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 Dispersion of Income

Chapter Author: Barry R. Chiswick

Chapter URL: http://www.nber.org/chapters/c3676

Chapter pages in book: (p. 143 - 178)

8

The Dispersion of Income

In this chapter the human capital earnings function—equation 6-10—is applied to generate and test the relation, with the U.S. states and the provinces of Canada as units of observation, between relative income (or earnings) inequality and the distributions of the rate of return from schooling, years of schooling, years of experience, weeks of employment during the year, and race. Using this function instead of ad hoc reasoning has the advantage of specifying the explanatory variables, the manner in which they enter the regression equation, and the economic interpretation of the regression coefficients.

The variance of the natural log of income is generated as the dependent variable. It is a commonly used measure of income inequality, and, devoid of units, facilitates interregional comparisons. There appears to be, moreover, greater social concern with relative than absolute inequality of income. Note that the explanatory power of the regressions, the adjusted coefficient of determination, is at least as high here as via ad hoc specifications.[1]

---

1. Using 1960 census data on family income inequality in the United States, Aigner and Heins obtained $\bar{R}^2$ of .76 to .85, and Al-Samarrie and Miller, of .82 to .84. Conlisk obtained $\bar{R}^2$ = .88, but only after experimenting with his estimating equation and deleting insignificant variables. See A. Al-Samarrie and H. P. Miller, "State Differentials in Income Concentration," *American Economic Review*, March 1967, pp. 59-72; D. J. Aigner and A. J. Heins, "On the Determinants of Income Equality," *American Economic Review*, March 1967, pp. 175-184; and J. Conlisk, "Some Cross-State Evidence

STATISTICAL IMPLEMENTATION OF THE MODEL

The Inequality Equation

Equation (6-10), which expresses the relation between earnings or income and the explanatory variables schooling, age, and weeks worked (employment), was written as:

$$\ln Y_i = X + r_i S_i + r_i' (A_i - S_i - 5) + \gamma \ln (WW_i) + U_i \qquad (8\text{-}1)$$

where U_i is the residual. As indicated in Chapter 6, it is assumed that r_i is independent of S_i and r_i' of $(A_i - S_i)$. The variance of a product of two independent random variables A_i and B_i is[2]

$$\text{Var}(AB) = \bar{A}^2 \text{ Var}(B) + \bar{B}^2 \text{ Var}(A) + \text{Var}(A) \cdot \text{Var}(B). \qquad (8\text{-}2)$$

Then, taking the variance of both sides of equation 8-1 results in:

$$\begin{aligned} \text{Var}(\ln Y) = {} & [(\bar{r} - \bar{r}')^2 + \text{Var}(r) + \text{Var}(r')] \text{ Var}(S) \qquad (8\text{-}3) \\ & + [(\bar{r}')^2 + \text{Var}(r')] \text{ Var}(A) + \gamma^2 \text{ Var}(\ln WW) \\ & + [2\bar{r}' (\bar{r} - \bar{r}') - 2 \text{ Var}(r')] \text{ Cov}(A,S) \\ & + [2\gamma (\bar{r} - \bar{r}')] \text{ Cov}(S,\ln WW) + [2\gamma\bar{r}'] \text{ Cov}(A,\ln WW) \\ & + [\text{Var}(r)] \bar{S}^2 + [\text{Var}(r')] (\bar{A} - \bar{S} - 5)^2 + \text{Var}(U), \end{aligned}$$

if it is assumed that r_i and r_i' are independent of each other.

Equation 8-3 expresses relative income inequality as a function of (a) the distributions of and intercorrelations among schooling, age, and employment (weeks worked), (b) the distributions of the rate of return from schooling (r_i), and (c) the distribution of the slope of the experience-log of income profile (r_i').[3] The data calculated for the states are the means and variances of schooling and age, the variance of the log of weeks worked, the covariance of schooling and age, and the average rate of return from schooling. A discussion of the definition and computation of these variables appears in Appendix A-2.

---

on Income Inequality," *Review of Economics and Statistics*, February 1967, pp. 115-118.

In my study, $\bar{R}^2$ is .88 for the income inequality of adult males, .92 for white males, and .86 for nonwhite males (see Tables 8-1, 8-4, and 8-6).

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

3. This equation was also used in Chiswick and Mincer, "Time Series Changes in Income Inequality in the United States from 1939, with Projections to 1985," *Journal of Political Economy, Supplement*, May-June 1972, s34-s66.

The regression equation is

$$\begin{aligned} \text{Var}(\ln Y) = b_0 &+ b_1\,(\bar{r}^2\ \text{Var}(S)) + b_2\ \text{Var}(S) \\ &+ b_3\ \text{Var}(A) + b_4\ \text{Var}(\ln WW) + b_5\ \text{Cov}(A,S) \\ &+ b_6\ \overline{S}^2 + b_7\,(\overline{A} - \overline{S} - 5)^2 + V' \end{aligned} \tag{8-4}$$

where the residual V' contains the intrastate residual variance (Var (U)) and the covariances of employment with age and schooling.[4] The slope coefficients can be interpreted as:

$$\begin{aligned} b_1 &= 1 \\ b_2 &= (\bar{r}')^2 - 2\bar{r}\bar{r}' + \text{Var}(r) + \text{Var}(r') \\ b_3 &= (\bar{r}')^2 + \text{Var}(r') \\ b_4 &= \gamma^2 \\ b_5 &= 2\bar{r}'\,(\bar{r} - \bar{r}') - 2\ \text{Var}(r') \\ b_6 &= \text{Var}(r) \\ b_7 &= \text{Var}(r'). \end{aligned} \tag{8-5}$$

The Variables

Equation (8-4) is a theoretical equation that relates the inequality of income to a set of human capital and employment variables. How the human capital analysis predicts the effects of each of the explanatory variables used in the actual statistical implementation of the equation is discussed below.

The Dependent Variables

Two dependent variables are under study here for the United States: the variance of the natural log of 1959 income of males, twenty-five and over, with income, and the variance of the natural log of 1959 earnings of males, fourteen and over, with earnings. Neither measure is ideal, since the model is developed for the variance of earnings of males who have completed their schooling, and since the explanatory variables are computed for males between the ages of twenty-five and sixty-four.[5]

Except for the inclusion of nonlabor income, the variance of the log of income is the more desirable dependent variable of the

4. Appropriate data are not available for weeks worked in Canada (see Appendix A-2). Because of the small sample size, the model is estimated in a stepwise manner for Canada.

5. See Appendix A-2.

two. Even the inclusion of nonlabor income does not cause serious difficulty in interpreting the results. First, the upper open-end interval of the income variable has $10,000 as the lower bound, and the mean of the interval is estimated by the Pareto equation. The distribution of income among those with large incomes (over $10,000 in 1959) does not affect the measure of inequality used here. Second, owners of nonhuman capital need not "live" in the same state as their capital. (Our model, however, relates income distribution among the residents of a state to human capital and employment in that state, so nonlabor income may be viewed as causing some measurement error in the dependent variable.)

The earnings data have a $10,000 lower bound to the upper open-end interval too, but the data include the earnings of young males, many of whom are also investing in schooling. Thus, using this measure of inequality may result in measurement errors correlated with the explanatory variables and hence generate biased coefficients. For example, with schooling level held constant, states with younger populations will have a lower value for the experience level of adults. Because of the younger population, these states will have a relatively greater proportion of the age fourteen-and-over labor force in the age group between fourteen and twenty-four. The larger the fraction of very young males in the labor force, the larger the inequality of earnings. Hence, the error in measuring the earnings inequality of adults is correlated with an explanatory variable—the level of experience.

The Rate of Return-Schooling Inequality Interaction Variable

No direct calculation of the rate of return from schooling can be made for each of the states for 1959 because of the lack of income (or earnings) data cross-classified by schooling and age. Two different approaches are employed to compute values for the rate of return, and both sets of estimates are used empirically. These are:[6] (a) r_c, the "regression" estimate based on all levels of schooling; and (b) r_m, the "overtaking age" estimate for high school.

The regression estimate for each state and race is the slope coefficient when the natural log of income is regressed on years of schooling, using microdata for the race-specific males in that state. The slope coefficient of schooling is a downward-biased estimate of the rate of return because years of labor market experience is

6. The regression estimate of the rate of return is developed and discussed in detail in Chapter 3. The estimation procedure for both rates of return appears in Appendix A-2.

an omitted variable.[7] It is, however, an average rate of return to all levels of schooling.

The overtaking age rate of return (r_m) is computed from the earnings of high school graduates ($\ln Y_{12}$) and grammar school graduates ($\ln Y_8$) at the "overtaking age." The overtaking age is the age at which the upward-rising age-earnings profile of a given level of schooling would cross the horizontal age-earnings profile that would exist if there were no investment in on-the-job training.[8] From knowing the earnings at this age and using the relation $\ln Y_{12} = \ln Y_8 + r_m (4)$, where 4 is the four years of high school, the term r_m can be computed. Because of data limitations, the procedure involves several approximations which generate considerable measurement errors.

A positive partial slope coefficient equal to unity is predicted for the rate of return-schooling inequality interaction term $\bar{r}^2 \operatorname{Var}(S)$. Because of the downward bias inherent in the regression estimate of the rate of return (r_c), the slope coefficient of the rate of return-schooling inequality interaction variable will be greater than unity when this rate of return is used. This bias above unity will be smaller for nonwhites than for whites because the rate of return is biased downward more for whites.[9] When the overtaking age rate of return is used, the regression slope coefficient may be less than unity (but still positive) due to errors of measurement.

