Wallace, L. A. Ihnen, and B. Gardner. Helpful comments were also given by Finis Welch.

The research reported herein was performed pursuant to grants from the National Science Foundation and the U.S. Department of Labor. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any granting agency.

between earnings or human wealth and the characteristics of an individual which determine it is specified and when the joint density of those characteristics in a population is known the distribution of that population can be deduced.

The empirical illustration entails application of the general framework to a particular set of data, the 1960 Census of Population, using an estimated earnings function. Both distributions of earnings and human wealth are predicted. The predicted distributions are then compared with actual distributions when possible.

Earnings as a cubic function of age, schooling, ability and their interactions are estimated using results from the NBER-Thorndike sample. 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 based on age and schooling data from the 1960 Census of Population and ability data within schooling classes from the NBER-Thorndike sample. Predicted earnings distributions are derived for the overall population, for schooling classes, for age groups, and for ability classes. Both the distribution of expected or mean earnings 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 distribution of human wealth of the 1960 Census population is then predicted. Human wealth is defined as the present value of the earnings stream from the end of schooling until age sixty-six. Human wealth is then a function of schooling, ability, and the rate of discount. It is assumed that everyone has the same rate of discount but rates from 3 to 7 per cent are considered. Human wealth distributions are predicted for the overall population, for schooling classes and ability classes.

The most striking result of the empirical estimation is that there is clearly a great deal more inequality in earnings than in human wealth. Also an interesting result is that at a discount rate above 5 per cent an increased schooling level of the population does not necessarily increase the expected value of human wealth and that at 6 per cent and above it actually declines monotonically. Detailed comparisons of age groups, schooling groups and ability groups are discussed.

FRAMEWORK FOR TRANSLATING THE JOINT DISTRIBUTION OF POPULATION CHARACTERISTICS INTO EARNINGS DISTRIBUTIONS

Life cycle earnings models allow the specification of a functional form for the earnings of an individual over his life cycle. Earnings at any point in the life cycle are determined by both formal schooling and past schooling investment behavior. Earnings are a function of, for example, the initial stock of human capital at the beginning of the life cycle, the human capital production parameters including the efficiency parameter β and the input coefficients β_1 and β_2, the rate of depreciation of the capital stock, the length of working life, the market interest rate, the rental rate of human capital, the price of direct educational imputs and of course, age. The characteristics may be summarized by the vector

$$C = (\beta, E_o, \beta_1, \beta_2, \delta, \pi, N, P, R) \quad (1)$$

The purpose of this section is to show how the overall earnings distribution of a population and earnings distributions of various subpopulations can be derived from the joint distribution of characteristics of that population. The joint distribution of characteristics can be transformed through the functional form for earnings into the various earnings distributions.

Any individual takes his characteristics as given, and not under his control, and invests in an optimal fashion accordingly. For any individual, earnings are thus a function of age with his characteristics acting as parameters. For a given population the individuals in it will differ in characteristics as well as age. The characteristics are then variable across individuals in the population as is age and the joint density function of age and the vector of characteristics.

Variation in earnings within the population can be considered as coming from two sources, variation in age and variation in characteristics. Individuals with different parameters and endowments are constrained to select different optimum age profiles. Individuals with the same parameters and endowments are constrained to select different optimum age profiles while individuals with the same parameters and endowments select the same investment patterns but still differ in age. These sources of variation are considered separately, then combined.

Variation in Earnings due to Age Only

Variation of earnings due to variation in age in the population is an important source of income inequality and can be analyzed by means of the life-cycle earnings function. The specification by means of the life cycle model of an earnings function allows the age profile of earnings for an individual to be identified and thus allows a much more refined analysis of the population income density function because the age distribution of the population can be explicitly accounted for.

First, consider the effect of the age distribution on the income of identical persons. That is, consider the population of all persons who have exactly the same initial endowment, production parameters, depreciation rate, working life, input price and rental rate. Let these persons differ in age. Since the individuals are identical in all characteristics except age, they all have the same age profile but are at different points along that profile. For each age there is a unique income, Y. Let Y(a,C) be the earnings function, where a is age[1] and C is a vector of parameters and initial endowments.

Consider the density of persons with respect to age. The density function $g_1(a)$, where a stands for age, is given by

$$\begin{cases} g_1(a) & 0 \le a \le N \\ 0 & \text{elsewhere} \end{cases} \quad (2)$$

Consider the two basically different regions of a with respect to the age profile.[2] (See Figure 1.)

$$a^* \le a \le a_{max} \quad (3)$$

where Y(a,C) is monotonically increasing with respect to a, and

$$a_{max} < a \le N \quad (4)$$

where Y(a,C) is monotonically decreasing with respect to a where:

---

[1]Note that age is different from work experience. Age begins at the time when the individual makes a decision by selecting an investment pattern, or begins to make his own investment decisions. Work experience begins when the individual enters the labor force, but a year of labor force participation does not necessarily imply a full year of experience especially if the work is not full time.

[2]Life cycle theories indicate age profiles demonstrating these characteristics. The monotonicity need not be limited to two regions, but this assumption simplifies the analysis. The shapes of the curves are assumed but are consistent with the empirical findings.

a^* = age where positive earnings begin

a_{max} = age where earnings are a maximum

N = age at the end of the work life

In the region a^* to a_{max} the individual is investing only a fraction of his potential earning and is realizing positive earnings and those earnings are rising monotonically. In the region a_{max} to N, the individual's earnings are declining monotonically due to low investment and a dominating depreciation of existing human capital. This last region need not exist if there is no depreciation.[3] In the region 0 to a^*, the individual is investing all of his potential earning in producing more human capital and is realizing no income.[4] Individuals in this age category will not be included in the discussion since earnings while in full-time schooling are not generally reported and the data used in a later section will be for persons reporting earnings from employment. Therefore, all earnings distributions are for persons reporting positive earnings and those persons still in full-time schooling are not included.

The objective is to transform the density of persons with respect to age into the density of persons with respect to income via the earnings function Y(a,C). (See Figure 1) To accomplish this, first solve for the inverse functional relationship mapping income into age. This can be done only for monotonically increasing or decreasing regions. We

---

[3]If the stock of human capital does not depreciate, earnings received as a rent on that stock must rise since the stock can only grow and investment in producing more is a declining function of age.

[4]Net earnings will be zero, but observed earnings may not be under imperfect loan markets.

defined Y(a,C) to be monotonically increasing in the region $a^* \leq a \leq a_{max}$ and monotonically decreasing in the region $a_{max} \leq a \leq N$. On each region we can then define the inverse function $a = Y^{-1}(Y)$. For any income in the range $Y(a^*, C) \leq Y < Y(N,C)$ there is exactly one inverse age in the region $a^* \leq a < a_d$.[5] However, for any income in the range $Y(N,C) \leq Y < Y(a_{max},C)$ there are two ages corresponding to that income.[6] These considerations must be dealt with in the transformation to obtain the income density function.

Given the foregoing specifications, the conditional probability density function for income is $g_3(Y/C)$. The probability density element for any age interval is transformed through the earnings function into the probability density element of the corresponding income interval. Age density is transformed uniquely from the age interval $a^* \leq a < a_d$ into the income interval $Y(a^*, C) \leq Y < Y(a_d,C)$. However, in the age interval $a_d \leq a \leq N$ two age intervals map into each income interval, one from the interval $a_d \leq a \leq a_{max}$ and another from the interval $a_{max} < a \leq N$. A doublecounting of income occurs then at higher levels of income, while a single counting occurs at lower levels of income with a discrete jump point at the income $Y(a_d, C)$. These statements are presented in statistical terms in Appendix A.

For a given density function of age a downturn in earnings at later years in the life cycle adds a doublecounting effect which introduces a

---

[5]The point a_d is the age at which income during the life cycle equals income at the end of the working life. Thus, all incomes above Y(N) occur twice, once while rising and once while declining.

[6]Again the slopes of the curves are assumed. It is possible to consider the case where $Y(N,C) \leq Y(a^*,C)$ but the above case was chosen for expositional purposes.

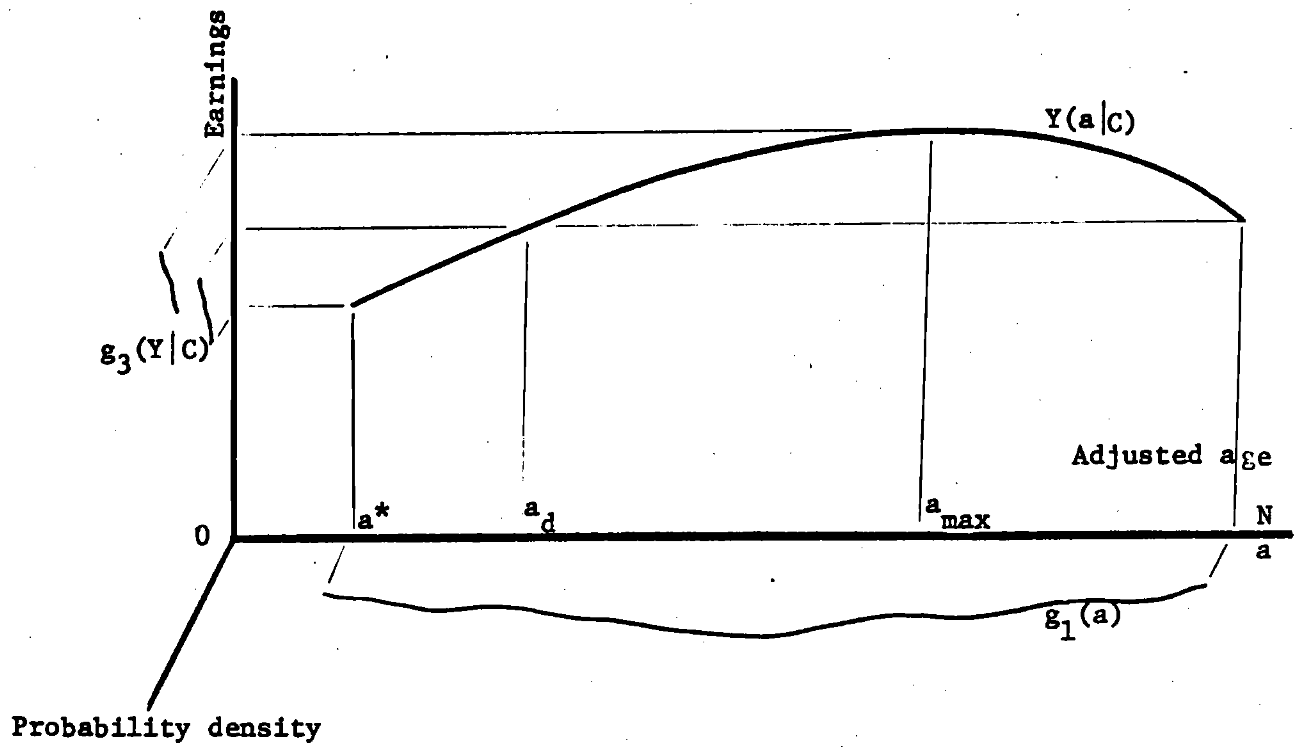

Figure 1. Transforming age density into income density

source of negative skewness or at least less positive skewness that is not present if the earnings function increases monotonically.

A second source of a tendency toward negative skewness can be noted at this point. If the earnings function rises at a decreasing rate and falls at an increasing rate as is drawn in Figure 6, then as the age interval moves toward a_{max} from a*, or from N, it maps into increasingly smaller income intervals thus transforming probability density from equal age intervals into increasingly smaller income intervals. This tends to increase the density of high income levels relative to low income levels for a given earnings function and a given age density function relative to what it would otherwise be. These sources of a tendency toward negative skewness have not been noted in the income distribution literature.

The moments of this conditional income distribution are functions of the moments of age distribution. The relationships are specified through the earnings function. The moments of the income distribution of persons with the same parameters and initial endowments are conditional upon and thus functions of those characteristics. It should also be noted that the moments of income for an entire population are functions of the moments of these conditional moments over characteristic space. See Appendix B for a more precise statement.

Variation in Earnings due to Variation in Endowment and Characteristics Only

Consider the effect of the distribution of persons with respect to characteristics on the income of persons the same age.[7] Each individual

[7] This statement implied using the joint density with respect to characteristics of all persons at the given age. Alternatively, it is possible to consider the distribution of earnings if everyone were at the specified age by using the overall joint of the entire population.

with different characteristics has a different income profile but is the same distance a into his profile. The distribution of income will derive from the distribution of persons by characteristics at that age (see Figures 2 and 3). There is a different density function for each age a, but there is a continuous array of density functions over all ages between 0 and N.[8]

The objective is to transform the joint density function of persons with respect to parameters and initial endowments into the density of persons with respect to income via the earnings function. The distribution of these characteristics is assumed independent of age. Density is transformed from an interval in n-dimensional characteristic space into a corresponding income interval through the earnings function. The moments of this conditional income distribution are functions of the moments of the characteristics. The relationship is specified through the earnings function. The moments of the income distribution of persons the same age is conditional upon and thus a function of age. For example, the popular belief that the variance of income increases with age as summarized by Mincer (1970) can be considered in this context. An empirical consideration of such matters appears in a later section. For a fuller exposition, see Appendix A.

Overall Variation in Earnings

Consider the combined effect of the variation of the population with respect to both age and parameters and initial endowments on the distribution of personal income. Probability density is transformed from

---

[8] Again only persons having completed full-time schooling are included in the age distribution.

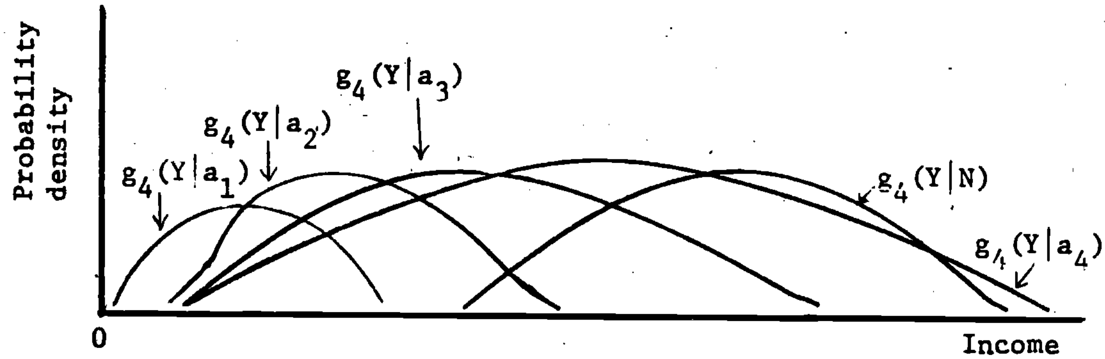

Figure 2. Conditional income density functions for several ages in two dimensions

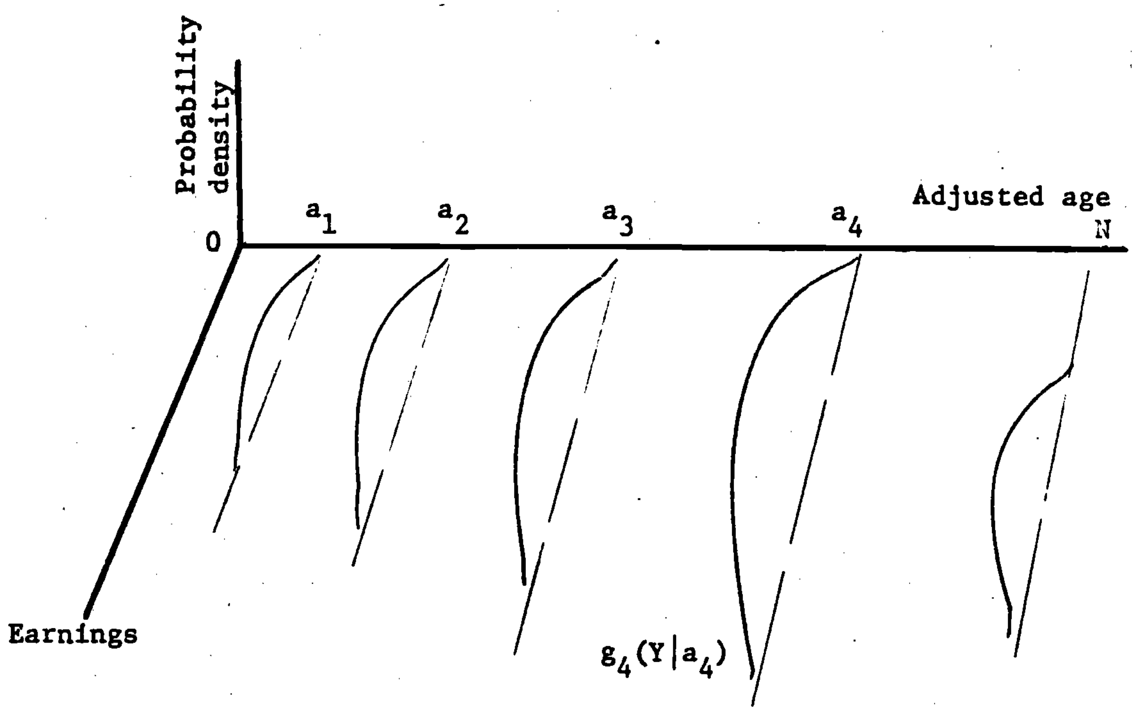

Figure 3. Conditional income density functions for several ages in three dimensions

intervals in (n+1) dimensional age-characteristic space into income intervals through the earnings function. This is considered as an extension of the previous framework.

In an earlier part of this section the density function of income at each age was derived (see Figure 3). Now combine these density functions to form the joint density function of income and age, $g_6(Y,a)$ in Figure 4. The conditional densities $g_4(Y|a)$ are in effect scaled down by the marginal density of age $g_1(a)$ through the relationship

$$g_6(Y,a) = g_4(Y|a) \cdot g_1(a). \tag{5}$$

The overall earnings distribution represented by the marginal density function of income Y is obtained by integrating the joint density function $g_6(Y,a)$ over the age dimension. As is illustrated in Figure 5, this sums all of the density from two dimensions into one. The contour of density is effectively flattened against the income density plane. Again the moments of the personal income distribution are functions of the moments of age and of the moments of the parameters and initial endowments. These relationships are estimated in a later section. For a more complete statement of these concepts see Appendix A.

Another way to view this same process is to reconsider Figure 2. Graphing the density functions $g_4(Y|a)$ each weighted by $g_1(a)$ in income-probability space results in a continuum of curves such that when the heights or functions are added overall a density function of income is obtained as is illustrated in Figure 6.

Thus, the density function for income is derived from the density functions of age and of parameters and initial endowments. The analysis proceeded through several stages, each of which yields useful information about the personal income distribution.

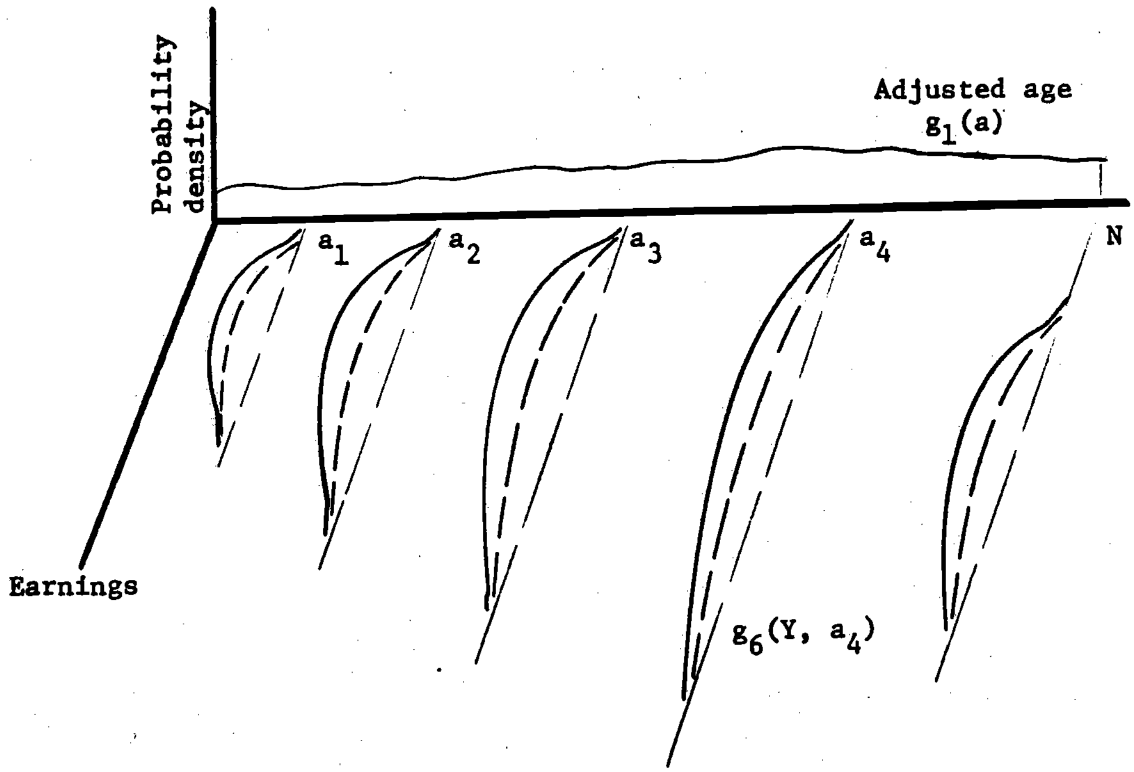

Figure 4. Scaling the distribution of income conditional on age to obtain the joint distribution of income and age

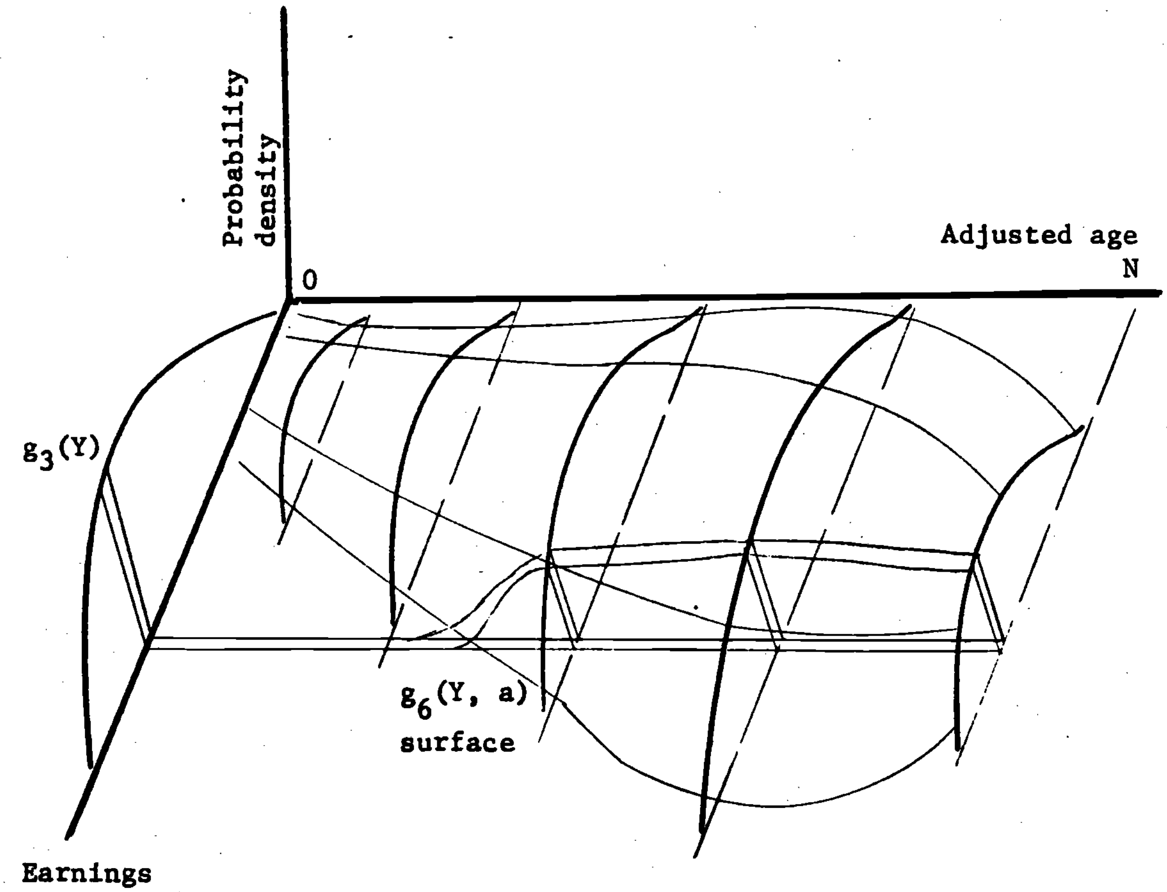

Figure 5. Summing the joint density of income and age over the age dimension to obtain the marginal density of income in three dimensions

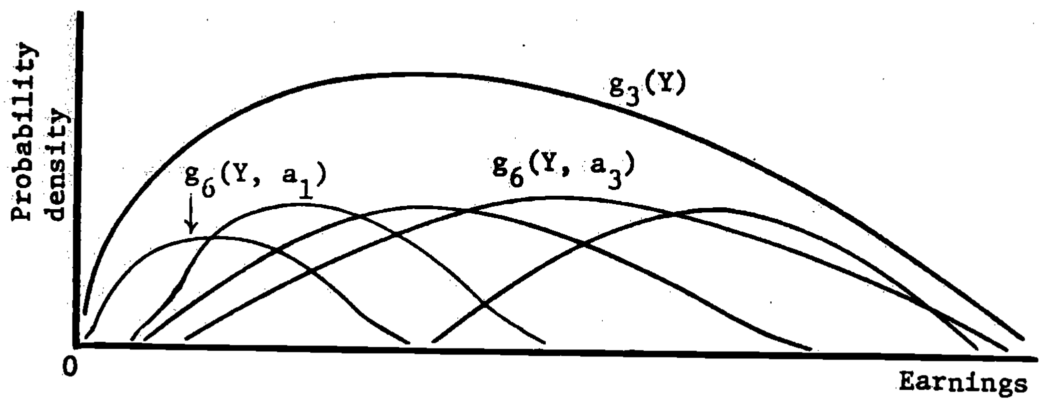

Figure 6 . Summing the joint density of income and age over the age dimension to obtain the marginal density of income in two dimensions

Variation in the Present Value of Net Earnings, Human Wealth

Consider the distribution of the population by human wealth.[9] Human wealth is defined as the present value of the stream of net earnings over the individual's life cycle at the age a = 0. Therefore,

$$W(C) = \int_{a^*}^{N} e^{-ra} Y(a,C)\, da. \tag{6}$$

Differences in wealth between individuals occur because of differences in their characteristics. If we know the joint density function of C, we can derive the density element of W(C) in a manner similar to that of income.

