SOCIAL SCIENCE RESEARCH 11, 103–126 (1982)

Sources of Black–White Earnings Differences

ERIC A. HANUSHEK

University of Rochester

An unmistakable fact of U.S. society is the disparity of incomes between blacks and whites. At an aggregate level, the income of a typical black family is some 60% of that for a typical white family.[1] Despite extensive research, it remains unclear what causes these differences. While race per se undoubtedly enters, it is clearly an oversimplification to label all differences simply as due to discrimination. The typical black and white workers differ in terms of schooling, experiences, job choices, residential location, and a myriad of other factors that might affect earnings. Knowledge of the quantitative importance of each of these possible sources of differences is essential whenever one considers policies that might be introduced to lessen the observed differences.

This paper attempts to decompose the observed earnings differences between blacks and whites into more fundamental factors. These differences can, very generally, arise from a variety of underlying factors—including differences in the schooling and experience levels of individuals (the focus of much active governmental policy), differences in the "quality" of schooling and experiences, differences in "general abilities" of the population, and differences in the rewards to these factors.

The problem is, however, more complicated than this. The rewards to any specific factors represent market outcomes that aggregate the supplies of individual characteristics and the demands for these. The demands for specific factors may well vary across labor markets, implying that the rewards, say for given amounts of schooling, will also vary across labor markets. Thus, the specific distribution of individuals across labor markets will enter into any aggregate earnings comparisons.

Support for this research was provided by the Institute for Research on Educational Finance and Governance of Stanford University. This paper benefitted from comments by Stanley Engerman and two anonymous referees.

[1] The precise ratio varies depending on the specific year, the comparison made (such as family income, individual income), and so forth. Yet, no matter what the comparison, the observed differences remain substantial.

0049-089X/82/020103-24$02.00/0

The panoply of possible explanations and factors entering into observed earnings differences far exceeds our current analytical abilities. As a result, all analytical efforts must concentrate upon a more limited range of factors. This analysis is no exception. This analysis focuses on the interaction among skill differences of workers, geographic location (or labor market), and race and the relative importance of each of these in determining aggregate earnings differences. The novelty of the analysis is allowing for interactions with specific labor markets instead of (as is more typical) averaging across different labor markets.

The generalization of the analysis to consider interactions with specific labor markets is, however, not without costs. Both because of data limitations and because of the necessity to simplify the analysis in other dimensions, the characteristics used to describe differences among individuals are quite parsimonious (schooling and experience levels).[2] This introduces some ambiguities into the analysis. If we look at individuals with the same measured characteristics such as same age and years of schooling, their earnings can differ even within a given labor market, either because of differences in unmeasured characteristics (such as quality of schooling) or because of discrimination. As discussed below, differentiating between "unmeasured quality differences" and discrimination is generally not possible. It is nevertheless possible to place bounds on the magnitude of such differences—regardless of which underlying explanation is true.

1. ALTERNATIVE ANALYTICAL FRAMEWORKS

This work builds upon general analyses of earnings determination. There is no single model of earnings determination, however; instead, there has been a variety of alternative approaches, each highlighting a different aspect of labor market operations. This research melds together the key elements of the major approaches followed in the past.

The currently dominant strand of research into earnings determination utilizes the framework of human capital. This approach concentrates upon supply-side decisions of individuals. Individuals make a series of investment decisions, such as schooling or on-the-job training decisions, with the expectation of higher earnings in the future. This approach has been pursued in a multitude of theoretical and empirical studies (see, e.g., the reviews by Mincer (1972) and Rosen (1977)). The focus of these studies is heterogeneity of workers as measured by such things as schooling differences, ability, and experience.

However, while human capital research is the dominant stream of research, it is not the only one. An alternative view, which historically

[2] As described below, some attempts are made to capture differences in school quality. Further, the analysis does control for employment status and work time.

preceded this research, concentrates not on differences among workers but on aggregate earnings differences as related to characteristics of workers' employment. In particular, aggregate earnings differences have been decomposed by employing industry, by occupation, and by location or region of employment. While these studies have seldom considered differences in worker characteristics,[3] the implicit notion is that these employment characteristics represent the key determinants of earnings differences. Although generally not explicit, the underlying notion seems to be that barriers in mobility of workers prevent adjustment to earnings differences and allow differences in demands for workers to be reflected in wages.

A third line of research concentrates explicitly upon the demand for workers. This research generally begins from consideration of production functions and develops the derived demand for different workers based upon the production technology. This line of modeling, which assumes fixed supplies of workers, has generally concentrated upon international wage differences or intertemporal wage differences for a given country (e.g., Dresch, 1975).

These research efforts have seldom been integrated.[4] Each of the separate research lines appears to offer some confirmation of the importance of the different perspectives. For example, virtually all human capital studies find that different characteristics of workers are highly correlated with earnings differences; each of the decompositions, based upon industry, occupation, or region, likewise find significant differences in earnings; and, finally, demand studies show systematic earnings variations which are consistent with underlying notions of production functions and the derived demand for labor. Yet, this joint consistency with the different perspectives should make one suspicious, since the various perspectives are quite inconsistent with each other.