Variance in Schooling

No prediction is offered on a priori grounds as to the sign of the coefficient on the variance in schooling when the rate of return-schooling inequality interaction term is held constant. Depending on the magnitudes of the component parameters, b_2 can

7. If the expression for weeks worked is deleted from equation 8-1, we have

$$\ln Y_i = X + rS_i + r' (A_i - S_i - 5) + U_i.$$

If experience $(A - S - 5)$ is now deleted, the slope coefficient of S is biased downward because r' is positive and the covariance of S and $(A - S - 5)$ is negative. $\operatorname{Cov}(S, A - S - 5) = \operatorname{Cov}(A,S) - \operatorname{Var}(S)$. The secular increase in years of schooling means the covariance of A and S is negative.

8. See Mincer, *Schooling, Experience, and Earnings*, Part 1.

9. The downward bias is due to the negative correlation of schooling with years of investment in experience. Since, for each level of schooling, nonwhites appear to have a flatter cross-sectional experience-income profile (lower r'), the downward bias of r_c will be smaller. See the last section of Chapter 6 for a discussion of racial differences in the slope of the experience-income profile.

be either positive or negative. An examination of microdata parameter estimates, however, suggests that b_2 is negative for all males and for white males.[10]

The slope of the variance in schooling is expected to be less negative for nonwhites. Although nonwhites may have lower rates of return from schooling and have flatter age-earnings profiles (lower $\bar{r}'$), they appear to have a higher coefficient of variation in rates of return from schooling than do whites.[11]

Variance of Age

The variance of age (with level and dispersion of schooling and level of age held constant) is expected to have a significant positive slope coefficient. The larger the dispersion in age, the larger the proportion of young and old workers, and the larger the dispersion in income. Very young workers have lower net earnings than prime age workers of the same level of schooling and employment because they have less previous investment in experience and are making larger current investments. On the other hand, older workers have higher weekly wages than prime age workers due to the accumulated effects of experience, but these are subject to depreciation and obsolescence.

The slope of the variance in age is expected to be lower and less significant for nonwhites than for whites because of the flatter nonwhite cross-sectional experience-earnings profile.

Covariance of Schooling and Age

The covariance of schooling and age is negative due to the secular increase in years of schooling. It varies across states due to migration and different secular trends in schooling. For the same overall levels and inequalities of schooling and age, a stronger

10. b_2 is negative if $b_2 = (\bar{r}')^2 - 2\,(\bar{r})\,(\bar{r}') + \sigma^2\,(r) + \sigma^2\,(r')$ is negative. For nonfarm white males, Mincer (in *Schooling, Experience, and Earnings*, 1972 mimeographed version) found that $\bar{r} = .11$ and that the coefficient of variation in the rate of return from schooling and postschooling human capital is approximately one-third. If the coefficient of variation of the slope of the age-earnings profile (r') is also one-third, b_2 is negative if

$$(\bar{r}')^2 - 2\,(.11)\,(\bar{r}') + (.11)^2\,(.33)^2 + (\bar{r}')^2\,(.33)^2 < 0,$$

since $\sigma^2\,(r) = [\bar{r}\,C.V.\,(r)]^2$. This holds if the slope of the experience-log of earnings profile $(\bar{r}')$ is equivalent to $.006 < \bar{r}' < .192$. Hence it is hypothesized that b_2 is negative.

11. See my "Racial Differences in the Variation in Rates of Return from Schooling," G. von Furstenberg et al., eds., *Patterns of Racial Discrimination*, Vol. 2, *Employment and Income*, D.C. Heath, 1974.

secular trend implies a more negative (algebraically lower) covariance of schooling and age. Since the square of the standard deviations of schooling and age are being held constant, interstate variations in the covariance really reflect differences in the correlation between schooling and age [Cov $(A,S) = R\ (A,S) \cdot SD\ (A) \cdot SD\ (S)$].

If we assume no depreciation of human capital with age and a uniform quality of schooling over time, an increase in the algebraic value of the correlation of schooling and age (i.e., a lower secular trend) increases the inequality of income if b_5 is positive (see equation 8-5). An examination of microdata parameter estimates suggests the hypothesis that the slope coefficient of the covariance of schooling and age (b_5) is positive.[12]

For the moment let us assume that there are no investments in postschool training ($\bar{r}'$ and $\sigma^2\ (r') = 0$). Age-earnings profiles would be horizontal and b_5, the coefficient of the covariance of schooling and age, would be zero. Suppose there were secular increases in the quality of schooling or in job opportunities. Younger workers would have higher earnings for the same number of years of schooling than older workers. For the same overall level and dispersion of schooling and age, income inequality would be greater if younger workers were concentrated in the higher schooling categories. Thus, given secular increases in schooling quality or job opportunities, the greater the algebraic value of the correlation of schooling and age (i.e., the weaker the secular trend), the smaller is the inequality of income.

Depreciation of human capital with age has a similar effect as secular improvements in school quality. For a given correlation of schooling and age, the greater the rate of depreciation of earnings with age, the larger is the inequality of incomes.

If we combine the effects of investment in postschool training, secular increases in the quality of schooling or job opportunities, and depreciation of earning potential with age, the sign of the slope coefficient of the covariance of schooling and age be-

12. Using microdata on the earnings of nonfarm white males, Mincer found $\bar{r} = .11$ and a coefficient of variation in rates of return from human capital of one-third (*Schooling, Experience, and Earnings*, 1972 mimeographed version). If the coefficient of variation in r' is also one-third, b_5 is positive if

$$2\,\bar{r}'\,(.11 - \bar{r}') - 2\,(\bar{r}')^2\,(.33)^2 > 0, \text{or } \bar{r}' < .10.$$

The average slope of the age-earnings profile is less than the slope coefficient of experience when log earnings is regressed on schooling, experience, and experience squared. Mincer found the slope of experience to be .08. Therefore, b_5 is expected to be positive.

comes unclear, a priori. The slope coefficient is more likely to be positive, the more important postschool training is relative to (a) secular changes in the quality of schooling or job opportunities for new labor force entrants and (b) the decline of earnings with age. The previous discussion of racial differences in these parameters suggests that the algebraic value of the slope coefficient (b_5) of the covariance term will be lower for nonwhites than for whites.

Variance of the Log of Weeks Worked

Our model generates relative income inequality as a function of the relative variance of weeks worked (employment). This is the first cross-sectional study of income inequality in which this variable has been used.[13] Previous cross-sectional and time series studies used either the fraction of the labor force or of the population reported as unemployed on a particular date.[14] The dispersion in weeks worked is an analytically superior variable—it gets to the heart of the employment-income relation. Suppose, for example, that weekly wages are independent of weeks worked and that all workers have the same level of skill, but 30 per cent of the labor force works forty weeks per year while the remaining 70 per cent works a full year. Thus, 30 per cent of the work force is unemployed approximately one-fifth of the year, which generates inequality in weeks worked and in annual income. Suppose we have another community where, again, all workers have equal weekly wages and the weekly wage is independent of the number of weeks worked. In this second community, however, everyone works forty weeks and everyone is unemployed approximately one-fifth of the year. In spite of the greater rate of unemployment in the second community, the inequality of annual income is smaller (in fact, zero in this extreme example because there is no dispersion in weeks worked).

In the United States, the unemployment rate and the variance in the log of weeks worked are highly correlated over the business cycle, but weakly correlated across states at any one moment in

13. See also Chiswick and Mincer, "Time Series Changes in Income Inequality."

14. See T. P. Schultz, "Secular Trend and Cyclical Behavior of Income Distribution in the United States: 1944-1965," in Lee Soltow, ed., *Six Papers on the Size Distribution of Wealth and Income*, NBER, 1969; and L. Thurow, "Analyzing the American Income Distribution," *American Economic Review*, May 1970, pp. 261-269. See also the articles cited in footnote 1, p. 143, by Conlisk, Al-Samarrie and Miller, and Aigner and Heins.

time.[15] Therefore, for a cross-sectional study of income inequality the relative variance in weeks worked is also statistically superior to the unemployment rate. For a time series study, however, the unemployment rate is a good proxy for the relative inequality of weeks worked.[16]

The regression slope coefficient of the variance in the log of weeks worked is γ^2, where γ is the elasticity of earnings with respect to the fraction of weeks worked. γ^2 shall be tested for statistical significance against the two hypotheses that the population values are 1.0 and 1.37. A coefficient of unity implies that weekly wages are uncorrelated with the number of weeks worked.[17] Recall that Mincer obtained $\gamma = 1.17$ (hence, $\gamma^2 = 1.37$) in a microdata analysis for nonfarm white males in the 1/1,000 sample of the 1960 Census.[18]

A third hypothesis is that the elasticity of earnings with respect to weeks worked (γ) is lower for nonwhite than for white males. The theoretical arguments supporting this hypothesis were developed above (see pp. 125-126).

Average Schooling

The statistical model assumes that neither level nor inequality of rates of return from schooling vary with the level of schooling. On this assumption, the partial slope coefficient of average schooling should be positive, and with the use of microdata, the relative inequality of income within schooling, age, and employment cells should increase with the level of schooling. Empirically, however, with age held constant, the absolute inequality of labor market income increases, but its relative inequality decreases, at higher levels of schooling.[19] This decline in relative income inequality may be due to the decline, with higher levels of schooling, in the relative

15. See M. Hashimoto, "Factors Affecting State Unemployment," Ph.D. dissertation, Columbia University, 1971.

16. One reason for this is that in recessions the number of weeks worked per year decreases more for workers who tend to work fewer weeks per year even under normal business conditions (primarily those with low levels of training). This is consistent with the human capital model of investment in job-specific training. (See Becker, *Human Capital*, Part 1 and Chiswick and Mincer, "Time Series Changes in Income Inequality.")

