

Essays and Data on
American
Ethnic Groups

Edited by
Thomas Sowell
with the assistance
of Lynn D. Collins

Ethnic Income Variations: Magnitudes and Explanations

ERIC HANUSHEK

Income variations among ethnic groups exist within a wider context of general income inequalities in the American economy, and among the various nations of the world. Indeed, similarly obvious and dramatic income disparities existed in ancient Rome. Interpreting interethnic differences in incomes is complicated by the fact that ethnic groups differ systematically in many ways other than ethnicity per se. This chapter attempts to sort out some of the factors affecting ethnic differences in the distribution of incomes.

As an illustration of persisting income differentials in the United States, table 1 shows the percent of aggregate income received by each fifth of all U.S. families (ranked by total income) for each decade since 1950. The top fifth of families, in terms of income, received over 40 percent of the total income while the bottom fifth received around 5 percent. And, this skewed distribution has remained virtually unchanged over the past two decades. Interestingly, such disparities exist even within sub-populations. Considering just families with a white head

Table 1

PERCENT OF AGGREGATE INCOME RECEIVED BY
EACH FIFTH OF U.S. FAMILIES: 1950, 1960, 1970

Income Rank	1950	1960	1970
Lowest Fifth	4.5	4.9	5.5
Second Fifth	12.0	12.0	12.0
Middle Fifth	17.4	17.6	17.4
Fourth Fifth	23.7	23.6	23.5
Highest Fifth	41.6	42.0	41.6

SOURCE: U.S. Bureau of the Census, *Statistical Abstract of the United States*, 1973 (Government Printing Office, 1973), p. 330.

for 1970, the top fifth captured 41 percent of white income while the bottom fifth managed only 5.8 percent; for the black population, the corresponding figures were 44 percent and 4.5 percent.¹

This aggregate income distribution can be collapsed to observe racial disparity in incomes. As table 2 shows, median family income of whites and nonwhites is nowhere near being equal. (While it would be desirable to analyze additional ethnic groups, this study will focus primarily upon black-white differences because of data availability. Where possible, "Spanish Americans"—defined as Mexican Americans and Puerto Ricans combined—will also be studied. However, it is not possible to go further in disaggregating ethnic groups.) Even though nonwhite income has improved relative to white income in recent years, nonwhite income remained less than two-thirds of white income. Moreover, such racial differences are found at points other than the median of the income distribution.

Interest in the racial and ethnic aspects of the distribution of income arises from several obvious factors. First, there are clear moral and legal interests that revolve around discriminatory treatment. Second, it is a type of comparison which we can more comfortably and more validly make than comparisons of individual incomes. At the individual level, personal characteristics such as innate ability, training, experience, attitudes, and personality along with luck and good fortune become overwhelmingly important.² But, for a racial or ethnic group as a whole these factors—often unmeasurable—either may be reasonably pre-

sumed to be similar across the group or the differences must be the subject of public policy consideration.

Nevertheless, the significance and interpretation of racial and ethnic differences in incomes depend importantly upon understanding the reasons behind observed differences. Most important among the potential underlying factors that have been considered has been education. The distribution of education has been labeled as both a significant cause of the current income dis-

Table 2

MEDIAN INCOME OF FAMILIES BY RACE OF HEAD,
1947 TO 1974

(Current Dollars)

Year	White	Negro and Other	Ratio Nonwhite to White
1947	\$3,157	\$1,614	.51
1948	3,310	1,768	.53
1949	3,232	1,650	.51
1950	3,445	1,869	.54
1951	3,859	2,032	.53
1952	4,114	2,338	.57
1953	4,392	2,461	.56
1954	4,339	2,410	.56
1955	4,605	2,549	.55
1956	4,993	2,628	.53
1957	5,166	2,764	.54
1958	5,300	2,711	.51
1959	5,643	2,917	.52
1960	5,835	3,233	.55
1961	5,981	3,191	.53
1962	6,237	3,330	.53
1963	6,548	3,465	.53
1964	6,858	3,839	.56
1965	7,251	3,994	.55
1966	7,792	4,674	.60
1967	8,274	5,141	.62
1968	8,937	5,590	.63
1969	9,794	6,191	.63
1970	10,236	6,516	.64
1971	10,672	6,714	.63
1972	11,549	7,106	.62
1973	12,595	7,596	.60
1974	13,356	8,265	.62

SOURCE: U.S. Bureau of the Census, "Money Income in 1974 of Families and Persons in the United States," *Current Population Reports*, Series P-60 (Government Printing Office, 1976).

tribution and a focal point for changing the distribution. Further, in looking at the racial distribution of income, many have pointed to the large racial differences in both quantity and quality of schooling. These differences are assumed to be an important cause for racial income discrepancies and are thought of as a natural and compelling place for governmental action, partially because a basic societal premise has been the right to free and equal educational opportunity.

Table 3 shows the degree of difference in schooling by age and race. The differences in median schooling are most dramatic for older members of the population, but they remain significant for even the 25- to 29-year-old group. Blacks consistently receive less schooling.

As dramatic as these differences appear, the educational disparity portrayed is an understatement of reality. Data for 1965 from the "Coleman report" show the equally dramatic quality disparity in schooling.³ As seen in table 4, twelfth-grade blacks in the rural South perform at the seventh-grade level of whites in the urban Northeast. Similar though not as large differences exist for other regions.

