Volume Title: Education, Income, and Human Capital

Volume Author/Editor: W. Lee Hansen, editor

Volume Publisher: NBER

Volume ISBN: 0-870-14218-6

Volume URL: http://www.nber.org/books/hans70-1

Publication Date: 1970

Chapter Title: An Interregional Analysis of Schooling and the Skewness of Income

Chapter Author: Barry Chiswick

Chapter URL: http://www.nber.org/chapters/c3279

Chapter pages in book: (p. 157 - 191)

EDUCATION AND THE DISTRIBUTION OF INCOME

AN INTERREGIONAL ANALYSIS OF SCHOOLING AND THE SKEWNESS OF INCOME

BARRY R. CHISWICK •
UNIVERSITY OF CALIFORNIA, LOS ANGELES

INTRODUCTION

THE skewness in a distribution indicates its deviation from perfect symmetry around the mean. It reflects the shape of a distribution. The shape of the personal distribution of income is relevant for discussions of the equity of the distribution, as well as for Engel curve and savings-and-investment analyses. The determination of the factors that generate the shape of the distribution is therefore a matter of considerable interest.

The skewness in the distribution of income was considered important in the past, although the scarcity of data restricted empirical studies. A. A. Young wrote in 1917 that skewness, not "concentration," is the relevant parameter for studies of the social desirability of the distribution of income.[1] Three years later, in the first edition of his *Economics of Welfare*, A. C. Pigou tried to reconcile the assumed normal distribution of ability with the positive skewness of income.[2] In spite of the rapid

NOTE: The author is indebted to Gary S. Becker and Linda B. Wedel for many helpful comments on earlier drafts of this paper. The research for this paper was supported, in part, by the Institute of Government and Public Affairs, University of California, Los Angeles.

[1] A. A. Young, "Do the Statistics of Concentration of Wealth in the United States Mean What They Are Commonly Assumed to Mean?" *Journal of the American Statistical Association*, Vol. 15 (March 1917), pp. 471–84.

[2] A. C. Pigou. *Economics of Welfare* (London. 1920). pp. 695–97.

TABLE 1

Annual Earnings During and After Training

Years of Training	Year							
	1	2	3	. . .	N-1	N	N+1	. . .
0	Y_o	Y_o	Y_o	. . .	Y_o	Y_o	Y_o	. . .
1	0	$Y_o(1+r)$	$Y_o(1+r)$	. . .	$Y_o(1+r)$	$Y_o(1+r)$	$Y_o(1+r)$	. . .
2	0	0	$Y_o(1+r)^2$	. . .	$Y_o(1+r)^2$	$Y_o(1+r)^2$	$Y_o(1+r)^2$	. . .
. . .				. . .				
N	0	0	0	. . .	0	0	$Y_o(1+r)^N$	. . .

increase in data on the distribution of income, the skewness of income has been ignored by most recent studies.[3]

The purpose of this paper is to refocus attention on this important aspect of the distribution of income. In particular, by viewing schooling as a form of capital, this paper attempts to ascertain the extent to which schooling can explain the observed positive skewness in the distribution of personal income. An a priori model is developed in Part I, and then employed in the empirical analysis of the United States and Canada in Part II. The final section consists of a summary and conclusion.

I. THE MODEL

Let us designate Y_N the perpetual annual earnings after N years of training are completed, and Y_o the perpetual earnings if there is no training. It is assumed initially that all persons are of equal ability, that the only private costs of training are earnings foregone, and that during the training period there are no actual earnings. Using these assumptions, Table 1 will help clarify the derivation of the relation between training and earnings.

A person without training would earn Y_o every year, as is shown in the first row of Table 1. A person who invested for one year is assumed to have foregone the amount Y_o, that is, no earnings were received during this year. This is shown by the zero in the second row of the first column. If a rate of return of r were received on his investment, he would earn $Y_1 = Y_o + r\,Y_o = Y_o\,(1 + r)$ in year two and all subsequent years, where $r\,Y_o$ is the perpetual return on the investment Y_o. This is shown in the second row of Table 1. If the rate of return were the same for all

[3] There are, however, some exceptions. Income skewness was considered explicitly in Gary S. Becker, *Human Capital and the Personal Distribution of Income* (Ann Arbor, 1967), Barry R. Chiswick, "Minimum Schooling Legislation and the Cross-Sectional Distribution of Income," *Economic Journal* (September 1969), pp. 495–507, Stanley Lebergott, "The Shape of the Income Distribution," *American Economic Review*, Vol. 49 (June 1959), pp. 328–47, Harold Lydall, *The Structure of Earnings* (Oxford, 1968), Thomas Mayer, "The Distribution of Ability and Earnings," *Review of Economics and Statistics*, Vol. 42 (February 1960), pp. 189–95, Herman Miller, "Elements of Symmetry in the Skewed Income Curve," *Journal of the American Statistical Association*, Vol. 50 (March 1955), pp. 55–71, and Jacob Mincer, "Investment in Human Capital and Personal Income Distribution," *Journal of Political Economy*, Vol. 66 (August 1958), pp. 281–302.

years of training, a person with two years of training would have received no earnings during years one and two and after that an amount equal to

$$Y_2 = (Y_o + rY_o) + r(Y_o + rY_o) = Y_o(1 + r)(1 + r) = Y_o(1 + r)^2$$

where $r(Y_o + rY_o) = rY_1$ is the perpetual annual return to the investment in the second year of $Y_o + rY_o = Y_1$. A person with N years of education would receive nothing during the first N years and

$$Y_N = Y_o + r(Y_o) + rY_o(1 + r) + \ldots + rY_o(1 + r)^{N-1}$$

or

$$Y_N = Y_o(1 + r)^N \tag{1}$$

after the investment period.

If the rate of return were not the same for all years of training, the factors could not be combined, and the post-investment income stream would be represented by

$$Y_N = Y_o \mathop{\pi}_{j=1}^{N} (1 + r_j), \tag{2}$$

where π is the mathematical symbol for product.

The assumptions that there are no direct costs of training and no actual earnings during the period of investment are not realistic. A year of schooling ordinarily leaves the summer free for working, and for some levels of schooling direct costs (i.e., tuition, school supplies, and other expenses necessitated by schooling) are far from negligible. Those engaged in on-the-job training usually receive positive incomes in excess of direct costs, although in the past, payments for appenticeship programs were quite common. The earnings equation may be modified to make the model more consistent with reality.

Let C_j equal the direct plus foregone-earnings costs of the investment in the jth year of training. Y_{j-1} is the income that would be received after $j-1$ years of training if no further investments were undertaken. Designate by K_j the ratio C_j/Y_{j-1}, that is, K_j equals the proportion of potential income that is invested during year j. We previously assumed that the only cost of education was a full year of foregone earnings, so that $C_j = Y_{j-1}$ and $K_j = 1$. Now K_j may differ from unity. If total costs were greater than potential earnings during the year of training, K_j would be greater than 1. If the potential earnings exceeded total costs, K_j would be less than 1.

The introduction of the investment-income ratio, K, modifies the earnings equation. If there were no investment, Y_o would still be earned. If in year 1, the amount $C_1 = K_1 Y_o$ were invested at a rate of return of r_1, the post-investment income would be

$$Y_1 = Y_o + r_1(K_1 Y_o) = Y_o(1 + r_1 K_1).$$

If N years of investments were undertaken,

$$Y_N = Y_o(1 + r_1 K_1)(1 + r_2 K_2) \ldots (1 + r_N K_N)$$

or

$$Y_N = Y_o \prod_{j=1}^{N} (1 + r_j^*), \tag{3}$$

where $r_j^* = r_j K_j$ is the "adjusted rate of return" to the jth year of education.[4]

Individual differences due to other forms of human capital, physical capital, and luck may be introduced into equation 3 by the inclusion of a residual U_i^*. Differences in "ability" may be introduced by permitting differences in rates of return to a given level of training. The earnings equation becomes

$$Y_{N,i} = Y_o \prod_{j=1}^{N} (1 + r_{ij}^*)\, U_i^*, \tag{4}$$

where r_{ij}^* is the adjusted rate of return to the ith individual for the jth year of training. Taking logarithms of both sides of equation 4, and using the relation Ln $(1 + a) \approx a$ when a is small, results in

$$\text{Ln } Y_{N,i} = \text{Ln } Y_o + \sum_{j=1}^{N} r_{ij}^* + U_i, \tag{5}$$

where $U_i = \text{Ln } U_i^*$ and the "approximately equal to" sign has been replaced by the symbol for "equal to." Differences in earnings at the zero investment level may be considered to be in the residual.