17. See discussion in Chapters 2 and 6.

18. See Mincer, *Schooling, Experience, and Earnings*, Part 2, 1972 mimeographed version.

19. See T. P. Schultz, "Long-Term Changes in Personal Income Distribution," *American Economic Review*, May 1972.

inequality of hours worked per week, the average level of the rates of return, or the variance in rates of return.[20]

Suppose rates of return from schooling decline with higher levels of schooling when the age and employment variables are held constant. Then a negative slope coefficient of the level of schooling could emerge in our regression analysis. A higher level of schooling implies a larger proportion of the population with some college schooling. If the effect of declining rates of return is operative, the downward bias in the coefficient of average schooling will be greater if the measure of the rate of return is the rate for high school (as is the overtaking age rate of return, r_m), rather than a weighted average rate for all levels of schooling (as is the regression estimate, r_c). Thus, if the coefficient of average schooling is algebraically lower when r_m rather than r_c is held constant, there will be indirect evidence of a decline in the rate of return at higher levels of schooling.[21]

Level of Experience

A rise in the level of experience results in an increase in the relative inequality of income if there are differences in the slope of the experience-income profile and if the other assumptions of the statistical model apply. However, the income data are for observed income, that is, income net of the opportunity cost of postschool training. Differential investments in postschool training tend to increase the inequality of observed income for young members of the labor force. The result may be the appearance of an inverted U-shaped relation between the (squared) level of experience and the relative inequality of income.

A less positive (or more negative) relation between experience level and income inequality will emerge as young males assume more importance in the data because of their low level and large

20. The relative inequality of weeks worked per year, but not of hours worked per week, is held constant in the empirical analysis.

21. Becker, Hanoch, Hansen, and Johnson all found lower rates of return for college graduates than for high school graduates. They did not, however, hold employment constant. Using the 1960 Census 1/1,000 sample for white males, Mincer found no significant evidence of a higher (or lower) rate of return at higher levels of schooling, holding experience and weeks worked during the year constant.

See Becker, *Human Capital*, Part 2; Hanoch, "An Economic Analysis of Earnings and Schooling"; W. Lee Hansen, "Total and Private Rates of Return to Investment in Schooling," *Journal of Political Economy*, April 1963, pp. 128-140; Mincer, *Schooling, Experience, and Earnings*, Part 2; and Johnson, "Returns from Investment in Human Capital."

inequality of income. Since the earnings data are for males fourteen years of age and over while the income data are for males twenty-five years of age and over, a year's increase in the mean level of experience is expected to have a smaller positive effect on earnings inequality than on income inequality. This expectation is strengthened by the rise in the level and inequality of nonlabor income with experience.

Race Composition and Region

Let us assume Y_w is the earnings of a white worker of a given level of years of schooling, years of experience, and weeks of employment. The earnings of a nonwhite in the same schooling-experience-weeks worked cell could be written as $Y_{NW} = Y_w\ (1 - d)$ where d is the per cent difference in earnings between white and nonwhite workers. The variable d is positive if nonwhites receive a lower quality of schooling or a lower level of investment in postschool training, work fewer hours per week, or are subject to direct labor market discrimination.

We can write

$$Y_i = Y_{i,w}\ (1 - d)^{NW_i} \tag{8-6}$$

where $NW_i = 1$ for nonwhites and $NW_i = 0$ for whites. Then,

$$\ln Y_i = \ln Y_{i,w} - (d)\ (NW_i), \tag{8-7}$$

if d is a small number, and if individual differences in d_i are placed in a residual,

$$\sigma^2\ (\ln Y_i) = \sigma^2\ (\ln Y_{i,w}) + (d^2)\ (p - p^2) + U \tag{8-8}$$

where p is the per cent of the population nonwhite and the covariance of $\ln Y_{i,w}$ and NW_i is in the residual U.[22] The variables which comprise Var $(\ln Y_w)$ are shown in equation (8-4). By adding the variable $p - p^2$ to equation (8-4) in the analysis of all males we can test for the effect of race composition on the inequality of income among adult males.

A region dummy variable ($NSD = 1$ for Southern states) is added to test for the existence of regional differences in inequality after controlling for the human capital, employment, and race composition variables.

The empirical analysis is performed not only for all males, but also for white and nonwhite males separately as in the previous

22. If p is the per cent nonwhite, it is the mean value of the dichotomous variable NW_i, and Var $(NW_i) = (\overline{NW^2}) - (\overline{NW})^2 = (\overline{NW}) - (\overline{NW})^2 = p - p^2$.

chapter. That is, the variables in equation (8-4) can be race-specific. When this is done, a variable for the nonwhite percentage of the adult male labor force can be added to the equation. For the analysis of white income inequality, this variable's partial slope coefficient measures the effect of the relative number of nonwhites in the states on the inequality of income *among* whites within schooling, experience, and weeks worked cells. It offers a similar interpretation when added to the equation for nonwhite income inequality.[23]

23. The variable per cent nonwhite in the state is added to the white (nonwhite) income inequality equation on the basis of the white (nonwhite) employee discrimination hypothesis.

If white employees act as if they had a taste for discrimination against nonwhite workers, those whites who work with nonwhites receive a positive compensating wage differential. This creates a component of inequality in the income of white males, holding other variables constant. Up to a point, white income inequality increases the more nonwhites there are in the state. A positive slope coefficient for the variable per cent nonwhite in the white income inequality equation would be consistent with the white employee discrimination hypothesis. (A similar analysis applies to the income inequality of nonwhites.) The statistical testing of these hypotheses is developed in detail in Barry R. Chiswick, "Racial Discrimination in the Labor Market: A Test of Alternative Hypotheses," *Journal of Political Economy*, November-December 1973, pp. 1,330-1,352, reprinted in G. von Furstenberg *et al.*, eds., *Patterns of Racial Discrimination*, Vol. II, *Employment and Income*, New York, D.C. Heath, 1974.

A thumbnail sketch of the model follows.

Let Y_i^* be the weekly wage of a white worker of a given level of skill (i.e., schooling and labor market experience) if he does not work with nonwhites. Let X_i be a dichotomous variable that takes the value of unity if he works with nonwhites and the value of zero if he does not. It is assumed that white and nonwhite workers are neither perfect substitutes nor perfect complements in production, and that some white workers have $X_i = 1$ and others $X_i = 0$. Let d_i be a market discrimination coefficient, the per cent increase in wages paid to the white worker if he works with nonwhites. The weekly wage of the ith white worker is

$$Y_i = Y_i^* (1 + d_i X_i).$$

Taking the natural log of both sides of the above equation and assuming d_i is small,

$$\ln (Y_i) = \ln (Y_i^*) + d_i X_i.$$

Evaluating the variance of both sides of this new equation across white males in the state,

$$\sigma^2 (\ln Y_i) = \sigma^2 (\ln Y_i^*) + \sigma^2 (d_i X_i) + 2 \operatorname{Cov} (\ln Y_i^*, d_i X_i).$$

If d_i and X_i are independent,

$$\sigma^2 (dX) = \bar{d}^2 \sigma^2 (X) + \bar{X}^2 \sigma^2 (d) + \sigma^2 (X)\, \sigma^2 (d)$$
$$= [\bar{d}^2 + \sigma^2 (d)]\, \sigma^2 (X) + [\sigma^2 (d)]\, \bar{X}^2 .$$

Summary

The human capital earnings function is converted into a model relating the relative inequality of earnings (or income) to the distributions of schooling, age, and employment.

The rate of return-schooling inequality interaction term is hypothesized to have a slope coefficient equal to unity if the rate of return is measured correctly. Due to data limitations two estimates of the rate of return are employed. For the regression estimate of the rate of return (r_c), the slope coefficient is expected to be biased upward (i.e., above unity), and more so for whites than nonwhites. When the overtaking age rate of return (r_m) is used, the slope coefficient is hypothesized to be positive but less than unity due to errors of measurement.

On the basis of parameter estimates from microdata, a negative partial effect is predicted for the variance in schooling for white males and for all males. A less negative effect is hypothesized for nonwhite males.

The slope coefficient of the inequality of age is predicted to be positive, but lower for nonwhites than for whites as a result of lower investments in postschool training by, and the secular decline in discrimination (in schooling as well as in the labor market) against, young nonwhites.

The partial effect of the covariance of schooling and age is not clear a priori. The stronger the upward secular trends are in schooling (i.e., the more negative the correlations of schooling and age), the smaller is income inequality within the framework of the model. However, secular uptrends in the quality of schooling or job opportunities for new entrants, as well as the depreciation of earning potential with age, tend to impart a negative partial effect between the covariance term and income inequality. The lower level of investment in training by nonwhite males, the greater secular improvement in the quality of schooling and job oppor-

---

Let us designate k as an index of labor market integration so that $k = \overline{X}/p$, where p is the per cent nonwhite, then $\overline{X} = kp$ and $\sigma^2(X) = (kp) - (kp)^2$.

After these adjustments, our equation can be rewritten as

$$\sigma^2\ (\ln Y) = \sigma^2\ (\ln Y^*) + \{k[\bar{d}^2 + \sigma^2(d)]\}\,p + (-\bar{d}^2 k^2)\,p^2 + U.$$

The coefficient of p is positive and that of p^2 is negative.

Differentiating the last equation with respect to p indicates that income inequality is a rising function of p as long as less than half of the whites in a state work with nonwhites. Thus, the white employee discrimination model hypothesizes that, holding $\sigma^2\ (\ln Y^*)$ constant, *white* income inequality in a state is a rising function of the nonwhite percentage of the male labor force (p).

tunities for young nonwhites, and the larger relative depreciation of earning potential with age for those engaged in less skilled jobs all point to a lower algebraic value for the slope coefficient of nonwhites.