Table 3

YEARS OF SCHOOL COMPLETED BY RACE AND AGE,
1972 (PERSONS 25 YEARS AND OLDER)

Race and Age	Percent of Population Completing							Median yrs.
	Elementary			High School		College		
	0-4 yrs.	5-7 yrs.	8 yrs.	1-3 yrs.	4 yrs.	1-3 yrs.	4 or more yrs.	
All Races	4.6	8.3	11.9	17.0	35.2	10.9	12.0	12.2
25-29 yrs.	.8	2.7	3.1	13.6	43.8	17.0	19.0	12.7
30-34 yrs.	1.4	3.3	4.7	16.7	43.9	13.5	16.5	12.5
35-44 yrs.	2.5	5.4	6.9	18.5	41.5	11.5	13.8	12.4
45-54 yrs.	3.4	7.3	10.8	18.6	38.6	10.6	10.6	12.3
55 and over	9.0	14.3	21.2	16.4	23.6	7.6	7.9	11.6
Negro	12.8	17.0	9.6	24.0	24.9	6.5	5.1	10.3
25-29 yrs.	1.3	5.6	4.9	24.1	42.9	13.0	8.2	12.3
30-34 yrs.	1.6	6.4	6.3	29.4	40.5	8.4	7.4	12.2
35-44 yrs.	5.4	12.7	8.4	31.9	28.7	7.3	5.7	11.2
45-54 yrs.	10.5	20.0	12.1	26.2	21.2	5.9	4.1	9.9
55 and over	29.7	27.8	12.3	14.5	10.1	2.6	2.9	7.2

SOURCE: U.S. Bureau of the Census, *Statistical Abstract of the United States: 1973* (Government Printing Office, 1973), p. 118.

Table 4

ACHIEVEMENT DIFFERENTIALS BY RACE AND REGION:
AVERAGE GRADE LEVELS BEHIND THE AVERAGE
WHITE IN THE METROPOLITAN NORTHEAST,
VERBAL ABILITY

Race and Region	Grade Levels Behind		
	Grade 6	Grade 9	Grade 12
White, Nonmetropolitan			
South	.7	1.0	1.5
Southwest	.3	.4	.8
North	.2	.4	.9
White, Metropolitan			
Northeast	—	—	—
Midwest	.1	.0	.4
South	.5	.5	.9
Southwest	.5	.6	.7
West	.3	.3	.5
Negro, Nonmetropolitan			
South	2.5	3.9	5.2
Southwest	2.0	3.3	4.7
North	1.9	2.7	4.2
Negro, Metropolitan			
Northeast	1.6	2.4	3.3
Midwest	1.7	2.2	3.3
South	2.0	3.0	4.2
Southwest	1.9	2.9	4.3
West	1.9	2.6	3.9
Mexican American	2.0	2.3	3.5

SOURCE: James S. Coleman et al., *Equality of Educational Opportunity* (Government Printing Office, 1966), table 3.121.1.

This chapter analyzes the strength of these underlying relationships and their implications for earnings distributions. First, a systematic appraisal of the relationship between schooling and earnings is undertaken. This appraisal relies upon the data from the 1970 Census of Population. Second, the analysis considers other basic factors which influence incomes—namely, experience and labor market locations. Finally, the racial and ethnic disparity in incomes is analyzed to determine the relative importance of input differences (such as in tables 3 and 4) and of the discriminatory aspects of labor markets.

I. MODELS AND DATA

The objective of this analysis is to gain a better understanding of the process of income determination. Such understanding, in turn, allows consideration of alternative strategies for changing the current distribution of income. The determination of individual incomes clearly cannot be understood by simply looking at the average incomes obtained by a person of a given ethnic group or given schooling class. The process is too complicated for that, and the apparent influence of any single factor is masked by the correlations among groups of factors. For example, while we found that blacks on the average earned less than whites, we also found that blacks on the average have less schooling, that blacks tend to be more concentrated in the South, and so forth. For that reason we must go to more complicated models and more sophisticated analytical techniques than simple comparisons of average earnings.

The basic conceptual model for looking at income determination relied upon in this analysis is an offshoot of the human capital model. This model, developed over the past 15 years, views schooling and training as an investment.⁴ Using the analogy of investing in physical capital, a person can increase his future earnings capacity by investing in himself. A corollary of this model is that earnings can be viewed as a function of the past history of human capital investments.⁵

There are, of course, many nuances to this basic model because there are many ways in which one can invest in himself. Investment can be made through formal schooling, on-the-job training, health care, or migration to a different region, to name a few. It is also possible to view the costs of such investments as coming from a variety of sources: tuition payments, direct cash outlays, or foregone earnings.

With varying allowance for some of these factors, this basic model relating earnings to schooling and other attributes has been estimated in a variety of situations. Relying chiefly upon cross-sectional data, these studies have produced estimates of the value of additional schooling and of other types of investments. These estimates, in turn, have been used: (1) to predict how additional schooling might affect the income distribution; and, (2) to say something about the relative value of schooling (and discrimination) across races and ethnic groups.

Without going into the details of either human capital models or the offshoots of such models, it is instructive to note that

these models are really reduced form models. The reasons why investments in schooling and training raise the future incomes of individuals are not addressed directly, but an underlying presumption is that the returns to additional schooling or investment reflect both the demands for skills and training and the relative supplies of these attributes. As such, the advisability of further investment from the individual's perspective must be conditioned by both the current and future balance of these underlying factors.⁶

Calculating the effect on earnings of an increase in schooling for the entire country and projecting these results into the future assumes that the historical relationship between supply and demand for skilled labor will be maintained. Further, it implies that it is reasonable to think of a national labor market with a single wage for individuals of different schooling categories. Both of these presumptions are open to considerable question.

Studies of schooling-earnings relationships relying upon cross-sectional data for different samples and different periods of time have yielded significantly different results. These differences have been explained in part by differences in aggregate labor market conditions; i.e., with different demand conditions for skilled labor, one finds different reduced form relationships between schooling and earnings.⁷

Past research into earnings models also has shown large differences in earnings for macrogeographical regions of the country such as the South. While there has not been much investigation of smaller regions, there is, on the other hand, little reason to presume that the gross regional measurements of past studies adequately capture geographical, or labor market variations in earnings relationships.⁸

In recognition of the potential importance of labor market differences—both in describing earnings relationships and in formulating policies with respect to earnings, this analysis is designed to allow labor market interactions with schooling-earnings relationships. The analytical design in this chapter follows directly from that previously considered by the author.⁹ Earnings data are analyzed within individual labor markets, or individual Standard Metropolitan Statistical Areas (SMSAs). By considering the relationship between earnings and various inputs to earnings for each SMSA separately, it is then possible to sort out the relative importance of input factors (such as schooling), of race and ethnic background, and of specific labor market conditions.