The inconsistencies arise in several areas. Human capital models assume perfect mobility of workers and competitive labor markets; the aggregate decompositions, however, rely upon limited mobility and barriers to competition (in the employment dimensions identified by the separate studies). Human capital models ignore any demand differences while other research brings this to the forefront. Aggregate decompositions, as noted, generally neglect differences in workers (or compositional differences in the labor force at the aggregate level), demand studies generally look at quite crude differences in the labor force (say

[3] An exception is Fuchs (1967).

[4] There are a few scattered exceptions to this statement. As noted, Fuchs (1967) considers both differences in individual characteristics of workers and regional or labor market differences. Johnson (1970) considers relative demands simultaneously with migration and regional differences. Nevertheless, there has been little effort to consider systematically the importance of the different perspectives.

schooling in two or three classes), and human capital studies consider at times quite extensive descriptions of individual worker differences. The result is that these separate analyses provide conflicting explanations of the source of individual earnings differences.[5]

When we consider racial differences in earnings, even more models come into the picture. Much of the attention has concerned the existence and potential magnitude of discrimination. On a theoretical level, a variety of alternative models have been suggested (see, e.g., Freeman, 1974). One class of models, following the initial work of Becker (1957), assumes competitive labor markets with discrimination entering through the preferences of white employers or employees. A second class of models assumes imperfect labor markets with whites being able to command some market power (e.g., Thurow, 1969, 1975). A third class concentrates upon information and "statistical" hiring decisions (e.g., McCall, 1973; Spence, 1973). A final class highlights the structure of the labor market per se and the possibility of institutional restrictions (see, e.g., the review by Cain, 1976).

The related empirical work on discrimination is, however, only loosely connected to the theoretical analyses. The essence of the empirical work is the estimation of empirical earnings functions which attempt to characterize differences among individuals in some detail. Once done, the question becomes whether or not one can detect differences across race—either in the intercept or various slope parameters of the earnings relationship. The difficulty in interpretation arises first from one's judgment about the adequacy of measurement of individual differences in skills. In particular, differences across race could simply reflect mean differences in characteristics not measured or poorly measured in the earnings estimation. For example, if school quality systematically differed by race and was not adequately measured, race differences could be observed even though "identical" workers of different races were paid exactly the same. Some attempts have been made to consider these issues with regard to school quality (e.g., Welch, 1972; Weiss, 1970) and with regard to ability (e.g., Griliches and Mason, 1972). Even beyond the measurement issues, there is a deeper issue of interpretation. Do observed differences arise from employer actions that are discriminatory (such as offering less training to blacks as suggested by Lazaer, 1979), or from

[5] Part of the differences in the studies may not be so much a reflection of inherent contradictions as simply consideration of different phenomena. For example, the focus of the demand studies is typically consideration of aggregate wage differences among countries or across extended time periods. In these, assumptions of fixed supply of labor (of specific types) might be appropriate, even though apparently contradictory to the focus of supply-side models of the human capital type. Nevertheless, even after making allowances for the different focuses of the studies, it seems difficult to neglect the evidence that each type of study provides for the others.

different investment strategies, in human capital terms, by blacks and whites? Even if blacks and whites follow different investment strategies, should we still attribute at least part of the outcomes in earnings to a backdrop of discrimination in the labor markets?

These are not the type of issues that are easily resolved. Available data are unlikely to allow any precise testing of the alternative theories of discrimination.

This work begins to integrate the alternative views of earnings determination with particular attention to differences in earnings between blacks and whites. The central empirical work involves estimation of earnings relationships, of a standard type, for different local labor markets, schooling groups, and race. This analysis, which allows for both individual differences and locally based demand differences, provides the basic data for investigation of the alternative factors that enter into aggregate racial differences in earnings. The next section describes the empirical models and the data, while the subsequent sections provide the empirical results.

2. EMPIRICAL MODELS AND DATA

Basic Models

The heart of this investigation is consideration of the interactions among race, characteristics of individual workers, and the reward structures of individual labor markets. This section describes the basic models and approaches. However, at the outset, it must be noted that the complexity of this task, combined with the large data requirements, requires analysis of quite simplified models of individual earnings. While the models actually estimated are widely used, they are clearly incomplete when compared to some of the more detailed investigations of individual earnings. The implications of this for interpretation of results are discussed in more detail below.[6]

The basic approach of this study closely follows much of the existing empirical research into individual earnings in the specification of the basic statistical models. It differs from previous work chiefly in consideration of samples for estimation and in the interpretation of the earnings models. The most common approach to the analysis of individual earnings involves finding a sample of data that simultaneously measures individual earnings and the characteristics of the individuals and then conducting a cross-sectional regression of earnings on the identified characteristics.

[6] All of the analysis will consider just earnings of males. This reflects both the inadequacy of models to describe the character of earnings by females and, relatedly, inadequacy in the underlying data. In particular, as will be apparent, actual labor force experience is not directly observed; instead, "potential" experience, or time out of school, is used in the estimation. For males, this is not as severe a problem as it is for females, where intermittent labor force participation is more prevalent. See Hanushek and Quigley (1981).