The predicted racial difference in the partial effect of the covariance of age and schooling is quite important for the statistical analysis. Where racial differences are hypothesized for the coefficients of other variables, values closer to zero are predicted for nonwhites. It might be argued, however, that the lower absolute value of observed slope coefficients for these variables is due to greater sampling variability or larger measurement errors in the nonwhite data. Thus, if the slope of the covariance of schooling and age is found to be negative and algebraically lower for nonwhites, there would be evidence that the lower nonwhite slope coefficients for the other variables are not entirely due to sampling and measurement problems.

Another explanatory variable generated by the statistical model is the variance of the log of weeks worked. Here a positive slope coefficient is hypothesized, which, if greater than unity, implies that males of a given age and schooling who work more weeks per year have higher weekly wages. Using microdata, Mincer found the elasticity of earnings with respect to weeks worked (γ) to be equal to 1.17. Thus, two hypotheses to be tested are whether the slope coefficient (γ^2) is equal to unity, or equal to 1.37.

Those with higher weekly wages work more weeks per year if the labor supply curve rises upward, their wages are higher because of greater investments in specific training, or the hours worked per week and weeks worked per year are positively correlated. On the other hand, those with higher weekly wages work fewer weeks per year if labor supply curves bend backward, or if the higher wage is an equalizing differential compensating for seasonal employment. Thus, a third hypothesis is that the slope coefficient γ^2 is lower for nonwhites than whites because of a smaller amount of investment in training and a greater concentration in seasonal occupations.

Assuming a constant level and inequality of rates of return from schooling across schooling levels, the statistical model predicts a positive slope coefficient of the squared level of schooling. If the level or inequality of rates of return declines with schooling, a negative coefficient could emerge. A lower coefficient of average schooling with the rate of return to high school (r_m) rather than a weighted average rate of return to all levels of schooling (r_c) held constant gives indirect support to the hypothesis that average rates of return decline with higher levels of schooling.

The partial effect of the (squared) level of experience on income inequality may be U-shaped. During initial periods of gathering experience, income inequality exists partly because of differential investments in training; in subsequent periods, partly because of the returns on previous investments. Even if absolute income inequality increases with experience, relative inequality may first decline and subsequently increase. Thus, the sign of the level of experience is not clear a priori. Because of the inclusion of young males in the earnings data, the coefficient of experience is expected to be algebraically lower for earnings than for income inequality.

The effect on overall income inequality (i.e., for all males) of racial differences in the incomes of whites and nonwhites within schooling, experience, and weeks worked cells is incorporated through the variance of the dichotomous variable NW_i, which takes the value of 1 for nonwhites and of zero for whites. The variance of this variable is $p - p^2$, where p is the nonwhite percentage of the adult male labor force. This variable is predicted to have a positive slope coefficient, which would be consistent with the hypothesis that nonwhites have lower incomes than whites within schooling, experience, and weeks worked cells. When separate analyses are performed for inequality among whites and nonwhites, a variable (p)—nonwhite percentage of the male labor force—can be added. This variable shows the effect of a state's racial composition on the income inequality of a particular race within that state.

A region dummy variable, *NSD* (where *NSD* = 1 for Southern states), is included to test for the existence of regional differences in income inequality when the human capital, employment, and race composition variables are held constant.

To summarize, the predicted signs for the human capital and employment variables are the following:

Variable	Symbol	Hypothesized Sign
Rate of return-schooling inequality interaction variable	r^2 Var(S)	+
Variance of schooling	Var(S)	−
Variance of age	Var(A)	+
Variance of log of weeks worked	Var(lnWW)	+
Covariance of age and schooling	Cov(A,S)	[a]
Average schooling squared	$(\mathrm{Av}(S))^2$	[a]
Average experience squared	$(\mathrm{Av}(EXP))^2$	[a]

[a]Sign is ambiguous.

EMPIRICAL APPLICATION

U.S. Males

The human capital income inequality model developed above is applied here to the income inequality of males twenty-five years of age and over and the earnings inequality of males fourteen years of age and over in 1959.[24] The analysis covers the fifty-one states (the District of Columbia is treated as a state) and the forty-nine coterminous states (with Alaska and Hawaii deleted from the data). Separate analyses are performed for the regression estimate and the overtaking age estimate of the rate of return from schooling.

Analysis with Regression Estimate of Rate of Return

Table 8-1 presents the regression results for income and earnings inequality for the fifty-one- and forty-nine-state data sets, using the regression estimate of the rate of return and the schooling, age, and weeks worked variables. The model's explanatory power is very high; the adjusted coefficient of determination is 88 per cent for income and 77 per cent for earnings. Nearly all of the slope coefficients differ significantly from zero at a 10 per cent level.

The slope coefficient of the rate of return-schooling inequality interaction variable (r_c^2 Var (S)) is highly significant. It is larger than unity in all four regressions, the difference from unity being significant for earnings but not quite significant for income.[25]

24. The means and standard deviations of the variances of the log of income and earnings for all males are:

	Mean	Standard Deviation
All States (fifty-one)		
Var($\ln Y$)	0.7867	0.1184
Var($\ln E$)	0.7743	0.1076
Non-South (thirty-four states)		
Var($\ln Y$)	0.7241	0.0795
Var($\ln E$)	0.7283	0.0902
Non-South (excl. Alaska and Hawaii)		
Var ($\ln Y$)	0.7228	0.0760
Var ($\ln E$)	0.7237	0.0809
South (seventeen states)		
Var($\ln Y$)	0.9119	0.0758
Var($\ln E$)	0.8662	0.0770

Source: Appendix A-2.

25. For a 10 per cent level of significance and 40 degrees of freedom, $t = 1.68$. The fifty-one-state sample has 43 degrees of freedom, while the forty-nine-state sample has 41 degrees of freedom.

TABLE 8-1

Regression Results, Income Inequality, All Males, with Regression Estimate of Rate of Return (r_c)

Independent Variable	Income Inequality 51 States	Income Inequality 49 States	Earnings Inequality 51 States	Earnings Inequality 49 States
r_c^2 Var(S)	1.2968 (7.28)	1.1239 (5.52)	1.5947 (7.43)	1.6243 (6.43)
Var(S)	-0.0071 (-1.88)	-0.0039 (-0.91)	-0.0174 (-3.82)	-0.0173 (-3.26)
Var(A)	0.0106 (4.33)	0.0110 (4.51)	0.0090 (3.08)	0.0091 (3.02)
Var(ln WW)	1.0378 (5.39)	0.9138 (3.95)	0.9208 (3.99)	0.8764 (3.05)
Cov(A,S)	-0.0078 (-2.22)	-0.0123 (-2.70)	-0.0085 (-2.02)	-0.0073 (-1.30)
(Av(S))^2	-0.0027 (-2.16)	-0.0027 (-2.16)	-0.0030 (-1.99)	-0.0028 (-1.76)
(Av(Exp))^2	-0.0005 (-2.06)	-0.0004 (-1.70)	-0.0010 (-3.38)	-0.0010 (-2.80)
Constant	-0.0816	-0.2067	0.6226	0.5601
N	51	49	51	49
R^2	0.8956	0.8982	0.8188	0.8003
$\overline{R}^2$	0.8785	0.8808	0.7893	0.7662

Note: Student t-ratio in parentheses.
Source: Appendix A-2.

Based on microdata parameters, the analysis earlier in the chapter predicted a negative partial slope coefficient of the variance in schooling with the rate of return-schooling inequality interaction variable held constant. The expected negative slope emerges in all four regressions. The coefficient is highly significant for the earnings data but presents a mixed picture for the income data. As to the variance in age, it has the expected positive slope coefficient and is highly significant. The size of the slope coefficient is not significantly larger for the income than the earnings data.

The relative inequality of weeks worked is also a highly significant variable. The slope coefficient (γ^2) does not differ significantly from a hypothesized value of 1.0—the t-ratios are all less than unity. If the population value of the elasticity of income with respect to weeks worked (γ) is predicted to be 1.17, we can test the hypothesis that γ^2 = 1.37. The observed γ^2 differ significantly from this value at a 10 per cent level (two-tailed test), but not at a 5 per cent level.[26]

It was argued in the first part of this chapter dealing with the statistical implementation of the model that a priori analysis does not predict a sign for the partial slope coefficient of the covariance of schooling and age. Empirically, in Table 8-1 the covariance term's slope is negative and generally significant. The negative slope of the covariance of schooling and age means that, with variances of schooling and age held constant, states with greater secular increases in years of schooling (i.e., a more negative correlation of schooling and age) have larger inequalities of earnings and income. This implies that the income effects of the secular improvement in schooling quality and job opportunities for young males and of depreciation of earnings potential with age outweigh those of investment in postschool training.[27] When separate analyses are performed for whites and nonwhites (see pp. 163-173 below), the slope of the covariance term is negative and significant only for nonwhites.

Empirically, in the analyses for all males, the squared levels of schooling and experience have significant negative slope coefficients. The negative slope for schooling level is not consistent with the model developed in the first part of this chapter, but would be if the model allowed for either a declining level or variance in marginal rates of return for higher levels of schooling. The negative slope for the level of experience implies that, with schooling distribution held constant, the relative inequality of income is lower in older age groups. As will be shown below, however, when the analysis is performed for white and nonwhite males separately, schooling level and experience level are generally not significant. This suggests an interaction between the effects of race and the effects of schooling and experience level on income inequality.