The implications of this general framework are as follows. Whether we are interested in studying the distribution of earnings or the distribution of human wealth, the life cycle income theories tell us that the distribution of the population by characteristics is important. These characteristics determine the optimal time spent in formal schooling, and after schooling investment decisions as well as the flow of actual earning and maximum human wealth.

---

[9] See T. W. Schultz (1971, p. 23) and an unpublished paper, the NBER Approach to Human Resource Problems, by Finis Welch, Research Associate, National Bureau of Economic Research, New York, 1970, p. 6, for further discussion of this concept.

AN EMPIRICAL ILLUSTRATION

This section will provide an illustration of the empirical use of the more general framework. An estimated earnings-age relationship dependent upon level of schooling and ability is described and earnings distributions as well as human wealth distributions are predicted from it. The estimated earnings function is based on the NBER-Thorndike sample of Air Force pilot and navigator school candidates in 1943. The predicted distributions are for employed males in 1960. This section may be characterized as answering "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?" Several caveats are in order to use one group to predict the other. Predicted and actual 1960 distributions are compared when possible.

A Specific Earnings Function and Estimates

It is well founded theoretically and empirically that earnings depend upon schooling, ability, and age or experience.[10] 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.[11] 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.[12] The estimated earnings function is

$$
\begin{aligned}
Y(A,S,B) &= 21108.50 - 3921.20A + 877.25S + 148.02SA \\
&+ 206.09A^2 - 794.20S^2 + 6.87SA^2 + 116.42S^1A - 7.82S^2A^2 \\
&-45197.00B + 11015.00BA + 4721.40BS - 1820.80BSA - 594.93BA^2 \\
&+ 1065.00BS^2 + 83.51BSA^2 - .\ BS^2A + 8.56BS^2A^2 \\
&+28134.00B^2 - 6738.40B^2A - 5035.20B^2S + 1435.20B^2SA + 371.38B^2A^2 \\
&- 240.65B^2S^2 - 72.59B^2SA^2 + 5.86B^2S^2A + 0.99B^2S^2A^2 \\
&- 2.99A^3 - 0.31A^3S + 0.15A^3S^2 + 9.09BA^3 - 1.04BA^3S \\
&- 0.17BA^3S^2 - 5.74B^2A^3 + 1.04B^2A^3S + 0.03B^2A^3S^2.
\end{aligned}
$$
[13]

---

[10] See for a review, Mincer (1970, JEL).

[11] Lillard "Human Capital LIfe Cycle of Earnings Models: A Specific Solution, and Estimation", NBER Working paper No. 4.

[12] Additional polynomial terms were added until they failed to significantly reduce error variance.

[13] $R^2 = .2759$. 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 resulting age-earnings profiles are presented in Figures 7 for various ability and schooling levels. Both schooling and ability raise earnings at every age in the life cycle after some initial period.[14] 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 8 for discount rates of 4 and 6 percent.[15]

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 bombadier programs in the last half of 1943. These volunteers were given initial screening tests and a set of seventeen tests to measure various abilities 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. The NBER sent to a subset of these a subsequent questionnaire in 1969 which included additional questions on earnings in later years and questions on schooling and initial job earnings.

---

[14] Again these results are discussed in detail in Lillard (1973).

[15] Due to the data limitations in age mentioned earlier, it was assumed that earnings remained constant over the life cycle after the peak earnings period.

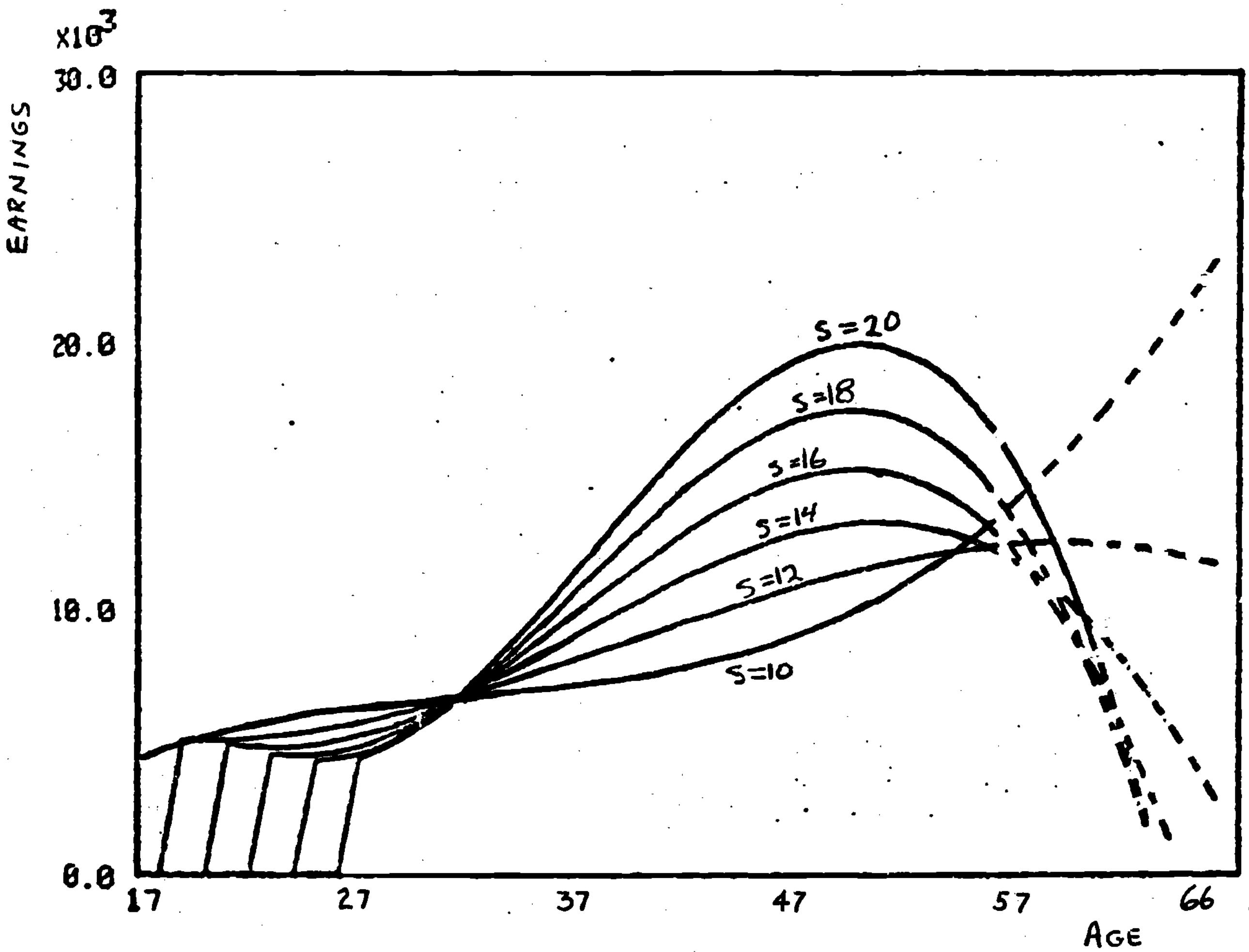

Figure 7.a. Cubic Estimated Age-Earnings Profiles Based on the NBER-Thorndike Sample for Several Schooling Levels at the Average Ability Level.

Note: All earnings are in 1957-59 dollars.

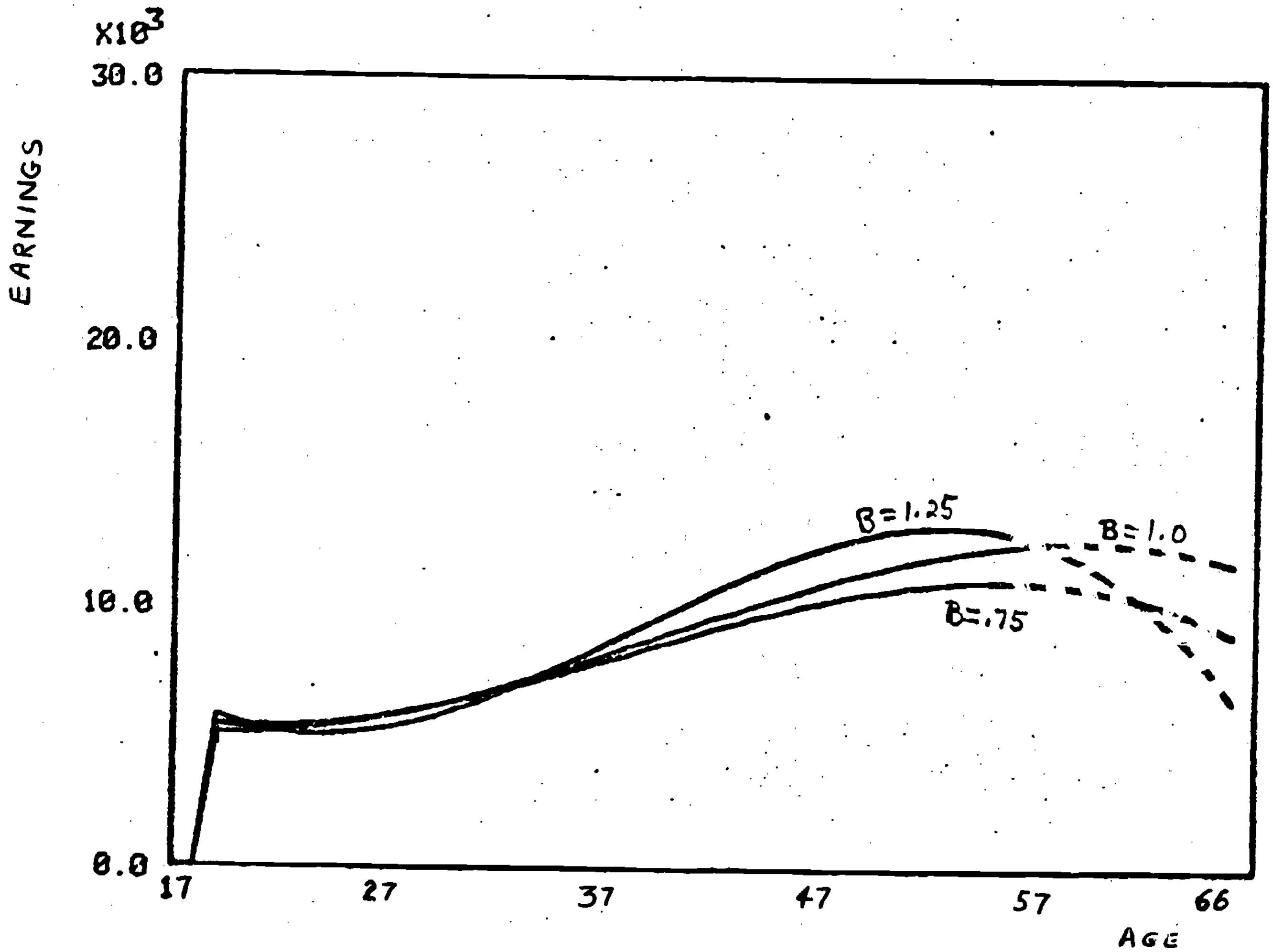

Figure 7.b. Cubic estimated age-earnings profiles based on the NBER-Thorndike sample for average ability and one standard deviation (.25) above and below, for high school graduates (S=12) and college graduates (S=16). S=12, cubic.

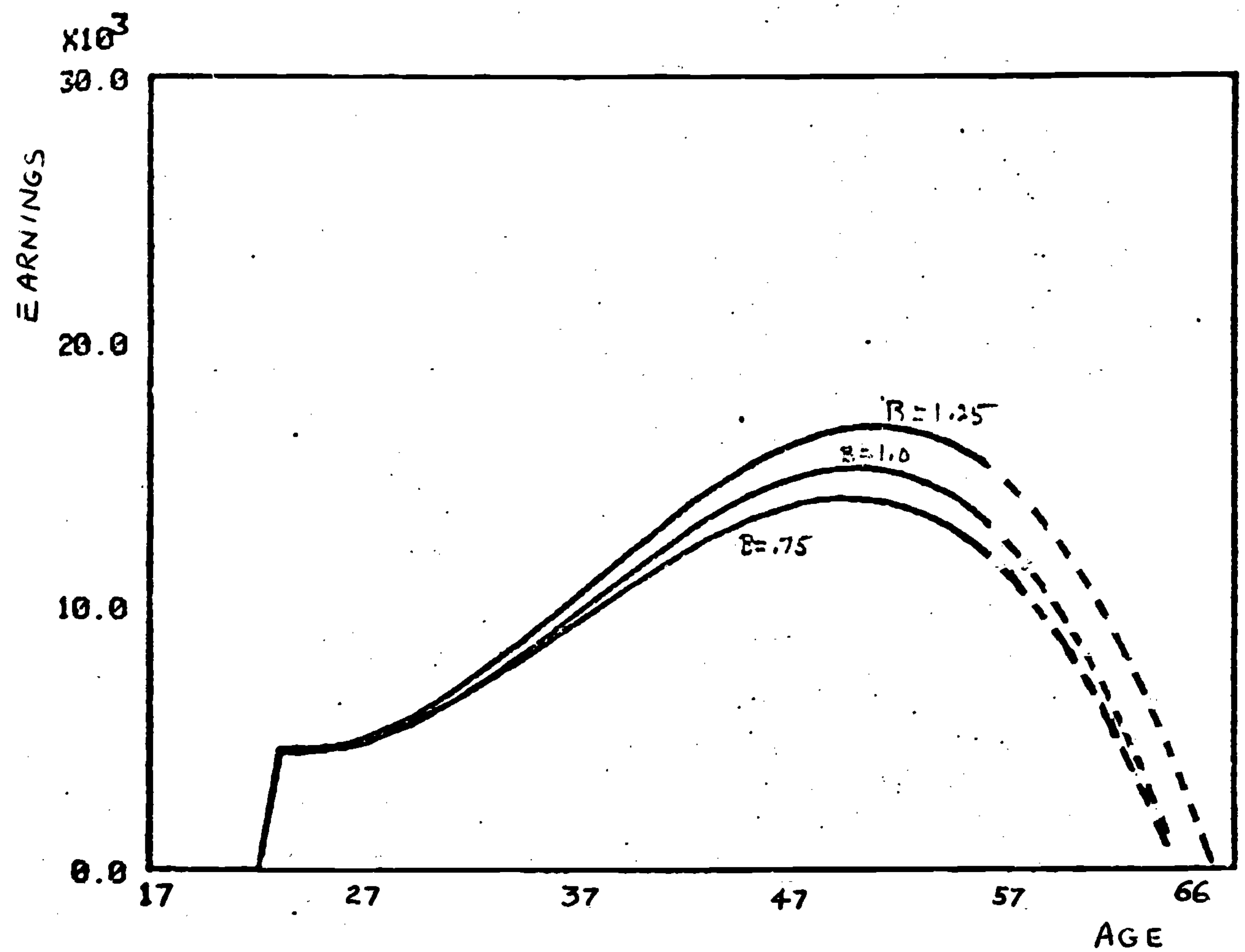

Figure 7.c. (continued, page 3). Cubic, S=16.

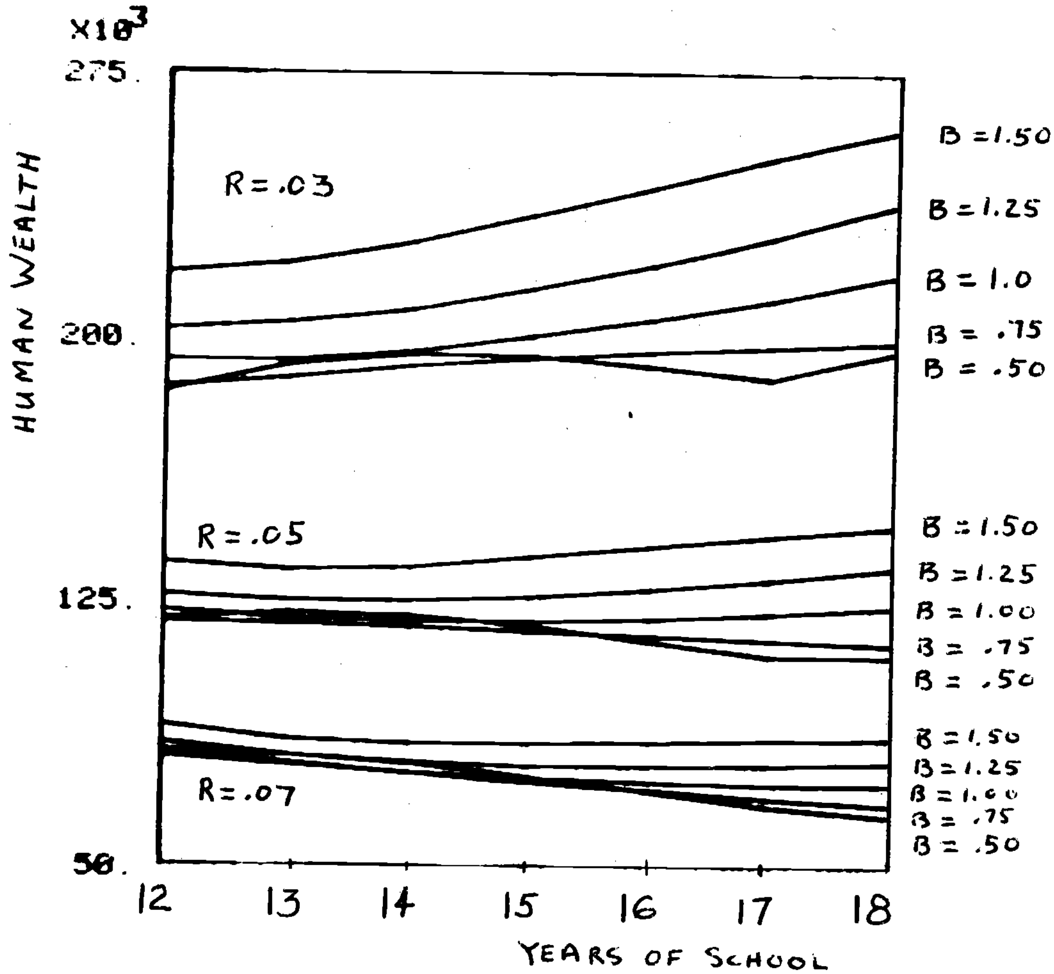

Figure 8.a. Present value of predicted observed earnings from the estimated quadratic age-earnings profiles based on the NBER-Thorndike sample. As a function of schooling. (N=66). Discounted at 3 per cent, 5 per cent, and 7 per cent.

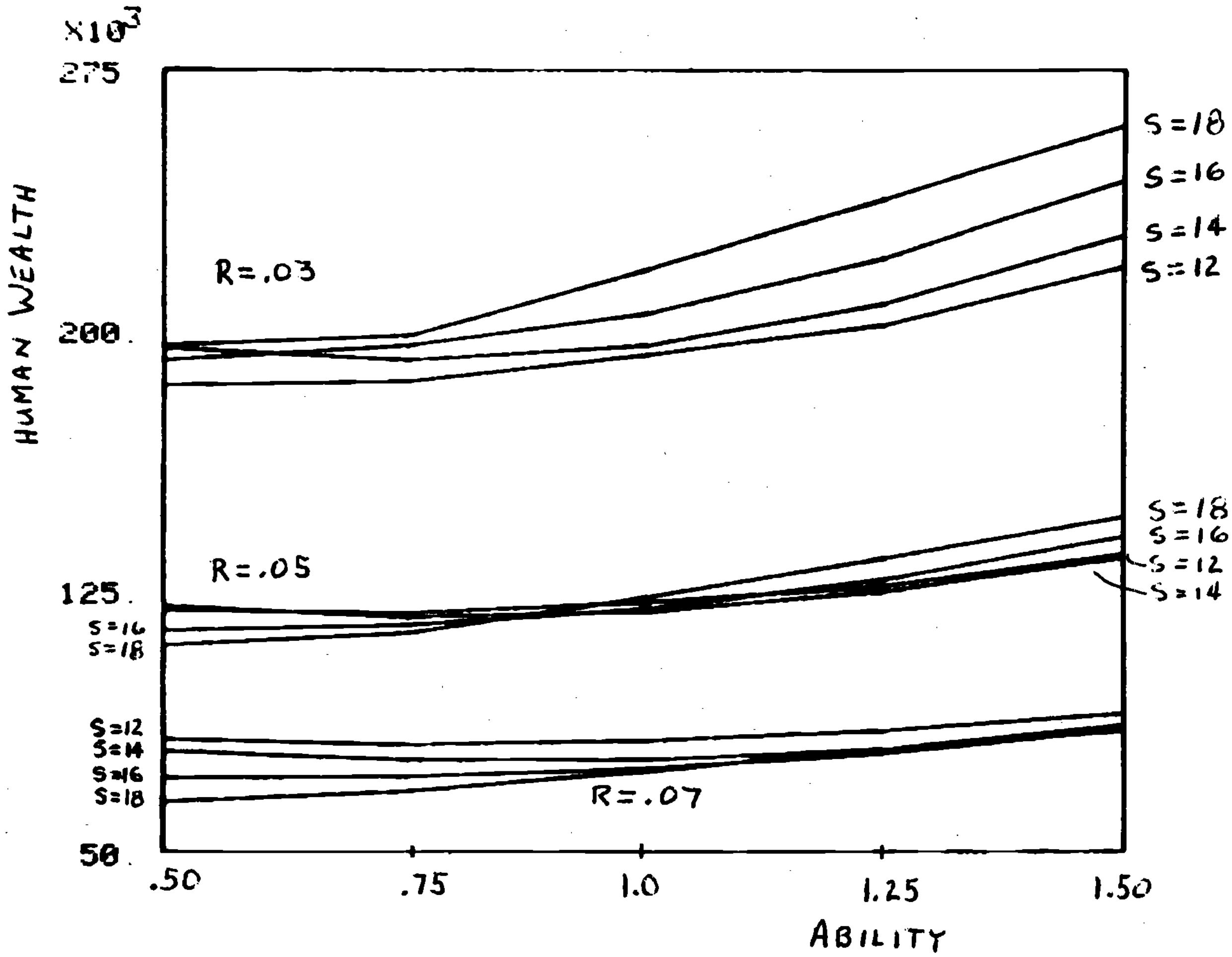

Figure 8.b. Present value of predicted observed earnings as a function of ability from the estimated quadratic age-earnings profiles based on the NBER-Thorndike sample. (N=66). Discounted at both 4 per cent and 6 per cent.

The data includes five separate approximately equally spaced points[16] 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.[17] First the sample includes a high ability group. All of the men completed high school of high school equivalency examinations, and passed the initial screening for the Air Force flight program. Their general health was better than the general population[18] in 1969. They were more homogenous 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 per cent 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 per cent outside the range nineteen to fifty-five. The cubic earnings equation is

---

[16]Any observation which might cause special problems is omitted. These include those individuals disabled, unemployed, in the military, or who is a pilot at his major occupation. Particular year observations for an individual are omitted if for example the year of initial job was questionable.

[17]Many of these comments originated with T. F. Juster who directed the data collection for the NBER.

[18]The model response was excellent with 57 per cent, 38 per cent were good, 3 per cent fair, and less than 1 per cent each were poor or nonresponse.

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.[19]

---

[19] When mean earnings predicted distributions are derived without this assumption (and negative earnings are not counted in relative frequencies) about 10 per cent of total frequency is lost primarily from high schooling, low ability, and late age groups. Three, fifteen, eighteen, and twenty-seven per cent are lost within schooling classes 12, 13-15, 16, and 17+, respectively. Four, eight, twenty, and thirty-eight per cent are lost within age groups 30-34, 35-44, 45-54, and 55-64, respectively. Thirty-three and fourteen per cent are lost within the lowest and second lowest (both below average) ability groups, respectively.

Deriving Earnings Distributions from the Estimated Earnings Function

The distribution of earnings derives from the distribution of the population with respect to age, ability, and time in schooling. The purpose of this section is to apply the principles of the general framework using the estimated earnings. Use of the framework is illustrated by using 1960 United States Census of Population data on the distribution of the United States 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.[20]

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 the earnings function is estimated based on a sample reporting earnings only from ages nineteen to fifty-five, all distributions are for that age group only. Since all persons in the NBER-Thorndike sample 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

---

[20]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).

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 1970 Census of Population.

	Years of Schooling				Age Marginal
	12	13-15	16	17+	
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

are employed.[21] 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. The distribution of ability by schooling class used is presented in Table 2. 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.[22]

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.[23] These relative frequencies are then summed into relative

---

[21]The age is extended to forty-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.

[22]Any assumption about how observations are distributions 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.

[23]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.