While the particular measured characteristics of individuals vary widely across studies, the core of the estimation almost always includes differences in schooling and labor market experiences across individuals. This analysis involves estimation of what, in fact, has become the "standard" earnings relationship (following the development of Mincer, 1974):

$$\log Y_i = a + b_s S_i + c_1 \mathrm{EX}_i + c_2 \mathrm{EX}_i^2 + U_i, \tag{1}$$

where $\log Y_i$ is the logarithm of annual earnings for individual i, S_i the years of schooling of individual i, EX_i the years of "potential labor market experience" defined as (age $-$ S $-$ 6) for individual i, EX_i^2 the potential experience squared for individual i, U_i the stochastic term in earnings of individual i, and a, b_s, c_1, c_2 the unknown parameters to be estimated.[7]

Local Labor Market Differences

A central concern here is the definition of appropriate samples for the estimation. While some interpretations, particularly the purely human capital analyses such as Mincer's (1974), attempt to interpret the relationships from a strictly supply-side view, the parameters of the earnings relationships must be thought of as reduced form parameters—parameters that include both supply- and demand-side factors. Further, since the estimated parameters are assumed to be constant across the population analyzed,[8] one must believe that the underlying structural relationships are the same for the entire sample. This would be violated if labor markets were "local" in the sense of having different underlying demand structures across local areas and if the sample data were drawn from different labor markets.

In fact, past research in earnings suggests that labor markets do indeed have a local nature. Geographic differences have been introduced in a

[7] The simplifications at this point are obvious. Equation (1) is not meant to capture all individual differences that are important in earnings determination. Instead, it is meant to portray the most significant systematic differences and to provide an overall characterization of human capital differences. Reliance upon such a simplified model is chiefly dictated by data availability. Since, as discussed below, an important element of this work is the analysis of labor market differences, it is necessary to have very large samples that contain geographic information. For this, the only acceptable data set comes from the Census of Population. But these data are limited in terms of information about qualitative differences among individuals.

[8] This is not completely necessary. One can think of the underlying model as having random coefficients, coefficients that differ across individuals. If these parameters are drawn from a common distribution and the parameters for an individual are independent of the individual's characteristics, one can interpret the estimation results as estimating the population mean of the parameter distribution. While estimation using ordinary least squares may be inefficient, such estimation will be consistent as long as the normal assumptions for OLS are also appropriate.

variety of ways: through introduction of regional dummy variables (e.g., South), through stratification by large regions (such as South and non-South), through state dummy variables or state stratifications, or through stratification by individual metropolitan areas. No matter how it is done, the estimated geographic differences are invariably significant.

This, by itself, might not be an entirely persuasive indication of problems with the labor market aggregation. Regional variations in labor market rewards, at any point in time, might not have any real substantive effect on the estimation if they simply reflected temporary, perhaps cyclic, differences; that is, if the basic underlying reward structure is the same, temporary fluctuations from year to year would have little serious impact on the estimation.[9] The justification for such an assumption typically relies upon simple theoretical models suggesting that regional variations should not exist. In particular, with competitive markets and free mobility of labor, individuals should migrate to high wage areas. This will drive down wages in those areas (and raise them in sending areas), thus leading to equality of earnings across regions.[10]

However, the available evidence suggests labor market differences do have an important effect on the estimation of Eq. (1). First, there is the previously cited evidence that virtually any measure of regional differences appears significant in earnings estimation. Moreover, there is a consistent pattern to these estimates; for example, earnings appear consistently lower in the South than in other areas. Second, this evidence is consistent with the aggregate decompositions of earnings. These analyses show differences that remain quite stable over time—suggesting more than simple cyclic variations about a common mean. Third, there is indirect evidence from the movement of labor itself. Most models of labor migration identify earnings differences as a key element in individual migration decisions.[11] If individuals migrate to obtain better earnings, one would not expect them to incur the substantial monetary and psychic costs involved if the earnings differences were to be short lived, i.e., when earnings differences at any point in time are merely temporary fluctuations. Moreover, the evidence on interregional migration is itself suggestive. At the state level, net migration over long periods of time is highly correlated (see Hanushek, 1981). This suggests that, if migration

[9] This essentially follows from the estimation of models with random coefficients; see footnote 8.

[10] An alternative theoretical argument rests on the free trade of goods across regions. This argument, developed in the factor price equalization theorems of international trade, indicates that labor earnings should be brought into balance across regions—much as it would be through the movement of labor itself.

[11] See Greenwood (1975) for a review of migration analysis and the models typically employed. As noted there, the exact specification differs across studies, but differences in earnings possibilities quite uniformly enter.

is chiefly motivated by earnings differences, the patterns of earnings differences themselves remain highly correlated over time.

Reconciling the theoretical arguments, which argue against long-term earnings differences, with the empirical evidence, which suggests that such differences do exist, is beyond the scope of this paper. However, the key to such conflicts undoubtedly lies in the assumptions of the theoretical models and the interpretation of them. The theoretical models generally assume competitive markets with no barriers to movement of labor (including no adjustment costs), no growth in the labor force, and no changes in demand for labor. Each of these simplifications is clearly inaccurate. Further, the arguments consider the static equilibrium that will be obtained after all adjustment has occurred, while saying nothing about the time path, or speed, of adjustment.[12]

Labor market differences are central to the empirical analysis here. The basic estimation of Eq. (1) is conducted for individual metropolitan areas (Standard Metropolitan Statistical Areas or SMSAs). The underlying presumptions are that the labor market for each individual metropolitan area is competitive and that the earnings parameters (a, b_s, c_1, and c_2) represent reduced form coefficients specific to a local area. While the earnings parameters of a given area may bear some relationship to those in other areas, through migration of firms and labor, they are nevertheless allowed to differ in accordance to local differences in supply and demand conditions.

