Volume Title: Income Inequality: Regional Analyses within a Human Capital Framework

Volume Author/Editor: Barry R. Chiswick

Volume Publisher: NBER

Volume ISBN: 0-870-14264-X

Volume URL: http://www.nber.org/books/chis74-1

Publication Date: 1974

Chapter Title: The Expanded Human Capital Model

Chapter Author: Barry R. Chiswick

Chapter URL: http://www.nber.org/chapters/c3674

Chapter pages in book: (p. 105 - 118)

PART C

Income as a Function of Schooling and Market Experience

6

The Expanded Human Capital Model

This chapter presents a model in which an individual's income can be related to his years of schooling, years of labor market experience, and level of employment (weeks worked) during the year. It is a brief exposition of what has come to be called the expanded "human capital earnings function."[1] The mean and variance of the function are computed and used empirically in Chapters 7 and 8 to explain interstate differences in the income distribution of the United States and Canada.

INVESTMENT IN TRAINING

If earnings were due only to investments in training and employment lasted the full year, there would be a simple relationship between earnings and training parameters. Suppose Y_0 is a full year's earnings of an individual with zero training, and he invests (direct and opportunity costs) $100k$ per cent of his potential income in year 1. If he does not undertake additional investments, and there is no depreciation of his stock of capital, his earning in year 2 and in all subsequent periods will be

$$Y_1' = Y_0 + r_1(k_1 Y_0) = Y_0(1 + r_1 k_1), \tag{6-1}$$

where r_1 is his average rate of return on the investment and $k_1 Y_0$

1. The expanded human capital earnings function was developed by Jacob Mincer in his *Schooling, Experience, and Earnings*, Part 1, NBER, 1974.

is the dollar value of the investment. If he invests in N periods of training, his earnings after training will be shown by the identity

$$Y'_{Nj} = Y_0 \prod_{j=1}^{N} (1 + r_j k_j), \tag{6-2}$$

where r_j is the average rate of return on the investment in the jth period and k_j is the fraction of potential earnings in year j that was invested.[2]

The earnings an individual could receive if he did not invest in training after N years is his gross or potential earnings (Y'_N). Net earnings (Y_N) are gross earnings after investment costs are deducted. That is,

$$Y_N = Y'_N - k_{N+1}(Y'_N) = Y'_N(1 - k_{N+1}),$$

where k_{N+1} is the fraction of potential (gross) earnings invested in year $N + 1$. Net earnings are equal to observed (or reported) earnings in a year only if all training costs are opportunity costs. If some training costs involve direct or out of pocket expenditures, gross earnings exceed observed earnings, which exceed net earnings.

During some years of formal schooling direct costs are substantial. During postschool investment in on-the-job training direct costs are negligible to the worker since the direct costs of the training provided by the firm are deducted from the "trainee's" wages.[3] The net earnings of a worker investing $100k_{N+1}$ per cent of his potential earnings are

$$Y_N = Y_0 \prod_{j=1}^{N} (1 + r_j k_j)(1 - k_{N+1}). \tag{6-3}$$

---

2. From the principle of mathematical induction, if a relationship holds for $j = 1$, and if, when it holds for $j = N$ it also holds for $j = N + 1$, then it holds for all values of j. The relation was shown to hold for $j = 1$ in the text.

If $Y'_N = Y_0 \prod_{j=1}^{N} (1 + r_j k_j)$, and an $N + 1$st year of investment is undertaken,

$$Y'_{N+1} = Y'_N + r_{N+1}(k_{N+1} Y'_N) = Y'_N(1 + r_{N+1} k_{N+1}) = Y_0 \prod_{j=1}^{N+1} (1 + r_j k_j).$$

Thus, the expression is valid for all values of j.

3. While firms provide training, they do not finance any general training (training useful in many firms) and finance only part of specific training (training useful only in that specific firm). See Gary Becker, *Human Capital*, 1974, Ch. 2, and Donald Parsons, "Specific Human Capital: An Application to Quit Rates and Layoff Rates," *Journal of Political Economy*, November-December 1972, pp. 1120–1143.