We can now add the race composition variable $p - p^2$ (where p is the percentage of nonwhite labor force males) and the South-non-South dummy variable (NSD = 1 in the South) to the regression equation. When these variables are included the adjusted co-

26. To test the null hypothesis $\gamma^2 = 1.37$, the t-ratios are:

Equation	Fifty-one States	Forty-nine States
Income	1.73	1.95
Earnings	1.95	1.72

The critical t-ratios for a two-tailed test, 40 degrees of freedom, are $t = 1.68$ at a 10 per cent level of significance and $t = 2.02$ at a 5 per cent level.

27. This is based on Chapter 7.

TABLE 8-2

Effect of Including Race and Region Variables on the Unadjusted (R^2) and Adjusted ($\overline{R}^2$) Coefficients of Determination for All Males

	Race and Region Not Included		Race and Region Included	
	R^2	$\overline{R}^2$	R^2	$\overline{R}^2$
	A: Using r_c			
Income inequality				
51 states	0.8956	0.8785	0.8970	0.8744
49 states	0.8982	0.8808	0.8994	0.8761
Earnings inequality				
51 states	0.8188	0.7893	0.8192	0.7795
49 states	0.8003	0.7662	0.8003	0.7541
	B: Using r_m			
Income inequality				
51 states	0.8142	0.7839	0.8350	0.7987
49 states	0.8543	0.8294	0.8614	0.8294
Earnings inequality				
51 states	0.7473	0.7061	0.7643	0.7127
49 states	0.7522	0.7098	0.7576	0.7016

Sources: Columns (1) and (2) from Tables 8-1 and 8-3. Columns (3) and (4), regressions not reported, see Appendix A-2.

efficient of determination ($\overline{R}^2$) is lower (see Table 8-2, rows 1 to 4). This means that the F-ratio for the inclusion of these two variables is less than unity; adding these variables to the regression equation does not significantly increase the model's explanatory power. The added variables have insignificant slope coefficients, and the slopes and t-ratios of the other variables are barely changed. Thus, in the analysis of income or earnings inequality for all males, after controlling for the regression estimate of the rate of return and the distributions of schooling, experience, and weeks worked, there are no systematic regional or race composition effects. Hence, the observed South-non-South difference in income and earnings inequality must be explained by the model's human capital and employment variables.

Analysis with Overtaking Age Rate of Return

We might ask, however, what would happen if the alternative measure of the rate of return—the overtaking age estimate (r_m)—were used in the empirical analysis. The regression results using r_m rather than r_c appear in Table 8-3. The model's explanatory power

TABLE 8-3

Regression Results, Income Inequality, All Males, with Overtaking Age Estimate of Rate of Return

Independent Variable	Dependent Variable			
	Income Inequality		Earnings Inequality	
	51 States	49 States	51 States	49 States
r_m^2 Var(S)	0.2218	0.1872	0.3780	0.3649
	(3.26)	(2.98)	(5.24)	(5.04)
Var(S)	-0.0001	0.0026	-0.0114	-0.0109
	(-0.03)	(0.54)	(-2.22)	(-1.98)
Var(A)	0.0114	0.0114	0.0093	0.0090
	(3.47)	(3.90)	(2.68)	(2.67)
Var(ln WW)	1.0136	0.8664	0.7507	0.7514
	(3.78)	(3.10)	(2.64)	(2.33)
Cov(A,S)	-0.0040	-0.0166	-0.0041	-0.0132
	(-0.86)	(-3.09)	(-0.83)	(-2.13)
(Av(S))^2	-0.0049	-0.0047	-0.0054	-0.0055
	(-3.06)	(-3.23)	(-3.17)	(-3.27)
(Av(Exp))^2	-0.0007	-0.0007	-0.0013	-0.0015
	(-2.03)	(-2.10)	(-3.60)	(-2.33)
Constant	0.2623	0.1299	1.1661	1.2321
N	51	49	51	49
R^2	0.8142	0.8543	0.7473	0.7522
$\bar{R}^2$	0.7839	0.8294	0.7061	0.7098

Note: Student t-ratio in parentheses.
Source: See Appendix A-2.

is approximately 10 per cent lower when we substitute r_m for r_c—at approximately 80 per cent for income and 71 per cent for earnings.

The estimated slope coefficients of the rate of return-schooling inequality interaction variable are significantly greater than zero, but significantly lower than unity. It is not clear whether the apparent downward bias in the slope is due to measurement errors or to a characteristic of the parameter r_m. This uncertainty will be resolved in the future when data become available permitting a direct calculation of the overtaking age rate of return by state. The slope coefficients and t-ratios of the other variables are quite similar to the results obtained using r_c (compare Tables 8-1 and 8-3).[28]

28. One difference is that the level of schooling has a more negative slope and a larger t-ratio when the overtaking age rate of return is used. This is consistent with the hypothesis that the level or variance of rates of return from schooling is larger for high school than for higher levels of education.

Summary

Our analysis of the income and earnings inequality of males, using the variables suggested by the model developed in the beginning of the chapter, shows that the human capital and employment variables have a very high explanatory power ($\overline{R}^2$)—88 per cent for income inequality and 77 per cent for earnings inequality. The rate of return-schooling inequality interaction variable, the variances of age, and the log of weeks worked all have the predicted positive slope coefficients and are always statistically significant. The negative partial slope coefficient for the variance of schooling, predicted on the basis of microdata analyses, is obtained here, and is generally significant for earnings but not for income. The covariance of schooling and age, as well as the squared levels of experience and schooling, have negative and usually significant slope coefficients.

The addition of variables for the race composition of the state and region generally results in no significant increase in the model's explanatory power, in only insignificant slope coefficients for race and region, and in minor changes in the slope coefficients and *t*-ratios for the human capital and employment variables. Thus, the human capital and employment variables explain the larger inequality of income observed in the South than in the North.

U.S. White Males

The foregoing analysis for all males is performed here separately for white males in the coterminous states. White males are defined as whites in the thirty-nine states for which a white-nonwhite breakdown of the data was provided in the 1960 Census of Population and for all males in the ten states for which no separate data exist.[29] The analysis, for the sets of forty-nine and thirty-nine states, employs the regression estimate of the rate of return computed from data for white males.

The regressions presented in columns 1 and 2 of Table 8-4 for income inequality indicate a high explanatory power, $\overline{R}^2 = .92$ for the forty-nine states, and $\overline{R}^2 = .93$ for the thirty-nine states. The significant variables in the regressions are the rate of return-schooling inequality interaction term and the variance of age and log of weeks worked. These variables have the expected positive slope.

29. These ten states, all non-Southern, contain a small proportion of nonwhites. They are Idaho, Maine, Montana, Nevada, New Hampshire, North Dakota, Rhode Island, Utah, Vermont, and Wyoming.

TABLE 8-4
Regression Results, Income Inequality, White Males

Independent Variable	Dependent Variable			
	Income Inequality		Earnings Inequality	
	49 States	39 States	49 States	39 States
r_c^2 Var(S)	1.9735	1.7600	1.7702	1.4914
	(11.46)	(9.97)	(7.10)	(6.68)
Var(S)	-0.0030	-0.0025	-0.0143	-0.0087
	(-0.91)	(-0.74)	(-2.97)	(-1.99)
Var(A)	0.0102	0.0131	0.0109	0.0146
	(5.48)	(6.58)	(4.03)	(5.77)
Var(ln WW)	0.8585	0.7749	0.9335	0.6178
	(4.14)	(3.36)	(3.11)	(2.12)
Cov(A,S)	-0.0054	0.0014	-0.0055	-0.0059
	(-1.45)	(0.31)	(-1.00)	(-1.00)
(Av(S))^2	0.0006	-0.0008	-0.0009	-0.0024
	(0.72)	(-0.81)	(-0.71)	(-1.97)
(Av(Exp))^2	-0.0000	-0.0002	-0.0008	-0.0008
	(-0.03)	(-0.86)	(-2.52)	(-2.66)
Constant	-0.8854	-0.8128	-0.0276	-0.1776
	(-4.83)	(-4.25)	(-0.10)	(-0.73)
N	49.0	39.0	49.0	39.0
R^2	0.9297	0.9411	0.7846	0.8676
$\overline{R}^2$	0.9177	0.9278	0.7478	0.8377

Note: Student t-ratio in parentheses.
Source: See Appendix A-2.

The coefficient of the interaction term is significantly greater than unity, which was not the case in the all-male analysis. This is not surprising, however, since the regression estimate of the rate of return is biased downward more for white males than for all males. The slope of the relative inequality of employment is not significantly less than unity, but it is significantly less than 1.37.[30]

Three variables change in the switch from all males to white males from generally significant negative (all males) to insignificant (white males). These are the square of the levels of schooling and experience, and the covariance of age and schooling. Recall that in the all-male analysis these variables retained their generally significant negative slopes even when the race composition variable was added. This suggests that the effect of race differences in the all-male analysis is more complex than would be indicated by an

30. For a comparison with a hypothesized population value of $\gamma^2 = 1.37$, the t-ratios are 2.5 and 2.6, respectively, in the analyses for the forty-nine and thirty-nine states.

interstate difference in intercepts depending on the proportion of nonwhites. Indeed, it suggests an interaction of the effect of race composition with the effect of schooling and age levels and the covariance of schooling and age on the inequality of income.

The model's explanatory power is lower for earnings inequality—$\bar{R}^2 = .75$ for forty-nine states, $\bar{R}^2 = .84$ for thirty-nine states (columns (3) and (4) in Table 8-4)—than income inequality, but it is still substantial. Once again, the rate of return-schooling interaction term and the inequalities of age and employment have significant positive slopes. The covariance of age with schooling is not significant. While the squared level of experience has a significant negative slope, the squared level of schooling is not significant at a 5 per cent level.[31]

When earnings inequality for whites is the dependent variable, the coefficient of the rate of return-schooling interaction term is significantly greater than unity. The coefficient of relative inequality of weeks worked is not significantly less than unity, and significantly less than 1.37 only in the thirty-nine-state analysis.[32]

Table 8-5 presents the analysis of white income inequality where race composition and region have been added to the human capital and employment variables.[33] A dummy variable Z is created where $Z = 1$ in the ten states (all non-Southern) where separate race data do not exist. A region dummy variable (NSD = 1 in the South) is also added.