Sample Selection

The estimation relies upon data from the 1/100 Public Use Sample of the 1970 Census of Population. This sample provides basic data on earnings, schooling, and age (which is transformed into potential experience) for individuals. The chief advantage of these data is information on the SMSA of residence for each individual. The chief disadvantage is the limited data about individual characteristics. Individuals are stratified by SMSA so that Eq. (1) can be separately estimated for each local labor market—that is, the parameters of the relationship are allowed to vary freely across SMSAs.

The empirical work also goes further in the elimination of restrictions on the earnings estimation. First, Eq. (1), as stated, implies that the marginal effect of different amounts of schooling is constant across schooling groups.[13] Additional flexibility in the earnings relationships is

[12] Indeed, some analyses suggest that regional differences are narrowing over time, even though they remain substantial.

[13] More precisely, since the model is specified is a semilogarithmic form, it implies that the proportional increase in earnings for a year of schooling is constant across all levels of schooling. In investment terms, b_s has an interpretation of the rate of return on a year of schooling, and the rate of return is assumed constant across levels of schooling.

allowed by also stratifying the data into two different schooling classes—high school diploma or less (S less than or equal to 12 years), and greater than high school education (S greater than or equal to 13). Note that the experience parameters are also allowed to differ by schooling class, reflecting either differing levels of on-the-job training investments in skills or differing average amounts of actual labor market experience for given amounts of potential experience (that is, differing average unemployment rates). Second, since a major focus of this paper is differences in earnings by race, the data are further stratified by race (black and white). Therefore, for each metropolitan area, a total of four earnings relationships are considered—defined by schooling class and race.

The data requirements for such an exercise are clearly large, and even the Public Use Sample data are insufficient to allow estimation of all of the relationships suggested for each SMSA. In particular, sample sizes become very small for many SMSAs, especially when one considers groups that do not appear too frequently, such as highly educated blacks. A somewhat arbitrary cutoff is imposed: Samples for any particular strata must include at least 25 observations of the group.

This analysis also looks at just one aspect of total earnings differences. The only earnings differences that are available from the Census data refer to annual earnings (in 1969). Annual earnings, however, are composed of two elements—wage rates times amount of work. The same observed annual earnings can arise from high wage and low amount of work time or a low wage and high amount of work time. The forces that affect work time may well differ from the forces that affect wage rates. In particular, we might believe that work time (unemployment plus length of work week) is governed importantly by cyclic factors specific to local areas while wage rates reflect more fundamental differences in labor market conditions. Therefore, the following analysis pertains just to earnings of individuals who stated they worked full time (greater than 35 hr per week) and full year (48 or more weeks per year). For this group, annual earnings comes close to measuring wage rates.[14]

The number of regions (SMSAs) and number of observations used in the estimation of the separate models are presented in Table 1. Since the ultimate objective is a consideration of differences in earnings between blacks and whites, we only consider the estimated earnings for regions that contain sufficient observations of both blacks and whites (of a given schooling group). Therefore, while earnings models could be estimated for whites in 147 separately identified SMSAs, most of the

[14] The full sample selection criteria were that individuals worked full time, full year; were not in school; had positive earnings; and were between ages 16 and 64. It was also required that individuals have a known state of birth; this was used in attempts to control for school quality, as described below.

TABLE 1
Numbers of SMSAs and Observations

Number	Total	Northeast	North central	South	West
SMSAs					
High school	84	16	14	47	7
College	18	4	5	7	2
Observations					
High school					
Black	13,397	3,345	3,269	5,865	918
White	84,520	26,951	23,468	24,156	9,945
College					
Black	1,600	388	448	425	339
White	28,185	9,200	7,416	6,451	5,118

analysis is restricted to the 84 SMSAs that also support estimation of black models for the high school strata and to the 18 SMSAs that support estimation of black models for the college strata. In terms of individual observations, there are a total of some 113,000 whites and 15,000 blacks. As noted, the blacks are considerably more concentrated in the high school group.

The table also describes the aggregate geographic distribution of the observations. The SMSAs are distributed across each of the census divisions, with the largest concentrations found in the South. (Note that the regional distribution is dictated by the relative locations of blacks, since only SMSAs contained substantial numbers of blacks are included.)

3. OVERALL EMPIRICAL RESULTS

The basic analysis calls for the estimation of 204 separate models like Eq. (1); this comes from 168 SMSA models for the high school group (84 white and 84 black) and 36 SMSA models for the college group (18 white and 18 black).

Interpreting the results from this extensive estimation is clearly difficult when done on a SMSA-by-SMSA basis. We therefore begin with an overall summary of the results.

The variance of individual earnings can be decomposed into a portion reflecting mean differences across SMSAs (between region variance) and a remainder reflecting variance in earnings within areas (within region variance). The within-region variance can be further decomposed into a portion explained by differences among individuals in the region (i.e., differences in schooling and experience) and a portion unexplained by these measured characteristics.