WEEKS WORKED

The model thus far assumes full employment of the worker during the year. Actual "net annual earnings" are lower than "net full-year employment earnings" to the extent that the worker is unemployed or absent from the labor force during part of the year. Employment will be expressed in terms of weeks worked, since this is the form used for employment data in the empirical analysis. If weekly wages were independent of the number of weeks worked, we could write

$$Y_N = Y_0 \left[\prod_{j=1}^{N} (1 + r_j k_j)\,(1 - k_{N+1}) \right] (WW) \tag{6-4}$$

where Y_N is now annual net earnings, Y_0 is the worker's earnings if he had not invested, the terms in square brackets represent the contribution of past and current human capital investments to his current net earnings, and (WW) is the fraction of weeks in the year $N + 1$ in which he works.[4]

Within the human capital framework, there are two reasons for believing that weekly wages and weeks worked are not independent of each other. First, those who have greater investments in training specific to a firm have higher wage rates and lower quit and layoff rates, and consequently work more weeks per year.[5] Second, those who have higher weekly wages because of greater investment in training (regardless of whether it is general or specific) face a higher opportunity cost of time and will work more during the year if the substitution effect outweighs the adverse effect on work effort produced by the increase in wealth due to the investment.[6] These points are part and parcel of the human capital analysis of earnings.

---

4. It is assumed that investments in training are proportional to time spent working.

5. See Becker, *Human Capital*, Part 1, and Parsons, "Specific Human Capital" for the relation between specific training and quit and layoff rates. Workers with lower levels of skill have more frequent spells of unemployment, and a slightly longer duration per spell. See R. Morganstern and N. S. Barrett, "Occupational Discrimination and Changing Labor Force Participation: Their Effects on Unemployment Rates of Blacks and Women," 1971, Mimeo.

6. Investments in human capital increase wealth only to the extent that the internal rate of return (r) on the investment exceeds the opportunity cost of the funds invested (R). For an investment of C dollars with a constant annual return and an infinite life the increase in wealth is $\frac{C\,(r - R)}{R}$.

Outside of the human capital framework there are additional reasons for a connection between weeks worked during the year and weekly earnings. First, if the higher weekly wage for the weeks actually worked is due to cyclical or seasonal sensitivity of employment in an industry, there will be a negative correlation between weeks worked and weekly wages, ceteris paribus.[7] Second, those whose weekly wages are higher for reasons other than investment in human capital or seasonal fluctuations (for example, due to ability, luck, or discrimination) will wish to work more weeks if their supply curve is upward rising. That is, if the price effect of the higher wage exceeds the wealth (income) effect, they will supply more labor to the market.

Finally, it has been observed that there is a positive correlation between weeks worked per year and hours worked per week.[8] Since those who work more hours per week have higher weekly wages, a positive correlation between hours per week and weeks per year, holding investment in human capital constant, results in a positive correlation of weekly wages with weeks worked. Note, however, that this is not a statistical artifact—this correlation is related to human capital analysis and labor supply theory. Those human capital and labor supply forces that encourage a greater labor supply in terms of weeks also encourage a greater labor supply in terms of hours.

Let us define γ as the elasticity of annual earnings with respect to the fraction of weeks worked per year. Human capital theory, upward rising labor supply curves, and the positive correlation of hours worked per week with weeks worked per year all predict that those with higher weekly wages work more weeks per year. This implies a γ greater than unity. The backward bending labor supply curve and cyclical or seasonal sensitivity of employment predict that those with higher weekly wages have a larger annual income but work fewer weeks per year. They predict a γ greater than zero but less than unity.[9] Empirically, Mincer found $\gamma = 1.17$, which was significantly greater than unity, for white nonfarm, nonstudent

7. The third of Adam Smith's five points which "make up for a small pecuniary gain in some employments, and counterbalance a great one in others" is "the constancy or inconstancy of employment in them." (*Wealth of Nations*, Modern Library Edition, 1937, p. 100.) The effect of seasonality of employment on weekly wages is analyzed in Chapter 2 of this study.

8. Victor Fuchs, *Differentials in Hourly Earnings by Region and City Size, 1959*, NBER, Occasional Paper 101, p. 4.

9. Define Y as observed annual earnings, Y' as full year employment earnings, and W as the number of weeks worked. If those with higher weekly wages work more weeks during the year, $Y'/52 > Y/W$, where $W < 52$. Since

males with earnings from the 1/1,000 sample of the 1960 Census of Population.[10]

Thus, the human capital earnings function is written as:

$$Y_N = Y_0 \left[\prod_{j=1}^{N} (1 + r_j k_j)(1 - k_{N+1}) \right] (WW)^{\gamma}, \qquad (6\text{-}5)$$

where γ is the elasticity of annual earnings with respect to the fraction of weeks worked. Weeks worked and γ are not simply standardizing variables, but, rather, integral parts of the human capital model and labor supply theory.