The sum of the adjusted rates of return $\sum_j r^*_{ij}$ can be rewritten as

$$\sum_j r_{ij} = \bar{r}_i^* N_i = \bar{r}^* N_i + (\bar{r}_i^* N_i - \bar{r}^* N_i), \tag{6}$$

where $\bar{r}_i^*$ is the ith person's average adjusted rate of return and $\bar{r}^*$ is the

[4] That is, the rate of return adjusted for the fraction of potential income that was invested.

average $\bar{r}_i^*$ for the population.[5] If it is assumed that deviations from the population's average adjusted rate of return appear in the residual U', equation 5 can be rewritten as

$$\text{Ln } Y_{N,i} = \text{Ln } Y_o + \bar{r}^* N_i + U_i'. \tag{7}$$

Training can be separated into two components, schooling and on-the-job training. Thus, the earnings equation becomes

$$\text{Ln } Y_{N,i} = \text{Ln } Y_o + \bar{r}_s^* S_i + \bar{r}_j^* J_i + U_i' \tag{8}$$

where S and J designate years of schooling and on-the-job training respectively.

The model can be used for interregional analyses of income inequality or income skewness. The development of the theory and empirical analyses for income inequality have been presented elsewhere.[6] The remainder of the present analysis is specifically concerned with an examination of the skewness of income.

Due to the scarcity of data for on-the-job training, the empirical analysis of Part II is for schooling alone. Consequently, the subsequent theoretical analysis focuses on schooling. On-the-job training is considered a component of the residual.

Let us first assume that income is derived solely from investments in schooling, and then investments in other assets and luck shall be

[5] In mathematical terms:

$$\bar{r}_i^* = \sum_{j=1}^{N_i} \frac{r_{ij}^*}{N_i} = \sum_{j=1}^{N_i} \frac{r_{ij}K_{ij}}{N_i},$$

where N_i is the number of years of training and

$$\bar{r}^* = \sum_{i=1}^{p} \frac{\bar{r}_i^*}{p}$$

where p is the size of the population.

[6] Gary S. Becker and Barry R. Chiswick, "Education and the Distribution of Earnings," *American Economic Review*, Vol. 56, No. 2 (May 1966), pp. 358–69; Barry R. Chiswick, "Human Capital and the Distribution of Personal Income" (unpublished Ph.D. dissertation, Columbia University, 1967); and Barry R. Chiswick, "The Average Level of Schooling and the Intra-Regional Inequality of Income: A Clarification," *American Economic Review*, Vol. 58, No. 3, pt. 1 (June 1968), pp. 495–500.

included.[7] If the ith person's income were derived solely from investments in schooling,

$$\text{Ln } Y_i = \bar{r}_i^* S_i. \tag{9}$$

In the population, the rate of return and level of schooling have positive means. Then, if it is assumed that $\bar{r}_i^*$ and S_i are independent and normally distributed, their product has a small positive skewness.[8] A small positive skewness in the natural logarithm of income implies a considerable positive skewness in income itself.[9]

If the cost of funds for investment in schooling were the same for all, those with higher marginal rates of return (e.g., those with greater ability) would invest more in schooling.[10] This produces a positive correlation between $\bar{r}_i^*$ and S_i. The positive correlation is reduced if, as seems plausible, those with higher apparent levels of ability have a lower cost of funds.[11] A positive correlation between the rate of return and the number of years of schooling increases the positive skewness of income.

Thus, income is positively skewed as long as the distributions of the rate of return or of schooling are not sufficiently negatively skewed or as long as the rate of return and the level of schooling are not sufficiently negatively correlated.

One problem with the formulation just presented is the current immeasurability of characteristics, other than average level, of the distribution of the rate of return from schooling. The formulation presented below permits an empirical analysis using the limited data that are available. In addition, it contains a residual, which includes the effects of differences in ability, of capital other than schooling, and of luck.

The earnings or income equation for schooling can be written as

$$\text{Ln } Y_i = \text{Ln } Y_o + \bar{r}^* S_i + [(\bar{r}_i^* - \bar{r}^*) S_i + U_{i,s}] \tag{10}$$

where the sum in brackets is the residual. The term $U_{i,s}$ includes the

[7] For an analysis of the effect of chance on the distribution of income, see Milton Friedman, "Choice, Chance and the Personal Distribution of Income," *Journal of Political Economy*, Vol. 61 (August 1953), pp. 277–90.

[8] C. C. Craig, "On the Frequency Function of XY," *Annals of Mathematical Statistics*, Vol. 7, No. 1 (March 1936), pp. 1–15.

[9] See also Mincer, *op. cit.*, and Becker, *Human Capital*, pp. 61–66.

[10] Becker and Chiswick, *op. cit.*, or Becker, . . . *Human Capital*, pp. 2–25.

[11] *Ibid.*

effects on income of on-the-job training, other forms of human capital, physical capital, and luck. The expression $\bar{r}_i^* - \bar{r}^*$ reflects individual differences in adjusted rates of return from schooling.

When the residual is neglected, a normal distribution of S_i produces a normal distribution in the natural logarithm of income. A normally distributed log of income implies that income itself is positively skewed. The greater the skewness of schooling, the greater the skewness of income. Income is positively skewed unless there is a sufficient amount of negative skewness in the distribution of schooling. Given current distributions of schooling, it seems reasonable to predict that if income were due solely to schooling, income would be positively skewed. This is consistent with the most distinctive and apparently universal characteristic of the distribution of income, its positive skewness.

The skewness of income is also a function of the level of the adjusted rate of return. There is reason to believe that, for a given distribution of schooling, the higher the rate of return the greater is the skewness in income. This is easily proved when income (Y) is log-normally distributed. The skewness of Y can be measured by

$$\mathrm{Sk}\,(Y) = \frac{(Z_3)^{1/3}}{(Z_2)^{1/2}}, \tag{11}$$

where Z_i is the ith moment around the mean.[12] Then,

$$\mathrm{Sk}\,(Y) = \left(\frac{Z_3}{(Z_2)^{3/2}}\right)^{1/3} = (\eta^3 + 3\eta)^{1/3}, \tag{12}$$

where $\eta^2 = e\sigma^2 - 1$, and σ^2 is the variance of the natural logarithm of Y.[13] If income were due solely to the rate of return from schooling and the level of schooling, and if for individuals these parameters were independent,[14] then

$$\sigma^2 = \mathrm{Var}(\mathrm{Ln}\ Y) = \bar{r}^2\,\mathrm{Var}(S) + \bar{S}^2\,\mathrm{Var}(r) + \mathrm{Var}(r)\,\mathrm{Var}(S). \tag{13}$$

Thus, *ceteris paribus*, the larger the adjusted rate of return, the larger is

[12] Kendall suggests this measure raised to the sixth power. M. G. Kendall, *The Advanced Theory of Statistics*, Vol. 1, 4th edition (London, 1948), p. 81.

[13] J. Atchison and J. A. C. Brown, *The Lognormal Distribution* (Cambridge, England, 1957), pp. 7–8.

[14] Leo Goodman, "On the Exact Variance of Products," *Journal of the American Statistical Association*, Vol. 55 (December 1960), pp. 708–13.

the skewness of income. In addition, this procedure predicts that, *ceteris paribus*, the skewness of income is larger at higher levels of schooling.