A comparison of the regressions with and without the race-

31. In the earnings inequality analysis, the slope coefficient of experience squared is expected to be less positive or more negative than in the income inequality analysis. The earnings data are for males fourteen and over with earnings, but the income data are for males twenty-five and older, and the explanatory variables are for males between twenty-five and sixty-four years of age. With schooling held constant, states with a younger population (i.e., lower level of experience for those between twenty-five and sixty-four) are likely to have a larger proportion of their population between the ages of fourteen and twenty-four. The level of experience may have a negative effect because the other explanatory variables do not adjust for the larger dispersions in hours per week, weeks worked per year, years of experience, and dollar investments in training of white males fourteen years and over than in the over-twenty-five group. That is, the variable "mean experience" may be capturing the effects of youths, tending to raise income inequality.

32. For a comparison with a hypothesized population value of $\gamma^2 = 1.37$, the t-ratios are 1.46 and 2.5 for the forty-nine-state and thirty-nine-state analyses, respectively.

33. The race variable is the percentage of nonwhite males between twenty-five and sixty-four years of age.

TABLE 8-5

Regression Results, Income Inequality, White Males, with Race and Region Variables

Independent Variable	Dependent Variable			
	Income Inequality 49 States		Income Inequality 39 States[a]	
	(1)	(2)[b]	(3)	(4)[c]
r_c^2 Var(S)	1.8747	1.8562	1.6827	1.7275
	(10.61)	(9.81)	(9.30)	(9.17)
Var(S)	-0.0066	-0.0071	-0.0057	-0.0051
	(-1.74)	(-1.83)	(-1.42)	(-1.26)
Var(A)	0.0095	0.0099	0.0126	0.0125
	(5.22)	(5.45)	(6.34)	(6.22)
Var(ln WW)	0.8627	0.9677	0.8132	0.8636
	(4.28)	(4.70)	(3.57)	(3.67)
Cov(A,S)	-0.0005	-0.0004	0.0005	0.0002
	(-1.62)	(-1.32)	(0.10)	(0.04)
$(\mathrm{Av}(S))^2$	0.0002	-0.0004	-0.0011	-0.0011
	(0.20)	(-0.38)	(-1.15)	(-1.18)
$(\mathrm{Av}(Exp))^2$	-0.0000	-0.0001	-0.0003	-0.0003
	(-0.19)	(-0.60)	(-1.05)	(-1.01)
p	0.1330	0.1848	0.1061	0.1426
	(1.79)	(2.23)	(1.49)	(1.73)
Z	—	-0.0209	—	—
		(-1.50)		
NSD	—	-0.0242	—	-0.0173
		(-1.22)		(-0.89)
Constant	-0.6788	-0.5687	-0.6554	-0.6590
N	49.0	49.0	39.0	39.0
R^2	0.9351	0.9410	0.9722	0.9729
$\overline{R}^2$	0.9221	0.9255	0.9305	0.9300

Note: t-ratio in parentheses.
Source: See Appendix A-2.
[a] Analysis is performed for states with $Z = 0$.
[b] Contains regional dummy variable *NSD* as well as Z.
[c] Contains regional dummy variable *NSD*.

region variables indicates that the added variables do not change the slope coefficients or the t-ratios of the human capital and employment variables. The adjusted coefficient of determination is 93 per cent. The variable for per cent nonwhite always has a positive slope coefficient. Under a type I error of 5 per cent, p is significant in the forty-nine-state analysis, and in the thirty-nine-state analysis when *NSD* is held constant.[34]

34. See footnote 23 of this chapter for an economic interpretation of the effect of the variable "per cent nonwhite."

The regional dummy variable (*NSD*) and the control dummy variable for separate race data (*Z*) are not separately significant under a two-tailed test, 10 per cent type I error.[35] Taken together, however, they are significant at a 5 per cent level.[36]

To sum up, the model relating human capital and employment parameters to the inequality of income explains statistically ($\bar{R}^2$) 92 per cent of interstate variations in income inequality and 75 to 84 per cent of interstate variations in earnings inequality for white males. The rate of return-schooling inequality interaction variable and the variances of age and log of weeks worked are the most important and significant variables. In addition, they carry the expected positive signs. The covariance of schooling and age and the levels of schooling and experience are generally not significant. The variance in schooling has the hypothesized negative slope, but is significant only in the forty-nine-state regressions.

A significant South–non-South difference in inequality among whites exists for the states before, but not after, controlling for the human capital and employment variables.

Regression analyses for white males—as well as nonwhite males and all males—were also performed for the states within the non-South and within the South. Although the number of degrees of freedom becomes quite small, especially in the South, the qualita-

35. The variance of the log of income for white males:

	Non-South (32 states)	South (17 states)
Mean	0.7206	0.8815
Standard Deviation	0.0762	0.0972

The student's t-ratio for the difference in mean variances between the South and non-South is $t = 5.7$. Thus, there are significant regional differences in inequality among whites before but not after controlling for the human capital, employment, and race variables. (See Appendix A-2.)

36. If R^2 (with *Z* and *NSD* not included) = .9351, then $\sigma^2(U)$ (with *Z* and *NSD* not included) = $(.0649)\sigma^2(\ln Y)$, and if R^2 (with *Z* and *NSD* included) = .9410, then $\sigma^2(U)$ (with *Z* and *NSD* included) = $(.0590)\sigma^2(\ln Y)$. The incremental explanatory power of *Z* and *NSD* is $R^2 = 1 - [\sigma^2(U)$ (with *Z* and *NSD* included)$]/[\sigma^2(U)$ (with *Z* and *NSD* not included)$] = 1 - (.0590)/(.0649) = 1 - .909 = .091$. Then $R = .301$, which is significant at the 5 per cent but not at the 2.5 per cent level. $R^2 = .091$ means that *Z* and *NSD* explain 9 per cent of the variation in the dependent variable otherwise unexplained after the human capital, employment, and race (*P*) variables are held constant. (See Table 8-5.)

tive findings are essentially the same as for the country as a whole.[37]

U.S. Nonwhite Males

Most interstate studies of nonwhite incomes have focused on the proportion of nonwhites above some poverty line or on the ratio of nonwhite to white incomes. There appears to be little previous research on regional differences in the inequality of the distribution of personal nonwhite income.[38] This section analyzes interstate differences in nonwhite income and earnings inequality for the coterminous thirty-nine states. The analysis is performed in the same manner as the thirty-nine-state analysis of white income and earnings inequality. The regression estimate of the rate of return is computed for nonwhite males in each state.

It would be useful at this point to review some expected differences in the effects of the explanatory variables on white and nonwhite income inequality developed at the beginning of this chapter. Accordingly, compared to whites, nonwhites appear to invest less in postschool training, may receive a lower rate of return from this training, and may have experienced a greater improvement in the quality of schooling and job opportunities in the decade or two prior to 1960 due to a decline in discrimination. These forces are consistent with the observed flatter experience-earnings profiles estimated from cross-sectional data for nonwhites than for whites. Then, the inequality of age will have a smaller and less significant effect on income inequality for nonwhite males.

Less investment in specific training by nonwhites implies a lower correlation of weekly wages with weeks worked. In addition, nonwhites appear to be subject to greater seasonality of employment, which implies a negative correlation of weekly wages with weeks worked. The elasticity of annual earnings with respect to the fraction of weeks worked (γ) is, therefore, expected to be

37. The regression results for the interstate analysis of white male and nonwhite male income inequality within the non-South and within the South are presented in Appendix B of my "Racial Discrimination in the Labor Market: A Test of Alternative Hypotheses," *Journal of Political Economy*, November-December 1973, pp. 1,330-1,352.

38. One study of regional differences in the inequality of income among nonwhites is Sharon M. Oster, "Are Black Incomes More Unequally Distributed?" in *The American Economist*, Fall 1970, pp. 6-20.

lower for nonwhites than for whites, and may even be lower than unity.[39]

The slope coefficient of the rate of return-schooling inequality interaction term will be less upward-biased for nonwhites than for whites. The slope coefficient, hypothesized to be equal to 1, was found to be greater than 1 for all males and white males, presumably because the regression estimate of the rate of return (r_c) is downward-biased. The downward bias is greater, the greater the correlation of schooling with experience and with returns from investments in experience. The flatter nonwhite age-earnings profile implies a smaller downward bias in r_c. Hence, less of an upward bias in the rate of return-schooling interaction term is expected for nonwhites.

The algebraic value of the covariance of schooling and age is expected to be lower for nonwhites than for whites if nonwhites invest in less postschool training and are more heavily concentrated in jobs where productivity declines with age. Also, if the quality of schooling and job opportunities for young nonwhites has been rising faster over time than for young whites, the slope of the covariance term will be lower for nonwhites.

Race and region variables are also introduced. The latter tests for regional differences in income inequality when the human capital, employment, and race variables are held constant. The race variable tests the effect of the proportion of nonwhites in the labor force on the income inequality among nonwhites when the other variables are held constant.[40]

Columns (1) and (2) in Table 8-6 present the regression results. The model has a high explanatory power; approximately 85 per cent of interstate differences in the inequality of income and earnings for nonwhites is attributable to the model's variables.[41]

39. A unitary elasticity implies a zero correlation between weekly wages and weeks worked. A positive (negative) correlation means an elasticity greater (less) than unity.