Table 2

Distribution of Ability Overall and By Schooling Level From the NBER-Thorndike Sample for Schooling Interval Midpoints in the 1960 Census of Population

Ability Index	Ability Conditional on Schooling — Years of Schooling: 12	14	16	18	Ability Marginal
0.05					.0
0.10	.0008				.0005
0.15	.0008		.0007		.0006
0.20	.0		.0		.0
0.25	.0008		.0		.0005
0.30	.0033	.0026	.0		.0024
0.35	.0	.0026	.0007		.0007
0.40	.0017	.0	.0043		.0015
0.45	.0075	.0026	.0029	.0026	.0054
0.50	.0150	.0026	.0050	.0026	.0099
0.55	.0233	.0158	.0093	.0052	.0182
0.60	.0366	.0184	.0143	.0157	.0279
0.65	.0450	.0342	.0228	.0184	.0373
0.70	.0616	.0500	.0314	.0281	.0521
0.75	.0849	.0763	.0342	.0367	.0720
0.80	.0874	.0658	.0527	.0525	.0750
0.85	.0916	.1053	.0670	.0840	.0907
0.90	.0883	.1079	.0656	.0630	.0871
0.95	.0891	.0605	.0720	.0840	.0804
1.00	.0741	.0974	.0805	.0735	.0798
1.05	.0774	.0605	.0834	.0604	.0729
1.10	.0583	.0632	.0770	.0840	.0642
1.15	.0391	.0474	.0606	.0551	.0451
1.20	.0358	.0500	.0563	.0682	.0446
1.25	.0258	.0263	.0499	.0630	.0326
1.30	.0208	.0342	.0513	.0525	.0306
1.35	.0100	.0237	.0385	.0262	.0181
1.40	.0100	.0184	.0356	.0289	.0168
1.45	.0025	.0132	.0221	.0289	.0098
1.50	.0025	.0132	.0185	.0210	.0086
1.55	.0042	.0	.0200	.0105	.0059
1.60	.0008	.0026	.0078	.0131	.0033
1.65	.0008	.0053	.0071	.0052	.0030
1.70			.0050	.0079	.0014
1.75			.0029	.0026	.0006
1.80			.0	.0	.0
1.85			.0007	.0026	.0003
1.90				.0	.0
1.95				.0026	.0003

frequencies of yearly earnings for intervals of a thousand dollars.[24]

The resulting predicted overall distribution of earnings and the predicted distribution for various subpopulations effectively represent distributions of mean earnings. However, only about 28 per cent of the variation is earnings is explained by variation in age, schooling, and ability.

Consider the problem of the distribution of earnings after correcting for variation not accounted for by variation in age, schooling, and ability. A very simple approximate estimating procedure is used here. The error variance of the estimating equation is 36593472. (standard error = 6049.25). It is assumed that the errors are identically and independently[25] distributed with mean zero and standard deviation 6049.25. 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. The interval in which the predicted earnings value

[24] The equal intervals of $1,000 is 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.

[25] Each individual observations error is distributed independently of age, schooling, ability and the error in any other observation.

falls receives an incremental relative frequency of .0662 times the relative frequency of that age, schooling, ability combination.[26] 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 years 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 Figure 9. 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

---

[26]Interval length is $1,000. All intervals 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.

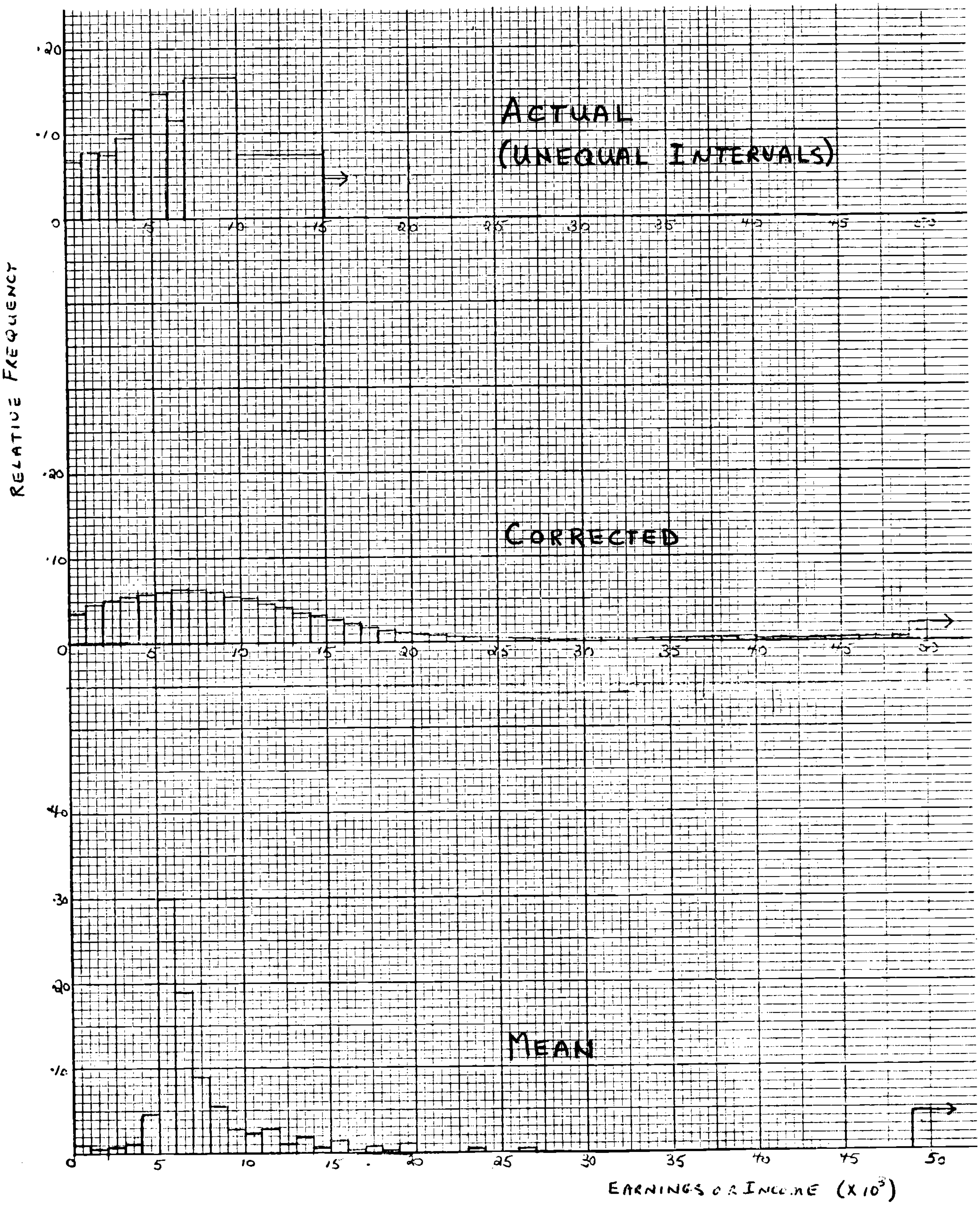

Figure 9. Predicted Mean Earnings, Predicted Corrected Earnings and Actual Income Distributions for employed males between the ages eighteen and sixty-four with at least a high school education, and the corresponding Lorenz Curves.

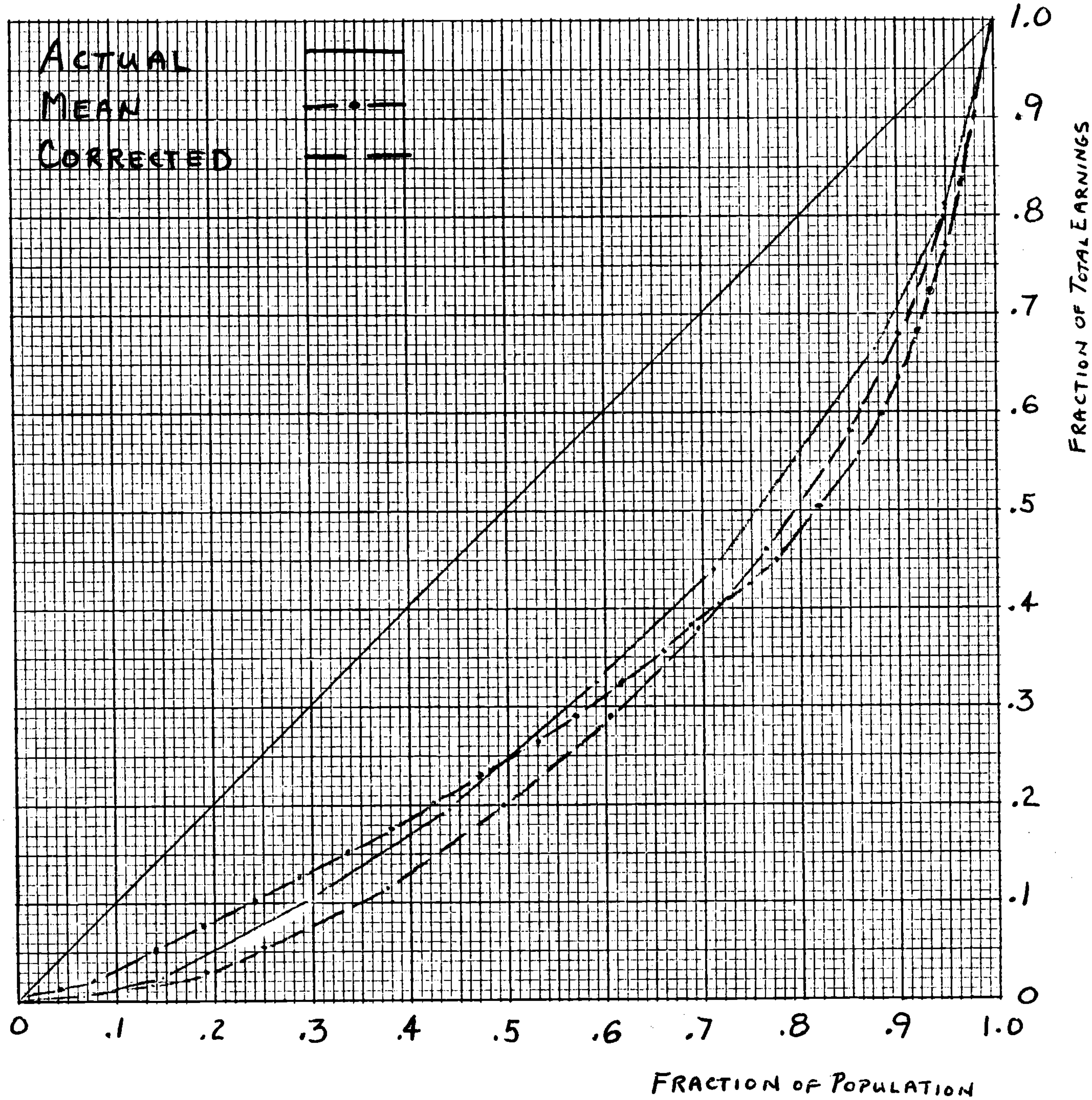

Figure 9. (continued, page 2)

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 per cent 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 different groupings. The interval midpoint 25,000 is used for the interval > 15 in the Census of Population while equal 1,000 intervals up to 50,000 are used for predicted distributions.

Several important differences remain. The 1960 Census figures are for total income while 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 while these persons are excluded in estimating the earnings function. This contributes largely to the large relative frequency of very low income at early ages. For example, 84 per cent of sixteen and seventeen year olds earned less than $1,000 as did 53 per cent of eighteen and nineteen year olds. These are likely to be mostly 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 are in Figure 10 and Table 4. Lorenz curves are plotted in Figure 10.

Table 3

Selected Statistics for Predicted Mean Earnings Distributions for Employed Males Eighteen to Sixty-six With at Least a High School Education.

	Mean	Std. Dev.	Coeff. of Var.	Skew.	Gini Coeff.	Median
Overall	10736.36	10811.56	1.01	3.55	.59	5622.25
By Schooling						
12	7856.08	3598.28	.46	3.12	.81	6665.54
13-15	10732.80	10451.32	.97	3.76	.61	5915.58
16	16407.45	15961.90	.97	1.77	.51	7758.43
17+	19968.13	19424.57	.97	1.00	.47	11332.00
By Age						
18-19	5678.17	501.43	.09	3.87	.97	5582.41
20-21	5618.79	416.75	.07	4.84	.98	5552.38
22-24	5534.43	581.58	.09	1.91	.97	5509.30
25-29	5759.62	1488.06	.26	1.91	.90	5586.29
30-34	7993.22	4812.79	.60	4.18	.75	6774.98
35-44	11685.52	10386.26	.89	3.73	.62	8046.59
45-54	15900.39	14600.20	.92	2.02	.55	8860.97
55-64	18389.76	16746.86	.91	1.43	.53	8942.47
By Ability						
<.75	6981.21	2167.35	.31	-0.33	.83	6749.76
.75-1.00	6109.06	1620.06	.27	-0.17	.86	5961.75
1.00-1.25	15263.06	12653.57	.83	2.62	.60	10540.84
>1.25	32665.95	17609.30	.54	0.04	.71	38554.34

Note: Skewness is measured by the square root of $E(X-\bar{X})^3 S^3$. Coefficient of variation is $S/\bar{X}$.

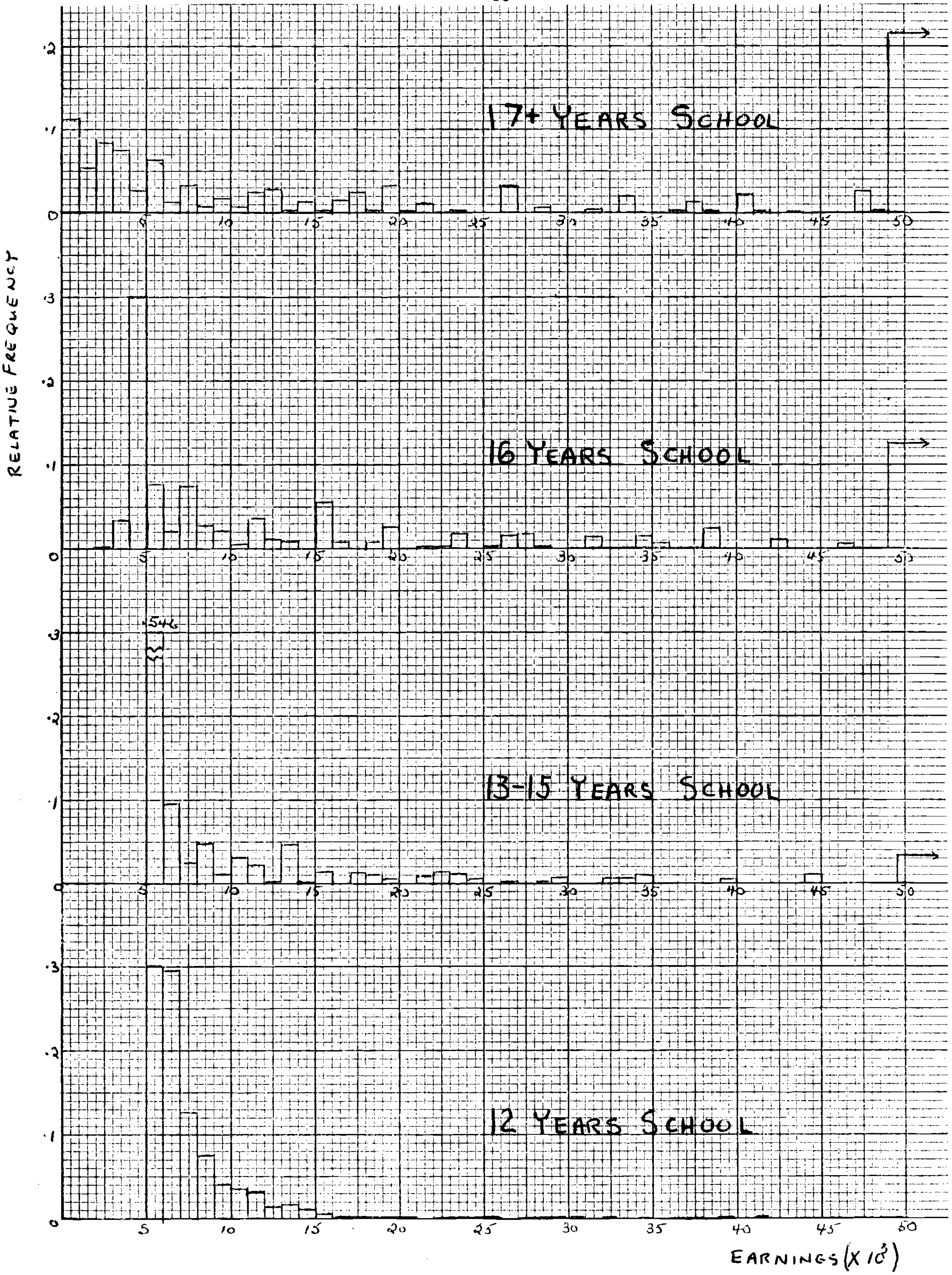

Figure 10. Predicted Mean Earnings distributions by schooling class, age group, and ability class, and the corresponding lorenz curves.

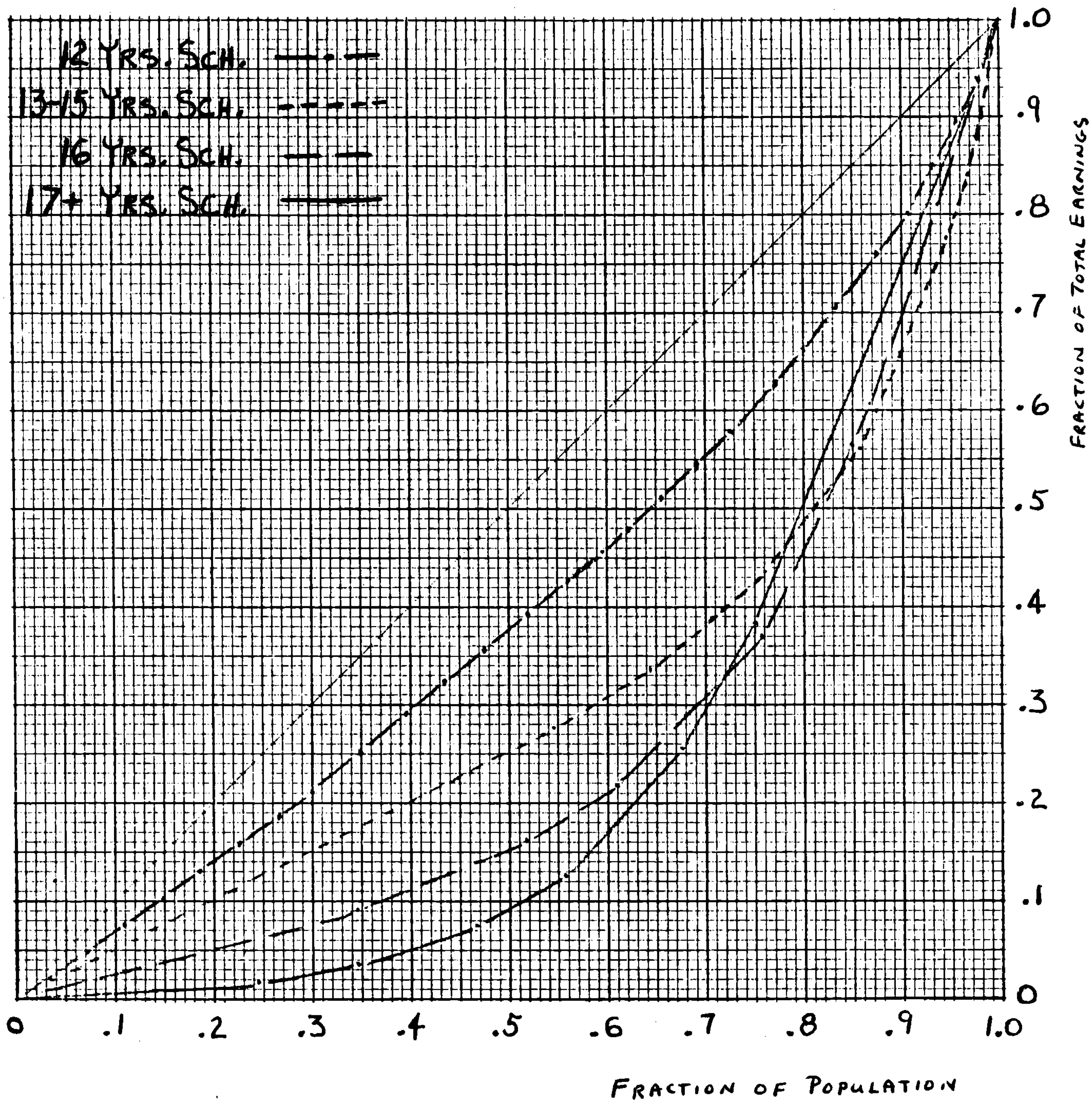

Figure 10. (continued, page 2)

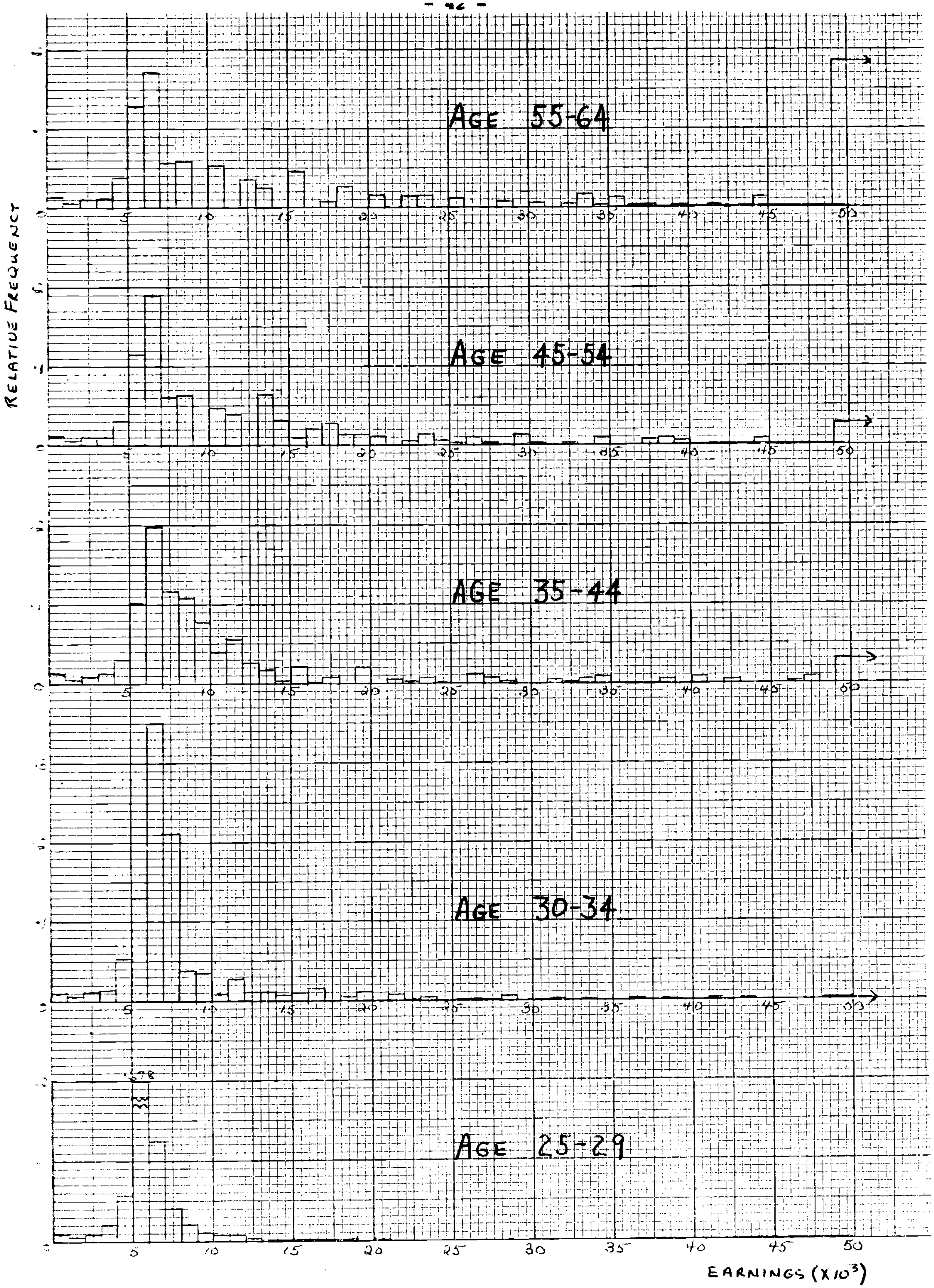

FIGURE 10. (CONTINUED, PAGE 3)

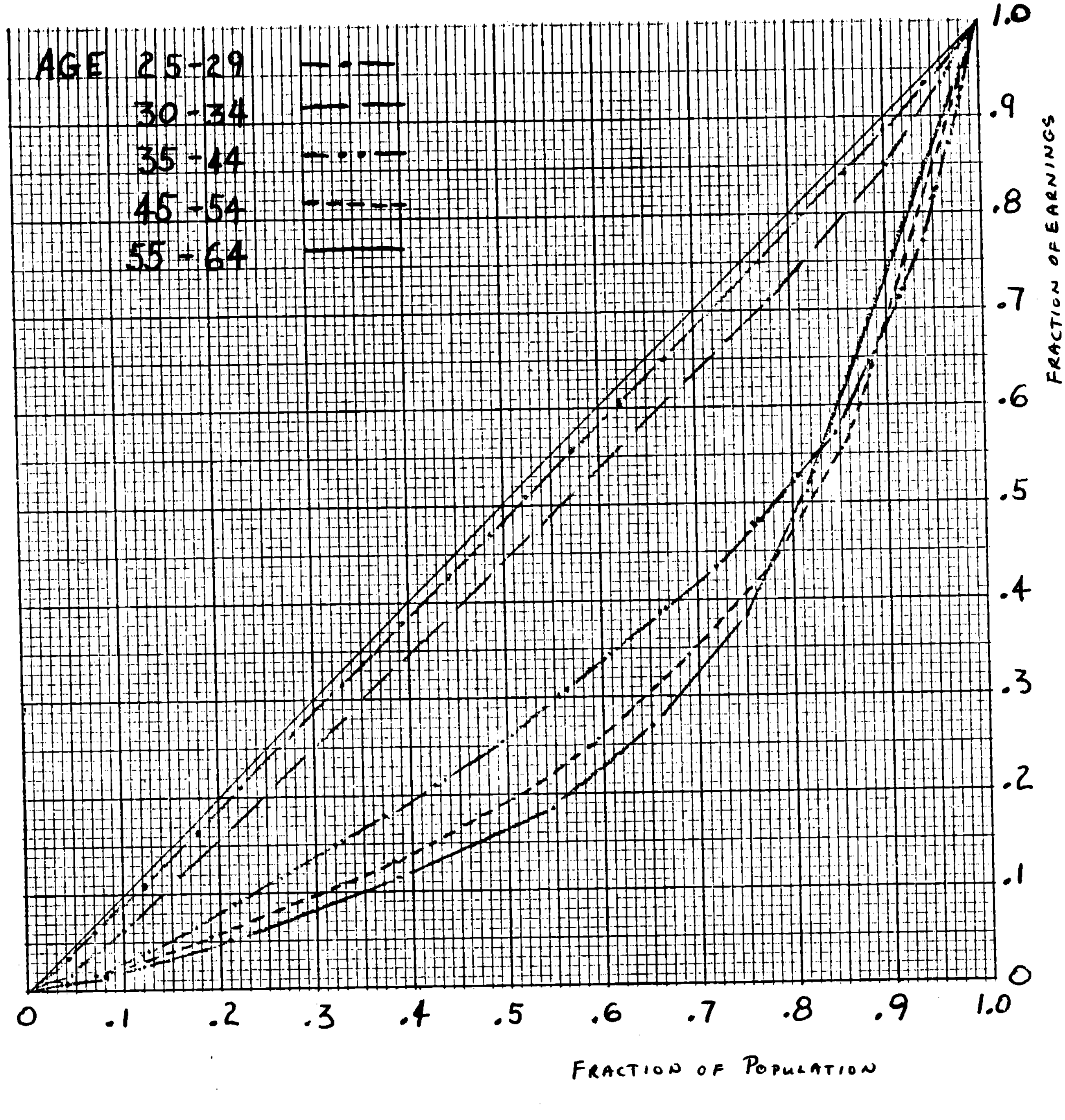

FIGURE 10. (CONTINUED, PAGE 4)