Observing earnings within individual labor markets allows testing of whether or not migration and trade between different geographical markets is sufficient to bring the markets into equilibrium with respect to payments to different types of labor; i.e., are the earnings of similar individuals in different regions the same? ¹⁰

The empirical analysis here relies upon the 1970 census data from the Public Use Samples. These data provide a picture of the earnings in 1969, the occupations, and the backgrounds of individuals. Further, the county group data from the Public Use Samples allow identification of individuals by Standard Metropolitan Statistical Area (SMSA) provided there are at least 250,000 people in the SMSA.

The basic criteria for inclusion in the analysis were that: (1) the individual was a male between 16 and 65 years old; (2) the individual had worked full time (35+ hours) and full year (48-52 weeks); (3) his earnings from wages, self-employment, and farming were positive; and (4) he had a known U.S. state of birth. These criteria were applied generally to eliminate certain sources of earnings variations which were not central to this study and thereby to simplify the task of modeling earnings of individuals. The analysis was restricted to males, in part, because of the widely observed differences in male and female career progressions and, in part, because attempts to develop models of female earnings have not been too successful. The age restriction mirrors the accepted working life prescribed by child labor laws and Social Security regulations and represents an attempt to avoid peculiar working relationships, particularly for individuals over 65. The full-time, full-year work restriction was introduced to minimize problems of interactions between work activity and wage rates.¹¹ Earnings were required to be positive both because the models considered apply best to permanent income for which negative values do not make much sense and because negative values are difficult to handle analytically.¹² Finally, since attempts will be made to correct for differences in school quality, it was necessary to know the state of birth for the individual in order to estimate where the education was received.

Within each of the SMSAs identified in the Public Use Sample, several subsamples of observations were created. Individuals were divided by race (into white, black, and Spanish samples) and by years of schooling (into schooling less than or equal to 12 years and schooling greater than 12 years). A subsample was

required to have at least 50 observations (meeting the criteria above) in order to be included in the analysis. Table 5 shows the number of regional subsamples created by race and schooling categories. This table also shows the geographical distribution of these SMSAs by the nine census regions of the country. Table 6 shows the corresponding number of observations in each of these samples. In total, there are over 175,000 observations for analysis.

The mean 1969 earnings for this sample are displayed in table 7.¹³ It is clear from this table that there are large and systematic differences in earnings by race, region, and schooling categories. For example, within the urban Northeast, the average white with a high school education or less earns \$1,900 more than a corresponding black even though they work the same amount; this is a full 27 percent more. Similar disparities are evident across regions and across schooling categories.

Corresponding to these income differences are a number of fundamental differences, as in amount of schooling. Table 8 shows the mean schooling levels by race and region for each of the samples. The "unadjusted" levels give the years of schooling completed in 1970. The "adjusted" values represent a crude attempt to allow for quality differences in the schooling received by each individual. These quality adjustments are based upon the 1965 nationwide testing done to calculate school quality for *Equality of Educational Opportunity* (the "Coleman report").¹⁴ These school quality data, reproduced in table 4 of this chapter, compare the verbal achievement scores by race and region to the performance of whites in the urban Northeast. For each individual, based upon state of birth, years of school completed, and race, a quality-adjusted score was calculated; for schooling past grade 12, the twelfth-grade adjustment was used.¹⁵ As table 8 vividly shows, the distribution of schooling, when quality adjusted, looks much more skewed than it does when not quality adjusted.

Using these data, for SMSAs, the relationship between earnings and human capital was estimated through regression analysis. Human capital, following the work of Mincer, was measured by schooling completed (or school quality), experience, and experience squared.¹⁶ The quadratic experience term permits varying returns to length of experience—a situation which would exist if there was on-the-job training at the beginning of a work career. Individual estimates of such earnings functions were obtained for each race/schooling group (with over 50 observa-

Table 5

NUMBER OF SAMPLED SMSAs BY RACE, SCHOOLING CATEGORY, AND CENSUS REGION

Census Region ^a	White			Black			Spanish		
	S ≤ 12 ^b	S > 12 ^b	Total ^c	S ≤ 12 ^b	S > 12 ^b	Total ^c	S ≤ 12 ^b	S > 12 ^b	Total ^c
Northeast	7	7	7	2		2			
Mid-Atlantic	21	21	21	7	2	7	1		1
E. North Central	25	25	25	10	3	11	1		1
W. North Central	8	8	8	2	1	2			
S. Atlantic	18	18	18	18	2	18	1		3
E. South Central	9	9	9	7		7			
W. South Central	14	14	14	7	1	9	5	1	6
Mountain	5	5	5				4		4
Pacific	18	18	18	2	2	3	8	5	9
All Regions	125	125	125	55	11	59	20	6	24

a. Regional divisions follow census definitions: *Northeast*—Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut; *Mid-Atlantic*—New York, New Jersey, Pennsylvania; *E. North Central*—Ohio, Indiana, Illinois, Michigan, Wisconsin; *W. North Central*—Minnesota, Iowa, Missouri, North Dakota, South Dakota, Nebraska, Kansas; *S. Atlantic*—Delaware, Maryland, D.C., Virginia, W. Virginia, North Carolina, South Carolina, Georgia, Florida; *E. South Central*—Kentucky, Tennessee, Alabama, Mississippi; *W. South Central*—Arkansas, Louisiana, Oklahoma, Texas; *Mountain*—Montana, Idaho, Wyoming, Colorado, New Mexico, Arizona, Utah, Nevada; *Pacific*—Washington, Oregon, California, Alaska, Hawaii.

b. S = Years of schooling completed.

c. The "total" samples for an ethnic group may be larger than the numbers for the schooling subgroup because of the observational cutoff of 50 sample points; i.e., there may be less than 50 in either schooling group but more than 50 if the samples are combined.