EMPIRICAL FORMULATION

The human capital earnings function in equation (6-5) relates annual net earnings (Y_N) to the number of periods of investment (N), the fraction of potential earnings which was invested in the past (k_j, $j = 1, \ldots, N$), the fraction of potential earnings invested in the current year (k_{N+1}), the fraction of weeks worked during the year (WW), and the relation between weekly wages and weeks worked (γ). This functional form, however, is not desirable for several reasons. First, with the exception of N and WW, these variables are not readily measurable. Second, linear models are easier to estimate empirically than nonlinear models. Third, if the variance of both sides of the equation is computed, the variance of income is a function of the variance of a product of terms. However, interest in income inequality is greater for relative than for absolute inequality. Relative inequality is devoid of units and therefore

---

$Y = Y'(ww)^{\gamma}$, where $ww = \frac{W}{52}$, $\frac{Y'}{52} > \frac{Y'(W/52)^{\gamma}}{W}$ or $(52)^{\gamma-1} > (W)^{\gamma-1}$. This holds if $\gamma > 1$.

If those with higher weekly wages work fewer weeks per year, $Y'/52 < Y/W$ where $W < 52$. Then, $\frac{Y'}{52} < \frac{Y'(W/52)^{\gamma}}{W}$ or $(W)^{\gamma-1} > (52)^{\gamma-1}$. This holds if $\gamma < 1$. However, if annual earnings are higher for those with more weeks of employment, $Y' > Y$ or $Y' > Y'(W/52)^{\gamma}$. For $W < 52$, this is true only if $\gamma > 0$. Thus, if those with higher weekly wages have higher annual incomes but work fewer weeks, $0 < \gamma < 1$.

10. This value—and further references to Mincer's data—are taken from the 1972 mimeographed version of his *Schooling, Experience, and Earnings*. The data in the 1974 version differ only slightly from the earlier ones and do not alter the substance of my arguments.

facilitates intercountry comparisons. Moreover, computing the variance of a product of terms is more difficult than computing a sum of terms.[11] Thus, the algebraic manipulations of equation (6-5) that follow are designed to convert the equation into a functional form which facilitates the theoretical and empirical analysis of income distribution.[12]

Taking the natural logarithm of both sides of equation (6-5), and using the relation that the natural log of one plus a small number is approximately equal to that small number,[13] we obtain (approximately)

$$\ln Y_N = \ln Y_0 + \sum_{j=1}^{N} r_j k_j + \ln (1 - k_{N+1}) + \gamma (\ln WW). \qquad (6\text{-}6)$$

The N years of training can be decomposed into S years of schooling followed by $N - S$ years of postschool training if it is assumed that schooling precedes on-the-job training. The $T = N - S$ years of postschool training can also be approximated as $T = A - S - 5$, where A is age and it is assumed the worker is investing in training in each of the years since the age at which he left school ($S + 5$ years). If we assume that during the years of schooling the opportunity costs and direct costs of schooling are approximately equal to the potential or gross earnings of students,[14] we can write $k_j = 1$ for the schooling years. Let us also assume that the rate of return from schooling is constant for an individual for all levels of schooling.[15] Then we can write

$$\sum_{j=1}^{N} r_j k_j = rS + \sum_{j=S+1}^{N} r_j k_j .$$

11. See Leo Goodman, "On the Exact Variance of a Product," *Journal of the American Statistical Association*, December 1960, pp. 708–713.

12. These manipulations of the postschool training variables were developed in Mincer, *Schooling, Experience, and Earnings*, Part 1, and reported in Chiswick and Mincer, "Time Series Changes in Income Inequality," *Journal of Political Economy, Supplement*, May-June 1973.

13. $\ln (1 + r_j k_j) \approx r_j k_j$. Since r_j is likely to be in the neighborhood of 10 to 20 per cent, and k_j is not likely to be greater than unity, $r_j k_j$ is sufficiently small to make the approximation quite close.

14. Becker, *Human Capital*, Chapter 4; and G. Hanoch, "An Economic Analysis of Earnings and Schooling," *Journal of Human Resources*, Summer 1967, pp. 310–329.

15. There is some empirical support for this assumption. Using data for white nonfarm, nonstudent males with earnings, Mincer found no evidence of a higher (or lower) rate of return for higher levels of schooling when weeks worked and experience were held constant. (*Schooling, Experience, and Earnings*, Part 2.)