When a residual term does exist, the conclusions of the last two paragraphs are necessarily valid only for the income predicted by schooling. If the residual is held constant, they are also valid for observed income. I shall demonstrate that several arguments support the hypothesis that the correlation between the skewness of residual income and the skewness of predicted income is positive. If this hypothesis is correct, the skewness of observed income has a positive simple correlation with the predicted skewness, and tends to have a positive simple correlation with the skewness of schooling. In addition, more stringent assumptions result in the hypothesis that observed income skewness is positively correlated with the average level of the rate of return and the average level of schooling.

There is reason to believe that several of the components of the residual also tend to produce skewness in income. The residual of equation 10 is divided into two components. The first, reflecting differential ability, is $(\bar{r}_i^* - \bar{r}^*)\, S_i = d_i\, S_i$. If d_i and S_i were independent, and if both were normally distributed, their product d_iS_i would have a symmetric distribution since the expected value of d_i is zero.[15] The residual is in terms of natural logs, and a lack of skewness in the logarithm of income implies a positive skewness of income itself. As long as the distributions of d_i and S_i are not sufficiently negatively skewed or negatively correlated with each other, the distribution of the antilog of d_iS_i is positively skewed.

We may use the same arguments for other forms of human capital as we used for schooling. As long as the distributions of these other investments, and the rate of return from these investments, are not sufficiently negatively skewed or negatively correlated, their component of residual income will be positively skewed. In addition, the distribution of wealth tends to be highly positively skewed and is likely to produce skewness in its component of the residual income.

There appear to be no simple additive relations for skewness. Although we cannot be certain that positive skewness in its components produces a positively skewed residual, this seems plausible.

Several a priori arguments lead to an expectation of a positive cor-

[15] Craig, *op. cit.*

relation across areas between the predicted skewness and the residual skewness. An increase in the positive skewness of schooling (S_i) results in an increase in the skewness of predicted income and in the skewness of the differential ability ($d_i S_i$) component of the residual. In addition, it seems plausible to assume that the skewness of investments in schooling and the skewness of investments in other assets are positively related. It also seems plausible to assume that the rate of return from schooling and from other forms of capital are positively related across regions. These factors tend to produce a positive relation between the skewness of predicted income and the skewness of residual income.

Since there tends to be greater positive skewness in the ownership of physical capital than in years of schooling, and since investments in these two assets tend to be positively correlated, the residual skewness and the income skewness should be larger than the predicted skewness.[16] The difference would be smaller if earnings rather than total income were under consideration.

The a priori analysis indicates that the higher the level of the rate of return from schooling, the greater is the skewness of income. In addition, the empirical analysis of Part II indicates that the rate of return is important in explaining interregional differences in income skewness. Thus, a brief analysis of the factors that influence interregional differences in the rate of return would be useful.

It is not clear whether poorer regions have a higher or a lower rate of return from schooling than wealthier regions. Although the cost of obtaining a given level of funds for investment in schooling may be higher in a poorer area, the demand for human capital is lower, and the net effect is ambiguous. However, for regions among which there is considerable migration, the poorer ones may have a higher rate of return due to the effects of interregional migration.[17]

Workers with a higher level of education tend to have more knowledge about opportunities elsewhere and perhaps a reduced attachment to place of origin. In addition, they are likely to be wealthier than less educated workers, and thereby find it easier to finance the direct and oppor-

[16] Becker developed a model based on rational investment decisions which predicts a larger positive skewness of investments in physical capital than in human capital. (Becker, . . . *Human Capital*, especially pp. 35–37.)

[17] On this point see Chiswick, *op. cit.* "The Level of Schooling . . . ."

tunity costs associated with migration. Finally, whereas interregional wage differentials may tend to increase in proportion to the wage level for higher levels of skill, costs of migration are not likely to increase in proportion. The less-than-proportionate rise in migration costs is attributable to the direct cost component of migration which is not likely to rise as rapidly as the wage level for increasing levels of skill. These factors provide educated workers with a greater incentive to migrate than less educated workers. This higher migration rate has been found in empirical studies.[18]

The greater mobility implies that educated workers in poor regions face more of a national market for their services than do less educated workers. In the poor regions, those with higher levels of education receive wages relatively closer to the wage for their skill in the wealthier regions, than do workers with less education. Therefore, the rate of return from schooling tends to be higher in poorer regions. This relation between average income and the rate of return appears in interregional analyses for the United States and for Canada.[19]

Thus, holding the distribution of schooling constant, the poorer regions of countries in which there is considerable migration are likely to have a larger skewness of income than the wealthier regions. In an international comparison, however, the rate of return may be uncorrelated with the level of income and the latter may be uncorrelated with income skewness.

The rate of return from schooling may, however, be positively related to the rate of growth of income. If so, *ceteris paribus*, a region experiencing rapid economic development would tend to have a large income skewness. Thus, larger income skewness in rapidly growing countries may be a consequence, and not necessarily a cause, of the rapid growth.

The effect on income skewness of interregional differences in the

[18] See Becker, *Human Capital*, p. 89 note and Rashi Fein, "Educational Patterns in Southern Migration," *Southern Economic Journal, Supplement*, Vol. 32, no. 1, pt. 2 (July 1965), pp. 106–24.

[19] Becker and Chiswick, *op. cit.*, and Giora Hanoch, "An Economic Analysis of Earnings and Schooling," *Journal of Human Resources*, Vol. 2 (Summer 1967), pp. 310–29. In addition, it is implied in *Second Annual Review: Toward Sustained and Balanced Economic Growth* (Ottawa: Economic Council of Canada, December 1965), p. 119.

distribution of schooling, due to differences in the level of income, is not clear a priori. The model predicts, *ceteris paribus*, that a higher level of schooling produces a larger income skewness. The level of schooling in a region tends to be negatively related to the skewness of schooling, and the analysis predicts that lower schooling skewness produces lower income skewness. Thus, the simple correlation between income skewness and the level of schooling need not be positive.[20]

A negative correlation between the level and skewness of schooling is found empirically and is easy to understand. The distribution of years of schooling tends to have a central tendency and two finite limits—namely, zero or some minimum required by law as the lower limit, and say twenty years as the upper limit. In almost all regions there would be some at each extreme. As the mean rises from a low to a higher level, the skewness of schooling tends to decline from a positive to a negative value. Therefore, regions with a higher level of income and schooling may tend to have a smaller skewness of schooling, with an ambiguous net effect on the skewness of income. Similarly, economic growth may raise the level but decrease the skewness in years of schooling, with an ambiguous net effect on the skewness of income.

In summary, the rate of return from schooling is likely to be higher in the poorer regions of a country, but uncorrelated with average income across countries. In addition, lower average levels of schooling or income are likely to be associated with higher positive skewness of schooling. There does not appear to be any direct relationship between the rate of return and the skewness of schooling. However, the above suggests that, due to their mutual intercorrelation with the level of schooling, the rate

[20] Let $X_0 =$ the skewness of income,
$X^1 =$ the level of schooling,
$X_2 =$ the skewness of schooling,
$X_0 = a_0 + b_{y1.2}X_1 + b_{y2.1}X_2$, and
$X_0 = a_1 + b_{y1}X_1$,
where the theory predicts that $b_{y1.2} > 0$ and $b_{y2.1} > 0$, and the empirical analysis indicates that $b_{12} < 0$. From statistical theory we know that $b_{y1} = b_{y1.2} + b_{12}\, b_{y2.1}$. [Arthur S. Goldberger, *Econometric Theory* (New York, 1964), pp. 194–95.]

If the magnitudes of b_{12} and $b_{y2.1}$ are sufficiently large and $b_{y1.2}$ is sufficiently small, b_{y1} will be negative. Thus, a negative correlation between the level and the skewness of schooling could change a positive partial relation between the level of schooling and income skewness into a negative simple correlation.

of return and the schooling skewness are positively correlated across the regions of a country, but are uncorrelated across countries.