Table 6

NUMBER OF OBSERVATIONS IN SAMPLED SMSAs BY RACE, SCHOOLING CATEGORY, AND CENSUS REGION^a

Region	White			Black			Spanish		
	S ≤ 12 ^b	S > 12 ^b	Total ^c	S ≤ 12 ^b	S > 12 ^b	Total ^c	S ≤ 12 ^b	S > 12 ^b	Total ^c
Northeast	7,374	3,987	11,361	131		166			
Mid-Atlantic	26,290	12,678	38,968	2,895	394	3,398	99		114
E. North Central	23,105	11,579	34,684	2,727	427	3,324	131		170
W. North Central	7,060	3,687	10,747	407	65	492			
S. Atlantic	10,848	6,533	17,381	2,895	237	3,316	57		206
E. South Central	4,213	2,040	6,253	859		937			
W. South Central	7,026	4,812	11,838	1,171	76	1,471	966	59	1,195
Mountain	2,089	1,717	3,806				483		599
Pacific	12,754	11,364	24,118	827	346	1,223	1,727	575	2,446
Total	100,759	58,397	159,156	11,912	1,545	14,327	3,463	634	4,730

a. Number of observations based on full-time, full-year workers contained in the 1/100 sample; see sampling criteria in text.

b. S = Years of schooling completed.

c. The "total" samples for an ethnic group may be larger than the numbers for the schooling subgroup because of the observational cutoff of 50 sample points; i.e., there may be less than 50 in either school group but more than 50 if the samples are combined.

SOURCE: 1970 Census of Population (Public Use Sample).

Table 7
1969 SAMPLE MEAN EARNINGS BY RACE, SCHOOLING,^a AND CENSUS REGION^b
 (Hundreds of dollars)

Region	White			Black			Spanish		
	S ≤ 12	S > 12	Total	S ≤ 12	S > 12	Total	S ≤ 12	S > 12	Total
Northeast	\$ 90	\$143	\$108	\$71	\$	\$61	\$	\$	\$
Mid-Atlantic	93	148	111	68	96	72	76		66
E. North Central	99	140	113	74	97	78	84		71
W. North Central	94	130	106	65	86	68			
S. Atlantic	89	142	109	57	85	60	81	126	100
E. South Central	83	127	97	51	62	52			
W. South Central	89	129	105	53	67	55	61	105	69
Mountain	92	125	107	60			68	104	75
Pacific	100	135	117	73	91	78	85	112	92

a. S = Years of schooling completed.

b. Means calculated for full-time, full-year workers in SMSAs; see sampling criteria in text.

SOURCE: 1970 Census of Population (Public Use Sample).

Table 8

1970 MEAN SCHOOLING—ADJUSTED AND UNADJUSTED^a

Census Region	White		Black		Spanish	
	Unadj.	Adj. ^b	Unadj.	Adj. ^b	Unadj.	Adj. ^b
Northeast	12.3	12.3	8.7	5.9		
Mid-Atlantic	12.2	12.2	10.6	7.4	9.0	6.2
E. North Central	12.2	11.8	10.5	7.3	7.9	5.6
W. North Central	12.2	11.9	10.5	7.4		
S. Atlantic	12.2	11.8	9.8	6.7	12.3	9.1
E. South Central	11.8	11.1	9.3	6.2		
W. South Central	12.3	11.7	9.8	6.5	9.5	6.8
Mountain	12.8	12.4			10.6	7.6
Pacific	13.0	12.6	11.2	7.7	11.3	8.2

a. Means calculated for full-time, full-year workers in SMSAs; see sampling criteria in text.

b. Schooling adjusted for estimated quality differences in elementary and secondary schools. Quality differences come from *Equality of Educational Opportunity*, p. 274; see table 4 above and footnote 15.

SOURCE: 1970 Census of Population (Public Use Sample).

tions) in each SMSA.¹⁷ The results are discussed in the next section.

II. PRELIMINARY RESULTS

Estimation of the earnings models for the individual regions and race/schooling subsamples produces a large number of individual results. The discussion here will concentrate on some of the more important summary conclusions.

Table 9 summarizes the overall explanatory power of the basic regional estimation. The first column (R^2) shows the "average" proportion of variance explained within each of the subregions and groups. The next three columns relate to the national variance in earnings (or, more precisely, log earnings). The within region explained variance plus the between region explained variance yields the total explained variance (col. 4).¹⁸ The within region explained variance tells how well the estimated SMSA models do at explaining the differences in individual earnings. The between region explained variance results from the stratification of the sample into different SMSAs; the size of this explained variance is related to the differences in mean earnings across SMSAs. For whites with a high school education or less, 11 percent of the variation in individual incomes is explained by differences in schooling and differences in labor force experience (i.e., the model estimated *within* each region); 3 percent of the variation in individual earnings is accounted for by geographical location (i.e., differences in mean earnings *between* SMSAs; this leaves 86 percent of the variation in earnings for this group unaccounted for by this analysis). As can be seen by the fourth column, this analysis accounts for between 15 and 30 percent of the variation in individual earnings. In other words, there remains a considerable amount in the income generation process that we do not understand, or that at least is not explained by this basic model.

Table 9 has another interesting aspect. The between region explained variance is a measure of how different the regions are in terms of earnings. For each racial group, the difference between regions is more important for those with a high school education or less than for those with more than a high school education. This is most pronounced for black and Spanish males. This is consistent with a hypothesis that the labor market for

Table 9
AGGREGATE EXPLANATORY POWER OF EARNINGS MODELS

Group	R-Squared ^b	Individual Income Variance Explained ^a			No. of SMSAs	No. of Observ.
		Within SMSAs	Between SMSAs	Total		
White						
$S \leq 12$	.115	.112	.031	.143	125	100,759
$S > 12$	.239	.233	.026	.259	125	58,397
Total	.216	.210	.030	.240	125	159,156
Black						
$S \leq 12$	.073	.067	.084	.151	55	11,912
$S > 12$	.147	.143	.027	.170	11	1,545
Total	.099	.091	.086	.177	59	14,327
Spanish						
$S \leq 12$	.172	.153	.110	.263	20	3,463
$S > 12$	.276	.270	.023	.293	6	634
Total	.224	.205	.086	.291	24	4,730

a. The calculations of regional variance explained relate to national variance in income for the given racial/schooling sample. deviations in the dependent variable.

b. R-squared is defined as the aggregate of explained sum of squares in each region divided by the aggregate of sum squared deviations in the dependent variable.

college-educated individuals is more national in character while the labor market for less-educated individuals is more local. Thus, the higher mobility of college-educated people tends to equalize the individual labor markets to a greater extent than is the case for less-educated and less-mobile individuals. This immobility could arise either because of differences in information or differences in the rewards to moving, relative to the costs of moving. There is, however, more to the study of geographic or labor market differences. As we shall see below, we are interested in the shape of the earnings function as well as the level (as given by the SMSA mean earnings for the race/schooling groups).