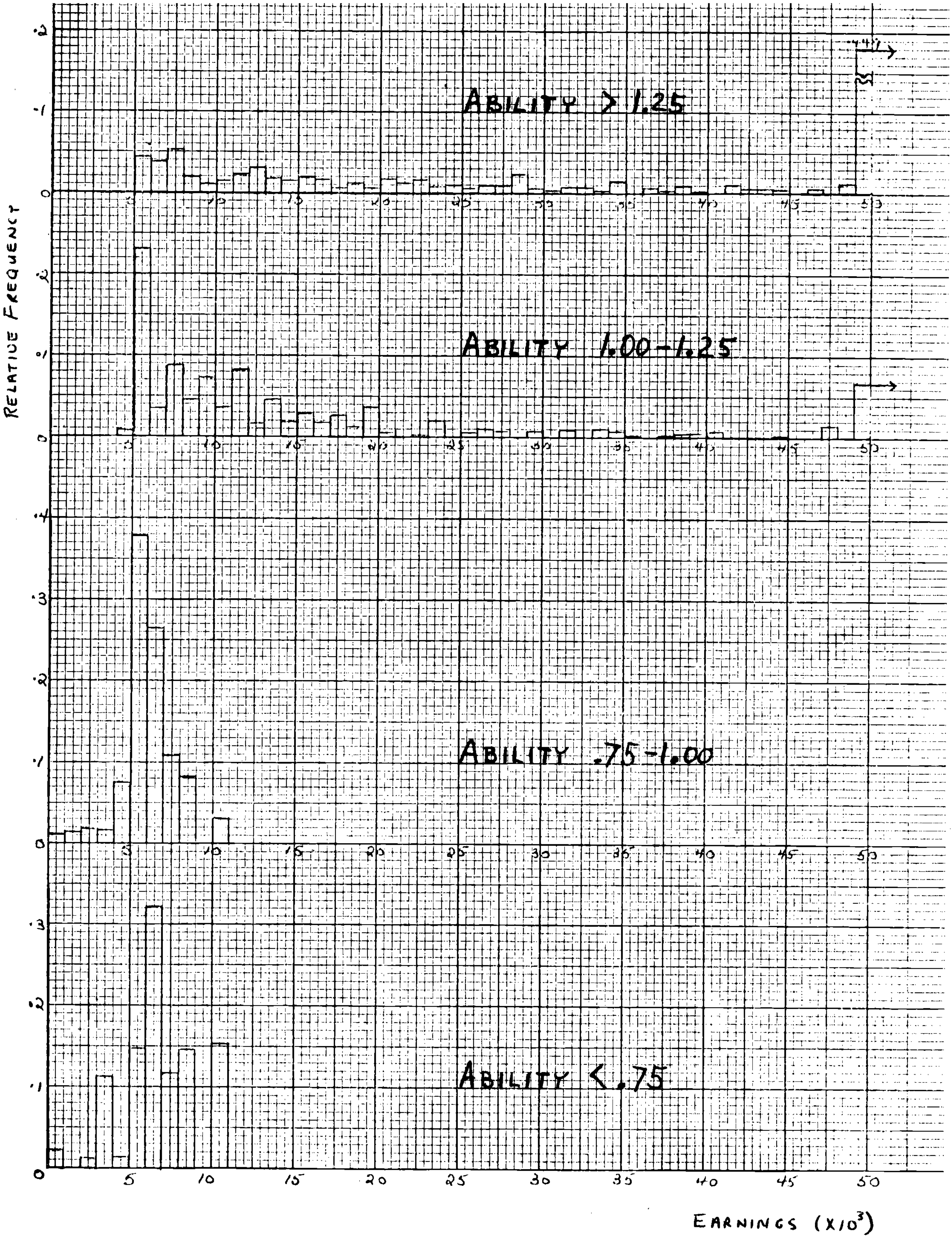

Figure 10. (continued, page 5)

Note: Mean Ability is 1.00 and the standard deviation is .25.

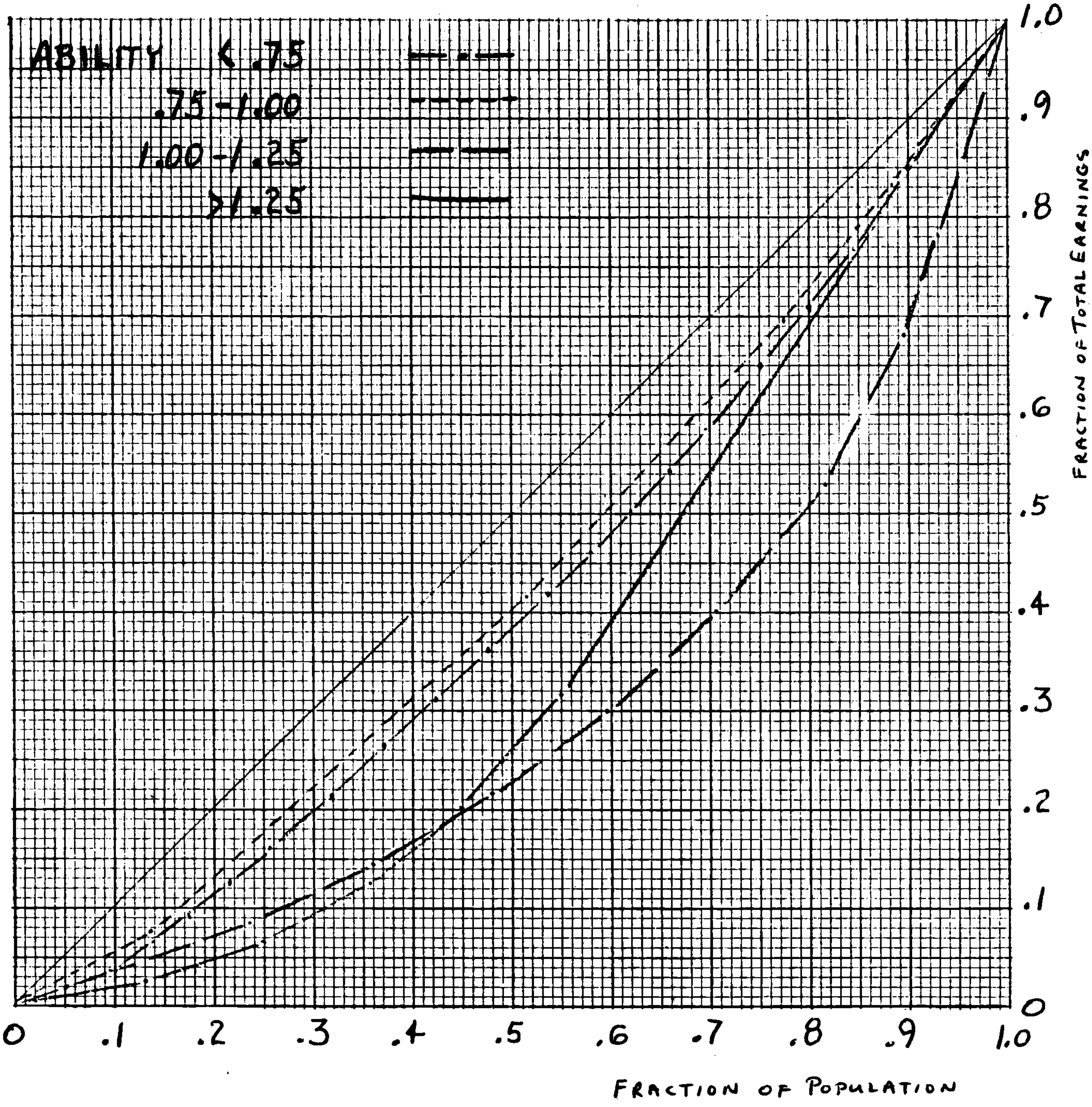

FIGURE 10, (CONTINUED, PAGE 6)

Table 4

Predicted Distribution of Expected Earnings in 1960 for Employed Males Eighteen to Sixty-four Years Old With At Least a High School Education for the Overall Population and by Schooling Class, by Age Group, and by Ability Class.

Earnings in $1000	Overall	Schooling Class			
		12	13-15	16	17+
0-1	.0110				.1102
1-2	.0052				.0525
2-3	.0085			.0007	.0840
3-4	.0120			.0364	.0755
4-5	.0411			.3102	.0273
5-6	.3031	.3030	.5461	.0762	.0641
6-7	.1914	.2960	.0983	.0203	.0125
7-8	.0901	.1282	.0256	.0741	.0327
8-9	.0573	.0750	.0495	.0291	.0094
9-10	.0297	.0403	.0109	.0221	.0196
10-11	.0272	.0351	.0304	.0048	.0039
11-12	.0293	.0312	.0219	.0374	.0250
12-13	.0127	.0144	.0019	.0115	.0275
13-14	.0207	.0165	.0472	.0098	.0016
14-15	.0079	.0116	.0004	.0	.0127
15-16	.0137	.0067	.0131	.0565	.0008
16-17	.0065	.0073	.0	.0075	.0145
17-18	.0087	.0062	.0131	.0	.0243
18-19	.0057	.0041	.0115	.0069	.0004
19-20	.0105	.0045	.0069	.0267	.0323
20-21	.0036	.0064	.0	.0	.0004
21-22	.0034	.0	.0090	.0036	.0102
22-23	.0042	.0017	.0126	.0039	.0
23-24	.0069	.0025	.0120	.0192	.0051
24-25	.0020	.0017	.0048	.0	.0
25-26	.0014	.0023	.0	.0014	.0
26-27	.0059	.0	.0034	.0147	.0334
27-28	.0024	.0004	.0	.0174	.0
28-29	.0022	.0015	.0034	.0006	.0056
29-30	.0023	.0007	.0090	.0	.0
30-31	.0005	.0009	.0	.0	.0
31-32	.0022	.0	.0	.0147	.0041
32-33	.0006	.0009	.0007	.0	.0
33-34	.0035	.0	.0067	.0	.0203
34-35	.0043	.0002	.0108	.0150	.0
35-36	.0010	.0	.0	.0083	.0
36-37	.0004	.0002	.0	.0	.0026
37-38	.0015	.0003	.0	.0	.0131
38-39	.0032	.0	.0	.0248	.0010
39-40	.0011	.0001	.0050	.0	.0
40-41	.0023	.0	.0	.0	.0226
41-42	.0002	.0001	.0	.0	.0015
42-43	.0013		.0	.0104	.0
43-44	.0001		.0	.0	.0005
44-45	.0025		.0115	.0	.0
45-46	.0		.0	.0	.0
46-47	.0008		.0	.0064	.0
47-48	.0021		.0	.0	.0209
48-49	.0001		.0	.0	.0005
>49	.0460		.0343	.1295	.2274

Table 4 (continued (page 2)

Earnings In $1000	Age Group 22-24	25-29	30-34	35-44	45-54	55-64
0-1		.0111	.0140	.0125	.0131	.0138
1-2		.0053	.0066	.0059	.0062	.0066
2-3	.0001	.0085	.0107	.0096	.0101	.0106
3-4	.0040	.0200	.0137	.0116	.0118	.0131
4-5	.0718	.0575	.0536	.0338	.0328	.0394
5-6	.8327	.6780	.1300	.1063	.1168	.1297
6-7	.0749	.1244	.3502	.1983	.1917	.1738
7-8	.0116	.0415	.2062	.1170	.0629	.0573
8-9	.0037	.0203	.0381	.1073	.0633	.0591
9-10	.0	.0112	.0349	.0776	.0	.0
10-11	.0011	.0076	.0077	.0420	.0439	.0423
11-12		.0076	.0285	.0568	.0399	.0
12-13		.0025	.0115	.0259	.0	.0345
13-14		.0012	.0105	.0178	.0534	.0258
14-15		.0013	.0075	.0029	.0314	.0
15-16		.0005	.0091	.0218	.0096	.0477
16-17		.0008	.0147	.0010	.0211	.0
17-18		.0003	.0	.0096	.0230	.0092
18-19		.0003	.0056	.0	.0135	.0271
19-20		.0	.0114	.0214	.0139	.0
20-21		.0003	.0	.0	.0112	.0182
21-22			.0096	.0066	.0	.0
22-23			.0031	.0045	.0054	.0160
23-24			.0046	.0075	.0141	.0167
24-25			.0	.0035	.0054	.0
25-26			.0011	.0	.0013	.0120
26-27			.0037	.0120	.0100	.0
27-28			.0008	.0070	.0013	.0
28-29			.0041	.0025	.0	.0097
29-30			.0	.0	.0123	.0
30-31			.0	.0	.0034	.0047
31-32			.0028	.0062	.0	.0
32-33			.0	.0005	.0004	.0047
33-34			.0013	.0062	.0	.0167
34-35			.0	.0073	.0111	.0012
35-36			.0	.0	.0	.0120
36-37			.0017	.0	.0	.0012
37-38			.0	.0	.0072	.0019
38-39			.0007	.0048	.0092	.0
39-40			.0	.0	.0059	.0004
40-41			.0	.0077	.0	.0
41-42			.0010	.0	.0	.0004
42-43			.0	.0044	.0	.0
43-44			.0003	.0	.0	.0
44-45			.0	.0	.0076	.0125
45-46			.0	.0	.0	.0
46-47			.0	.0027	.0	.0
47-48			.0	.0071	.0	.0
48-49			.0003	.0	.0	.0
>49			.0003	.0305	.1131	.1816

Note: The ability index is distributed with mean 1.00 and standard deviation .25.

Earnings In $1000	Ability Class			
	<.75	.75-1.00	1.00-1.25	>1.25
1	.0220	.0102		
2	.0	.0139		
3	.0132	.0184		
4	.0625	.0175		
5	.0125	.0754	.0085	
6	.1483	.3791	.2314	.0435
7	.3221	.2640	.0349	.0399
8	.1184	.1080	.0883	.0527
9	.1474	.0829	.0456	.0190
10	.0	.0	.0722	.0100
11	.1535	.0306	.0355	.0150
12			.0830	.0225
13			.0163	.0319
14			.0455	.0180
15			.0188	.0145
16			.0295	.0199
17			.0191	.0186
18			.0285	.0053
19			.0117	.0135
20			.0385	.0045
21			.0058	.0180
22			.0	.0123
23			.0031	.0175
24			.0201	.0071
25			.0	.0099
26			.0053	.0068
27			.0113	.0090
28			.0076	.0095
29			.0	.0235
30			.0071	.0058
31			.0	.0039
32			.0092	.0080
33			.0	.0080
34			.0104	.0029
35			.0076	.0141
36			.0024	.0
37			.0	.0074
38			.0030	.0024
39			.0044	.0092
40			.0055	.0008
41			.0084	.0
42			.0	.0100
43			.0	.0064
44			.0	.0069
45			.0040	.0038
46			.0	.0
47			.0	.0068
48			.0106	.0
49			.0	.0121
>49			.0671	.4491

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 5. Relative frequency distributions are presented in Figure 11 and Table 6. Lorenz curves are plotted in Figure 11.

Table 5

Selected Statistics for Predicted Mean Earnings Distributions Corrected for Unexplained Variation.

	Mean	Std. Dev.	Coeff. of Var.	Skew.	Gini Coeff.	Median
Overall	11933.57	10975.44	.92	3.18	.65	9061.61
By Schooling						
12	9189.30	5897.97	.64	0.95	.65	8437.22
13-15	11951.87	10739.64	.90	2.99	.56	9098.92
16	17225.48	15287.94	.89	1.76	.53	11400.00
17+	20827.39	17843.29	.86	1.07	.53	13246.37
By Age						
18-19	7534.56	4702.64	.62	0.57	.65	7010.70
20-21	7496.54	4687.64	.63	0.58	.65	6977.01
22-24	7447.63	4674.96	.63	0.58	.65	6921.01
25-29	7629104	4812.89	.63	0.61	.65	7073.13
30-34	9401.41	6537.79	.70	1.43	.63	8389.98
35-44	12671.68	10797.32	.85	2.77	.57	9764.00
45-54	16528.72	14230.82	.86	1.98	.55	11840.37
55-64	18851.26	15999.89	.85	1.44	.54	12739.88
By Ability						
<.75	8475.90	5132.90	.61	0.53	.66	7974.93
.75-1.00	7858.59	4881.44	.62	0.57	.65	7330.64
1.00-1.25	15848.62	12769.66	.81	2.34	.58	12234.19
>1.25	31748.09	16885.72	.53	0.14	.70	39271.31

Note: Skewness is measured by the square root of $E(X-\bar{X})^3/S^3$. Coefficient of Variation is $S/\bar{X}$.

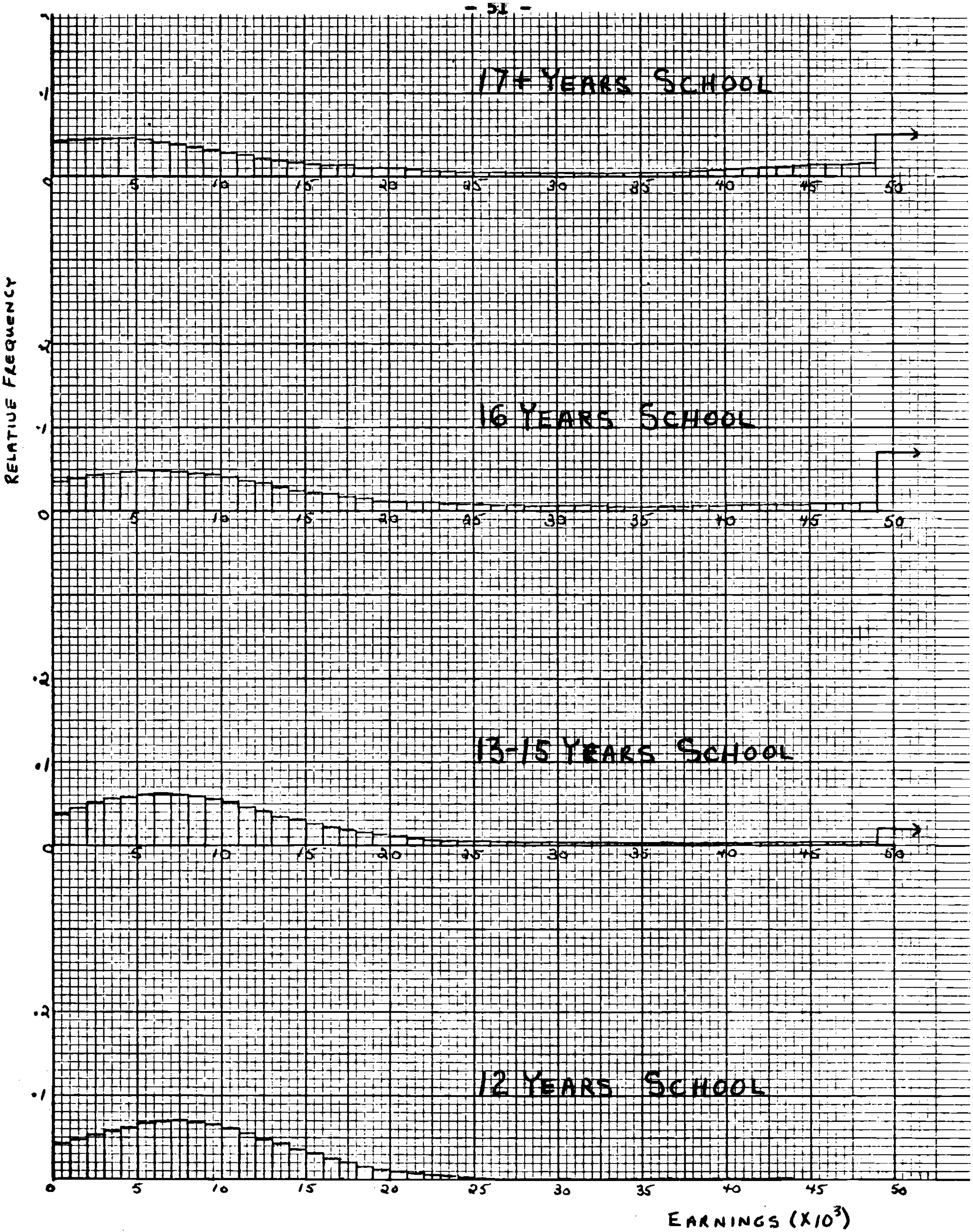

Figure 11. Predicted Earnings distributions corrected for unexplained variation by schooling class, age group, and ability class, and the corresponding lorenz curves.

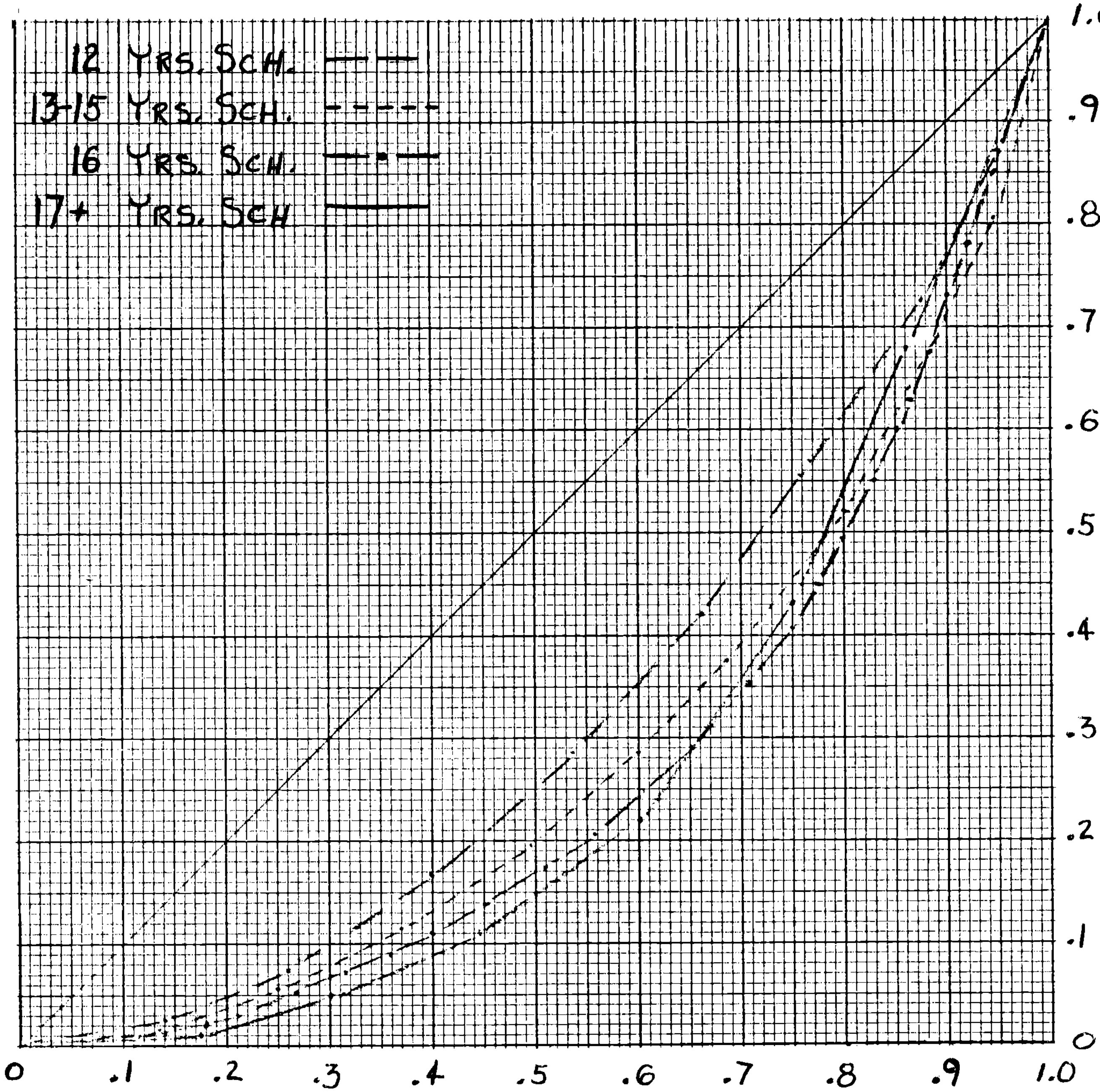

FIGURE 11. (CONTINUED, PAGE 2)