Results of the decomposition of earnings variance for all SMSAs avail-

TABLE 2
Decomposition of Variance in Individual Earnings (Proportions of Individual Variance)

	Between region	Within region		Total explained[b]
		Total	Explained[a]	
High school				
Black	.110	.890	.071	.173
White	.038	.962	.105	.139
College				
Black	.026	.974	.169	.191
White	.027	.973	.209	.230

[a] Proportion of within-region variance explained is calculated as 1 minus total within-region residual sum of squares/total within-region variance in earnings.

[b] Total explained variance is calculated as between-region variance plus proportion of within-region variance explained times proportion of total variance within regions.

able for the estimation are displayed in Table 2.[15] Several things are worth noting. First, a substantial proportion of the total earnings for the less educated black group reflects mean differences in earnings across the 84 metropolitan areas (i.e., 11%). While mean differences are less important for the other groups, they still exist. Second, the earnings model of Eq. (1) explains between 7 and 21% of the individual variance within regions. Even though this appears modest, it must be remembered that the populations in the separate samples are much more homogeneous than usual. Through sample design and stratification, any variance in earnings related to differing unemployment rates, race, aggregate schooling class, or geographic differentials has already been eliminated.

The overall character of the separate earnings models can be seen in Table 3, in which the mean values of each of the estimated coefficients for the sampled SMSAs as a whole and for the individual census divisions are listed. (The estimates for the census divisions are based upon aggregations of the individual SMSA estimates within each division, as shown in Table 1. While individual SMSAs remain the primary unit of analysis, the aggregation-to-census division is presented to summarize some of the overall variation in earnings relationships.)[16]

Consider first the estimated schooling coefficients. These have an interpretation of a rate of return to additional years of schooling; that is, the

[15] Data on white earnings from all 147 SMSAs are used, not just those which also support estimation of black earnings.

[16] Note that the aggregations-to-census divisions are no longer "representative" of the population because they rely upon the SMSA stratifications and the existence of at least 25 observations for each SMSA. This, however, is unlikely to cause major biases in the results.

TABLE 3
Mean Estimated Coefficients for Eq. (1) (Weighted by Observations)

	Total	Northeast	North central	South	West
Schooling (b_s)					
High school					
Black	.039	.042	.031	.045	.024
White	.049	.048	.048	.057	.034
College					
Black	.117	.110	.132	.103	.125
White	.105	.110	.098	.102	.108
Experience (c_1)					
High school					
Black	.031	.030	.026	.036	.030
White	.044	.040	.045	.048	.045
College					
Black	.034	.034	.031	.033	.041
White	.062	.062	.063	.061	.065
Experience squared (c_2)[a]					
High school					
Black	−.049	−.045	−.039	−.056	−.051
White	−.071	−.063	−.073	−.077	−.074
College					
Black	−.076	−.061	−.071	−.075	−.102
White	−.116	−.117	−.117	−.109	−.119

[a] Coefficients multiplied by 100.

coefficient times 100 is the percentage increase in schooling associated with an additional year of schooling. For the country as a whole, a black with 12 or less years of schooling can expect earnings to increase by 3.9% for each added year of schooling; a similar white can expect a 4.9% increase. For both races, the return to additional secondary schooling is least in the West and greatest in the South. For the college-educated group, however, the added earnings from additional schooling is somewhat higher for blacks than for whites.[17] Further, the West is no longer the worst area of the country in terms of returns to schooling.

In terms of added earnings with experience, white earnings consistently rise faster than black earnings. This effect is clearest for the college group but is still seen for the high school group. At the same time, white

[17] Note that the earnings models across schooling groups have not been constrained to be equal at the break points. They may also be subject to rather large sampling errors at the extremes. Table 3 indicates that the returns for an additional year of schooling between 11 and 12 years are 4–5%, while the returns between 13 and 14 years are 10–12%. This discontinuity seems large, perhaps too large. Note, however, that comparisons across the two regimes must take into account differences in the other coefficients, most importantly the intercept.

earnings are also more peaked (i.e., the negative coefficient on the quadratic term in experience is greater for whites).

It should also be pointed out that these estimates perhaps misstate the true racial differential of labor market experience. The experience measure used, again, is potential experience and is the same as actual labor market experience only if the individual is fully employed from the time of leaving school. Since unemployment rates are significantly different for blacks and whites, the estimated coefficients do not accurately estimate the earnings effects of actual labor market experience. Nevertheless, even if we adjust the coefficients for mean differences in employment probabilities, we find that the picture is changed little.[18]

These estimates indicate significant differences in the estimated earnings relationships by race, schooling group, and metropolitan area.[19] The exact implications of these differences for the earnings of blacks and whites is, nevertheless, complicated. Across the metropolitan areas for the study, the aggregated characteristics of workers differ (e.g., the average amount of schooling differs), the rewards to different characteristics differ (as indicated by the estimated coefficients), and the regional distribution of blacks and whites differs. Therefore, the next section attempts to disentangle the influences of these separate factors.