When looking at earnings functions, most attention is naturally centered upon the returns to additional schooling. The government has traditionally taken an interest in the schooling of the population, and schooling is seen as the most legitimate way for the government to attempt changes in the income distribution. The analysis here concentrates on the earnings function for whites and blacks. Earnings of Spanish Americans (both Mexican American and Puerto Rican) are included with the other analyses, but the sample sizes for this group are generally quite small, and the estimates tend to be unreliable.

For comparisons of the estimated relationship between earnings and schooling, we will concentrate upon the value of an additional year of schooling. Specifically, we will look at the percentage increase in earnings that would be expected from one more year of schooling or, alternatively, one more year of quality equivalent schooling. This provides a summary measure of the value of additional schooling and allows comparisons to be made across race and schooling groups. At the same time, we shall also be looking at the effect of individual labor market conditions on the return to additional schooling.

Table 10 provides some insights into the effect of labor market conditions on earnings of whites. Each of the SMSAs which were sampled are placed in one of the nine census regions so that the range of estimated schooling-earnings relationships can be viewed. Clearly, even within the census regions, there is wide disparity in the returns to schooling. The returns to schooling, or percentage increase in earnings that is related to an additional year of schooling, differ within regions by at least 2.9 percent (for individuals with a high school or less education in the Northeast) but go as high as almost 19 percent in the Pacific

Table 10

RANGE OF ESTIMATED PERCENTAGE INCREASES IN EARNINGS PER YEAR OF
SCHOOLING BY REGION AND SCHOOLING CLASS: WHITES

Census Region	No. of SMSAs	S ≤ 12			S > 12		
		Min.	Max.	Range	Min.	Max.	Range
Northeast	7	2.7	5.6	2.9	8.0	11.0	3.0
Mid-Atlantic	21	2.3	9.3	7.0	.2	13.5	13.3
E. North Central	25	2.2	7.5	5.3	2.2	13.2	11.0
W. North Central	8	3.0	6.4	3.4	6.0	12.6	6.6
S. Atlantic	19	2.6	8.9	6.3	6.7	14.5	7.8
E. South Central	8	3.7	7.9	4.2	4.4	11.9	7.5
W. South Central	14	2.4	9.6	7.2	-.2	15.1	15.3
Mountain	5	-2.8	5.1	7.9	5.3	12.0	6.7
Pacific	19	-2.6	16.2	18.8	4.1	13.0	8.9

region. Some of this variance arises from the sampling distribution of the estimates, and there are even a few negative estimates for the effect of additional schooling.¹⁹ Nonetheless, since the standard errors of these estimates tend to range between .5 and 1.5 for all but the smallest SMSAs, the effect of individual labor market conditions on the schooling-earnings relationship cannot be disregarded.

For black and Spanish individuals, the lesser number of SMSAs precludes meaningfully looking at the ranges of estimates within census regions. However, similar variance among SMSAs appears for these groups. With blacks, the estimated percentage increase in earnings arising from an additional year of schooling ranges from negative to 10.4 percent when schooling completed is 12 years or less and from 8.2 to 11.4 percent for more than 12 years of schooling. With the Spanish samples, the similar ranges are 3.0 to 8.8 percent for high school or less schooling and 7.4 to 17.4 percent for post-secondary schooling. (The problem of large sampling errors is more critical for the black and Spanish samples because more of the SMSAs tend to be close to the 50 observation cutoff and, thus, the estimated relationships are not as precise.)

Previously, we noted the importance of mean differences across SMSAs. Particularly for less educated minorities, a significant portion (8-11 percent) of the variance in individual earnings was explained by overall labor market differences. We now find that the shape of the earnings function also differs significantly by region.

In order to understand better the relationship between schooling, school quality, and earnings, we will shift to a discussion of the mean effect on earnings of increasing schooling. However, throughout these discussions the variance in earnings relationships across regions should not be forgotten. The range of predicted effects within a region-race-schooling group is often larger than the observed differences between racial groups or regions.

Table 11 presents the estimated mean relationships between schooling and earnings for different aggregations of individuals. The values given are the average estimated percentage effect on earnings of an increase of one year in schooling (or quality adjusted schooling) where each SMSA value is weighted by the number of observations in the SMSA. There are several striking features in this table.

Table 11

ESTIMATED PERCENTAGE INCREASE IN EARNINGS
PER YEAR OF SCHOOLING
(Weighted Average of Individual SMSAs)

Sample	No. of SMSAs	Schooling	School Quality ^a
White			
S ≤ 12	125	4.44	4.55
S > 12	125	9.81	9.67
Total	125	7.55	
Black			
S ≤ 12	55	3.66	5.08
S > 12	11	9.02	8.73
Total	59	5.02	
Spanish			
S ≤ 12	20	5.22	7.12
S > 12	6	11.88	11.89
Total	24	6.82	
All Individuals			
S ≤ 12	125	5.24	
S > 12	125	10.11	
Total	125	7.96	

a. Results indicate average coefficients when model is estimated using adjusted schooling, defined in footnote 15, instead of actual schooling.