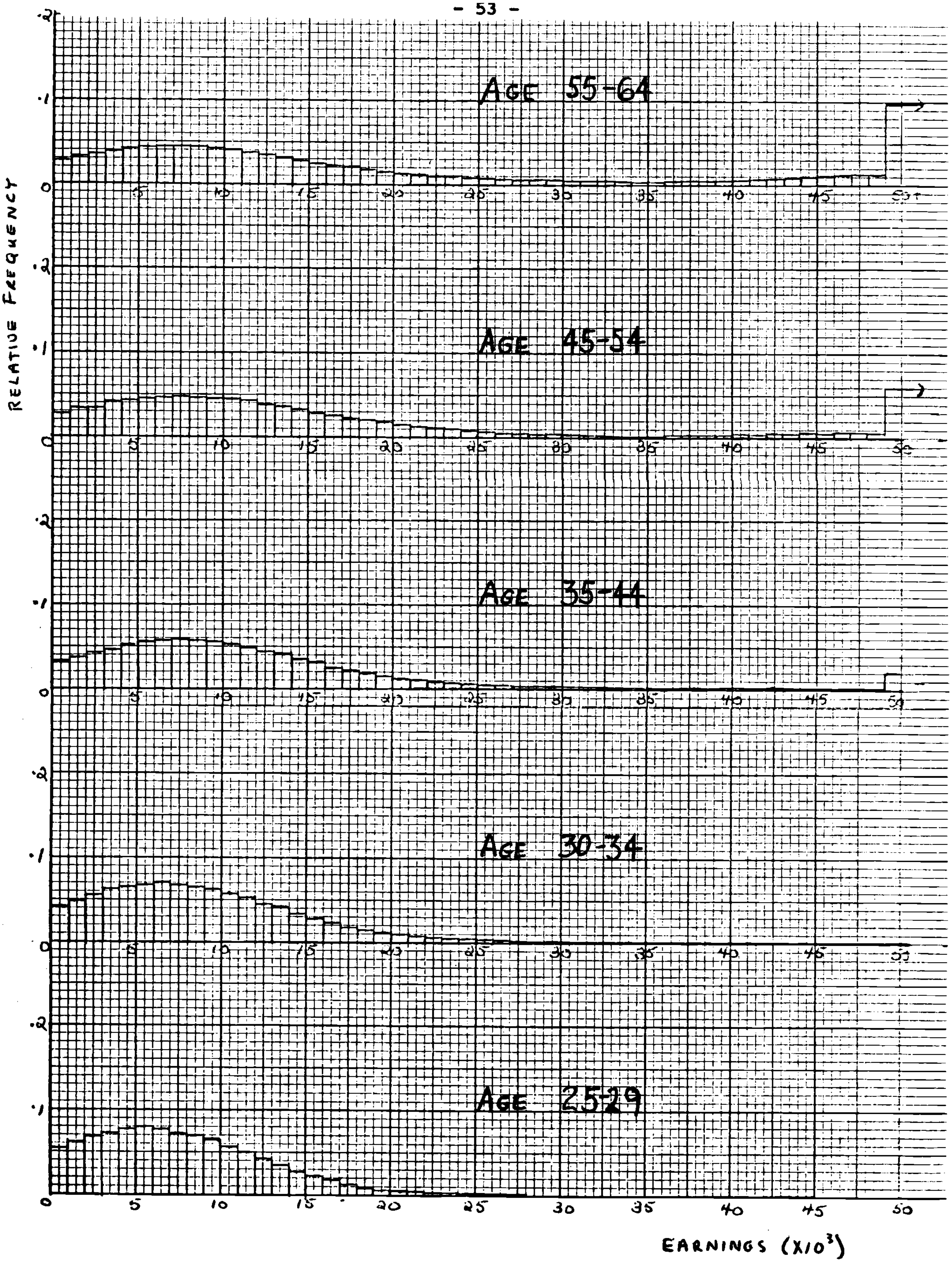

Figure 11. (continued, page 3)

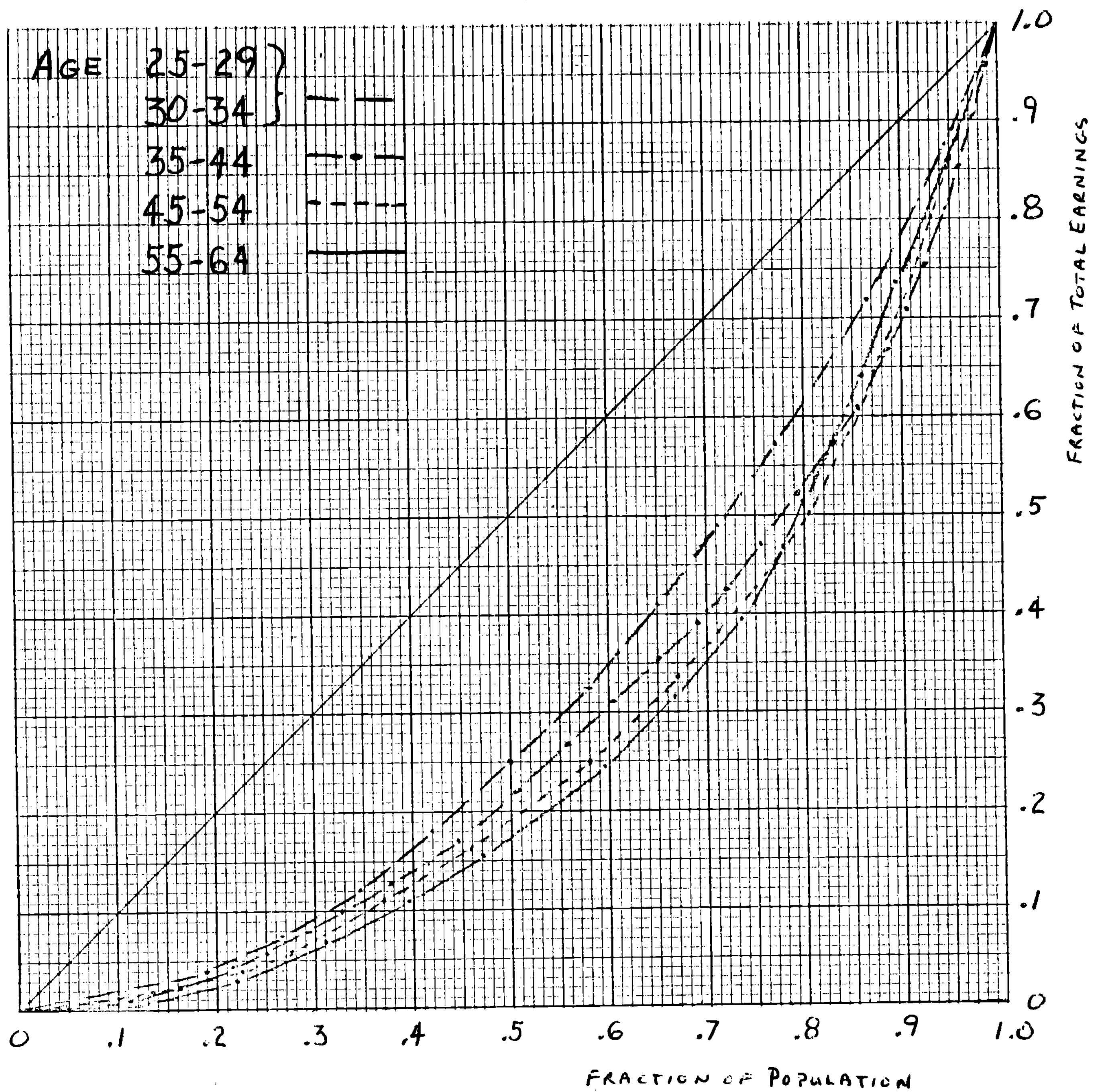

FIGURE 11. (CONTINUED, PAGE 4)

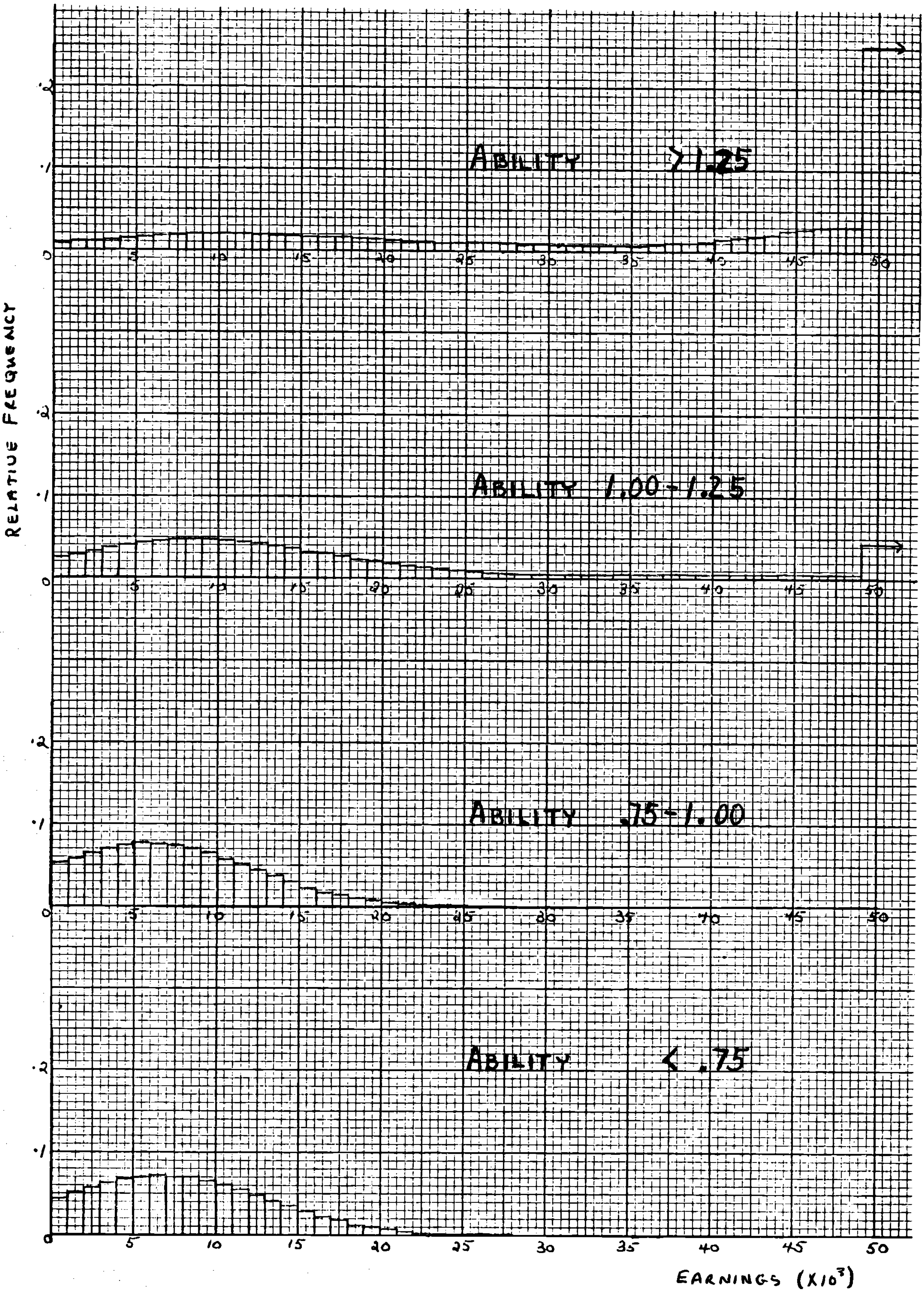

FIGURE 11. (CONTINUED, PAGE 5)

Note: Mean Ability is 1.00 and the standard deviation is .25.

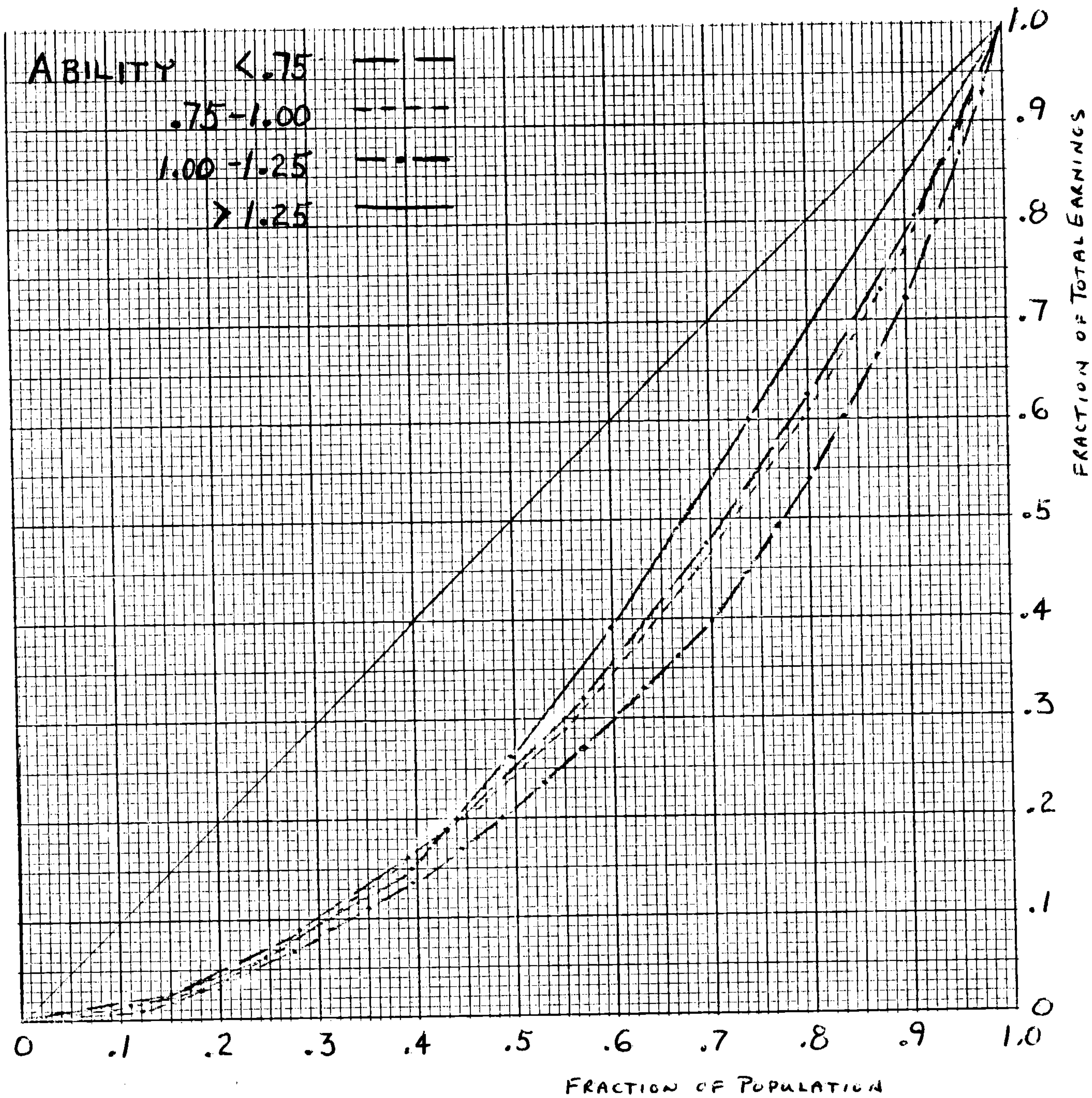

FIGURE 11. (CONTINUED, PAGE 6)

Table 6

Predicted Distribution of Earnings Corrected for Unexplained Variation for the Overall Population, by Schooling Class, by Age Group, and by Ability Class.

Earnings in $1000	Overall	Schooling Class 12	13-15	16	17+
0-1	.0399	.0407	.0398	.0347	.0424
1-2	.0456	.0473	.0455	.0392	.0443
2-3	.0510	.0537	.0512	.0430	.0457
3-4	.0557	.0596	.0559	.0456	.0460
4-5	.0590	.0639	.0587	.0480	.0452
5-6	.0614	.0673	.0614	.0488	.0438
6-7	.0621	.0689	.0615	.0491	.0416
7-8	.0618	.0690	.0613	.0480	.0389
8-9	.0600	.0677	.0592	.0459	.0357
9-10	.0568	.0644	.0556	.0433	.0326
10-11	.0526	.0600	.0514	.0401	.0293
11-12	.0479	.0548	.0467	.0365	.0262
12-13	.0427	.0488	.0414	.0328	.0233
13-14	.0374	.0425	.0362	.0293	.0207
14-15	.0322	.0364	.0313	.0259	.0185
15-16	.0274	.0306	.0267	.0228	.0165
16-17	.0230	.0253	.0226	.0200	.0149
17-18	.0192	.0205	.0190	.0177	.0136
18-19	.0158	.0165	.0160	.0155	.0121
19-20	.0131	.0131	.0135	.0138	.0112
20-21	.0108	.0103	.0115	.0123	.0102
21-22	.0090	.0080	.0098	.0114	.0094
22-23	.0077	.0065	.0089	.0100	.0088
23-24	.0063	.0049	.0072	.0091	.0082
24-25	.0052	.0036	.0063	.0085	.0077
25-26	.0045	.0028	.0057	.0080	.0072
26-27	.0039	.0022	.0051	.0075	.0068
27-28	.0035	.0018	.0047	.0071	.0064
28-29	.0032	.0014	.0043	.0068	.0062
29-30	.0029	.0012	.0040	.0065	.0059
30-31	.0027	.0010	.0037	.0063	.0057
31-32	.0025	.0008	.0034	.0061	.0056
32-33	.0024	.0007	.0032	.0062	.0057
33-34	.0022	.0005	.0030	.0059	.0056
34-35	.0022	.0004	.0029	.0059	.0058
35-36	.0021	.0004	.0027	.0059	.0060
36-37	.0021	.0003	.0026	.0060	.0064
37-38	.0021	.0002	.0025	.0060	.0068
38-39	.0022	.0002	.0025	.0062	.0075
39-40	.0023	.0002	.0025	.0065	.0083
40-41	.0024	.0001	.0026	.0068	.0093
41-42	.0025	.0001	.0026	.0072	.0103
42-43	.0027	.0001	.0027	.0077	.0116
43-44	.0029	.0001	.0028	.0082	.0129
44-45	.0031		.0029	.0087	.0141
45-46	.0033		.0030	.0091	.0152
46-47	.0035		.0031	.0095	.0162
47-48	.0036		.0031	.0098	.0169
48-49	.0035		.0030	.0097	.0170
>49	.0268		.0213	.0737	.1322

Table 6 continued (page 2)

Earnings In $1000	Age Group					
	22-24	25-29	30-34	35-44	45-54	55-64
0-1	.0570	.0557	.0427	.0332	.0285	.0282
1-2	.0643	.0627	.0490	.0384	.0327	.0321
2-3	.0712	.0693	.0550	.0437	.0369	.0359
3-4	.0763	.0743	.0603	.0485	.0406	.0392
4-5	.0785	.0767	.0645	.0525	.0437	.0418
5-6	.0804	.0786	.0669	.0556	.0460	.0436
6-7	.0785	.0772	.0682	.0577	.0476	.0446
7-8	.0765	.0752	.0677	.0584	.0482	.0447
8-9	.0714	.0707	.0659	.0581	.0431	.0441
9-10	.0646	.0644	.0626	.0564	.0469	.0424
10-11	.0573	.0575	.0579	.0535	.0450	.0403
11-12	.0496	.0502	.0525	.0499	.0426	.0375
12-13	.0414	.0424	.0466	.0455	.0396	.0346
13-14	.0337	.0350	.0406	.0408	.0365	.0314
14-15	.0270	.0283	.0346	.0359	.0331	.0282
15-16	.0209	.0223	.0291	.0312	.0298	.0253
16-17	.0157	.0171	.0241	.0267	.0266	.0225
17-18	.0115	.0127	.0198	.0227	.0236	.0200
18-19	.0082	.0093	.0160	.0190	.0207	.0177
19-20	.0056	.0066	.0129	.0159	.0182	.0156
20-21	.0039	.0047	.0104	.0133	.0159	.0139
21-22	.0025	.0031	.0085	.0112	.0138	.0124
22-23	.0022	.0026	.0069	.0095	.0121	.0111
23-24	.0005	.0010	.0057	.0082	.0106	.0100
24-25	.0001	.0004	.0045	.0069	.0090	.0088
25-26		.0002	.0035	.0060	.0079	.0080
26-27		.0001	.0030	.0054	.0069	.0073
27-28		.0001	.0026	.0048	.0062	.0067
28-29		.0001	.0022	.0044	.0056	.0062
29-30			.0019	.0041	.0050	.0058
30-31			.0017	.0038	.0046	.0054
31-32			.0015	.0036	.0043	.0051
32-33			.0013	.0034	.0042	.0050
33-34			.0011	.0032	.0039	.0047
34-35			.0009	.0031	.0035	.0046
35-36			.0008	.0030	.0039	.0046
36-37			.0007	.0029	.0040	.0047
37-38			.0006	.0029	.0041	.0049
38-39			.0005	.0029	.0044	.0053
39-40			.0005	.0029	.0047	.0056
40-41			.0004	.0029	.0051	.0064
41-42			.0004	.0029	.0055	.0072
42-43			.0003	.0030	.0061	.0081
43-44			.0003	.0031	.0065	.0091
44-45			.0002	.0032	.0072	.0100
45-46			.0002	.0032	.0075	.0109
46-47			.0002	.0032	.0081	.0118
47-48			.0001	.0032	.0084	.0124
48-49			.0001	.0031	.0334	.0125
>49			.0006	.0215	.0335	.0998

Table 6 continued (page 3)

Note: The ability index is distributed with mean 1.00 and standard deviation .25.

Earnings In $1000	Ability Class			
	<.75	.75-1.00	1.00-1.25	>1.25
0-1	.0452	.0523	.0243	.0083
1-2	.0518	.0593	.0286	.0100
2-3	.0580	.0658	.0330	.0116
3-4	.0637	.0710	.0371	.0134
4-5	.0680	.0744	.0406	.0149
5-6	.0709	.0764	.0439	.0164
6-7	.0724	.0762	.0460	.0176
7-8	.0718	.0744	.0478	.0186
8-9	.0702	.0709	.0484	.0193
9-10	.0663	.0654	.0481	.0198
10-11	.0615	.0591	.0469	.0199
11-12	.0555	.0521	.0453	.0198
12-13	.0491	.0446	.0427	.0194
13-14	.0422	.0373	.0400	.0188
14-15	.0354	.0304	.0369	.0181
15-16	.0291	.0242	.0336	.0174
16-17	.0233	.0188	.0304	.0166
17-18	.0182	.0142	.0272	.0157
18-19	.0138	.0104	.0241	.0150
19-20	.0103	.0075	.0214	.0142
20-21	.0075	.0052	.0188	.0135
21-22	.0053	.0036	.0165	.0128
22-23	.0037	.0026	.0147	.0123
23-24	.0026	.0014	.0126	.0116
24-25	.0014	.0006	.0111	.0111
25-26	.0008	.0003	.0098	.0106
26-27	.0004	.0001	.0688	.0102
27-28	.0003	.0001	.0078	.0097
28-29			.0071	.0093
29-30			.0064	.0089
30-31			.0059	.0084
31-32			.0055	.0082
32-33			.0052	.0085
33-34			.0049	.0081
34-35			.0047	.0082
35-36			.0046	.0085
36-37			.0645	.0091
37-38			.0044	.0100
38-39			.0044	.0113
39-40			.0045	.0129
40-41			.0646	.0148
41-42			.0647	.0170
42-43			.0050	.0196
43-44			.0052	.0224
44-45			.0054	.0250
45-46			.0056	.0274
46-47			.0058	.0296
47-48			.0059	.0312
48-49			.0058	.0316
>49			.0422	.2517

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 7. Relative frequencies are presented in Figure 12 and Table 8. Lorenz curves are plotted in Figure 12.

Table 7

Selected Statistics for the Actual Distribution of Earnings.