The schooling parameters under study can be directly influenced by government educational policies, with accompanying effects on the skewness of income. For example, minimum schooling laws compel those who would otherwise invest in years of schooling below the legal minimum to increase their investments. This increases the skewness of schooling and also the skewness of income.[21]

The analysis also indicates that the skewness of income in a region depends in part on the correlation across individuals between the rate of return from schooling ($\bar{r}_i^*$) and the level of schooling (S_i). Government subsidies designed to increase the level of ability or reduce the cost of schooling to low-ability, poor students (e.g., the Head Start Program or scholarships to students from poor families) reduce the correlation between the rate of return and the level of schooling. *Ceteris paribus*, these policies reduce the skewness of income. Scholarships for high-ability, wealthy students have the opposite effect.

Not all of the hypotheses suggested here lend themselves to empirical testing. The scarcity of compatible international income data prevent an intercountry analysis. In addition, adequate data are not available for a time-series study. Thus, Part II consists of cross-sectional interregional analyses for the United States and for Canada.

The hypotheses tested below are:

1. Schooling parameters can produce a considerable amount of skewness in income.

2. The observed income skewness and the predicted income skewness are positively correlated, and have positive partial correlations with the rate of return, the average level of schooling and the skewness of schooling.

3. The residual skewness is positively correlated with the predicted skewness and the observed skewness.

4. Negative correlations between the level of schooling and the rate of return and schooling skewness may produce biased simple correlations between schooling and income parameters.

[21] The effects of minimum schooling laws on the distribution of income are analyzed in Chiswick, "Minimum Schooling Legislation . . .".

II. EMPIRICAL ANALYSIS

Cross-classified data for schooling and income exist for the regions and states of the United States and for the provinces of Canada. This permits the direct calculation of the skewness of earnings or income (i.e., observed skewness), the skewness of schooling, and the average level of schooling. Although average rates of return from schooling have been calculated for a number of regions and countries in recent years, their number is still too small for an effective interregional analysis.

Cross-classified data on schooling and income are used to estimate the average adjusted rate of return from schooling. A least squares linear regression analysis is performed using the equation

$$\text{Ln } Y_{S,i} = (\widehat{\text{Ln}}\, Y_o) + \hat{r}\, S_i + \hat{U}_i, \tag{14}$$

where $\hat{r}$ and $(\widehat{\text{Ln}}\, Y_o)$ are the regression estimates of the average adjusted rate of return and the average zero schooling level of income respectively. U_i is the residual whose squared deviation from the regression line is minimized. The regression approach appears to generate estimates of rates of return from schooling that are lower than the internal rates of return calculated by others. The bias appears for all regions, and may not alter the relative ranking of rates of return.[22]

If the antilogs of both sides of equation 14 are taken,

$$Y_i = e_{\mathrm{L}}{}^{\hat{n}\, Y_o}(1 + r)^{S_i} e^{\hat{U}_i}, \tag{15}$$

where $e^{\widehat{\text{Ln}}\, Y_o}\,(1 + \hat{r})^{S_i}$ is predicted income and $e^{\hat{U}_i}$ is the residual income. The measure of skewness used in this study is equation 11 which is a pure number, and equals zero for a symmetric distribution. Since Ln Y_o is assumed constant within each region, it has no effect on the predicted skewness. The predicted skewness is the skewness of $(1 + \hat{r})^{S_i}$. The residual skewness is the skewness in the residual income $e^{\hat{U}_i}$.

Since we are exploring the relation between schooling and the personal distribution of income, the characteristics of the entire population of a region are not relevant. Students should be removed because the model was developed for those who completed their investments. Wives

[22] For an analysis of regression estimates of rates of return, including a comparison with estimates of internal rates of return, see Chiswick, *Human Capital* . . ., chapter 2.

TABLE 2

Skewness Parameters in the United States for Earnings of Males Aged Twenty-Five to Sixty-Four[a]

	Skewness				Adjusted Rate of Return
	Observed Income (1)	Predicted Income (2)	Residual Income (3)	Schooling (4)	(5)
1. U. S.	1.69	0.85	1.64	+0.31	.08
2. U. S. White	1.67	0.80	1.64	+0.24	.07
3. North	1.69	0.83	1.63	+0.42	.06
4. North White	1.67	0.80	1.62	+0.39	.06
5. South	1.75	0.92	1.79	+0.57	.09
6. South White	1.71	0.84	1.76	+0.42	.08

[a]Definitions:

1. Skewness:

$$SK = \frac{(Z_3)^{1/3}}{(Z_2)^{1/2}}$$

where Z_i is the i^{th} *moment of a variable around its mean.*

2. Regression equation:

$$\text{Ln } Y_{S,i} = (\hat{Ln}Y_o) + \hat{r}\, S_i + \hat{U}_i\ ;$$

3. Observed Income: $Y_{S,i}$;
4. Predicted Income: $e^{\hat{Ln}Y_o}(1 + \hat{r})^{S_i}$;
5. Residual Income: $e^{\hat{U}_i}$

The probabilities presented in subsequent tables represent the chance that sample estimates of *R* will be greater than the values given. The probabilities are based on the number of degrees of freedom equal to, or nearest to, the number of observations minus two. (From R. A. Fisher and F. Yates, *Statistical Tables*, London, 1938, Table VI, p. 36-37.)

SOURCE: *U. S. Census of Population: 1960, Subject Reports. Occupation by Earnings and Education* (Washington: Bureau of the Census), Tables 1, 2 and 3.

should be excluded, since their labor force behavior is strongly influenced by their husbands' income and the number and age distribution of their children. The aged also should be excluded; many of them have low labor force participation due to ill health, discrimination, and pension income which often specifies maximum earnings.

The desired group can be approximated by restricting the data to males aged twenty-five to sixty-four. Although the model was developed for an infinite working life, the reality of a finite work life does not alter the model's basic stucture or its predictions. In addition, where possible

TABLE 3

Means and Standard Deviations of Skewness Parameters for the States (Income of Males Aged Twenty-Five and Over) and the Provinces (Income of Nonfarm Males Ages Twenty-Five to Sixty-Four)[a]

Regions	Skewness				Adjusted Rate of Return (5)
	Observed Income (1)	Predicted Income (2)	Residual Income (3)	Schooling (4)	
1. U. S. total–	1.14	0.93	1.26	−0.20	0.10
51 states	(0.09)	(0.11)	(0.09)	(0.48)	(0.01)
2. Nonsouth –	1.11	0.90	1.22	−0.32	0.10
34 states	(0.07)	(0.10)	(0.07)	(0.46)	(0.01)
3. South –	1.21	0.98	1.35	+0.03	0.12
17 states	(0.07)	(0.09)	(0.07)	(0.46)	(0.01)
4. U. S., white –	1.13	0.92	1.25	+0.00	0.10
51 states	(0.07)	(0.10)	(0.09)	(0.52)	(0.01)
5. Nonsouth, white –	1.10	0.91	1.21	−0.19	0.10
34 states	(0.07)	(0.09)	(0.08)	(0.51)	(0.01)
6. South, white –	1.17	0.95	1.33	+0.40	0.11
17 states	(0.06)	(0.11)	(0.07)	(0.26)	(0.01)
7. Canada –	1.42	1.16	1.55	+0.81	0.09
11 provinces	(0.13)	(0.09)	(0.20)	(0.25)	(0.01)

[a]See notes to Table 2. Standard deviations are in parentheses.

SOURCES: *United States Census of Population: 1960, Vol. 1, Characteristics of the Population,* Parts 2-52 (Washington: Bureau of the Census), Table 138; *Census of Canada: 1961* (Ottawa: Dominion Bureau of Statistics), Table A.11 for the provinces, unpublished.

and practicable, separate calculations are made to remove the effects of differences between races.