First, looking at the unadjusted schooling column, the difference in returns to schooling for the high school and less group and the more than high school group is dramatic. For each racial group, the percentage increase in earnings for the more educated group is at least double that of the less educated group; for both blacks and whites, each year of advanced schooling increases earnings by over 5 percent more than does a year of secondary schooling. The existence of a stable differential in the earnings of college educated individuals relative to others has been noted in the past.²⁰ The estimated difference here surprisingly seems even larger than those differences given for past periods.

Second, the effect of quality-adjusting the quantity of schooling each individual receives is noteworthy.²¹ The quality adjustment is, as explained above, quite crude; years of school are adjusted for estimated 1965 quality differences based upon region of birth. This adjustment is most appropriate for younger

individuals with 12 years or less schooling than it is for older individuals and individuals with some college, because the errors in estimating region of education and quality of schooling actually received will be less for the former individuals. The explained variance (R^2) is virtually identical for the adjusted and unadjusted models. The quality adjustment has a very significant effect on the aggregate earnings relationship portrayed. In unadjusted terms, the percentage increase per year of schooling for whites appears generally above that for blacks. When the quality adjustment is introduced, this conclusion no longer holds for the lesser educated individuals. Per year of "constant quality" schooling, black earnings increase by 5.1 percent while white earnings increase by only 4.5 percent. In each of the educational subgroups and for both adjusted and unadjusted, the returns to an additional year of schooling are higher for Spanish than for either white or black.²² The quality adjustment is considerably weaker for post-secondary education, and, as might be expected, the adjustment has less effect on the estimated post-secondary functions.²³

The estimated returns to schooling can also be aggregated over the census regions to provide additional information about the geographical distribution of returns to schooling. Table 12 displays the weighted average of estimated percentage increases in earnings per year of schooling for each of the nine regions. For less educated whites ($S \leq 12$) the returns are highest in the Mid-Atlantic region and in the South (S. Atlantic, E. South Central, and W. South Central). Patterns within the other racial and schooling groups are, however, less discernible.

Table 12 also indicates that, in all but two of the fourteen instances where black-white regional comparisons are possible, the average white returns are higher than the average black returns. Thus, the finding for nationally aggregated returns to schooling holds for almost all regions of the country.

The picture is somewhat different when quality adjustments are made to the schooling data. Table 13 displays the unadjusted and adjusted estimates by region for the high-school-or-less samples. In the adjusted estimates the pattern of high returns to school in the South still holds for whites. However, a similar pattern is now clearly visible for blacks. Moreover, in six of the eight regional black-white comparisons, the black earnings increase is higher than whites for the adjusted coefficients. The interpretation of the adjusted schooling models should be re-empha-

Table 12
ESTIMATED PERCENTAGE INCREASE IN EARNINGS PER YEAR OF SCHOOLING
(Weighted Average of Individual SMSAs)

Census Region ^a	White			Black			Spanish		
	$S \leq 12$	$S > 12$	Total	$S \leq 12$	$S > 12$	Total	$S \leq 12$	$S > 12$	Total
Northeast	4.0	9.9	8.0	-1.9		3.7			
Mid-Atlantic	4.8	10.4	8.3	4.0	9.4	5.5	5.6		7.3
E. North Central	4.2	8.9	6.8	3.5	10.6	4.9	3.0		5.7
W. North Central	4.7	10.1	7.2	4.0	4.2	5.1			
South Atlantic	4.9	10.3	7.8	3.7	9.0	5.2	-3.0		6.6
E. South Central	6.2	9.6	8.2	3.6		4.3			
W. South Central	4.8	9.6	7.6	3.6	5.3	4.2	6.4	17.4	8.0
Mountain	3.4	9.7	6.6				5.3		7.1
Pacific	3.1	9.8	7.1	3.9	8.4	5.1	5.0	11.3	6.2
Total	4.4	9.8	7.6	3.7	9.0	5.0	5.2	11.9	6.8

a. Regions follow Bureau of Census definitions.

Table 13
ESTIMATED PERCENTAGE INCREASE IN EARNINGS PER YEAR OF SCHOOLING—
SCHOOLING $\leq$ 12 YEARS
(Weighted Average of Individual SMSAs)

Census Region ^a	White		Black		Spanish	
	Sch.	Sch. Qual.	Sch.	Sch. Qual.	Sch.	Sch. Qual.
Northeast	4.0	4.0	-1.9	-3.8		
Mid-Atlantic	4.8	4.8	4.0	5.3	5.6	7.5
E. North Central	4.2	4.4	3.5	4.7	3.0	4.0
W. North Central	4.7	5.0	4.0	5.1		
South Atlantic	4.9	5.0	3.7	5.3	-3.0	-3.9
E. South Central	6.2	6.6	3.6	5.4		
W. South Central	4.8	5.1	3.6	5.5	6.4	8.7
Mountain	3.4	3.6			5.3	7.4
Pacific	4.4	4.5	3.7	5.1	5.2	7.1

a. Regions follow Bureau of Census definitions.

sized, however. These coefficients can no longer be interpreted as rates of return since the costs of achieving a "quality-adjusted" year of schooling are radically different across ethnic groups. Thus these comparisons relate more to the labor market valuation of "equal" inputs than to rewards for investment in human capital or to overall inequities in income distribution.

The regional estimates for the Spanish samples are based upon considerable fewer SMSAs and sampled individuals. Therefore, it is not possible to discuss confidently the regional pattern of these estimates.

III. SOME CONCLUSIONS

The process of income determination has received a fair amount of consideration in the last decade. And, yet, there remain significant amounts of uncertainty and controversy over the factors which determine income and over which potential government policies might be most appropriate to effect changes in the distribution of income.

This chapter presented some new evidence about income determination. In particular, data from the Public Use Sample of the 1970 Census of Population were used to investigate the relationship between earnings and some of the underlying factors affecting earnings. The study concentrates upon the earnings of males between the ages of 16 and 65. Earnings in 1970 are related to the individual's schooling and estimated labor force experience. Further, the analysis allows for differences in earnings functions by Standard Metropolitan Statistical Area (SMSA)—our attempt at capturing differences among local labor markets.