	Mean	Std. Dev.	Coeff. of Var.	Skew.	Gini Coeff.	Median
Overall	6387.63	5204.50	.81	2.14	.52	5885.10
By Schooling						
12	5408.75	3991.17	.74	2.52	.65	5931.97
13-15	6132.73	5288.12	.86	2.03	.58	5209.30
16	8175.46	5943.80	.73	1.75	.65	6414.63
17+	9946.95	7348.90	.74	1.10	.62	7256.52
By Age						
18-19	1401.03	1341.32	.96	4.10	.57	1000.00
20-21	2317.54	1821.16	.79	2.85	.61	1891.60
22-24	4033.45	2374.13	.59	1.64	.69	4111.11
25-29	5133.36	2977.93	.58	2.63	.72	4887.09
30-34	6752.81	4170.18	.62	2.45	.71	5941.48
35-44	7845.95	5304.58	.68	2.00	.68	6469.79
45-54	8419.13	6289.26	.75	1.64	.63	6536.00
55-64	8299.36	6822.15	.82	1.50	.59	6166.66

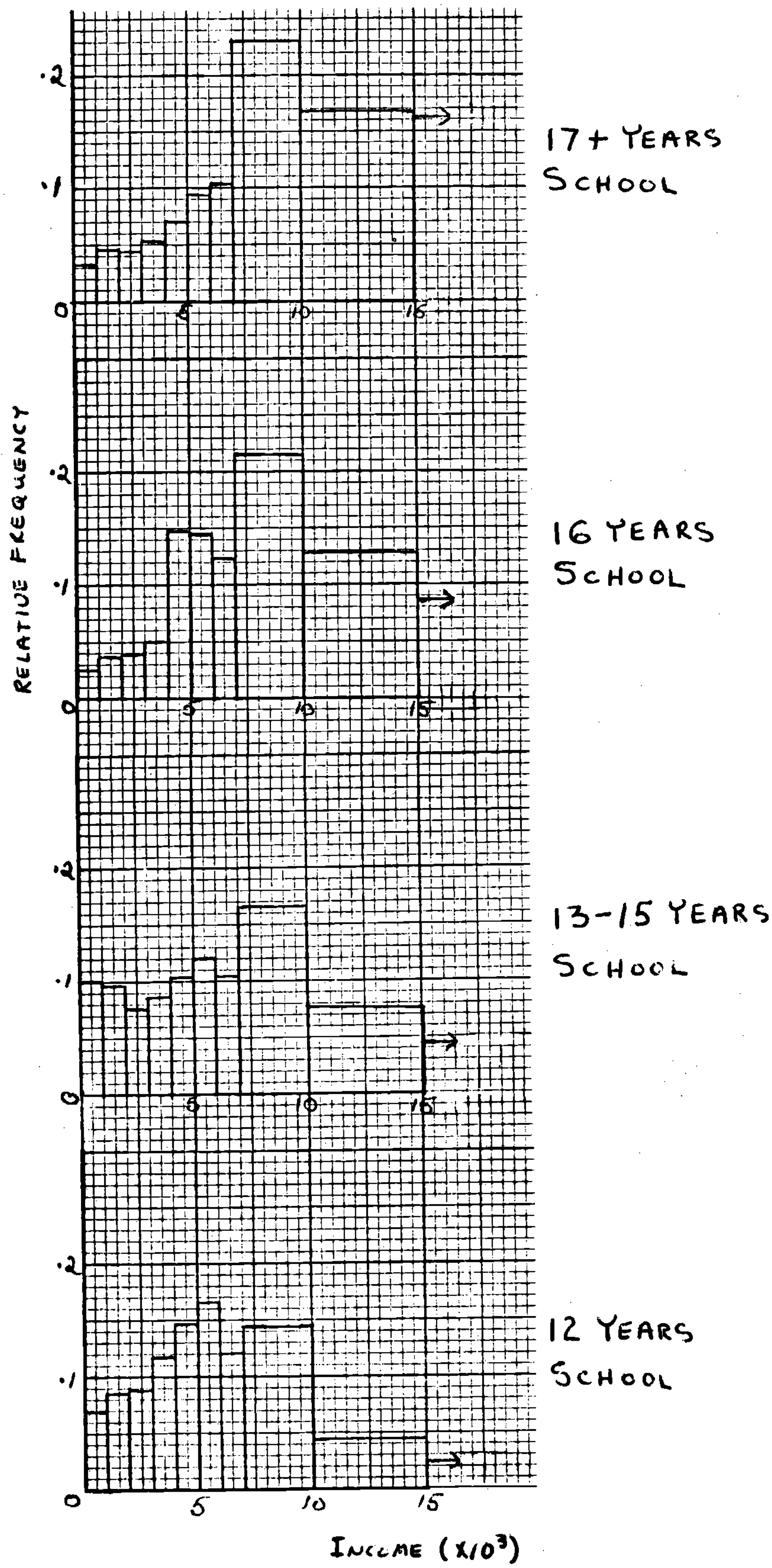

Figure 12. Actual distributions of total income reported in the 1960 Census of Population by schooling and age group, and the corresponding lorenz curves.

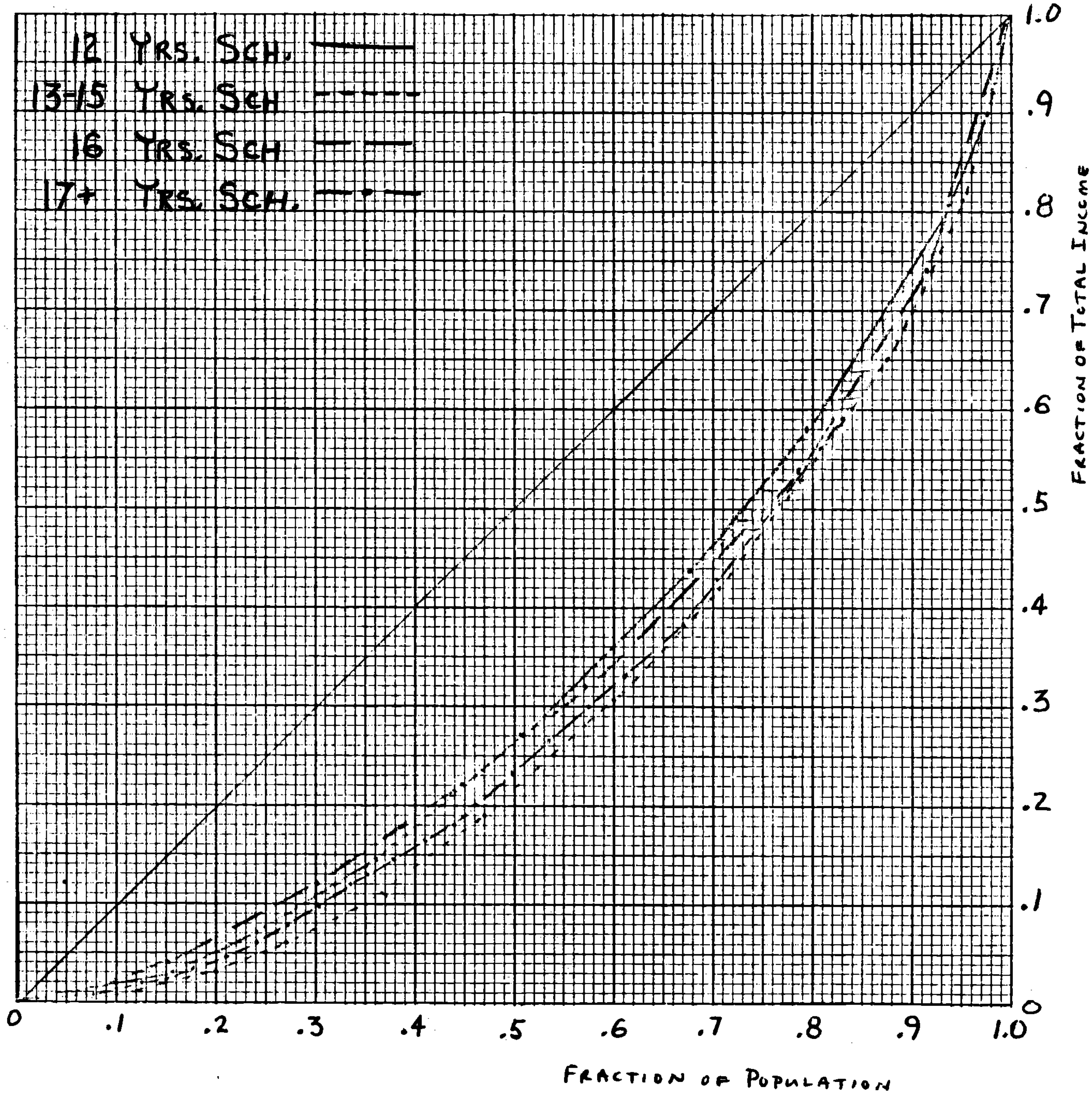

FIGURE 12. (CONTINUED, PAGE 2)

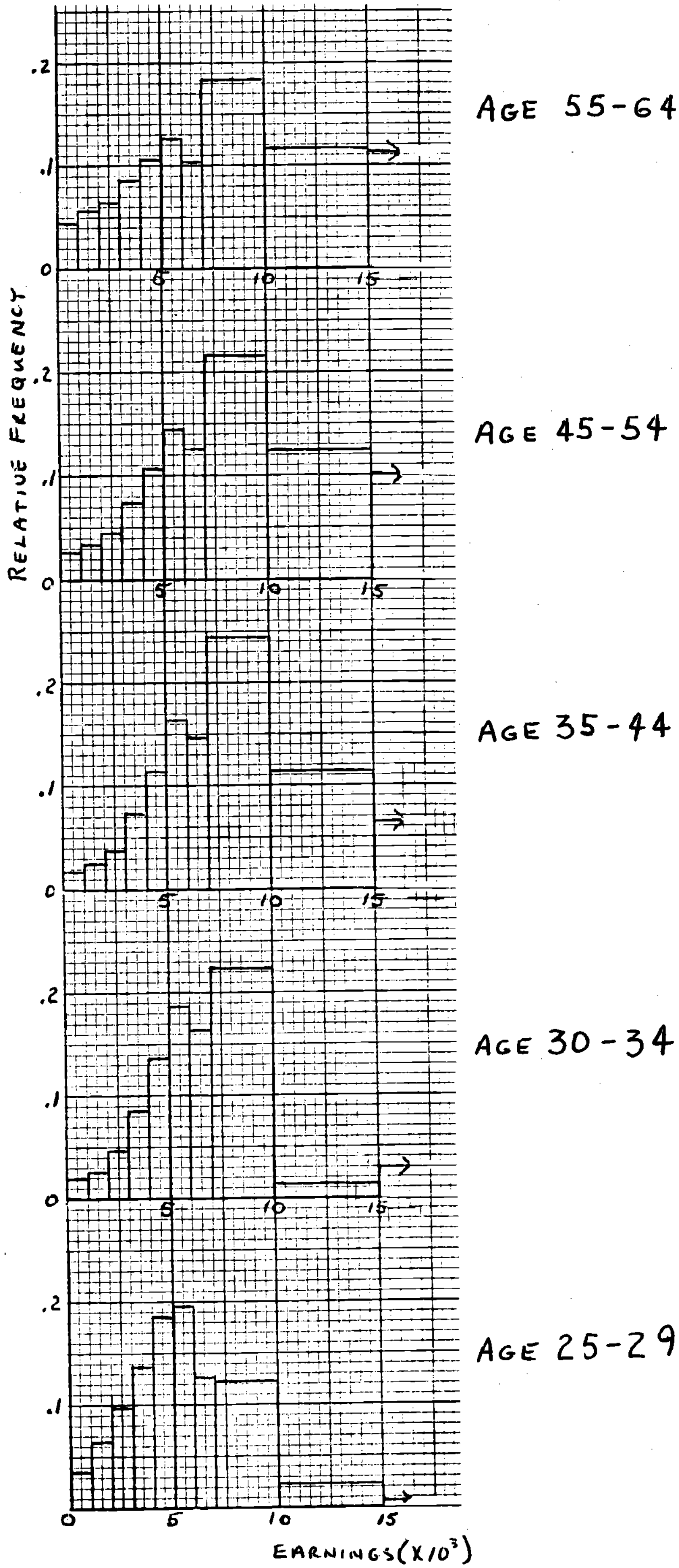

FIGURE 12. (CONTINUED, PAGE 3)

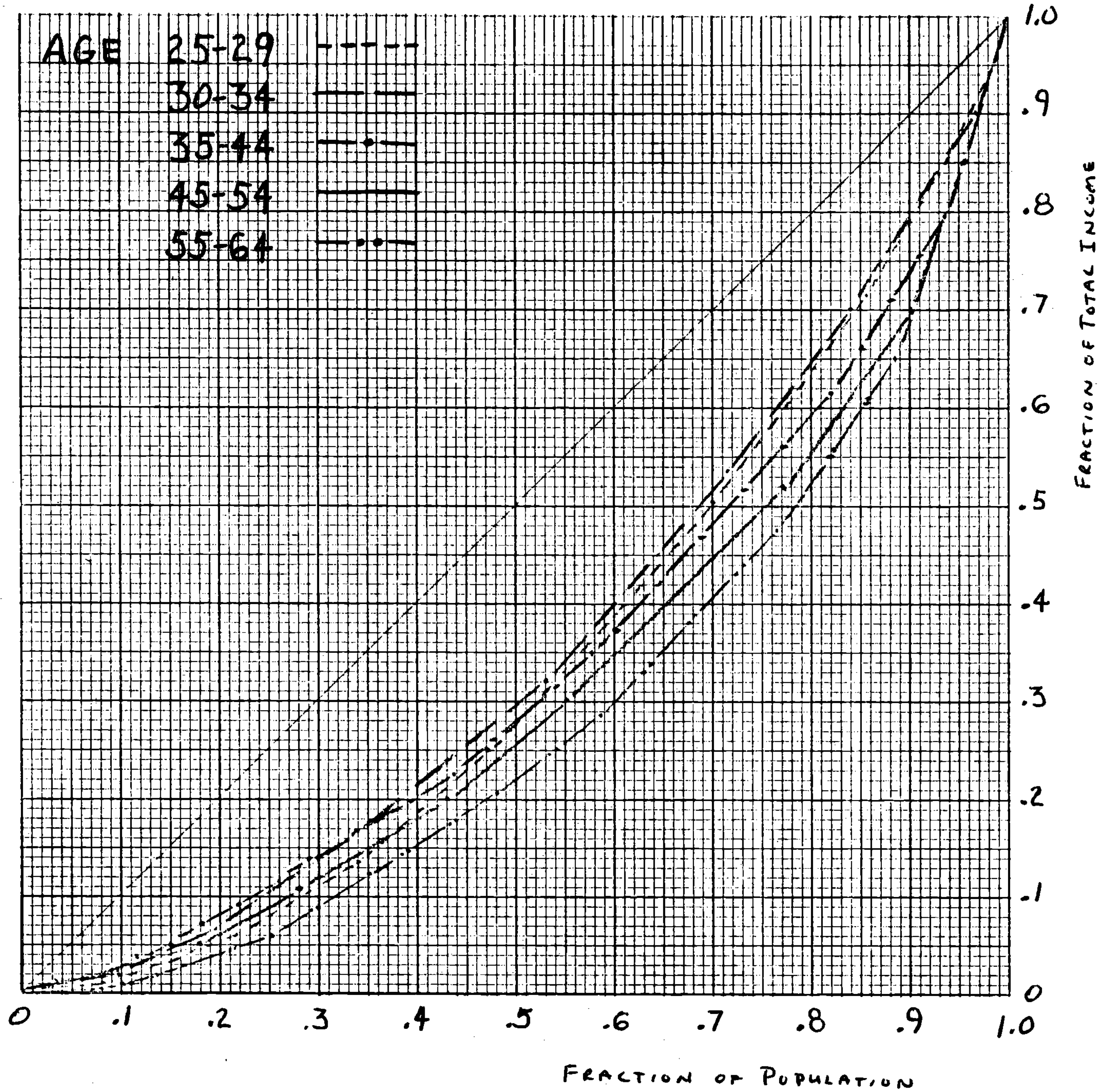

FIGURE 12. (CONTINUED, PAGE 4)

Table 8
Actual Distribution of Earnings in 1960, for Employed Males Eighteen to Sixty-four Years Old With at Least a High School Education, for the Overall Population, by Schooling Class, and by Age Group.

Earnings in $1000	Overall	Schooling Class			
		12	13-15	16	17+
0-1	.067	.070	.103	.026	.033
1-2	.078	.085	.099	.038	.047
2-3	.075	.089	.076	.039	.045
3-4	.096	.119	.087	.050	.051
4-5	.131	.147	.108	.049	.071
5-6	.148	.165	.129	.047	.094
6-7	.116	.120	.108	.123	.101
7-10	.167	.143	.167	.215	.230
10-15	.074	.042	.077	.129	.169
>15	.049	.021	.047	.085	.161

Earnings In $1000	Age Group					
	22-24	25-29	30-34	35-44	45-54	55-64
0-1	.081	.035	.020	.019	.028	.044
1-2	.134	.064	.028	.025	.034	.058
2-3	.126	.097	.049	.038	.045	.064
3-4	.134	.139	.088	.071	.074	.084
4-5	.225	.186	.138	.114	.108	.105
5-6	.155	.196	.188	.163	.144	.126
6-7	.078	.126	.162	.149	.125	.102
7-10	.055	.121	.224	.244	.218	.182
10-15	.010	.024	.074	.113	.123	.118
>15	.002	.008	.030	.063	.100	.115

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

Both the mean and corrected earnings distributions display the general characteristic of the actual distribution but tend to "overstate" earnings. Both the center and dispersion are larger in the predicted distributions. 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.

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 within age and schooling classes. The mean increases within higher schooling classes for both predicted and actual distributions. However, the median falls from high school to some college then rises in the mean and actual distributions but rises continuously throughout in the corrected distribution. 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. This dip in mean earnings is clearly evident in the age-earnings profiles in Figure 7. 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 dip slightly then rise very 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[27] of individual subgroup means, whether grouped by age, schooling, or ability.

Dispersion is overstated in the predicted distributions whether measured by the standard deviation or the coefficient of variation. The exceptions are that both measures are understated in the mean distribution for ages below tweny-nine and the high school group. 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. The standard dip then rises as ability increases. The coefficient of variation, or standard deviation relative to the mean, moves in a less stable way. As the schooling level increases it first rises than declines in the corrected and actual distributions, that is, the increased variance is more than offset by the increased mean. In the mean distribution the coefficient of variation doubles from high school to some college then remains constant. As age increases the coefficient of variation rises continuously in the predicted distributions while it forms a U-shape in

---

[27]The weights are obviously the relative frequency of the subgroups.

the actual distribution. The coefficient of variation is noticeably higher in the ability group average to one standard deviation above average and fairly constant outside that interval. 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.[28]

Another characteristic of earnings distributions widely discussed in the literature is concentration represented by the Lorenz curve and its summary statistic the Gini coefficient.[29] 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 per cent) than among high income holders relative to the other population. This says nothing about location of high and low, only about the concentration

---

[28] Both the average,of variances, and the variance, of averages, are calculated weighted by the relative frequency of the subgroups. Formally $E(Y) = E\ [Var(Y)\ subgroup\ ] + Var\ [E(Y)\ subgroup\ ]$.

[29] The Gini coefficient is the area under the Lorenz curve relative to the area of the triangle, one-half a larger Gini coefficient implies less concentration. The extremes are one when every individual gets and equal share of total income and zero when one individual holds total income.

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, the population with the largest positive skew will have its Lorenz curve above the other in the lower earnings region.

The actual distribution of income may be said to be unambiguously less concentrated than the corrected predicted distribution. The predicted mean distribution Lorenz curve crosses both the others from above, in Figure 9, it has the largest skew, and its Gini coefficient halfway between the other two. Even though the actual distribution has only one-half the standard deviation of the two predicted distributions its concentration is the least. Between the two predicted distributions which are approximately the same variance and close to the same concentration, within six one-hundreths, the distribution with the largest skew intersects the other from above. These same basic relationships hold within subgroup distributions.

Within schooling groups the largest range in concentration is noted in the mean distributions with a range from .81 in the lowest schooling class to .47[30] in the highest schooling class, declining monotonically.

---

[30] The Gini coefficient is overstated for both the highest schooling and highest ability groups in the predicted distributions due to the open-ended earnings interval > 49,000. The actual distribution Lorenz curve and Gini coefficient are also approximaate due to the wide unequal intervals for high income levels.

The corrected predicted and actual distributions range from .65 to .53 and from .65 to .58, respectively. Generally the Lorenz curves indicate increased inequality at the lower earnings levels as schooling increases with a convergence or crossing of the curves at higher earnings levels.[31]

Within age groups again the largest range in concentration is in the predicted mean distribution. Mean earnings are very concentrated, Gini of .90 to .98, for groups eighteen to twenty-nine and the correction for unexplained variation has the largest impact on this group bringing the Gini coefficient very close to the actual value. Concentration is predicted to decline with age but rises then declines in the actual distribution. This difference may be partly to the inclusion of more employed students at early ages with low incomes. The corrected and actual Lorenz curves and Gini coefficients are quite close except for generally less concentration in the corrected distribution.

Within ability classes only predicted distributions are available. Concentration in mean earnings is greater among below average individuals than above average individuals and is least concentrated for those within one standard deviation above the mean. For the corrected distribution roughly the same pattern holds but with slightly less concentration.

---

[31] Again higher and lower are relative with a distribution and are not comparable, with Lorenz curves, between distributions.

The predicted distributions tend to indicate more 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.

Deriving Predicted Human Wealth Distributions from the Estimated Earnings Function

The purpose of this section is to apply the principles of the general framework to predict the distribution of human wealth over all, by schooling class and by ability class for several interest rates. 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 an earnings somewhere between the net and gross values depending upon what fraction of investment is obtained on-the-job with direct inputs deducted from earnings. The empirical measure of human wealth is then the integral of the discounted estimated earnings function with respect to age from the end of formal schooling to the end of working life. 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.[32]

---

[32] This may not be the case in reality to the extent that the estimation error is explained by factors other than schooling, age, and ability which affect the lifetime pattern of earnings. The error term may not be correlated with age while factors which explain error or reduce error variance may be correlated with age. These factors would then affect human wealth other than in a transient way.

Human wealth is then a function of the level of schooling, ability, and the rate of discount (r) and the length of working life (N).

$$HW(S, B, R, N) = \int_{A=S}^{N} e^{-RA} \cdot Y(A, S, B)\, dA$$

The same schooling and ability distribution and midpoints are used as before. The predicted human wealth distributions are calculated for discount rates of 3, 4, 5, 6, and 7 per cent and assuming the end of working life is age sixty-six. Everyone in the population is assumed to have the same discount rate and the same working life, but individuals differ in schooling and ability. Following the general framework density is transformed from two-dimensional (schooling, ability) - space into human wealth-space through the integral function.

Selected statistics are presented in Table 9, relative frequency histograms and selected Lorenz curves are presented in Figure 13, and predicted distributions are presented in Table 10 for the overall population, by schooling group, and by ability class for interest rates from 3 to 7 per cent.

The most striking result is that there is much less inequality in human wealth than in earnings. The overall coefficient of variation drops from above 80 per cent to less than 10 per cent and the overall Gini coefficient rises from less than 65 per cent to at least 90 per cent for the interest rates considered.

Table 9

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.

	Disc. Rate	Mean	Std. Dev.	Coeff. of Var.	Skew.	Gini Coeff.	Median
Overall							
	.03	197994.50	14242.55	.072	1.64	.916	193782.00
	.04	154147.56	9736.00	.063	2.70	.957	150490.54
	.05	121943.88	6236.64	.051	3.23	.960	120222.24
	.06	97965.38	4651.35	.047	2.66	.953	97846.84
	.07	79948.06	4479.09	.056	1.09	.923	80847.16
By Schooling							
12	.03	190934.50	6881.71	.036	1.74	.975	189523.48
13–15		197873.06	8932.07	.045	1.89	.971	194327.66
16		211125.00	14513.05	.069	0.87	.951	206880.26
17+		221834.38	16480.59	.074	0.01	.945	221089.98
12	.04	150781.38	5501.02	.036	4.72	.975	149527.46
13–15		153052.25	6152.99	.040	2.16	.973	150484.84
16		160407.81	11277.45	.070	1.36	.946	157394.96
17+		167744.63	15504.52	.092	1.03	.928	165086.68
12	.05	121389.25	3850.12	.032	8.25	.979	120418.52
13–15		120215.44	4525.48	.038	2.00	.970	118660.54
16		123504.06	8480.42	.069	1.22	.939	120880.40
17+		126852.31	11740.12	.093	1.07	.917	125089.98
12	.06	99121.31	3078.66	.031	11.28	.977	98727.44
13–15		96294.44	2960.65	.031	2.28	.973	95422.40
16		96145.69	6226.45	.065	1.34	.929	94457.80
17+		97282.94	8851.46	.091	0.97	.898	96454.56
12	.07	82418.44	2562.40	.031	13.72	.973	82023.86
13–15		77941.94	1961.16	.025	2.60	.972	77352.30
16		75866.38	4662.79	.061	1.39	.906	74452.18
17+		75367.19	6651.07	.088	0.98	.857	74454.56

Table 9 continued (page 2)

	Disc. Rate	Mean	Std. Dev.	Coeff. of Var.	Skew.	Gini Coeff.	Median
By Ability							
<.75	.03	194900.25	16810.82	.086	2.37	.950	190960.58
.75-1.00		191568.25	6135.33	.032	1.43	.937	190076.86
1.00-1.25		204161.00	11879.20	.058	1.29	.960	199334.64
>1.25		236671.75	13848.72	.059	1.01	.962	246143.32+[1/]
<.75	.04	157804.68	25482.29	.161	2.85	.916	148745.06
.75-1.00		149568.50	2698.65	.018	1.75	.946	149085.50
1.00-1.25		157914.38	6512.29	.041	1.42	.969	155848.84
>1.25		194348.44	25671.42	.132	0.69	.901	187563.88
<.75	.05	127560.94	22495.98	.176	2.85	.897	120365.54
.75-1.00		118804.31	1739.81	.015	0.15	.762	118923.04
1.00-1.25		124224.13	3537.41	.028	0.96	.975	123722.84
>1.25		149105.81	18270.29	.123	0.96	.900	143322.84
<.75	.06	104715.25	20496.79	.196	2.74	.860	99247.58
.75-1.00		96179.25	2953.72	.031	0.99	.680	96881.88
1.00-1.25		98936.69	2535.08	.026	0.49	.971	99136.20
>1.25		115564.31	11799.28	.102	0.98	.903	112021.02
<.75	.07	87362.44	18656.63	.214	2.61	.807	82986.46
.75-1.00		78773.44	3892.44	.049	1.18	.815	80473.46
1.00-1.25		80394.19	3531.62	.044	0.36	.936	82015.90
>1.25		90919.94	8080.24	.089	0.84	.892	89031.98

[1/] The statistics for the highest ability class are calculated as if the highest open ended income interal 246,000+ were of length 2000 as the lower intervals are. This is however clearly not true. For r = .03, for example, this last interval for the highest ability class contains 53.8 percent of the operations. This is important only for the statistics relating to r = .03. At all other discount rates the percent in this final interval is less than 8 percent.