One final issue, discussed more fully below, must be mentioned before leaving the overall discussion of the estimated relationships. Clearly, the estimated models are very simple: There are many other factors which almost surely enter systematically into the determination of earnings. For example, much of the analysis of earnings has considered the issues related to omission, or poor measurement, of "ability"—where ability is meant to imply systematic skill differences among individuals. This is but one example of a possible problem in model specification. While some attempts have been made to expand the list of descriptors for individuals, lack of data from the Census of Population precludes going very far.[20]

[18] In 1970, the unemployment rate for black males was 7.3%, compared with 4.5% for white males. We can obtain an estimate of the expected amount of actual labor market experience by multiplying potential experience times the expected probabilities of employment, as calculated from the percentages above. If we do this transformation, we obtain mean estimated experience coefficients of .033 and .046 for blacks and whites, respectively, in the high school group and .037 and .065 for the respective college groups.

[19] Formal statistical tests of coefficient differences indicate that they are significantly different across groups. Given the fairly large samples for the separate estimations, however, statistical tests for differences are not very powerful. Further, it should be noted that the differences across individual SMSAs are larger than those across the aggregated census divisions summarized in Table 3.

[20] One area of attention was school quality differences. This was approached in two ways. First, following the analysis of Weiss (1970), data on regional differences in achievement were introduced. These data from Coleman et al. (1966) provided estimates of grade level equivalents on standardized reading scores by race and region. Number of years of

4. DECOMPOSITION OF RACIAL DIFFERENCES

We can now return to consideration of the sources of black–white earnings differences. The distribution of worker characteristics (schooling and experience), the rewards to these characteristics, and the location of blacks and whites enter into the determination of aggregate earnings. The first two columns of Table 4 show mean earnings for blacks and whites in the two schooling classes in the separate census divisions. The differences in mean earnings between blacks and whites in both schooling classes and across the census divisions are, as has been widely recognized, substantial.

At the same time, there are also significant differences in the average characteristics of workers. Within each schooling class, the mean black years of schooling always fall substantially below that for whites, and black workers tend to be more inexperienced.[21] This is shown in the middle columns of Table 4.

The final two columns give some indication of the wage differentials for "similar" blacks and whites. These columns present the estimated present value of earnings for workers with exactly 12 or exactly 16 years of schooling (for high school and college groups, respectively). In this, the separate regression estimates are used to calculate the expected earnings of a worker for each year of experience (and the fixed level of schooling). These are then aggregated and discounted at 5% to give an estimate of the earnings to be expected for an individual who remains in a specific region throughout his working life. This therefore summarizes the entire earnings profile estimated for each region. The results indicate significant differences in earnings over the lifetime. For the average black high school graduate, lifetime earnings (in 1969 dollars) fall $38,000 short of those for the average white high school graduate. For college graduates, this differential rises to some $54,000.[22]

The previous discussion uses the estimated earnings relationships to analyze the expected differences in wages for a "typical" black and white worker: one with the same quantity of schooling who is fully employed throughout a lifetime. A different way of viewing the earnings relationships is to analyze the overall mean earnings of blacks and whites

---

schooling was transformed into quality equivalent years based upon the region in which an individual grew up. Second, a series of dummy variables for region grew up was introduced in an attempt to directly estimate quality differences. The first estimation was indistinguishable from that presented in terms of explained variance or significance of schooling coefficients. The second estimation did not provide any consistent estimates of regional school quality differences.

[21] As noted previously, the gap in actual labor market experience is larger than that portrayed in Table 4 because of higher black unemployment rates.

[22] Note that these calculations entirely eliminate any possible effects of migration. Therefore, they must be interpreted with some care.

TABLE 4
Mean Earnings, Schooling, and Experience (Weighted by Observations)

	Earnings[a]		Schooling		Experience[b]		Present value[c]	
	Black	White	Black	White	Black	White	Black	White
High school								
All	$5435	$8373	9.4	10.6	24.6	25.5	$ 99,917	$138,056
Northeast	6087	8548	9.9	10.7	23.9	26.3	107,210	137,992
North central	6467	8897	9.8	10.6	24.8	25.3	109,839	146,394
South	4488	7517	8.8	10.2	25.0	24.7	89,043	128,807
West	6584	8912	10.2	10.9	24.5	25.6	107,484	141,030
College								
All	$8132	$12,654	14.7	15.5	16.4	18.7	$151,772	$205,774
Northeast	8147	13,010	14.8	15.6	17.1	18.9	150,993	210,427
North central	8627	12,645	14.6	15.3	16.7	18.4	160,331	207,306
South	7662	12,391	15.0	15.5	15.4	18.2	135,733	198,580
West	8086	12,375	14.5	15.3	16.7	19.2	161,461	204,256

[a] Earnings means are geometric means of annual earnings for full-time, full-year workers.
[b] Experience is calculated as (age-years of schooling minus 6).
[c] Present values of lifetime earnings streams are calculated from regression estimates for each SMSA, assuming a 5% discount rate. Calculations are based upon exactly 12 years of schooling for the high school class and exactly 16 years for the college class.

and to estimate the separate components of these differences. This is done in Tables 5 and 6. These tables combine the information about differences in average worker characteristics with the information about the earnings relationships in the different metropolitan areas.