The United States data are for adult males and come from the *1960 Census of Population.* For the regions and the United States as a whole, the data are for earnings, and the sample consists of males aged twenty-five to sixty-four with earnings in 1959. At the state level, however, the absence of a cross classification of earnings by schooling in the census necessitated the use of income rather than earnings. Therefore, the sample for the states consists of males aged twenty-five and over with income in 1959. The analyses are performed for males and for "white males," where "white males" are defined as whites in the United States and the regions. In the interstate analysis, however, "white males" are defined as whites only for the sixteen states (including the District of Columbia) with 10 per cent or more nonwhites plus New York State, and as all males in the remaining thirty-four states. Tests indicate that the inclusion of nonwhites in the latter thirty-four states has a negligible effect on the magnitude of the parameters. The Canadian data are for the ten provinces and the Yukon Territory and come from unpublished tables of the *1961 Census of Canada.* The data are for income and the sample consists of nonfarm males aged twenty-five to sixty-four with income in 1960.[23]

An examination of the data for the United States and Canada in Tables 2 and 3 reveals that the predicted skewness is large, but it is smaller than the observed skewness and the residual skewness.[24] Thus,

[23] The original data and the calculation of the adjusted rates of return for the United States and Canada are discussed in greater detail in Chiswick, *Human Capital . . .*, chapter 3.

[24] The small size of the observed and residual skewness in the states compared to the regions of the United States seems surprising. The measure of skewness defined in equation 11 is sensitive to the average income of the upper open-end interval. This was estimated from the Pareto equation and its value is sensitive to the grouping of high incomes. Since the point representing the upper open-end interval is considerably above the fitted regression line, the sensitivity of the mean to the value of this interval is greatest for the income and residual skewness and least for the predicted skewness.

Comparisons between the two major regions of the United States, among the states, and among the Canadian provinces can be made because the same grouping is used within each type of area or level of aggregation. It is not valid, however, to compare the states to the regions or provinces unless the same grouping is used. Thus, since the lower bound of the upper open-end interval is $25,000 for the United States, the North and the South, and $10,000 for the states, a comparison between the states and regions would be misleading unless the data were reorganized.

by itself, the distribution of schooling produces a considerable amount of skewness in the distribution of income.

Table 2 indicates that the total South and white South have higher values for the four measures of skewness and for the rate of return than do the total non-South and white non-South. The non-South and the United States results are nearly the same. The removal of nonwhites reduces the magnitude of all of the parameters. The most significant changes are in the skewness of schooling and the skewness of predicted income.

TABLE 4

Correlation Matrix for Skewness in the Fifty-one States for Income of Males Aged Twenty-Five and Over[a]

	Skewness				Adjusted Rate of Return
	Observed Income (1)	Predicted Income (2)	Residual Income (3)	Schooling (4)	(5)
1. Predicted	0.42				
2. Residual	0.86	0.41			
3. Schooling	0.47	0.82	0.41		
4. Adjusted rate of return	0.47	0.77	0.64	0.62	
5. Average schooling	−0.66	−0.58	−0.61	−0.61	−0.66

PROBABILITY	*R*
.05	.23
.025	.27
.01	.32

[a]Probabilities based on fifty degrees of freedom. See notes to Table 2.

SOURCE: *United States Census of Population: 1960*, Vol. 1, *Characteristics of the Population*, Parts 2-52 (Washington: Bureau of the Census), Table 138.

An examination of Tables 3 through 6 indicates that the states with higher rates of return and larger schooling skewness tend to have larger predicted skewness and observed skewness. The predicted skewness and residual skewness are positively correlated among all the states and within the non-South. Their insignificant negative correlation in the South may be explained by the negative correlation between the schooling skewness and the residual skewness. The observed skewness is positively correlated with the predicted and the residual skewness.

The interstate correlations are weaker when separate analyses are

TABLE 5

Correlation Matrix for Skewness in the Thirty-Four Nonsouthern States for Income of Males Aged Twenty-Five and Over[a]

	Skewness				
	Observed Income (1)	Predicted Income (2)	Residual Income (3)	Schooling (4)	Adjusted Rate of Return (5)
1. Predicted	0.30				
2. Residual	0.76	0.36			
3. Schooling	0.43	0.80	0.45		
4. Adjusted rate of return	0.08	0.70	0.50	0.61	
5. Average schooling	−0.53	−0.27	−0.43	−0.38	−0.06

PROBABILITY	R
.05	.30
.025	.35
.01	.41

[a]Probabilities based on thirty degrees of freedom. See notes to Table 2.

SOURCE: *United States Census of Population: 1960*, Vol. 1, *Characteristics of the Population*, Parts 2-52 (Washington: Bureau of the Census), Table 138.

performed for the South and the non-South. This is due to the significant differences between the South and non-South.

Similar relationships are found when nonwhites are excluded from seventeen states. Tables 7 through 9 present correlation matrices for all, non-Southern and Southern states after adjustments for nonwhites. The skewness of observed, predicted, and residual income and the adjusted rate of return are positively correlated with each other. The skewness in schooling is positively correlated with these parameters in the country for whites and in the white non-South, but negatively correlated in the white South.

TABLE 6

Correlation Matrix for Skewness in the Seventeen Southern States for Income of Males Aged Twenty-Five and Over[a]

	Skewness				Adjusted Rate of Return (5)
	Observed Income (1)	Predicted Income (2)	Residual Income (3)	Schooling (4)	
1. Predicted	0.20				
2. Residual	0.78	-0.07			
3. Schooling	0.22	0.78	-0.13		
4. Adjusted rate of return	0.30	0.87	0.19	0.45	
5. Average schooling	-0.37	-0.85	-0.09	-0.76	-0.74

PROBABILITY	*R*
.05	.39
.025	.46
.01	.53

[a]Probabilities are based on fifteen degrees of freedom. See notes to Table 2.

SOURCE: *United States Census of Population: 1960*, Vol. 1, *Characteristics of the Population*, Parts 2-52 (Washington: Bureau of the Census), Table 138.

Table 10 contains a correlation matrix for the Canadian provinces. The adjusted rate of return is positively related to the observed and the predicted skewness. The predicted skewness is positively related to the observed skewness and the skewness of schooling. The residual skewness is highly positively correlated with the observed skewness and weakly negatively correlated with the skewness of schooling. This may explain the insignificant negative correlation between the skewness of schooling and of observed income.

The correlation matrices indicate that across the states and the prov-

TABLE 7

Correlation Matrix of Skewness Parameters for the Fifty-One States of Which Seventeen Are for Whites, for Income of Males Aged Twenty-Five and Over[a]

	Skewness				Adjusted Rate of Return (5)
	Observed Income (1)	Predicted Income (2)	Residual Income (3)	Schooling (4)	
1. Predicted	0.50				
2. Residual	0.83	0.38			
3. Schooling	0.39	0.34	0.49		
4. Adjusted rate of return	0.49	0.72	0.68	0.41	
5. Average schooling	−0.40	−0.59	−0.32	−0.12	−0.37

PROBABILITY	R
.05	.23
.025	.27
.01	.32

[a]Probabilities are based on 50 degrees of freedom. See notes to Table 2.

SOURCE: *United States Census of Population: 1960*, Vol. 1, *Characteristics of the Population*, Parts 2-52 (Washington: Bureau of the Census), Table 138.

inces, the skewness of observed, predicted, and residual income tend to be positively correlated. In addition, they tend to be positively correlated with the regression estimate of the adjusted rate of return and the skewness of schooling, when the latter are positively correlated with each other. When the rate of return and schooling skewness are not positively correlated, it is the effect of the rate of return which dominates.

The correlation matrix tables indicate that the average level of schooling tends to be negatively related to the predicted and the observed

TABLE 8

Correlation Matrix of Skewness Parameters for the Thirty-Four Non-Southern States of Which Three Are for Whites, for Income of Males Aged Twenty-Five and Over[a]

	Skewness				Adjusted Rate of Return
	Observed Income (1)	Predicted Income (2)	Residual Income (3)	Schooling (4)	(5)
1. Predicted	0.50				
2. Residual	0.83	0.44			
3. Schooling	0.30	0.45	0.34		
4. Adjusted rate of return	0.30	0.76	0.51	0.32	
5. Average schooling	-0.42	-0.76	-0.39	-0.02	-0.54

PROBABILITY	*R*
.05	.30
.025	.35
.01	.41

[a]Probabilities are based on thirty degrees of freedom. See notes to Table 2.