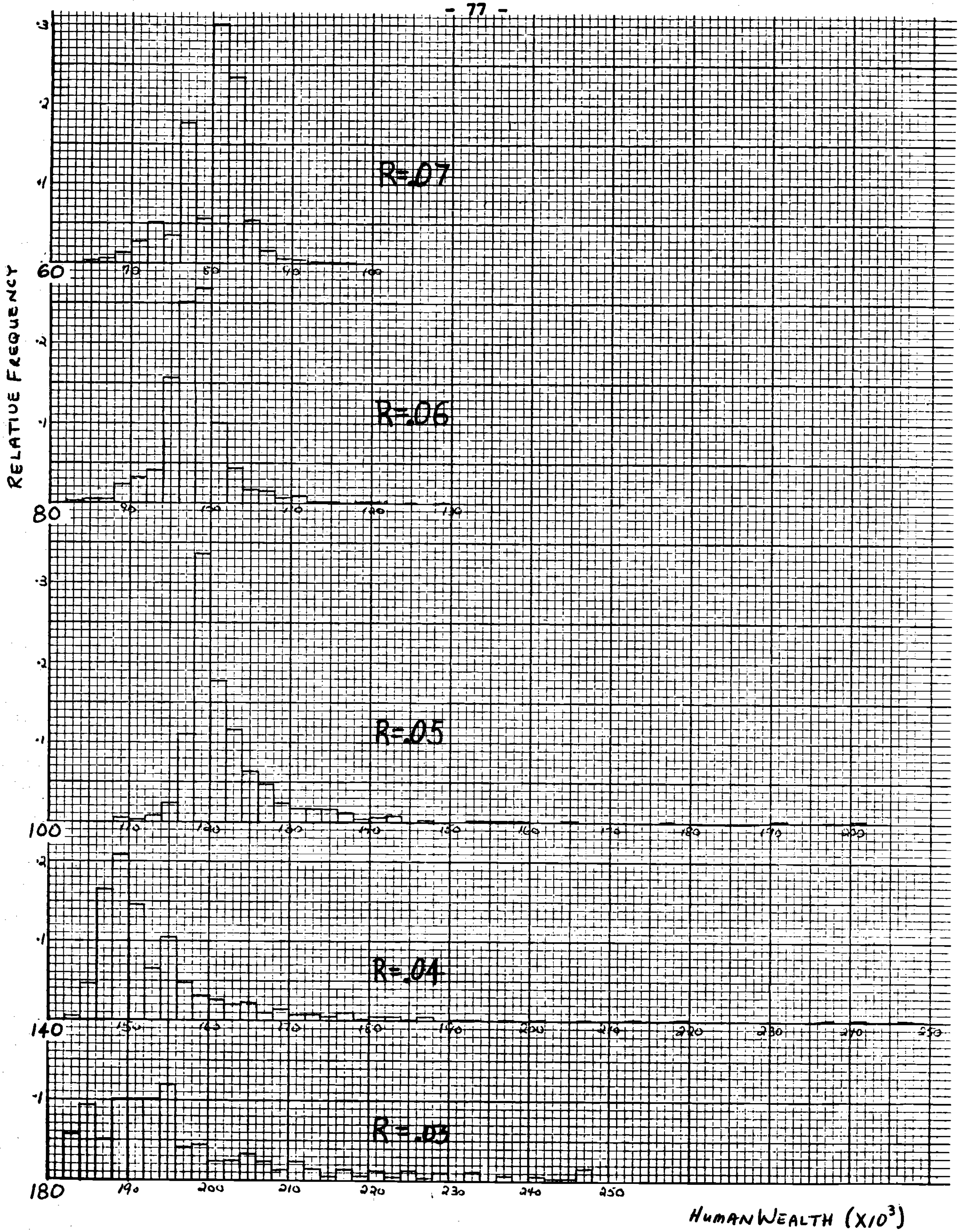

Figure 13. Predicted Distribution of Human Wealth for the Overall population, by schooling class and by ability class for several rates (*R*) of interest and the corresponding lorenz curves.

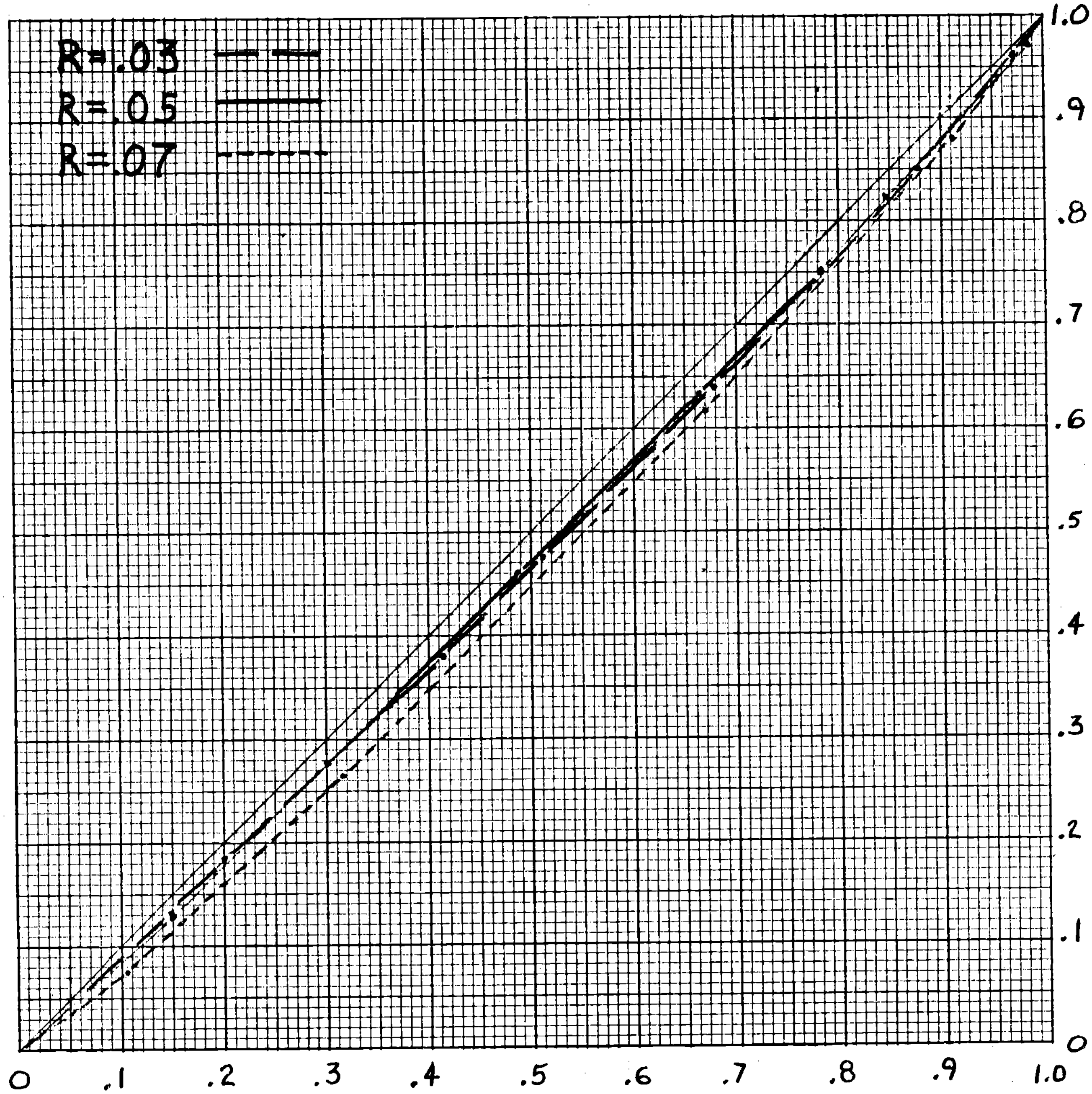

FIGURE 13. (CONTINUED, PAGE 2). LORENZ CURVES OVERALL.

NOTE: THE LORENZ CURVE FOR R=.04 APPROXIMATELY COINCIDES WITH THAT OF R=.03 AND THE CURVE FOR R=.06 LIES BETWEEN THOSE OF R=.05 AND R=.07 AND ARE OMITTED FOR CLARITY.

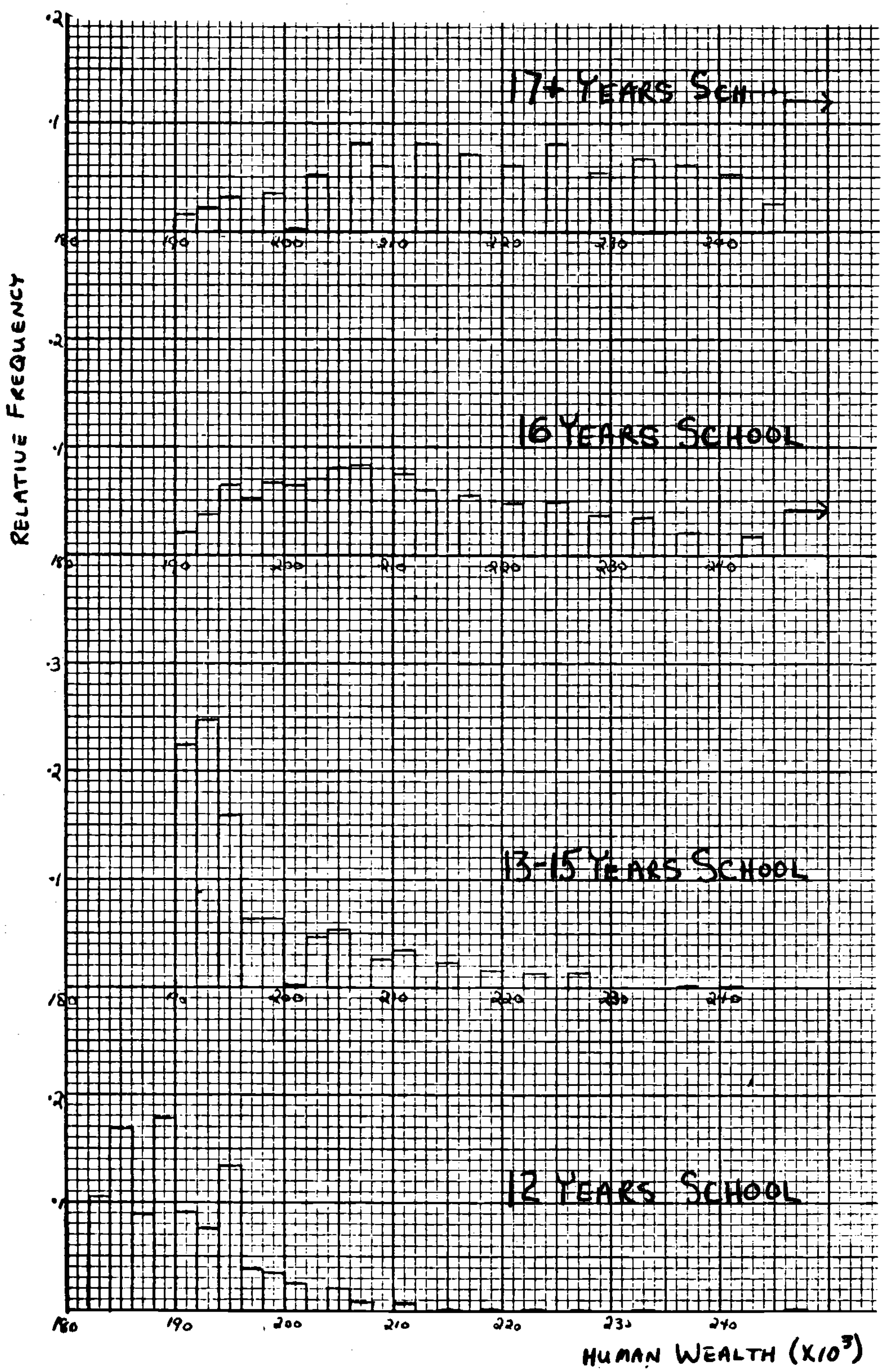

FIGURE 13. (CONTINUED, PAGE 3, R=.03)

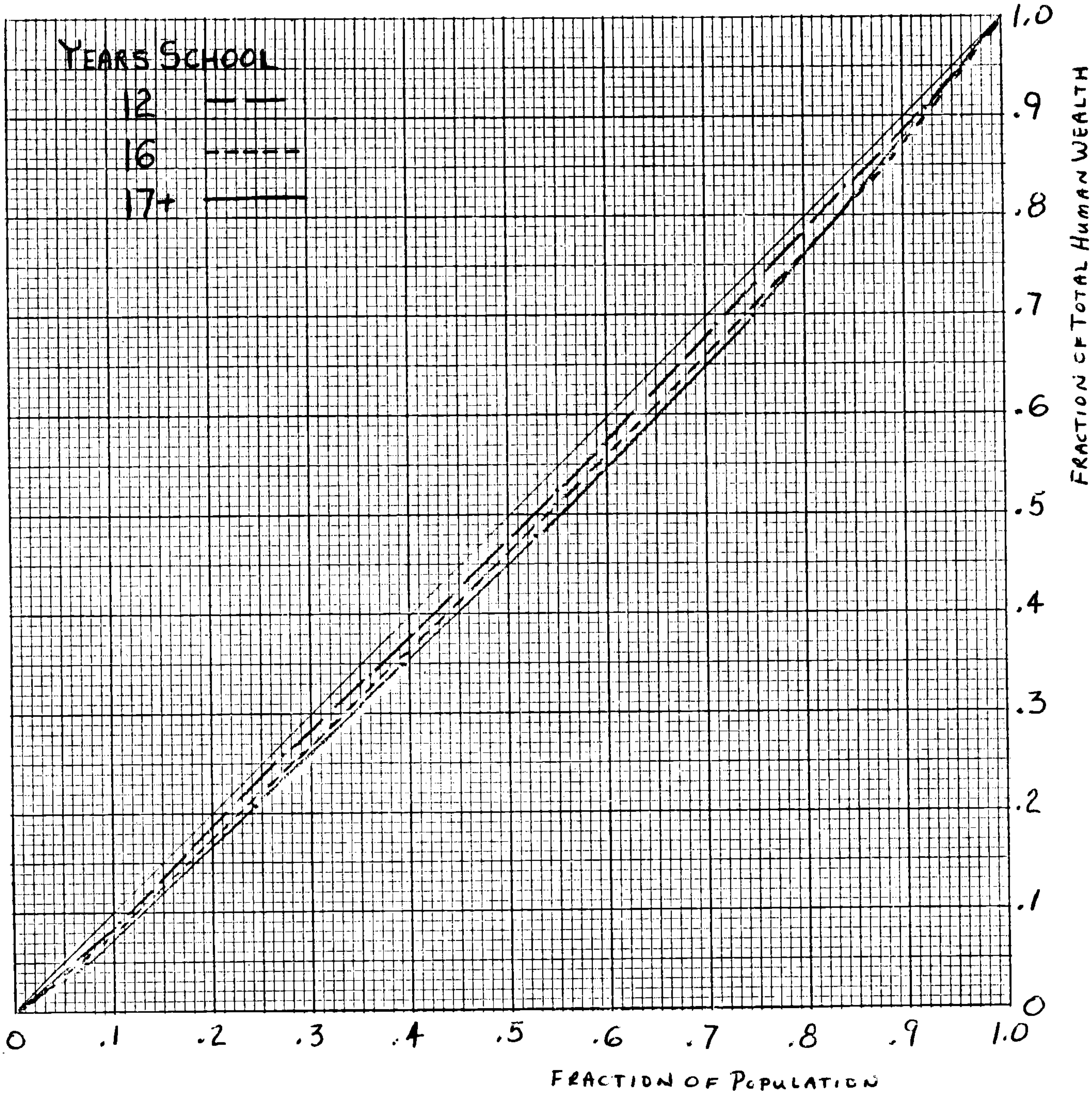

FIGURE 13, (CONTINUED, PAGE 4, R=.03)

NOTE: THE LORENZ CURVE FOR 13-15 YEARS OF SCHOOLING IS OMITTED FOR CLARITY OF PRESENTATION, IT LIES BETWEEN THE CURVES FOR 12 AND 16.

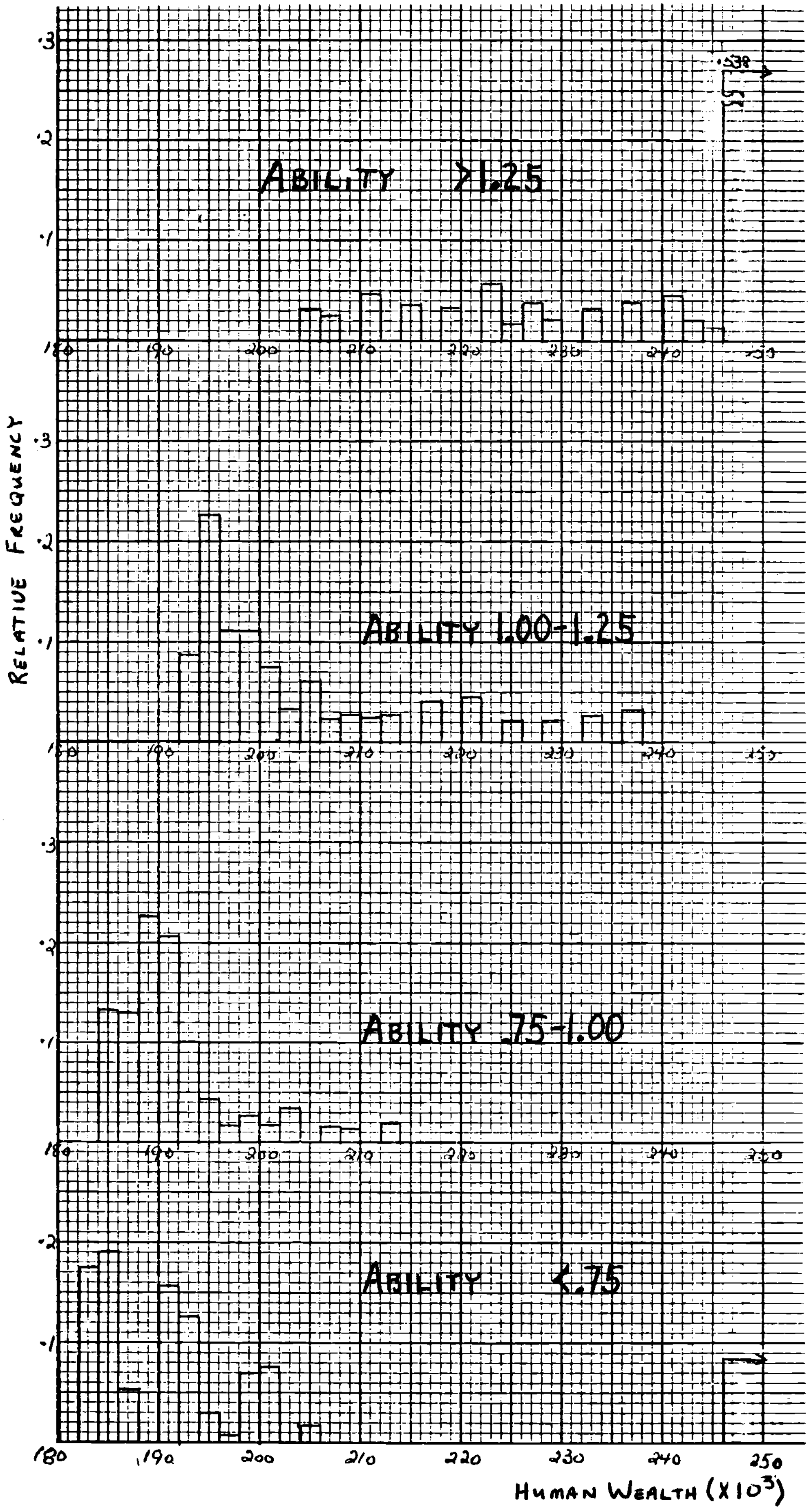

FIGURE 13. (CONTINUED, PAGE 5, R = .03)

Note: Mean Ability is 1.00 and the standard deviation is .25.

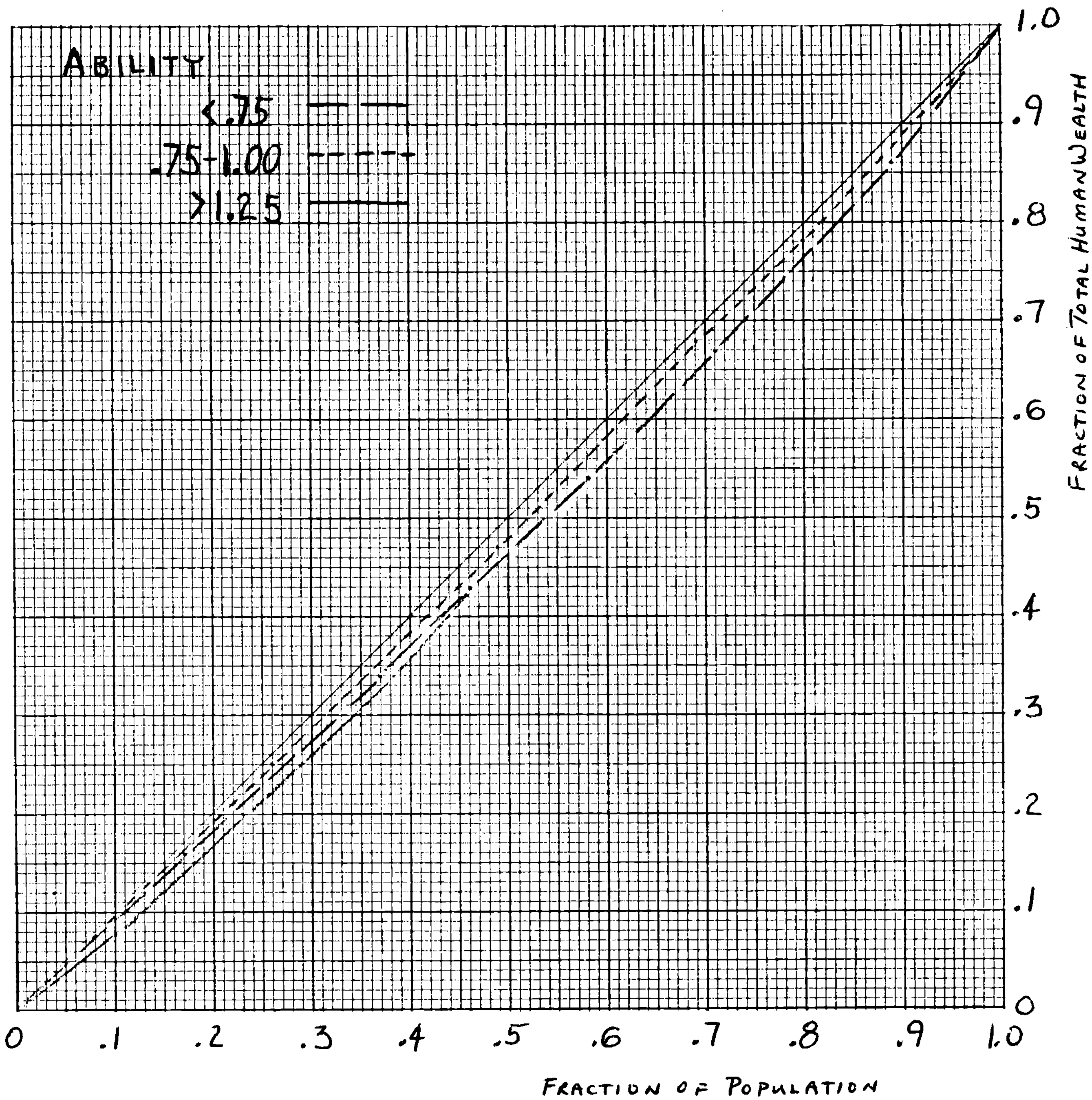

FIGURE 13. (CONTINUED, PAGE 6, R = .03)

NOTE: THE CURVE FOR >1.25 APPROXIMATELY COINCIDES WITH THAT OF <.75 WHERE IT IS NOT SHOWN. THE CURVE FOR 1.00–1.25 IS OMITTED FOR CLARITY, IT LIES BETWEEN THE CURVES FOR <.75 AND .75–1.00.