In Table 5, it is assumed that whites are distributed regionally in the same proportions as the observed blacks. In Table 6, on the other hand, it is assumed that blacks are distributed in the same proportions as the observed whites. Therefore, both tables hold the regional distributions of whites and blacks constant across races. The impact of different regional distributions of blacks and whites can thus be seen through a comparison of the two tables.

Tables 5 and 6 begin with the actual mean earnings for blacks and whites, divided by schooling class (columns 1 and 2). The two tables differ only in the weighting of the estimates: Table 5 weights by the actual metropolitan distribution of blacks, while Table 6 weights by the distribution of whites. For blacks with 12 or less years of schooling, average annual earnings are \$5,435, with the highest earnings found in the West.[23] For blacks with some college, average earnings are \$8,132, with the highest average earnings in the North Central region. If, on the other hand, blacks were distributed across the sampled SMSAs in the same proportions as whites, the average earnings of a black with 12 or fewer years of schooling would be \$5,729 (from Table 6), while average earnings for college-educated blacks would fall slightly. For whites, average observed earnings based upon the actual white distribution are found in Table 6. For the two schooling classes, the average actual earnings are \$8,373 and \$12,654. The observed averages of whites would differ only slightly if they were distributed across SMSAs in the same proportions as blacks. For whites, the highest average earnings are found in the West for the high school category and in the Northeast for the college category; however, the variation across regions for the college category (as also indicated by Table 2) is rather small.

Quite consistently, the typical black with at least some college earns less than the typical white in the lower schooling group. This is the case even though, as shown in Table 4, these workers differ by some 4 years of schooling.

The next two columns provide counterfactual estimates of the expected annual income of (a) a black who was paid according to the local black earnings functions but had the characteristics of the average white in the

[23] It must be remembered that the populations used for these calculations are based upon the sample definitions used in the earnings estimation. Sampled observations must meet the selection criteria and, importantly, only SMSAs with 25 or more individuals in a given race/schooling category are included. Therefore, the populations are not truly representative of the entire population.

TABLE 5
Relative Black–White Earnings (Black Regional Distribution)

	Observed mean earnings		Predicted mean earnings		Relative earnings		
	Black (1)	White (2)	Black coefficients White means (3)	White coefficients Black means (4)	(1)/(2)	(3)/(2)	(4)/(2)
High school							
All	$5435	$8299	$5744	$7687	0.65	0.69	0.93
Northeast	6087	8658	6386	8133	0.70	0.74	0.94
North central	6467	9032	6666	8564	0.72	0.74	0.95
South	4488	7618	4846	6877	0.59	0.64	0.90
West	6584	9108	6812	8687	0.72	0.75	0.95
College							
All	$8132	$12,742	$8854	$11,422	0.64	0.69	0.90
Northeast	8147	13,145	8966	11,832	0.62	0.68	0.90
North central	8627	12,730	9575	11,508	0.68	0.75	0.90
South	7662	12,690	8267	11,107	0.60	0.65	0.88
West	8086	12,376	8580	11,251	0.65	0.69	0.91

TABLE 6
Relative Black–White Earnings (White Regional Distribution)

	Observed mean earnings		Predicted mean earnings		Relative earnings		
	Black (1)	White (2)	Black coefficients White means (3)	White coefficients Black means (4)	(1)/(2)	(3)/(2)	(4)/(2)
High school							
All	$5729	$8373	$5993	$7841	0.68	0.72	0.94
Northeast	6161	8548	6440	8044	0.72	0.75	0.94
North central	6377	8897	6546	8459	0.72	0.74	0.95
South	4511	7517	4828	6830	0.60	0.64	0.91
West	6536	8912	6769	8553	0.73	0.76	0.96
College							
All	$8027	$12,654	$8776	$11,301	0.63	0.69	0.89
Northeast	7985	13,010	8895	11,615	0.61	0.68	0.89
North central	8567	12,645	9480	11,432	0.68	0.75	0.90
South	7413	12,391	7983	10,807	0.60	0.64	0.87
West	8152	12,375	8630	11,191	0.66	0.70	0.90

SMSA[24] and (b) a black with the characteristics of the average black in the region but being paid according to the local white earnings function.[25] These estimates disentangle the effects on average earnings of differences in worker characteristics and differences in the rewards for given characteristics. In both schooling categories, the average black has fewer years of completed schooling and less experience than the average white. Given the regional distribution of blacks and the pattern of local earnings relationships (column 3), raising average black schooling and experience to that of whites would increase average black earnings for the nation as a whole to $5,993 (high school) and $8,776 (college). These average predicted earnings, however, remain substantially below those for whites (with the same average characteristics and regional distribution). On a regional basis, the largest improvement is in the South, where the disparities in average schooling levels between whites and blacks is largest.

The effects of differences in the earnings functions, or rewards for specific factors, between blacks and whites can be found in column 4. In these calculations, the regional distribution of blacks and the average characteristics of black workers are held constant, and average black earnings based upon the white earnings functions are calculated. Here the differences are much more dramatic. For the nation as a whole, average black earnings are predicted to rise to $7,687 and $11,422 for the high school and college groups, respectively. For the high school group, the rise would be even somewhat larger if blacks also were distributed across SMSAs in the same proportions as whites (Table 6).