To evaluate the postschool training expression $\sum_{j=S+1}^{N} r_j k_j$ we must make some assumption concerning the temporal behavior of k_j, the fraction of gross earnings invested. There are several reasons for believing that k_j declines over time.[16] First, if labor market experience raises the productivity of time devoted to employment more than that devoted to the production of additional training, the opportunity cost of time invested in experience rises with additional experience. This decreases the profitability of additional investment. Second, additional experience reduces the length of the remaining working life and consequently the profitability of investment. Finally, if an investment is profitable, it is more profitable (i.e., has the highest net present value) the earlier it is undertaken. The decline with experience of the fraction of potential earnings that is invested is consistent with the observed concavity of the age-earnings profile.

Thus, it is assumed that k declines monotonically over a lifetime. For simplicity's sake, we assume that k_j declines linearly: $k_j = k_0\ (1 - T/T^*)$, where T is the number of years of postschool training, T^* is the number of years of positive net investment, and k_0 is the fraction of potential income invested in the initial ($T = 0$) year of postschool training. T^* is a large number. Converting to continuous time,

$$\int_0^T r_j k_j \, dT = (r_j k_0)T - \left(\frac{r_j k_0}{2T^*}\right) T^2 , \tag{6-7}$$

and the effect on the log of earnings of past investments in experience is a parabolic function of years of experience (T).

By assuming k_j declines linearly with experience, the term $\ln (1 - k_j)$ can be evaluated as a function of experience (T) by a Taylor expansion. Using a Taylor expansion evaluated around T^* taken to the third term, and ignoring the remainder,

$$\ln (1 - k_{T+1}) = -k_0 \left(1 + \frac{k_0}{2}\right) + \left(\frac{k_0}{T^*}\right) (1 + k_0)\, T + \left[\frac{-k_0^2}{2\,(T^*)^2}\right] T^2 . \tag{6-8}$$

Let us assume that variations across individuals in Y_0, k_0, T^* and γ appear in a residual U_i. Then, using the subscript i to desig-

16. See Yoram Ben-Porath, "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, August 1967, pp. 352-365; and Gary Becker, *Human Capital and Personal Income Distribution*, Ann Arbor, 1967.

nate individuals, combining the several previous steps, and rearranging terms,

$$\ln(Y_i) = \left[\ln Y_0 - k_0 \left(1 + \frac{k_0}{2}\right)\right] + r_i S_i + \left[r_i^* k_0 + \frac{k_0}{T^*}(1 + k_0)\right] T_i - \left[\frac{r_i^* k_0 T^* + k_0^2}{2(T^*)^2}\right] T_i^2 + \gamma (\ln WW_i) + U_i, \tag{6-9}$$

where r_i is the rate of return from schooling and r_i^* is the rate of return from postschool training.

Equation (6-9) expresses the log of income as a linear function of years of schooling, years of experience, years of experience squared, and the log of weeks worked. While data are generally not available for dollar investments in schooling or postschool training, data for investments measured in years exist in abundance. Indeed, much of public policy appears to be framed in terms of years of investment rather than dollar investments.

In this monograph the variable "experience squared" (T^2) is deleted for two reasons. First, in the analysis of income level, the mean levels of experience and experience squared are so highly correlated that multicollinearity becomes a problem. Second, in the analysis of income inequality, retaining experience squared necessitates computing the third and fourth moments of experience, but the additional explanatory power is not likely to be large. However, deleting the term experience squared biases the slope coefficient of experience downward. When experience squared is deleted, the coefficient of experience is designated r_i'. The term r_i' is the slope of the experience log income profile.

The variable experience (T) is replaced by age minus schooling minus five. Public policy can change the distribution of schooling independently of age, but if it does so, the distribution of experience is necessarily altered. There is also more concern with the distribution of earnings by age group than by experience group. Thus, it would be desirable to express earnings as a function of schooling and age (A). Fortunately, this is easy to do if we assume $T = A - S - 5$. The assumption, however, that an individual is continuously acquiring experience in the labor market after leaving school means that this formulation of the model is more relevant to an analysis of the income of males than to that of females, since the latter frequently have long periods of absence from the labor force.[17]

17. For an analysis of the effect of labor market experience on the earnings of women, see Jacob Mincer and Solomon Polachek, "Family Investments in Human Capital: Earnings of Women," *Journal of Political Economy, Supplement*, March 1974, pp. S76–S109.