The findings of this analysis are:

1. There appear to be significant differences in earnings among individual labor markets.
2. Less educated—and less mobile—individuals appear more dependent upon local labor market conditions, and thus geographic location has a stronger overall earnings effect (i.e., mean difference) for this group than on earnings of more educated individuals; this is especially true for black and Spanish males.
3. There is also a significant variance in the shape of the earnings relationship across labor markets; the estimated percentage increase in earnings per year of schooling is widely differ-

ent among SMSAs, and these differences are larger than those between racial groups.

4. The returns for a year of post-secondary education are at least double those for a year of elementary or secondary education.

5. In terms of average increases in earnings per year of schooling, whites receive a higher return on schooling than blacks; this holds across regions and schooling categories.

6. When a crude quality adjustment is made for the schooling input into earnings, there are significant changes in the results; blacks at the elementary and secondary levels appear to receive higher returns per year of quality-equivalent schooling than comparable whites.

7. On a regional basis, the returns to elementary and secondary schooling appear highest in the Mid-Atlantic and Southern regions; there is no discernible pattern to the returns over macro-regions for post-secondary education.

These findings leave the strong impression that we must better understand the workings of local labor markets. Local conditions appear to have a powerful impact upon earnings—perhaps more important than race, and at least equal to several years of schooling. The similarity of these findings for females, for less than full-time, full-year workers, and for other minorities needs confirmation, but this analysis dictates a change in our way of viewing the earnings process.

NOTES

1. U.S. Bureau of the Census, *Statistical Abstract of the United States: 1973* (Government Printing Office, 1973), p. 330.

2. There have been some analyses which have concentrated on individual differences and not group or class differences. Most notable is Christopher Jencks et al., *Inequality* (Basic Books, 1972).

3. The "Coleman report" is the massive governmental study of American primary and secondary education which was undertaken to assess the racial and ethnic differences in education. James S. Coleman et al., *Equality of Education Opportunity* (Government Printing Office, 1966).

4. There has been a large amount of research in this area. The foundations of the analysis are found in the works of T. W. Schultz (for example, "Investment in Human Capital," *American Economic Review*, March 1961) and of Gary S. Becker (for example, *Human Capital*. [National Bureau of Economic Research, 1964]). A survey of human capital research can be found in Jacob Mincer, "The Distribution of Labor Incomes: A Survey with Special Reference to the Human Capital Approach," *Journal of Economic Literature*, March 1970; see also his *Schooling, Experience, and Earnings* (National Bureau of Economic Research, 1974).

5. It is important to note that the human capital model does not depend upon the specific mechanism by which the earnings of the individual increase. Schooling may increase the productivity of the individual by adding skills which are valuable in work. On the other hand, schooling may simply identify, or "screen," more productive individuals. While the former model usually seems to be implied in the human capital literature, the model works perfectly well in the latter world. Discussions of the screening hypothesis can be found in Paul Taubman and Terrance Wales, *Higher Education and Earnings: College as an Investment and a Screening Device* (McGraw-Hill, 1974); Kenneth Arrow, "Higher Education as a Filter," *Journal of Public Economics* (July 1973); and Michael Spence, "Job Market Signaling," *Quarterly Journal of Economics* (August 1973). The main implication of the screening model is that private returns to schooling may diverge from social returns to schooling. When we are interested in the distribution of income, we are generally talking about the private returns to schooling, and thus it is not necessary to distinguish between these alternative explanations. One caveat is necessary, however. The different models of the role of education may have differing implications in the long run when dynamic matters are important.

6. Again, as long as schooling screens on the basis of productivity differences, it is still possible to talk about demands and supplies of underlying characteristics. Thus, the importance of the underlying structural characteristics holds even in a screening world.

7. See, for example, James Smith and Finis Welch, "Black-White Male Earnings and Employment: 1960-1970" R-1666-DOL (Santa Monica: The Rand Corporation, 1975).

8. Most earnings studies have included some regional measures. The only studies going into more detail have been F. Welch, "Measurement of the Quality of Schooling," *American Economic Review* (May 1966), which considers rural male earnings by state; and E. Hanushek, "Regional Differences in the Structure of Earnings," *Review of Economics and Statistics* (May 1973), which analyzes metropolitan area differences in earnings for young males.

9. E. Hanushek, "Regional Differences."

10. Economic theory predicts that either the movement of factors of production (e.g., migration of labor) or trade in finished goods will tend to bring about equality in the relative payments to factors of production. There is some evidence that the returns to capital are roughly equal; see M. Straszheim, "An Introduction and Overview of Regional Money Capital Markets," in J. F. Kain and J. R. Meyer (eds.), *Essays in Regional Economics* (Harvard University Press, 1971). This should, according to factor price equalization theorems, imply absolute equality of payments to other factors of production, in particular, different skill categories of labor.

11. This restriction will lead to an understatement of the returns to human capital investment because length of work along with the wage rate is one way of securing returns to capital investment. Nevertheless, because of the possibility of interactions with hourly earnings (see Mincer, *Schooling, Experience, and Earnings*) and because of the lack of data on other key factors such as part-time work by students, this factor was eliminated from the modeling efforts through sample design.

12. As will be discussed below, the actual models analyzed view the logarithm of income as a function of input factors. The logarithm of a negative number is not defined.

13. The entire analysis from this point on concentrates on earnings rather than income. Earnings depend upon current labor force activity and, therefore, are the subject of most theoretical analysis. Income other than earnings include: Social Security and government railroad retirement; dividends, interest, rental income, and royalties; public assistance and welfare payments; unemployment and workmen's compensation; government pensions and veterans' payments; and private pensions, annuities, alimony, etc. Earnings are by far the most important part of total income, representing 88 percent in 1971. The distribution of total income by source, of course, varies considerably by income level. See Bureau of Census, "Money Income in 1971 of Families and Persons in the United States," *Current Population Reports*, Series P-60, No. 85 (Government Printing Office, 1972), p. 25.