40. A positive slope coefficient for this variable is consistent with the hypothesis that nonwhite workers act as if they had a taste for discrimination against white workers. See footnote 23 above.

41. For the thirty-nine states, the adjusted coefficients of determination ($\overline{R}^2$) are:

	White	Nonwhite
Var(ln$Y$)	.93	.86
Var(ln$E$)	.84	.84

Source: Tables 8-4 and 8-6.

TABLE 8-6

Regression Results, Income Inequality, Nonwhite Males

Independent Variable	Dependent Variable			
	Income Inequality (1)	Earnings Inequality (2)	Income Inequality[a] (3)	Income Inequality[b] (4)
r_c^2 Var(S)	0.9431	0.9125	1.1169	1.0884
	(3.02)	(2.71)	(3.32)	(3.00)
Var(S)	-0.0013	-0.0036	-0.0011	-0.0010
	(-0.34)	(-0.88)	(-0.30)	(-0.25)
Var(A)	0.0033	0.0024	0.0023	0.0024
	(2.13)	(1.41)	(1.31)	(1.31)
Var(ln WW)	0.5952	0.6704	0.5081	0.5188
	(5.75)	(5.99)	(4.15)	(3.92)
Cov(A,S)	-0.0124	-0.0109	-0.0111	-0.0107
	(-3.08)	(-2.50)	(-2.70)	(-2.40)
(Av(S))^2	0.0004	0.0004	0.0005	0.0005
	(0.55)	(0.53)	(0.72)	(0.70)
(Av(Exp))^2	-0.0000	0.0000	0.0000	0.0000
	(-0.28)	(0.30)	(0.28)	(0.17)
p	—	—	-0.1160	-0.1243
			(-1.30)	(-1.28)
NSD	—	—	—	0.0059
				(0.24)
Constant	-0.0599	-0.0165	0.0329	0.0314
N	39.0	39.0	39.0	39.0
R^2	0.8843	0.8711	0.9436	0.9437
$\overline{R}^2$	0.8582	0.8420	0.8612	0.8567

Note: t-ratio in parentheses.
Source: See Appendix A-2.
[a]Includes the variable per cent nonwhite (p).
[b]Includes the variables per cent nonwhite (p) and the North-South dummy variable (NSD).

The rate of return-schooling inequality interaction term has a significant positive slope. The slope coefficient is not significantly lower than unity, but is significantly lower than the white slope.[42]

The variance in age has an insignificant effect for earnings inequality but is significant for income inequality. It will be seen below, however, that there, too, the coefficient becomes insignificant ($t = 1.31$) when the racial composition and region variables

42. The student t-ratio is $t = (b_w - b_n)/[\text{Var}\ (b_w) + \text{Var}\ (b_n) - 2\ \text{Cov}\ (b_w,b_n)]^{1/2}$. The correlation of b_w on b_n is positive but less than unity. For the income analysis and the two extremes of $R(b_w,b_n) = 0$ and $R(b_w,b_n) = 1$, $t = 2.23$ and $t = 5.28$, respectively. There are 62 degrees of freedom. (Source: Tables 8-4 and 8-6.)

are added. The slope coefficient of the variance in age is significantly lower for nonwhite males than for white males.[43] The levels of schooling and experience have no effect on state differences in income or earnings inequality.

As to log of weeks worked, the partial effect of the variance is positive and significantly less than unity. The magnitude of γ is smaller than that for whites (compare Table 8-4 with 8-6), but the difference does not appear to be significant.[44] The implied elasticity of earnings with respect to weeks worked for nonwhites is .77 for income and .82 for earnings. The results imply a tradeoff between weekly wages and weeks worked per year.

The covariance of age and schooling, producing an insignificant coefficient for white males, here has a highly significant negative effect for both the earnings and income analyses. The coefficient for nonwhites is lower than the coefficient for whites, especially in the income equation.[45] This suggests that it was not solely greater sampling or measurement error in the nonwhite than in the white data that led to the less significant or nonsignificant coefficients observed in the other explanatory variables in the nonwhite analysis.

Columns (3) and (4) of Table 8-6 contain the regression analysis for nonwhites when race and region variables are added to the income inequality analysis.[46] The variance in age becomes insignificant, and, although the significance of the covariance of

---

43. Since the slope coefficients of the variance in age for whites and nonwhites are likely to be positively correlated, the assumption of $\mathrm{Cov}(b_w, b_n) = 0$ gives a downward bias to the t-ratio. However, when this assumption is made for the income inequality of white and nonwhite males (thirty-nine states) the t-ratio for the difference in slope coefficients of the variance in age is $t = 3.92$. (Source: Tables 8-4 and 8-6.)

44. If it is assumed that the correlation between the white and nonwhite slope coefficient is zero, the t-ratio for the difference in slope coefficients is $t = 0.71$. If the correlation is assumed equal to unity, $t = 1.41$. With 62 degrees of freedom, under a one-tailed test (because of the hypothesis of a lower coefficient for nonwhites), the latter t-ratio is significant at a 10 per cent level. (Source: Tables 8-4 and 8-6.)

45. Since the covariance of the slope coefficients b_w and b_n is likely to be positively but not perfectly correlated across the states, the t-ratio is downward biased if it is assumed $R(b_w, b_n) = 0$. For the income inequality analysis of white and nonwhite males (thirty-nine states), $t = 2.29$ when it is assumed $\mathrm{Cov}(b_n, b_w) = 0$. For earnings inequality, $t = 0.71$ when $R(b_w, b_n) = 0$ and $t = 3.33$ when $R(b_w, b_n) = 1$. (Source: Tables 8-4 and 8-6.)

46. No separate nonwhite data were presented in the census for the ten states for which the dummy variable Z was given a value of 1 in the white analysis.

age and schooling declines, it remains negative and significant. The slope of the relative inequality of employment variable remains greater than zero, but less than unity.

The race and region variables are not significant either separately or as a set. There seems to be no correlation between the relative number of nonwhites in a state and the income inequality among nonwhites, and no regional difference in nonwhite income inequality after controlling for interstate differences in human capital and employment variables.

The mean level of the nonwhites' inequality of income in the thirty-nine states with separate race data is 0.69, compared with 0.80 for whites. The difference is statistically significant. We could ask what the nonwhite inequality would be if nonwhites had the same mean value of the explanatory variables as whites, or if they had the same values for the slope coefficients. Following are the results of such projections for the mean value of income inequality in the thirty-nine states:[47]

(a) Nonwhite means, nonwhite coefficients (observed):	0.6909
(b) White means, nonwhite coefficients (predicted):	0.6949
(c) Nonwhite means, white coefficients (predicted):	0.7742
(d) White means, white coefficients (observed):	0.8005

For nonwhites the effect of racial differences in the mean values of the explanatory variables is small. Three-fourths of the racial difference in income inequality is due to differences in the slope coefficients of the regression analysis.

Weeks worked during the year show significantly greater inequality for nonwhites than for whites. Since nonwhites have a smaller inequality of annual income, it follows that they necessarily have a smaller inequality of weekly wages. This smaller inequality in weekly wages appears to be largely a consequence of flatter cross-sectional experience-earnings profiles and of the lower rate of return from schooling.

To summarize, the human capital and employment model is successful in explaining nonwhite income and earnings inequality—85 per cent of interstate differences in nonwhite inequality is explained by the explanatory variables. The rate of return-schooling inequality interaction term, the covariance of schooling and age, and the inequality of weeks worked are all significant variables. The inequality of age has a weak positive effect due to the fairly

47. See Table 8-7 for (b) and (c); (a) and (d) are the observed average income inequalities.

TABLE 8-7

Predicted Values of Nonwhite Income Inequality under Alternative Assumptions

	Predicted Income Inequality A[a]		Predicted Income Inequality B[b]	
	White Coefficient	Nonwhite Means	Nonwhite Coefficient	White Means
r_c^2 Var(S)	1.7600	0.0693	0.9431	0.1475
Var(S)	-0.0025	14.6772	-0.0013	13.0627
Var(A)	0.0131	117.9030	0.0033	118.2932
Var(ln WW)	0.7749	0.2440	0.5952	0.1192
Cov(A,S)	0.0014	-13.7768	-0.0124	-10.3256
$(\text{Av}S)^2$	-0.0008	65.9635	0.0004	107.9427
$(\text{Av}(Exp))^2$	-0.0002	849.1903	-0.0000	759.1740
Constant	-0.8128	1.0000	-0.0599	1.0000
Predicted Value	0.7742		0.6947	

Sources: Tables 8-4 and 8-6, and Appendix A-2.
[a]White coefficients (from Table 8-4, col. 2) and nonwhite means.
[b]Nonwhite coefficients (from Table 8-6, col. 3) and white means.

flat age-income profile. The differences in the slope coefficients of the explanatory variables in the white and nonwhite analyses are in the expected directions. Three-fourths of the racial difference in income inequality is explained by different slope coefficients. In spite of larger inequality in weeks worked during the year, intrastate inequality is smaller for nonwhites than for whites. This is due to the flatter nonwhite experience-earnings profile and lower rate of return from schooling.

Canada

Utilizing the 1961 Census of Canada, we can see how the distributions of schooling and age affect the income inequality of Canada's nonfarm male population between the ages of twenty-five and sixty-four.[48] Our analysis is performed for the ten provinces and the Yukon Territory—referred to as the eleven provinces. Although the small number of degrees of freedom reduces the meaningfulness of tests of significance, the analysis is

48. Appropriate data are not available for the calculation of the variance in the log of weeks worked. See Appendix A-2 for a discussion of the Canadian data.