With these modifications, equation (6-9) becomes

$$\ln Y_i = X + r_i S_i + r_i' (A_i - S_i - 5) + \gamma (\ln WW_i) + U_i, \quad (6\text{-}10)$$

where X is the intercept and U_i is a residual. The residual, U_i, reflects individual differences in earnings for given levels of schooling, rates of return from schooling, age, and employment. It includes the effects of differences in the dollar amounts of post-school training in each year of experience, the nonpecuniary aspects of jobs, nonlabor income (if this is included in the income concept), luck, and errors of measurement, among other variables. We assume that the residual is a random variable.

For the purpose of human capital analysis, equation (6-10) has several desirable features. First, the available data sources permit us to measure investments in human capital in terms of years of schooling and years of labor market experience rather than dollar investments, and equation (6-10) relates income to years of training. Second, since income is more closely approximated by a log normal than a normal distribution, the structure in equation (6-10) will have residuals which are more homoscedastic than the structure in equation (6-5). Finally, there appears to be more interest in the relative than in the absolute level and inequality of income, and equation (6-10) is better suited to an investigation of relative income. Thus, this equation shall serve as the basic human capital earnings function in my analysis of the income level and inequality of adult males.

In the following two chapters the mean and variance of both sides of the human capital earnings function will be computed. Computing the mean of a product of two variables ($r_i S_i$ and $r_i' T_i$) is easy. This is not true, however, of the variance of a product of two random variables, unless these variables are statistically independent.[18] For the analysis of income inequality, the plausible assumptions will be made that r_i is independent of S_i and r_i' is independent of T_i. There is both theoretical and empirical support for these assumptions.[19]

For a nontechnical analysis, see also Council of Economic Advisers, *Economic Report of the President, 1974*, Washington, D.C. 1974, pp. 154-161.

18. Goodman, "On the Exact Variance of a Product," December 1960.

19. The theoretical support relies on the model dealing with supply and demand for funds to be invested in human capital. Individuals with greater "training ability" have, for a given cost of funds, a higher average and marginal rate of return, and thus tend to invest more. Those with lower levels of wealth, holding "training ability" constant, invest more but have a lower average and marginal rate of return. If greater wealth and greater "ability" are positively correlated, the relation between level of investment and marginal and average rates of return is ambiguous. Empirical support is found in the

RACIAL DIFFERENCES IN THE AGE-INCOME PROFILE

Much of the analysis of racial differences in the level of income in Chapter 7 and of income inequality in Chapter 8 will focus on the parameter r' in equation (6-10), the slope of the cross-sectional age-log-of-income profile, when schooling and weeks worked are held constant. A cross-sectional profile represents the income of persons at different ages at a moment in time (for example, the year 1959). This is shown by the curve BB in Figure 6-1. The slope of the profile is the per cent change in income as we look at older persons.

There is abundant evidence from many different data sets that the slope of the cross-sectional age-log-of-income profile is flatter for nonwhite males than for white males.[20] A flatter cross-sectional profile can occur for one or both of the following reasons. First, it may be that nonwhites have a flatter cohort profile. A cohort profile is obtained by observing the income of a group (cohort) of individuals as they age. Three cohort profiles (A_1A_1, A_2A_2, A_3A_3) for three different age groups are shown in Figure 6-1. A nonwhite cohort could have a flatter age-income profile than a white cohort because of the former's smaller investments in postschool training or their lower rates of return from this training. This may occur if nonwhites invest in less postschool training because of a lower level of wealth or in a poorer quality of schooling, if they are discriminated against in training opportunities, or if nonwhites who acquire postschool training are subject to more wage and occupational discrimination than less well-trained nonwhites. Ceteris paribus, a set of flatter cohort profiles would translate into a flatter cross-sectional profile.

Second, nonwhites may have a flatter cross-sectional profile

---

absence of a significant partial effect for schooling squared when the log of earnings is regressed on schooling, schooling squared, experience, and the log of weeks worked. See Becker, *Human Capital and the Personal Distribution of Income;* Becker and Chiswick, "Education and the Distribution of Earnings," *American Economic Review*, May 1966, pp. 358-369; and Mincer, *Schooling, Experience, and Earnings*, Part 2.