14. James S. Coleman et al., *Equality of Educational Opportunity* (Government Printing Office, 1966). Approximately 570,000 students spread among grades 1, 3, 6, 9, and 12 were given a battery of standardized achievement and ability tests.

15. Grade level differences for verbal ability tests (*Equality of Educational Opportunity*, table 3.121.1) were interpolated to give differences by race and region for each year of schooling from 1 to 12. Region of schooling was assumed to be region of birth. Since urban-rural distinctions for birthplace are unavailable, urban school quality differences were used throughout. All schooling above high school was adjusted using the twelfth-grade adjustment for the given race-region cell. Previous uses of this type of adjustment can be found in Randall Weiss, "The Effects of Education on the Earnings of Blacks and Whites," *Review of Economics and Statistics* (February 1970).

16. See Jacob Mincer, *Schooling, Experience, and Earnings*. Experience is estimated by age-schooling-6. This assumes all time outside of school was spent in the labor force. It would be preferable to include actual labor market experience instead of "potential" experience, as is done here. Unfortunately this information is not available in the census data. A discussion of the difference between actual and "potential" experience is contained in E. Hanushek and J. Quigley, "Explicit Tests of the Human Capital Model and Intertemporal Adjustments in Relative Wages," ISPS Working Paper No. 767 (Yale University, 1976).

17. The actual form of the models estimated was $\log Y = a_0 + a_1E + a_2E^2 + a_3S$ where Y = total earnings; E = experience = age-schooling-6; S = years of school completed. The logarithmic form has commonly been used; see the explanation by Mincer, "The Distribution of Labor Incomes." In the logarithmic form, the estimated coefficients times 100 can be interpreted as (approximately) the percentage increase in earnings attributable to a one unit increase in the explanatory variable—e.g., one year more of schooling.

18. The individual variance explained is related to the R^2 shown in the table. The total individual earnings variance can be decomposed into within and between region variance. R^2 tells what percentage of the within

region variance is explained by the models so that within-region explained variance is R^2 times the proportion of total variance within regions. Thus, for example, in the total "white" line of table 9, 3 percent of the variance in individual earnings is between regions—leaving 97 percent within regions. The models explain 21.6 percent of the within region variance (R^2), or $21.6 \cdot .97 = 21.0$ percent of the total variance is within-region explained variance.

19. Since these estimates result from analysis of census sample information, they are subject to sampling errors. Some of these errors can be large enough to yield negative estimates; i.e., additional schooling is predicted to lower earnings. The estimation problems are most severe when there are fewer observations. Thus, the black and Spanish samples, which tend to have fewer observations per SMSA than the white samples, are more affected by estimation problems.

20. The constancy of earnings differentials have been noted in a number of places. See, for example Zvi Griliches, "Notes on the Role of Education in Production Functions and Growth Accounting," in W. Lee Hansen (ed.), *Education, Income, and Human Capital* (National Bureau of Economic Research, 1970), or Finis Welch, "Education in Production," *Journal of Political Economy* (January-February 1970). This constancy is perplexing given the large increase in educational attainments over time. Since education is becoming more abundant, the returns to education would be expected to fall, *ceteris paribus*. Recent work by Richard Freeman, "Overinvestment in College Training?" *Journal of Human Resources* (Summer 1975) suggests this constancy may be disappearing in the 1970s. However, these data suggest no fall in returns to college through 1969, and perhaps even an increase.

21. The interpretation of the "adjusted" return to schooling is quite different from the "unadjusted" return. In the unadjusted model, the coefficient on schooling (which has been the focus of attention throughout) can be interpreted as a rate of return on an investment in schooling if certain conditions are met. The most important condition is that the cost of the investment is simply the amount of earnings foregone by going to school rather than working. This rate of return can then be compared with rates of return on other investments such as capital.

In the adjusted model, such interpretations are no longer possible. "Quality years" of schooling no longer represent the costs of schooling. It may take twice as many chronological years to achieve a given number of quality years. Because adjusted years are so different from chronological years, the estimated returns in these two models should also not be compared with each other.

It is difficult within the sample to answer the question of "which measure is better." This choice would presumably best be made on an explained variance criterion, but the near identity of R^2 's across the two alternative forms does not allow a reasonable choice.

22. For the total samples, the estimated white returns to schooling are higher than for Spanish. This primarily reflects the different weighting due to different school completion distributions. Because of the significant differences in the estimated earnings functions for the two schooling categories, the "total" models should not, on statistical grounds, be esti-

mated, and these results will not be emphasized here.

23. The adjustment of post-secondary schooling was based upon quality differences in the twelfth grade. This is a less reasonable adjustment because: (1) region of birth is likely to be a less reliable indicator of post-secondary school location than it was for primary and secondary, and (2) it assumes constant quality of post-secondary institutions. To the extent that it is a constant adjustment to the schooling values within a SMSA racial grouping, the estimated returns to schooling will be unchanged in the quality adjusted instance as compared to quality unadjusted.

Discrimination in the Academic Marketplace

RICHARD B. FREEMAN

Special features of the academic world make it possible to separate racial or ethnic discrimination from various other sources of group income differentials to a greater extent than in many other occupations. The elusive qualitative variables which make it difficult to separate differentials from discrimination are somewhat less elusive among academic personnel. For example, the quality of an academic individual's training is at least crudely indicated by (1) the level of his highest degree,¹ and by (2) the ranking of the institution granting the degree, as compiled from the respective academic disciplines themselves by the American Council on Education. A widely recognized indicator of performance on the job is also available in the number of publications, in a profession where much emphasis is placed on research "productivity" ("publish or perish").

The available data permit several important questions to be asked about the extent and trend over time of ethnic discrimination among college and university professors, deans, and other academic individuals. First, there can be a raw measure of gross annual income differentials between white academics and academics from black, Oriental, and other ethnic minority backgrounds. Second, there are measures of the extent to which the various groups differ in such major income-determining variables as degree level and quality, amount of publication, and experience—with "discrimination" being confined within the differences