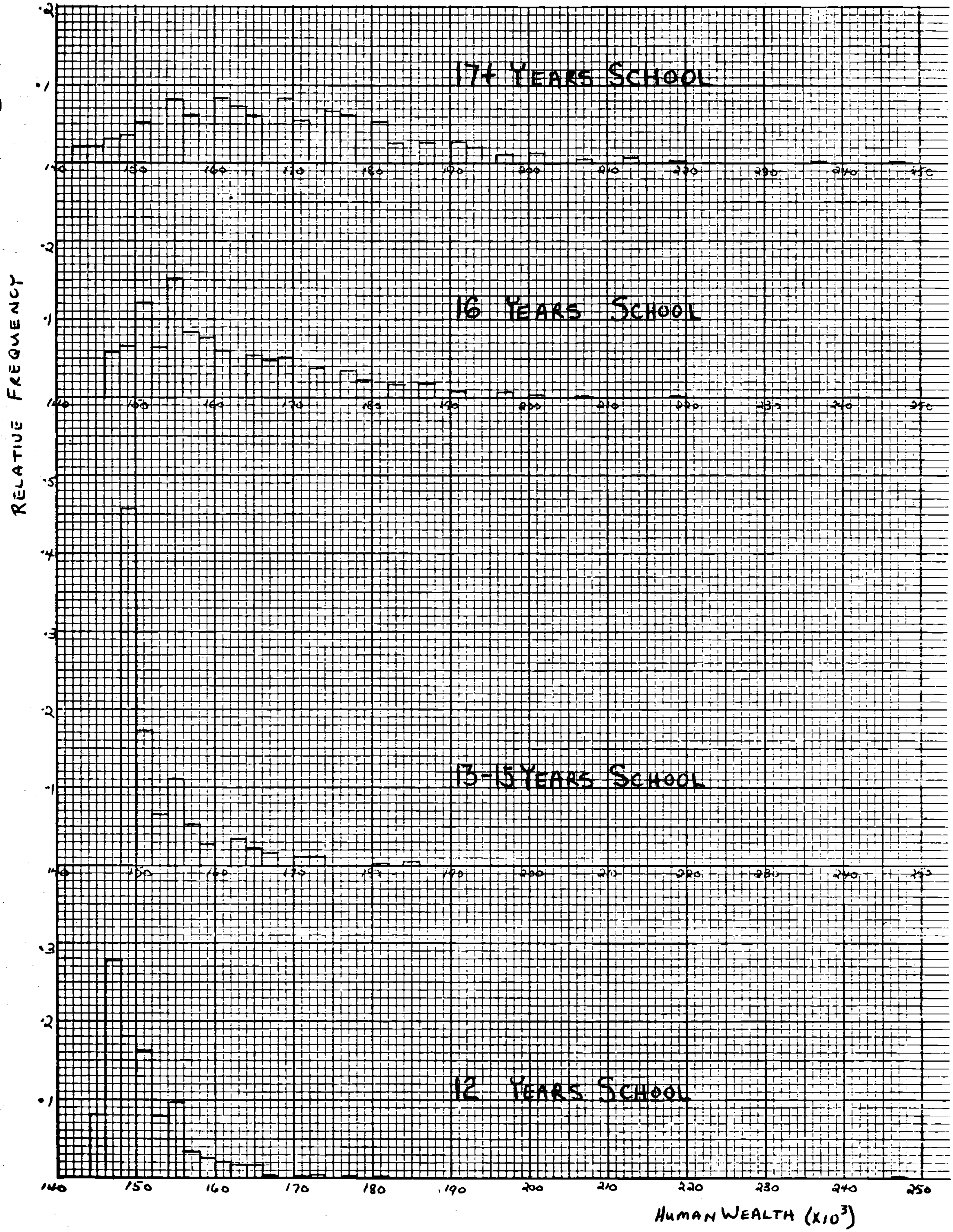

Figure 13. (continued, page 7, R=.4)

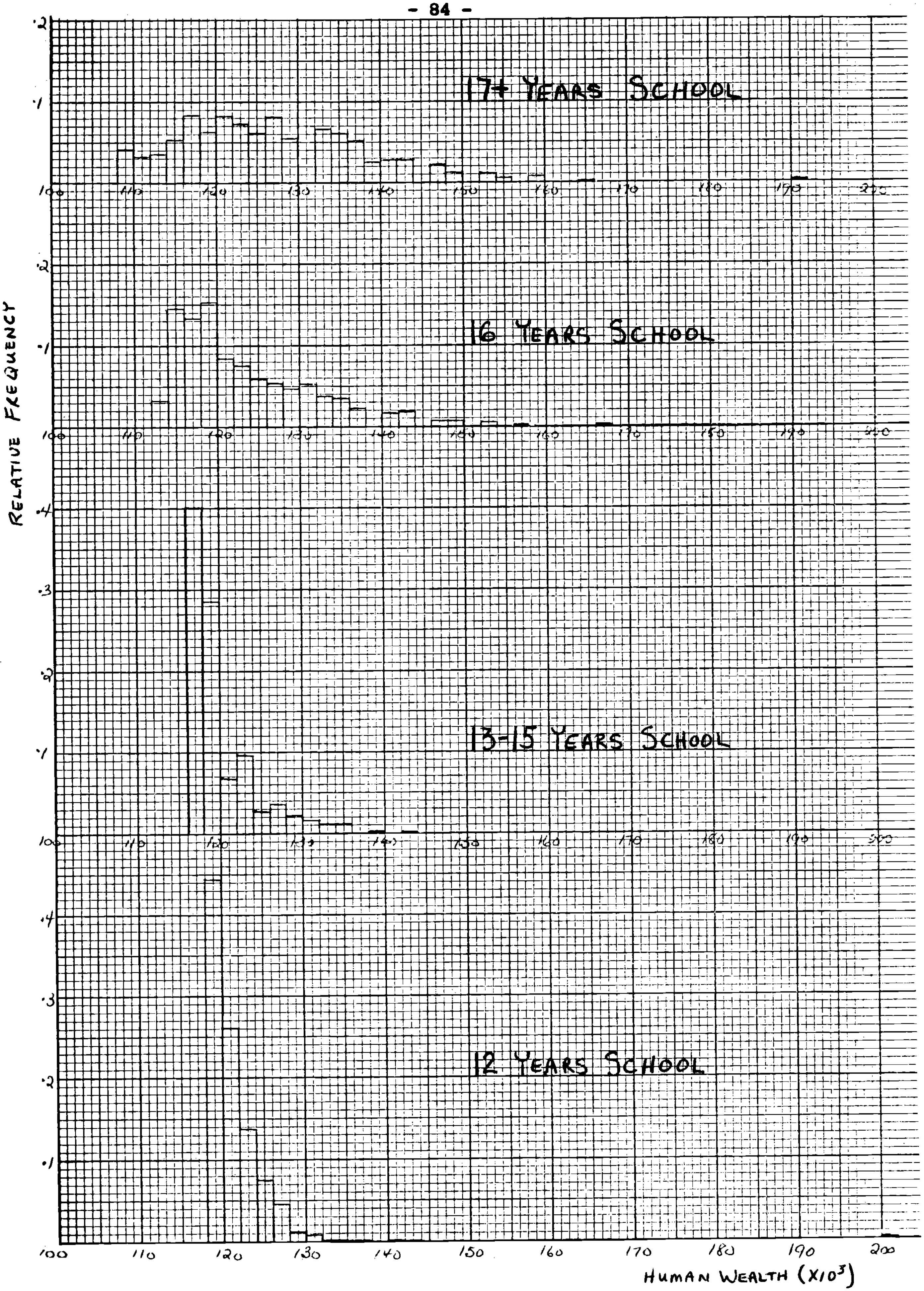

FIGURE 13. (CONTINUED, PAGE 8, R=.05)

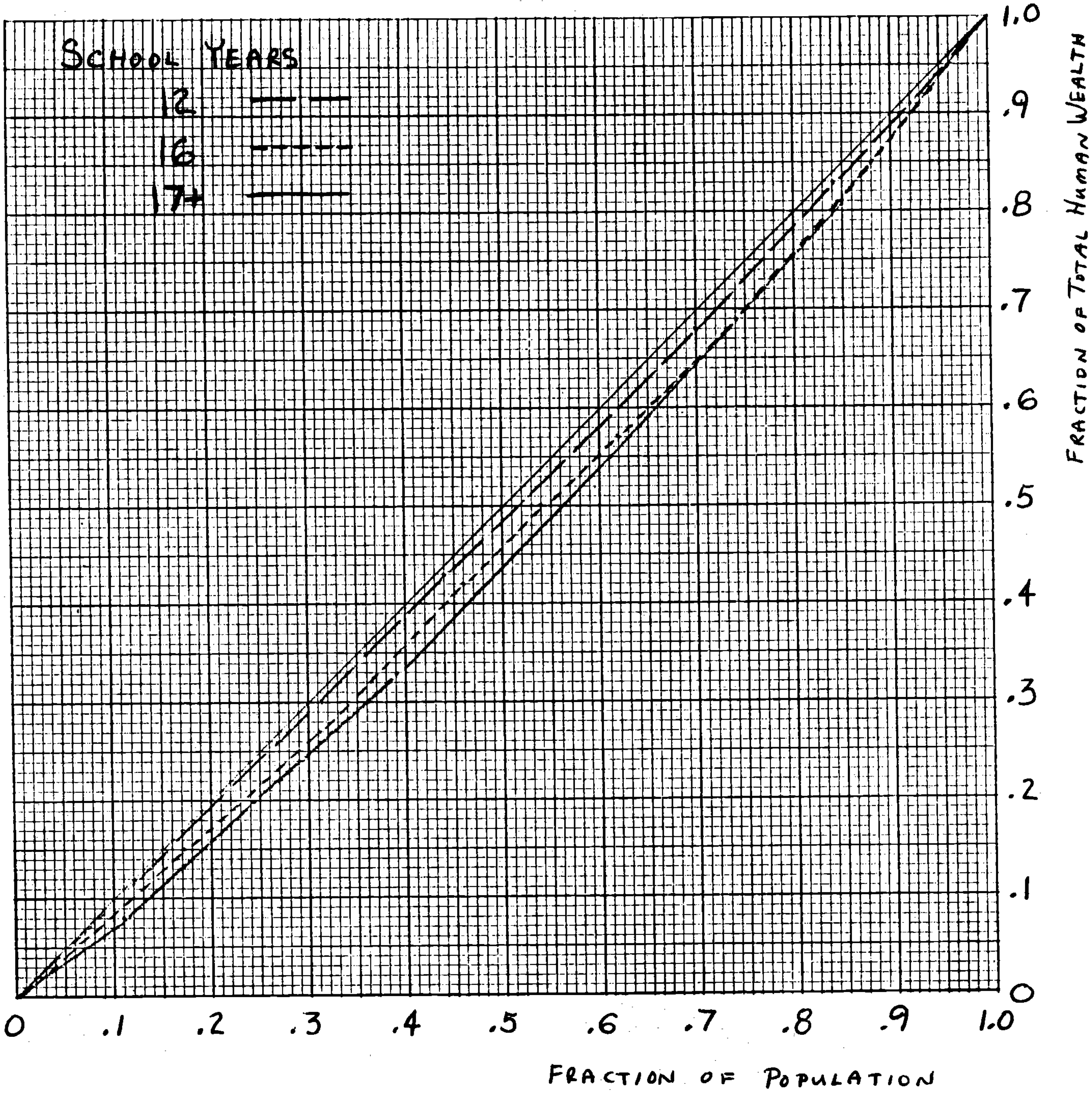

FIGURE 13. (CONTINUED, PAGE 9, R=.05)

NOTE: THE CURVE FOR 13-15 IS OMITTED FOR CLARITY. IT LIES BETWEEN THE CURVES FOR 12 AND 16.

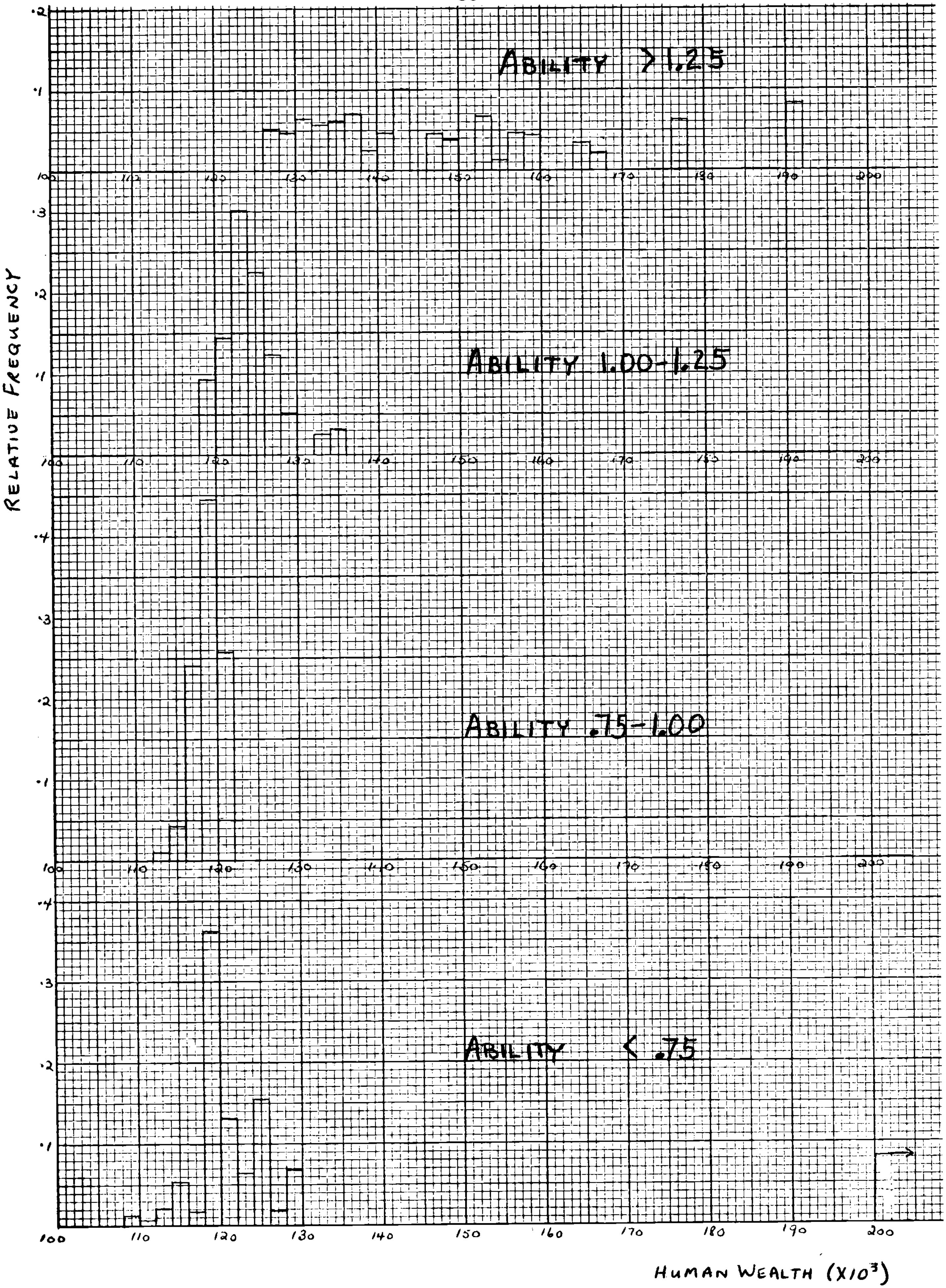

FIGURE 13. (CONTINUED, PAGE 10, R=.05)

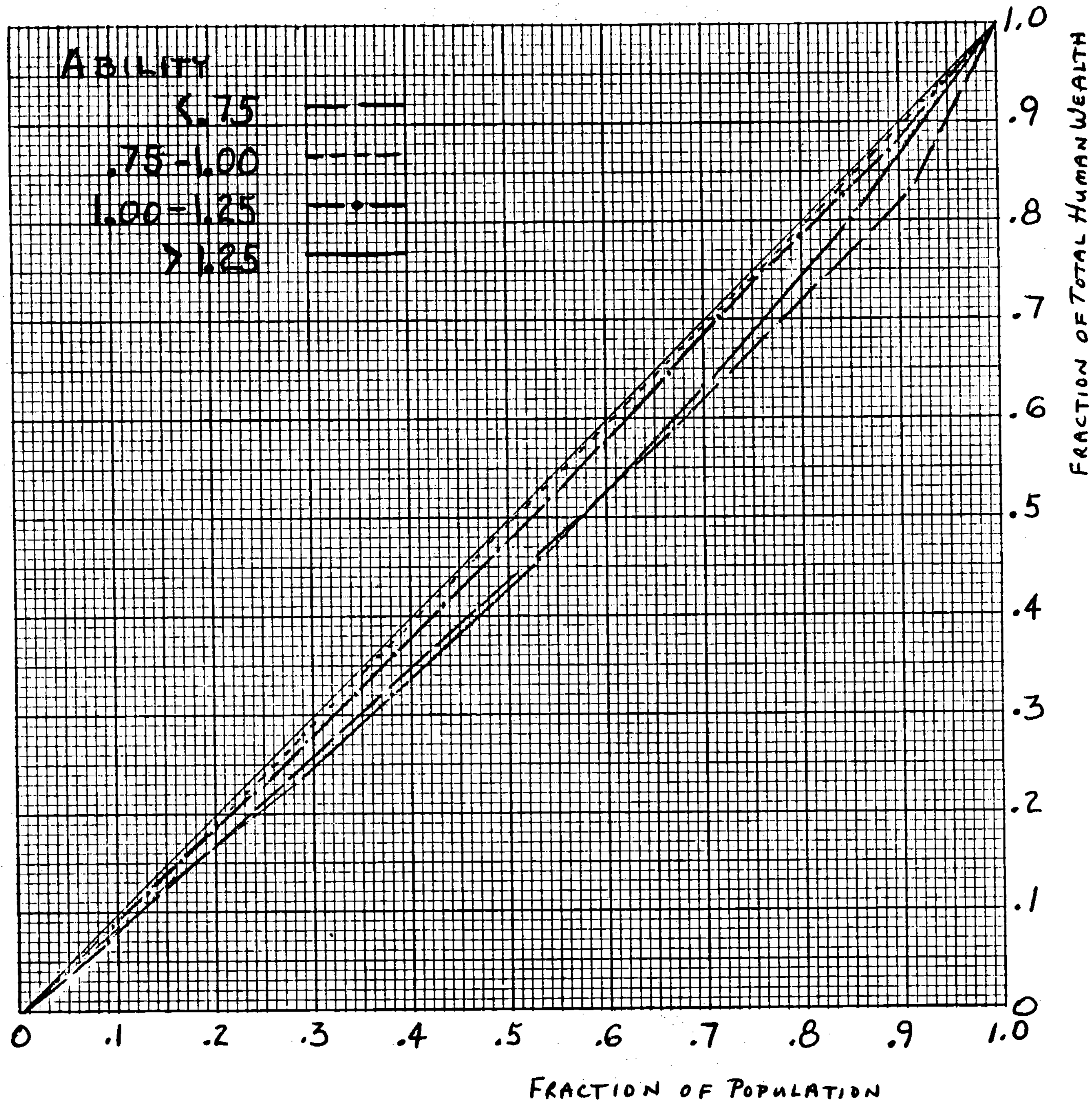

FIGURE 13. (CONTINUED, PAGE 11, R=.05)

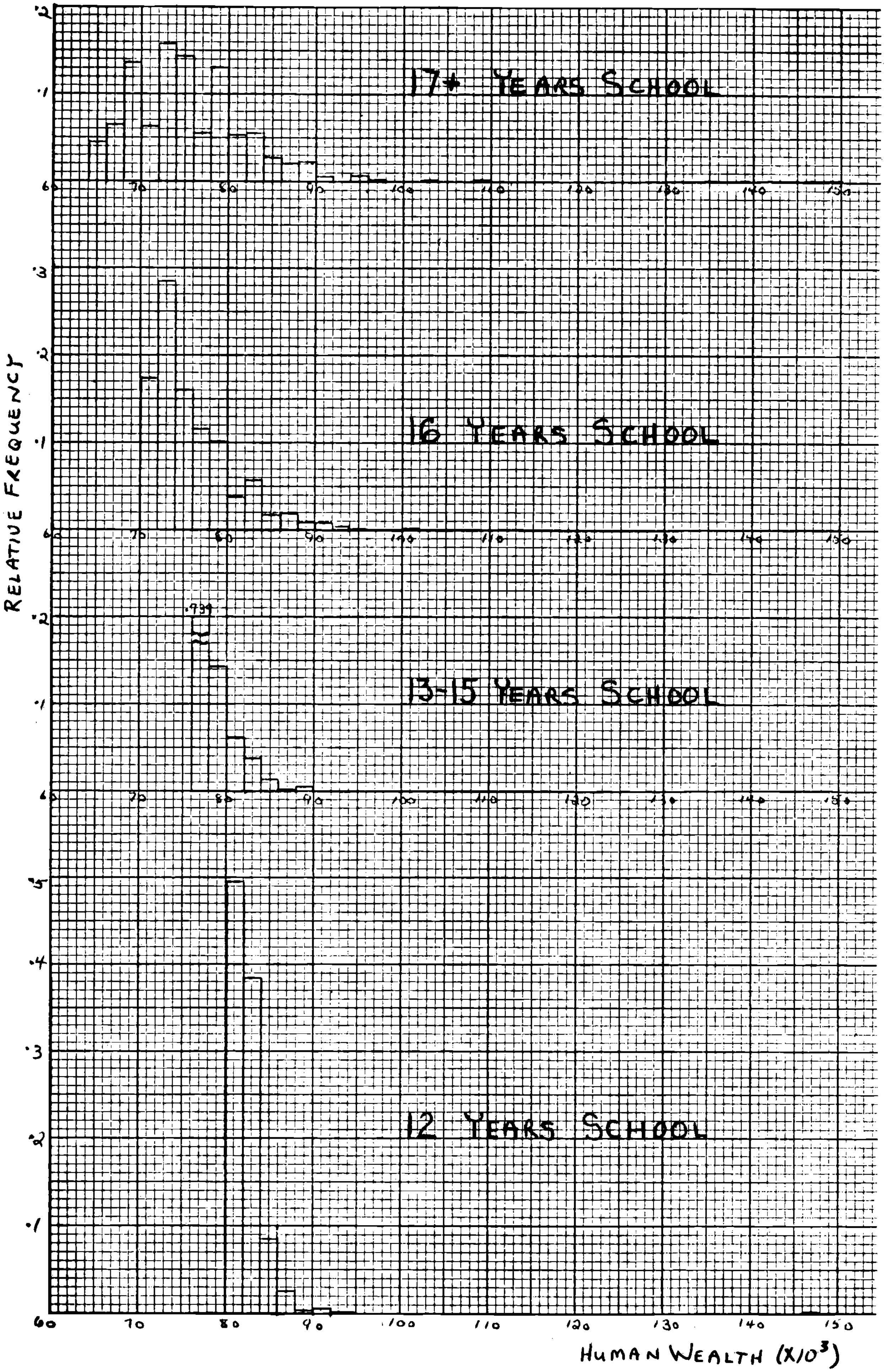

FIGURE 13. (CONTINUED, PAGE 12, R=.07)

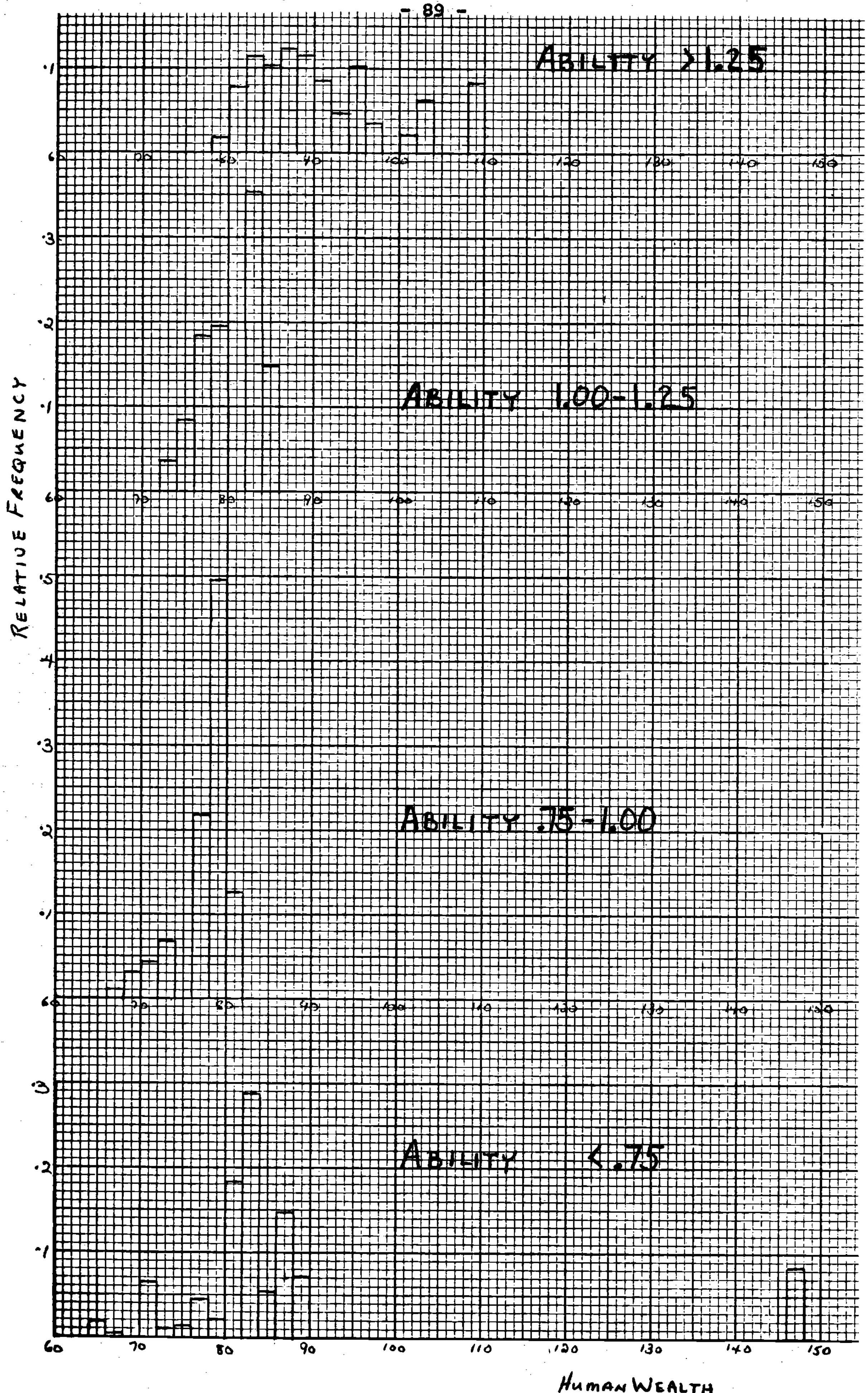

FIGURE 13. (CONTINUED, PAGE 13, R=.07)

Table 10

Predicted Distribution of the Present Value of Mean Earnings for the Overall Population of Employed Males in 1960 With at Least a High School Education, for Schooling Classes and Ability Classes, Discounted at Several Rates.

Disc. Rate	Interval Mid-point in $1000	Overall	Schooling Class 12	13-15	16	17+
.03	183	.059045	.104913			
	185	.095128	.169026			
	187	.050142	.089092			
	189	.101220	.179850			
	191	.104718	.092423	.226316	.022096	.015748
	193	.102016	.074938	.247368	.037063	.023622
	195	.122053	.135720	.160526	.066287	.031496
	197	.042058	.039134	.063158	.052744	.0
	199	.046076	.036636	.063158	.066999	.036745
	201	.023465	.025812	.002632	.065574	.002625
	203	.024266	.0	.047368	.071989	.052493
	205	.032937	.020816	.052632	.080542	.0
	207	.024308	.009992	.0	.083393	.083990
	209	.011902	.0	.026316	.0	.062992
	211	.022464	.009992	.034211	.076978	.0
	213	.015862	.0	.0	.060584	.083990
	215	.006471	.002498	.023684	.0	.0
	217	.014287	.0	.0	.056308	.073491
	219	.005345	.002498	.018421	.0	.0
	221	.012186	.0	.0	.049893	.060367
	223	.005157	.004163	.013158	.0	.0
	225	.014716	.0	.0	.051319	.083990
	227	.003283	.000833	.013158	.0	.0
	229	.010253	.0		.038489	.055118
	231	.0	.0		.0	.0
	233	.011676	.000833		.035638	.068241
	235	.0			.0	.0
	237	.009571		.002632	.022096	.062992
	239	.0		.0	.0	.0
	241	.006354		.005263	.0	.052493
	243	.002293			.018532	.0
	245	.002614			.0	.026247
	247	.018136	.000833		.043478	.123359