TABLE 8-8

Regression Results, Income Inequality, Canada

Independent Variable	Dependent Variable: Var(ln *Y*)					
	10 Provinces			11 Provinces		
	(1)	(2)	(3)	(4)	(5)	(6)
r_c^2 Var(S)	4.2189 (4.43)	4.1401 (4.94)	2.5866 (0.81)	4.1970 (4.63)	4.1505 (5.25)	2.7452 (1.77)
Var(S)	—	—	-0.0303 (-0.69)	—	—	-0.0290 (-0.89)
Var(A)	—	-0.0057 (-0.57)	0.0039 (0.12)	—	-0.0013 (-0.40)	0.0021 (0.23)
Cov(A,S)	—	-0.0238 (-2.07)	-0.0238 (-1.42)	—	-0.0216 (-2.18)	-0.0237 (-1.65)
$(\mathrm{Av}(S))^2$	—	—	-0.0053 (-0.62)	—	—	-0.0050 (-0.96)
$(\mathrm{Av}(Exp))^2$	—	—	-0.0008 (-0.61)	—	—	-0.0008 (-0.71)
Constant	0.2636	0.7390	1.2554	0.2679	0.2947	1.3683
df	8	6	3	9	7	4
R	0.8426	0.9121	0.9221	0.8394	0.9088	0.9222
$\overline{R}^2$	0.6737	0.7479	0.5511	0.6718	0.7513	0.6260
R^2	0.7100	0.8319	0.8503	0.7046	0.8259	0.8505

Note: t-ratios in parentheses.
Source: Appendix A-2.

still of considerable interest. The findings for Canada are similar to those obtained for the United States.

Of the three models presented in Table 8-8, the abbreviated version (which includes only the rate of return-schooling inequality interaction term, the variance of age, and the covariance of age and schooling) has the highest explanatory power—75 per cent—after adjusting for degrees of freedom. This high explanatory power was obtained without including the dispersion in employment (weeks worked) during the year, a highly significant variable in the U.S. analysis that, were it available, would presumably be important in Canada as well.

In the abbreviated version of the model (columns (2) and (5)), the expected positive effect of the rate of return-schooling inequality interaction variable emerges.[49] The covariance of age and schooling has a negative effect. However, the inequality of age has no effect.

49. The lack of significance of this variable in the full regression (columns (3) and (6)) is due to the large increase in its standard error, presumably due to multicollinearity and the small sample size (see Table 8-9).

TABLE 8-9

Correlation Matrix for Canada
(11 observations)

	Var(ln *Y*)	r^2 Var(*S*)	Var(*S*)	Var(*A*)	Cov(*A*,*S*)	$(\overline{S})^2$
r^2 Var(*S*)	0.8225					
Var(*S*)	-0.1883	-0.3883				
Var(*A*)	-0.0434	0.1929	-0.3682			
Cov(*A*,*S*)	-0.3207	-0.0392	-0.3831	-0.0322		
$(\overline{S})^2$	-0.5965	-0.6617	0.1555	0.1291	-0.2601	
$(\overline{A}$-$\overline{S}$-$5)^2$	0.4043	0.6604	-0.6170	0.5070	0.3298	-0.7055

Source: Appendix A-2.

SUMMARY

The final chapter of my study is devoted to interregional differences in the relative inequality of income. The model is tested empirically for the states of the United States and the provinces of Canada on the theoretical framework of a human capital model of income generation. The human capital earnings equation presented in Chapter 6 relates the natural logarithm of annual income of an individual to his level of schooling (S), rate of return from schooling (r), age (A), and log of weeks worked in the year ($\ln WW$). In the first part of Chapter 8 the variance of both sides of the equation is computed, and hypotheses are developed concerning the relationship between the explanatory variables and the dependent variable.

The variance of the log of income or earnings—a commonly used measure of inequality—is the dependent variable. The independent variables are: a rate of return-schooling inequality interaction term (r^2 Var (S)), the inequality of schooling (Var (S)), the inequality of age (Var (A)), the covariance of schooling and age (Cov (A,S)), the relative inequality of weeks worked (Var ($\ln WW$)), and the squared levels of schooling ($\overline{S}^2$) and years of labor market experience [$\overline{EXP}^2 = (\overline{A} - \overline{S} - 5)^2$].

The arguments behind the hypothesized signs of the regression slope coefficients are reviewed. The rate of return-schooling variance interaction term, the variance of age, and the relative variance of weeks worked are expected to have positive effects on income inequality, but of a lower magnitude for nonwhite males than for white males. Several factors influence the coefficient of the covariance of schooling and age; although no hypothesis is offered as to the sign, the algebraic value is predicted to be lower for nonwhites than for whites.[50]

Two measures of the dependent variable are employed for the United States: the variance of the log of income of males twenty-five years of age and over, and earnings of males fourteen years of age and over. Neither measure is a perfect fit to the theoretical concept of income—the labor market income (earnings) of males who have completed their schooling. The low bound to the upper open-end interval ($10,000 in 1959) reduces the effect of large nonlabor incomes on relative income inequality. The independent variables are defined for males between twenty-five and sixty-four years of age.

In the analysis for all males, the human capital and employment variables have a high explanatory power (R^2): 88 per cent for income inequality and 77 per cent for earnings inequality. The rate of return-schooling variance interaction term and the inequalities of age and log of weeks worked have the predicted positive effects, and are significant. The covariance of age and schooling and the squared levels of schooling and experience have negative slope coefficients. Variables designed to reflect the racial composition of males (per cent nonwhite) in a state and region (South or non-South) are not significant, and have no effect on the slope coefficients of the human capital and employment variables. Thus, interstate differences in the independent variables explain the observed larger income inequality in the South.

Applied to an analysis of the income and earnings inequality of white males, the model shows a high explanatory power ($\bar{R}^2$)—92 per cent for income inequality and approximately 80 per cent for earnings inequality. The rate of return-schooling variance interaction variable and the inequalities of age and log of weeks worked have significant positive effects. The covariance of schooling and age and the squared levels of schooling and experience, which show significant negative effects for all males, are generally not significant for white males. The elasticity of earnings with respect to the fraction of weeks worked does not differ from unity.

When a race composition and a South–non-South dummy variable are added to the analysis of white income inequality, the former has a positive and significant effect, while the latter has no effect. The explanatory variables account for the larger inequality of white income in the South.

---

50. This is the only variable where the model hypothesizes that the nonwhite slope may be further from zero than the white slope.

A separate analysis is performed for interstate differences in income and earnings inequality of nonwhite males. Once again, the model has a high explanatory power, 85 per cent. For nonwhites, the rate of return-schooling inequality interaction variable and the inequality of weeks worked have positive and highly significant effects, while the variance of age has a positive but generally insignificant effect. The hypothesized lower nonwhite slopes are found as expected, and the white-nonwhite differences are significant, except for the employment variable. The estimated elasticity of income (or earnings) with respect to weeks worked is significantly less than unity, 0.77 for income and 0.82 for earnings. This implies a tradeoff between weekly wages and weeks worked per year.

The slope coefficient of the covariance of age and schooling is negative and significantly lower than the coefficient found in the analysis for whites. This racial difference has an important implication: whereas the lower nonwhite slopes for several variables predicted by the human capital model are also consistent with greater measurement or sampling error in the nonwhite data, this interpretation is not tenable for the slope of the covariance of age and schooling. This provides additional support for the efficacy of the human capital analysis of white and nonwhite incomes.

The slope coefficients of the squared levels of schooling and experience are not significant in the analysis for nonwhites.[51] This suggests that the significant negative slopes obtained in the all-male analysis are due to the interaction of race with the model's variables, and that merely controlling for the proportion of nonwhites within a state is not sufficient for holding constant the effects of racial differences in the variables under study. Race composition of a state and region of the state is found to have no effects on nonwhite income inequality.

51. Except for the negative slope in respect to experience level, these slopes were insignificant also in the analysis for white males. The negative slope for experience in the earnings analysis (earnings for males fourteen years of age and over) is interpreted as being a consequence of large investments in human capital for those males with some earnings in the age group fourteen to twenty-four—many of whom are students. The relative inequality of earnings in this group is large and the level of earnings is lower than for males between twenty-five and sixty-four. The independent variables, however, are defined for males in the twenty-five to sixty-four group. Of these the most likely to capture the youth effect on income inequality is the level of experience, which will be lower in states with a younger population. Thus, the significant negative slope for whites—but not for nonwhites—for the squared level of experience in the earnings regression is consistent with a steeper experience-income profile for whites.

The observed income inequality within the states is smaller for nonwhites than for whites. It is not due to differences in the distributions of the explanatory variables. Three-fourths of the racial difference stems from the smaller values of the regression slope coefficients. Since the inequality in weeks worked is larger for nonwhites, it follows that nonwhites have a smaller intrastate inequality of weekly wages. This smaller inequality in weekly wages is primarily caused by a flatter experience-earnings profile and lower rate of return from schooling.

The findings of the interprovincial analysis of income inequality for nonfarm males in Canada are similar to those for nonwhites in the United States. The model's best explanatory power is 75 per cent. The rate of return-schooling inequality interaction term has a positive effect, the inequality of age has no effect, and the covariance of schooling and age has a negative effect.

The theoretical analysis in the first part of this chapter suggests that the algebraic value of the effect on income inequality of the covariance of schooling and age would be lower, the smaller the amount of, and rate of return from, investments in postschool training and the greater the secular increase in schooling level and the depreciation of earning potential with age. These forces also tend to reduce the effect of the inequality of age on income inequality, and may well cause the age effects on income inequality for nonwhite U.S. males and Canadian males.

Thus, based on the findings of this chapter, the human capital approach appears to be highly successful in explaining regional and racial differences in the inequality of income among males.