SOURCE: *United States Census of Population: 1960*, Vol. 1, *Characteristics of the Population*, Parts 2-52 (Washington: Bureau of the Census), Table 138.

skewness. This is contrary to the model's prediction, *ceteris paribus*, that the level of schooling is positively related to the skewness of income. The analysis of Part I and Tables 4–10 indicate that the level of schooling tends to be negatively correlated with the estimate of the average adjusted rate of return and the skewness of schooling. The data in Table 11 are intended to test the hypothesis that for the regions of the United States and Canada, the observed simple negative correlation between the level of schooling and the skewness of income is due to the effects of

TABLE 9

Correlation Matrix of Skewness Parameters for the Seventeen Southern States of Which Fourteen Are for Whites, for Income of Males Aged Twenty-Five and Over[a]

	Skewness				Adjusted Rate of Return
	Observed Income (1)	Predicted Income (2)	Residual Income (3)	Schooling (4)	(5)
1. Predicted	0.38				
2. Residual	0.72	0.07			
3. Schooling	-0.19	-0.30	-0.19		
4. Adjusted rate of return	0.41	0.71	0.55	-0.34	
5. Average schooling	-0.37	-0.43	-0.19	-0.11	-0.20

PROBABILITY	*R*
.05	.39
.025	.46
.01	.53

[a]Probabilities are based on fifteen degrees of freedom. See notes to Table 2.

SOURCE: *United States Census of Population: 1960*, Vol. 1, *Characteristics of the Population*, Parts 2-52 (Washington: Bureau of the Census), Table 138.

TABLE 10

Correlation Matrix for Skewness in the Provinces for Income of Nonfarm Canadian Males Aged Twenty-Five to 64[a]

	Skewness				
	Observed Income (1)	Predicted Income (2)	Residual Income (3)	Schooling (4)	Adjusted Rate of Return (5)
1. Predicted	0.49				
2. Residual	0.96	0.44			
3. Schooling	-0.07	0.60	-0.12		
4. Adjusted rate of return	0.66	0.75	0.64	0.28	
5. Average schooling	-0.45	-0.67	-0.35	-0.23	-0.58
		PROBABILITY		*R*	
		.05		.52	
		.025		.60	
		.01		.69	

[a]Probabilities based on nine degrees of freedom. See notes to Table 2.

SOURCE: *Census of Canada: 1961* (Ottawa: Dominion Bureau of Statistics), Table A.11 for the provinces, unpublished.

Notes to Table 11

[a]Row 1, partial slope or correlation coefficient; row 2, Student's *t* ratio (in parenthesis). The levels of significance of the coefficients are in row 3. *NS* = not significant at 5.0 per cent level. Degrees of freedom equal the number of observations minus two (for the simple correlation coefficient) or minus four (for the partial regression coefficients). Significance levels from H. M. Walker and J. Lev, *Statistical Inference* (New York, 1953), p. 465 and p. 470.

SOURCES: *United States Census of Population: 1960, Vol. 1, Characteristics of the Population,* Parts 2-52 (Washington: Bureau of the Census), Table 138; *Census of Canada: 1961* (Ottawa: Dominion Bureau of Statistics), Table A.11 for the provinces, unpublished.

TABLE 11

Average Schooling and the Skewness of Observed and Predicted Income in the United States and Canada[a]

Region– No. of Obser-vations	Row No.	I Skew–Obs. Inc. Partial Slope Coeff. (t values) Significant at: (per cent) $\bar{S}$	$\hat{r}$	Skew (S)	II Skew–Pred. Inc. Partial Slope Coeff. (t values) Significant at: (per cent) $\bar{S}$	$\hat{r}$	Skew (S)	III Skew.–Obs. Inc. Simple Cor. Coeff. Significant at: (per cent) $\bar{S}$	III Skew.–Pred. Inc. $\bar{S}$	IV Change in Coefficient of S: Less Neg +, More Neg –, No Change 0 Skew.–Obs. Inc.	Skew.–Pred. Inc.
U.S., Total– 51	1	-0.065	0.093	0.020	0.009	3.388	0.126	-0.661	-0.583	0	+
	2	(-3.822)	(0.097)	(0.763)	(0.700)	(4.763)	(6.320)				
	3	0.05	NS	NS	NS	0.05	0.05	0.05	0.05		
U.S., white 51	1	-0.029	1.643	0.034	-0.055	4.145	0.013	-0.403	-0.588	+	0
	2	(-2.092)	(2.153)	(1.851)	(-4.014)	(5.491)	(0.691)				
	3	2.50	2.50	5.00	0.05	0.05	NS	0.50	0.05		
Nonsouth total– 34	1	-0.063	-1.205	0.061	-0.005	3.272	0.134	-0.527	-0.269	+	0
	2	(-2.368)	(-0.952)	(1.897)	(-0.206)	(2.620)	(4.244)				
	3	2.50	NS	5.00	NS	1.00	0.05	0.50	NS		
Nonsouth, white– 34	1	-0.063	-0.269	0.040	-0.102	3.062	0.057	-0.445	-0.764	+	0
	2	(-2.423)	(-0.203)	(1.768)	(-6.257)	(3.690)	(3.980)				
	3	2.50	NS	5.00	0.05	0.05	0.05	1.0	0.05		
South, total– 17	1	-0.046	0.070	-0.026	+0.000	4.602	0.096	-0.373	-0.852	0	+
	2	(-0.898)	(0.030)	(-0.390)	(0.016)	(6.499)	(4.771)				
	3	NS	NS	NS	NS	0.05	0.05	NS	0.05		
South, white – 17	1	-0.021	1.459	-0.030	-0.039	5.087	-0.059	-0.372	-0.434	0	0
	2	(-1.316)	(1.162)	(-0.482)	(-1.821)	(3.091)	(-0.727)				
	3	NS	NS	NS	5.00	0.50	NS	NS	5.00		
Canada– 11	1	-0.022	7.035	-0.155	-0.036	3.060	0.138	-0.449	-0.672	0	+
	2	(-0.398)	((2.032)	(-1.039)	(-1.486)	(2.053)	(2.161)				
	3	NS	5.00	NS	NS	5.00	5.00	NS	2.50		

other variables, namely, the skewness of schooling and the rate of return from schooling.

Table 11 contains the results of a multiple regression of the skewness of observed income (column I) and predicted income (column II) on the level of schooling, the rate of return, and the skewness of schooling. Column III contains the simple correlations of the level of schooling with observed skewness and predicted skewness. Column IV is a summary which indicates whether holding the skewness of schooling and the rate of return constant decreases the negative (sign +), increases the negative (sign −) or does not change the significance (sign 0) of the correlation of level of schooling with observed and predicted skewness.

The correlation never becomes more significant in a negative direction, but frequently becomes less negative. In only two instances (among the states of the United States and the South for predicted income), the negative simple correlation becomes a positive partial correlation, but both positive correlations are insignificant. Thus, it appears that when the rate of return and skewness of schooling are held constant, the magnitude of the negative correlation between the level of schooling and income skewness is reduced, but there is no significant change in sign. Note, however, that for the total and white non-South the positive simple correlations between the rate of return and the skewness of observed income become insignificant negative partial slope coefficients. It is not clear why the level of schooling does not behave as expected while the rate of return generally follows the predicted pattern.

Testing the hypotheses developed from the theoretical analysis reveals that:

1. Schooling alone can produce a considerable amount of skewness in the distribution of income.

2. The observed income skewness and that predicted by schooling are positively correlated. When the rate of return and the skewness of schooling are positively correlated, each tends to be positively correlated with the predicted and the observed skewness. When they are not positively correlated, it is the rate of return which dominates. Indeed, the rate of return appears to be more important than the skewness of schooling in explaining income skewness.