20. See, for example, Jacob Mincer, "On-the-Job Training: Costs, Returns, and Implications," *Journal of Political Economy, Supplement*, October 1962, pp. 50-79; Thomas Johnson, "Returns from Investment in Human Capital," *American Economic Review*, September 1970, pp. 546-560; Finis Welch, "Black-White Differences in the Returns to Schooling," *American Economic Review*, December 1973, pp. 893-907; and Council of Economic Advisers, *Economic Report of the President, 1974*, Washington, D.C., 1974, pp. 150-154.

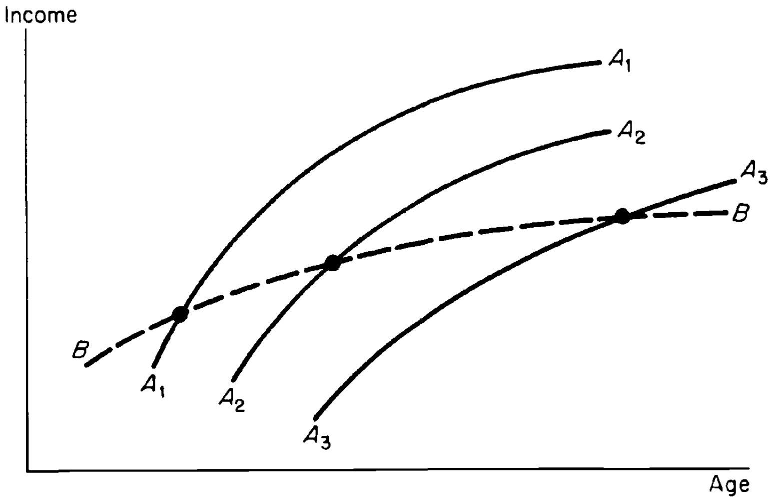

FIGURE 6-1
Hypothetical Cross-Section and Cohort Age-Income Profiles for Males

than whites because the height of the cohort profiles is rising faster over time for young nonwhites than for young whites. This could occur if there has been a secular decline in discrimination against nonwhites, particularly young nonwhites, in the labor market (in wages or in occupation) or in the quality of schooling. Recent studies suggest that there has been, indeed, a decline in labor market and school quality discrimination.[21]

---

21. The ratio of nonwhite to white expenditures per student in public schools in the South decreased dramatically from 1890 to 1915, but increased thereafter. For example, in Georgia the ratios of nonwhite to white expenditures per student were as follows:

Year	Ratio
About 1890	0.67
1915	0.21
1930	0.22
1945	0.38
1954	0.61

See Richard B. Freeman, "Labor Market Discrimination: Analysis, Findings, and Problems," paper presented at Econometric Society meeting of December 1972, Table 7. Similar evidence is reported in Finis Welch, "Black-White Differences in Returns to Schooling," *American Economic Review*, December 1973, pp. 893-907. In addition, Welch reports in the same article that in the 1960s the ratio of the nonwhite to white rate of return from schooling is higher for younger cohorts. This could be due to reduced discrimination in schooling or in the labor market.

TABLE 6-1

Income of Nonwhite Males as Percentage of Income of White Males, 1949, 1959, and 1969
(per cent)

Type of Income by Age Group	1949	1959	1969
Annual income:			
25-34 years	57	57	65
35-44 years	48	52	56
45-54 years	46	49	53
55-64 years	45	48	51
Weekly income:			
25-34 years	61	61	67
35-44 years	52	57	58
45-54 years	48	52	55
55-64 years	47	51	53

Note: Data for 1949 and 1959 relate to races other than white, data for 1969 relate to blacks only.
Sources: Council of Economic Advisers, *Economic Report of the President, 1974*, Washington, D.C., 1974, Table 39, p. 152.

Table 6-1 presents data on the income of nonwhite males as a percentage of the income of white males by age group for the three years 1949, 1959, and 1969. The ratio of the nonwhite to white cross-sectional age-income profile is obtained by reading down the columns of the table. The data suggest a flatter cross-sectional profile for nonwhites than for whites. Reading across the rows reveals a rise in the ratio of nonwhite to white incomes over time (with age held constant). That is, cohort profiles for nonwhites are rising at a faster rate than for whites.

If we read down the diagonals of Table 6-1 we follow a cohort as it ages. The data suggest that, as a cohort ages, the ratio of nonwhite to white income does not decline, or declines at a much slower rate than the cross-sectional profile. Thus, the flatter nonwhite cross-sectional profile is only partially due to a possibly somewhat flatter cohort profile.