3. The residual skewness tends to be positively correlated with the observed and the predicted skewness.

4. The average level of schooling tends to be negatively related to the rate of return from schooling and to the schooling skewness. Thus, a priori, it is not clear how changes in the distribution of schooling due to a higher level of income affect the skewness of income. Empirically, the significance of the simple negative correlation between the level of schooling and the skewness of observed and predicted income is reduced when the rate of return and the schooling skewness are held constant.

III. SUMMARY AND CONCLUSIONS

The a priori analysis of Part I and the empirical analysis of Part II indicate that income skewness can be related to rate of return and schooling parameters, and that these parameters are important for explaining interregional differences in income skewness. In particular, it has been demonstrated that the distribution of schooling by itself tends to produce a considerable positive skewness in the distribution of income. This suggests that income would have a considerable positive skewness even if human capital were the only source of income.

The model predicts that the skewness of income in a region has a positive partial correlation with the level of the rate of return from schooling, the skewness of schooling, the level of schooling, the skewness of predicted income, and the skewness of the residual income. In general, the predicted signs were found for simple correlations for all of the explanatory variables, except the level of schooling. It was argued, and demonstrated empirically, that the level of schooling tends to be negatively correlated with the rate of return and the skewness of schooling across the regions of a country. The partial correlation between income skewness and the level of schooling, when the rate of return and the skewness of schooling were held constant, was less negative than the simple correlation.

Since the rate of return from schooling tends to be higher in the poorer regions of a country, the analysis predicts, *ceteris paribus*, that poorer regions have a larger income skewness. The theory also suggests that the lower level of schooling in a poorer region should produce, *ceteris paribus*, a lower income skewness. Lower levels of schooling, however, tend to be related to larger schooling skewness, and therefore, larger income skewness. A priori, the net effect of these factors is ambiguous.

Empirically, income skewness is larger in poorer regions of the United States and Canada.

No international empirical analyses were performed. Economic theory does not suggest any clear relationship across countries between the level of income and the rate of return. In addition, the effects of a lower level of schooling and a larger schooling skewness in poorer countries tend to offset each other. Therefore, no prediction is offered as to whether, across countries, income skewness is related to the average level of income. If, however, rapid rates of economic growth are associated with high rates of return from investments, rapid economic development may generate a larger income skewness.

COMMENTS

MARY JEAN BOWMAN
UNIVERSITY OF CHICAGO

The remarks that follow will center around three topics: (1) meanings and measurements of skewness and of "inequality" in the analysis of the "shape" of an income distribution; (2) Chiswick's empirical analysis and statistical interpretations; (3) his attempt to integrate his empirical analysis with a Becker-style human-investment decision theory, as the theoretical starting point of his work.

CONCERNING THE "SHAPES" OF INCOME DISTRIBUTIONS AND THEIR MEASUREMENT

In his opening paragraph, Chiswick argues the importance of studying determinants of the skewness of a distribution, on two main grounds: the relevance of "shape" (asymmetry) for discussions of the equity of the distribution, and its relevance for "Engel curve and savings-and-investment analyses." I fully agree with his emphasis on the importance of asymmetry in the distribution of incomes. Furthermore, I have no objection to his particular choice of a measure of skewness. But this is only one among many possibilities, and I am disturbed by the lack of any mention of the fact that such measures can be various, and that they do not necessarily give identical rankings because skewness (however measured) may be raised, or lowered, by quite diverse kinds of changes in the shapes of the various segments of the distribution. The omission of any comment at all on this point is the more disturbing in view of the fact that Chiswick asserts that while skewness was considered important in the past, "it has been ignored by most recent studies." He mentions a few exceptions, but he sweeps aside, as though they did not exist, recent preoccupations with poverty and the tailing out of incomes to the bottom, the increasing use of the Gibrat coefficient or of variance in the logarithms as an inequality measure, the increasing tendency to compare results of alternative measures of "inequality" that incorporate nonsymmetrical weighting systems, and the renewed concern with "welfare" concepts in the selection of measures to describe the forms of income distributions generally.[1] It is true enough that interest in the Pareto measure of "inequality" has declined, and that Pareto's index was in fact a measure of an important component of skewness—the stretching out of the upper tail. And it is equally true that the use of third-moment measures of skew, though they pile out of the computers along with other univariate statistics, have received comparatively little attention as interesting end products in themselves. But this does not signal a lack of concern in recent literature with the asymmetries of income distributions.

Basically, Chiswick's lack of attention to just what his measure of skewness means seems to go back to neglect of three interconnected facts. First, "shape" is clearly a matter of an entire distribution, and no single parameter can give us an adequate picture of "shape." This does not mean that single parameters are not useful—and, indeed, indispensable. But it does mean that we must (and for simple measures, as of central tendency, we do) interpret them in the light of underlying characteristics of the distribution they are summarizing. Second, the only inequality measures that are free of skewness components are those that weight

[1] For the most recent such contribution (to my knowledge) see D. J. Aigner and A. J. Heins "A Social Welfare View of the Measurement of Income Equality," *Review of Income and Wealth*, Series 13, No. 1, March 1967, pp. 12–25.

deviation of an income from the mean (or of any pair of incomes from each other) by their arithmetic differences. Even the standard deviation is in fact a skew-biased measure of dispersion when the underlying distribution is not symmetrical—which is of course one of the reasons we so often switch to the use of logarithms of incomes (tacitly assuming a log-normal income distribution, or a good approximation thereto). And third, to repeat, skewness, by any measure that attempts to describe the asymmetry of an entire distribution, may rise or fall as a result of quite different sorts of changes in component parts of the distribution. Thus, for example, the skewness of the income distribution in the United States may decline either because there is a lesser tailing out at the top or because there is more tailing out at the bottom (as modal incomes rise leaving the bottom still very low); conversely, skewness could become greater either because the top incomes are pulling further away from the mode or because the lowest incomes are rising toward it. A recognition of these facts might have helped Chiswick interpret some of his results, even if he did not choose to go on to look at "shape" using other indicators of asymmetry in his statistical analysis.

THE EMPIRICAL ANALYSIS AND FINDINGS

Although the heart of Chiswick's paper is in his effort to build a rate-of-return human-investment decision model into the analysis and interpretation of income skewness, it will be easier to see just what he has done if we begin by stripping away all the theoretical paraphernalia to ask what his statistical operations are and what they tell us. So stripped, Chiswick's mean "adjusted rate of return," $\bar{r}^*$, becomes simply the slope coefficient $\hat{r}$ in regressions of the natural log of earnings (or incomes) on schooling. Starting from that point, Chiswick derives his predicted skewness of income from the skewness of the schooling term (taking antilogs) of $(1 + \hat{r})^{S_i}$. As he shows, his equation states that *ceteris paribus* income will be more skewed (1) the higher the slope coefficient $\hat{r}$, (2) the higher the mean level of schooling $\bar{S}$, and (3) the more skewed the distribution of schooling. Using various sets of states and the Canadian provinces in a series of regressions, he then presents correlation matrices including these three attributes along with the skewness of observed, predicted, and residual incomes. He gets his expected positive zero-order correlation with observed income for $\hat{r}$ in all cases and for skewness in schooling in all except the subsets of the seventeen Southern states (fourteen white) and the Canadian provinces. However, correlations of observed income skew with average schooling level, $\bar{S}$, emerge very substantially negative, exceeding those with $\hat{r}$ in four of the seven zero-order matrices, and never coming closer to zero than −.30, though in one case the correlation between $\hat{r}$ and observed income skew was only .08. In multiple regressions, taking observed skewness as the dependent variable, the negative coefficients of $\bar{S}$ are reduced, but remain significant at 2.5 per cent or better in all sets except those for the Southern states only and for the Canadian provinces. On the other hand, in these regressions, the partial coefficients for $\hat{r}$ were nonsignificant except for the set that used all fifty-one states but whites only (significant at 2.5 per cent) and that for the Canadian provinces (significant at 5 per cent). When the "predicted skewness" is the dependent variable in multiple regressions, the relative effectiveness of $\bar{S}$ and $\hat{r}$ as explanatory variables is reversed, with $\hat{r}$ coming through much more strongly, but this is clearly not a test of his model.

In revising his paper Chiswick has become much more cautious in his statements concerning the "dominant importance" of $\hat{r}$, and is less inclined to dispose of the question of the "wrong" (negative) signs on $\bar{S}$, with the observation that they are less negative in a multiple than in a zero-order correlation. Furthermore, he is quite explicit about the importance of the negative zero-order correlations between levels of schooling on the one hand, skewness of the schooling distribution and the value of $\hat{r}$ on the other—and the multicollinearity problems involved. Nevertheless, he still stresses the good performance of $\hat{r}$ and seems to regard that of $\bar{S}$ as puzzling. I wonder if he would still be puzzled if he stepped outside of the model developed in this paper to use his insights (and he has such insights) to try another, complementary approach. A simple preliminary example might be the use of path coefficient analysis or a two-stage regression in which $\hat{r}$ would be treated in the first instance as a function of $\bar{S}$. I am not suggesting that Chiswick should have done yet another analysis before presenting the results of this one. What I do suggest is that his findings thus far indicate that if study of determinants of skewness is worth while, then it would also be worth while to try some other, complementary approaches. This is the more true in view of the fact that there may be very little resemblance empirically between $\hat{r}$ ($= \bar{r}^*$) and an internal rate of return—the base on which his theoretical construct is built.

THE RELATIONSHIP BETWEEN $\hat{r}$ AND AN INTERNAL RATE OF RETURN TO SCHOOLING

There are several reasons why $\hat{r}$ and the internal rate of return are different things: (a) Working life is not infinite, as Chiswick in effect assumes. We may dismiss this as unimportant, however, in view of the fact that he is concerned with schooling only. (b) Cross-section income data are by no means the same thing as the income prospects of an individual through time. But let us set that also aside, accepting the cross-section data as an approximation to expected income streams of the future. (c) Much more important is the simplifying assumption that this is a world in which all that men learn they learn in school, that they forget none of it, that none of it becomes obsolete, and so on. I shall come back to this, but for the moment let us accept it too. Even under such circumstances, the "adjusted rate of return" is an odd animal. Let us take a look at it. Under these circumstances, any given incremental investment by an individual in the jth year of schooling yields a permanent incremental income stream. Omitting subscripts specifying an individual, let us designate the incremental investment as I_j and the associated incremental lifetime income (rental value) stream as W_j. Under the specified conditions, the internal rate of return will equal the mean rental value ratio to costs, which is of course W_j/I_j. If we take Y_j as the potential earnings if the individual were to go directly into the labor market, and assume that k_j is an adjustment factor (as in Chiswick) such that $I_j = k_j\, Y_j$, we can write $r_j = W_j/k_j\, Y_j$. Now Chiswick's "adjusted rate of return" is $r_j^* = r_j k_j = W_j/Y_j$. But note that by this time we have eliminated any assessment of costs from the denominator. What we have, of course, is simply the ratio of the income increment associated with the jth year of schooling to mean lifetime income were schooling to stop at the lower schooling level. To call this a "rate of return," adjusted or not, is to contribute, in my judgment, to the deterioration of the language of economics—though I must admit that Chiswick has respectable company in this usage. Semantics aside, the important question is, of course, the magnitude of k, and its stability or variability among the various subpopulations to which the analysis is applied.

(d) Returning to the real world, from which Chiswick's observations were in fact obtained, it is clearly not so that all learning is in school (or at least before leaving school), and that there is no subsequent obsolescence of skills or learning of them—with or without cost. The magnitude of this effect is substantial. I have shown elsewhere[2] that for U. S. data the internal rates of return as conventionally computed from cross-section data equal about half or less of the ratio of an average individual's mean incremental annual income to investments in the associated increment of schooling. Chiswick's finding (in his dissertation) that his $\hat{r}$ values are not much changed when computed within age categories, excepting for a reduction of the estimate in the youngest age category, does not answer this challenge.

(e) Since Chiswick (correctly) uses observed incomes, without discounting, we might expect a priori that his estimates of rates of return would exceed the internal rates of return computed in conventional human-investment decision models. Yet, in fact, he obtained lower values for $\hat{r}$ than those found in internal rate-of-return comparisons. This reflects the fact that this foregone income or cost proxy is the mean income of men *of all ages* who are in the next lower level of schooling or, when he does the analysis by age groups, the mean incomes of men in the designated age categories. Those incomes incorporate returns to postschool learning. Although, in revising his paper, Chiswick has agreed that this is likely to cause a cost overestimation, and hence downward bias in his "rate-of-return" estimates, it seems to me that he still fails to appreciate the broader implications. Those implications are critical in any interpretation of his $\hat{r}$ as though it were a legitimate measure of $\bar{r}^*$, and especially in an analysis that attempts to explain and interpret variations in the skewness of income distributions.

[2] *The Assessment of Human Investments as Growth Strategy*, Joint Economic Committee, 90th Congress, 2nd Session, 1968, pp. 84–99.

DUNCAN FOLEY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

This paper sets out to study the relation between the positive skewness of distributions of income and the distribution of schooling which is presumably one of several variables that affect income. In studying this problem, the author leans very heavily on the assumption that the relation between schooling and income takes a particular form: that the logarithm of income depends linearly for an individual on years of schooling. Given this functional form, much of the paper follows analytically. As the author points out, this log-linear specification transforms a zero or small negative skewness in the distribution of schooling into a positive skewness in the distribution of income.

Since this assumption is of such fundamental importance, it is somewhat disturbing that it is not tested in any very direct way. The author takes the "rate of return" formulation for this function without question. Rates of return are very convenient ways to summarize the performance of any investment *ex post*, but it is not self-evident, at least to me, that the only possible functional relation between income and years of schooling is log-linear. One curious consequence, for example, of these specifications is that the marginal product of schooling increases with years of schooling.

Other functional forms might also be fitted to the data to see if there is any good empirical basis for preferring the log-linear hypothesis. The consequences of other functional forms on skewness would also be an interesting problem in this connection.

The author indicates at one point that educational policy might be employed to alter the skewness of the distribution of income. He bases this comment on his finding that the distribution of schooling affects the distribution of income in a particular way. But we see from his original formulation that the distribution of schooling affects the distribution of income through the distribution of the rate of return to schooling. This fact is obscured in the empirical part of the paper when variations in rates of return are lumped into the residual because of lack of data.

Given the author's earlier model, policy recommendations concerning the relation between schooling and income distributions depend crucially on the distribution of rates of return, that is, on the distribution of part of his residual. These recommendations would be clearer and more strongly founded if they were based explicitly on some statistically testable property of the residual. The paper as it stands contains no detailed study of the residual in the regression to back up the policy argument.

What is a good strategy for achieving a coherent economic explanation of the distribution of income? It seems to me that this paper offers substantial food for methodological thought. Certainly the forces working on the distribution of income are numerous and their interaction complex. The author mentions on-the-job training and the distribution of wealth in addition to schooling, and there are many other phenomena of importance, such as discrimination, the distribution of natural talent and temperament, and so on, affecting the distribution of income. Is it possible to deduce a great deal about the sources of skewness in the income distribution without a fairly detailed and well articulated model which takes several of these factors into account?

The author chooses to test his model only very indirectly, by looking at the relation between the skewness parameters he predicts on the basis of the distribution of schooling and observed skewness. If the model he has set up and begun to estimate correctly reflects an important relation, it should predict many features about the distribution of income besides skewness. To approach the problem in this way would involve specifying the residual distribution more carefully and testing hypotheses about the residual itself, in addition to looking at somewhat indirect consequences of the model, such as predicted skewness.