The results of these estimates are summarized in the final three columns of Tables 5 and 6. These columns provide comparisons with observed mean white earnings (by schooling class and area of the country). The first is the ratio of actual black earnings to those of whites (while setting the geographical distribution the same for blacks and whites); the following two rely upon the predicted earnings for blacks from the two counterfactual cases.

Concentrating upon Table 5, we see that actual black earnings are 65% of white earnings for the high school group (64 for the college group). The adjustment for differences in characteristics indicates that black workers would receive 69% of what white workers receive if they had the same levels of schooling and experience but were paid according to black earnings schedules. However, they would have 90% *or more* of white earnings if they could be paid according to the white earnings

[24] All calculations are specific to SMSA and schooling class. Therefore, for example, we consider the characteristics of the average white in the high school group within each SMSA.

[25] A person with the average black characteristics paid according to the local black earnings function would have an expected income exactly equal to the observed mean black income, and similarly for white means and white earnings functions.

functions (but still had their lower observed levels of schooling and experience). In other words, equalizing differences in characteristics would close 11–15% of the racial gap in incomes, but equalizing the payments for the measured characteristics would close 70–80% of the gap.

Across the different census divisions, the picture is quite consistent. Different reward structures, not differences in worker characteristics, account for the vast majority of the differences in mean earnings for blacks and whites.

Table 6 merely reweights the earnings by the white distribution of workers across local areas. While there are some small differences between Tables 5 and 6, they do not affect any of the previous findings by very much.

The previous estimates are based upon the complete earnings functions for the different metropolitan areas. This includes both "base" earnings levels (i.e., the intercept value) and returns to specific worker characteristics (i.e., the schooling and experience coefficients). In Table 7 a distinction between these two components is made, and similar relative earnings estimates, in which intercept differences are separately identified, are presented. A comparison of predictions to actual white mean earnings for all of the sampled SMSAs (using both the black and white worker distributions) is also made.

The intercept differences represent broadly based differentials affecting all workers, regardless of their schooling and experience levels. A comparison of each of the pairs in a given column indicates the separate

TABLE 7
Estimated Earnings Relative to Actual White Earnings (All regions)

	Black schooling and experience coefficients		White schooling and experience coefficients	
Schooling level and regional distribution	Black means	White means	Black means	White means
High school (black distribution)				
Black intercepts	.65	.69	.84	.91
White intercepts	.72	.76	.93	1.00
High school (white distribution)				
Black intercepts	.68	.72	.90	.96
White intercepts	.71	.75	.94	1.00
College (black distribution)				
Black intercepts	.64	.69	.73	.81
White intercepts	.78	.85	.90	1.00
College (white distribution)				
Black intercepts	.63	.69	.75	.84
White intercepts	.76	.83	.89	1.00

effect of such diffuse differentials, holding constant both worker characteristics and returns to these. From this, it appears that such "across the board" differences are quite modest for the lower schooling group but substantial for the higher schooling group. For example, using the black geographic distribution, the observed relative earnings of the average black and white worker in the high school category is .65; if the average black had the same base earnings as whites, this would rise to .72; but, adding in the same returns to schooling and experience, this rises to .93 (as noted before). On the other hand, for the average black in the college category, over half the rise from the observed .63 to the .90 that would be obtained with the white reward structure comes from intercept, or base level, differences. In other words, the differential returns to schooling and experience are much more significant for blacks with lower schooling levels, while pervasive differences are more significant for blacks with higher schooling levels. (This difference between the high school and college groups is also seen in Table 3 in terms of the estimated schooling and experience coefficients for blacks and whites.)

For the high school group, the effect of black–white differences in base earnings is about the same as the effect of differences in mean schooling and experience levels, and these effects are dominated by black–white differentials in the returns to schooling and experience. However, for the college group, the effects of base level differences and of differences in the return to schooling and experience are about the same, and both are significantly greater than the effects of differences in mean characteristics.

At this point, we must return to issues of specification of the earnings models. The models capture differences in rewards that are related to the measured characteristics of workers. To the extent that the measured characteristics do not index differences that are important across workers, strict interpretation of the different coefficients may be misleading.[26] Again, consider the simple example of school quality. If, because of school quality differences, every year of schooling by blacks involves less "learning" than years by whites, a white and black with the same measured years of schooling would have systematic differences in skills. In this case, we would expect a smaller schooling coefficient for blacks than for whites, even if the monetary rewards for actual learning were the same. Because this seems like a real possibility, it is not reasonable to conclude that the differences in earnings parameters reflect pure discrimination. Further, the latter calculations that distinguish between base level, or intercept, differences and the differential returns to schooling and experience do not eliminate this ambiguity. For example, "learning

[26] For a general discussion of these issues, see Griliches (1977).

differences" that were proportional to quantity of schooling would appear in the intercept.

5. SOME CONCLUSIONS

The previous analysis has considered how interactions among race, characteristics of workers, and the structure of earnings for different individuals lead to the observed aggregate differences in black and white earnings. The basic conclusion is that differences in rewards, or payments to different individual characteristics, between blacks and whites are the major source of differences in aggregate earnings. Many people would perhaps argue that this was obvious.[27] But, if obvious, it is strangely at odds with a variety of policies. Policies such as providing freer access to schools or improving school retention for blacks are directed at equating the characteristics of black and white workers in the schooling dimension. And much of the attention to migration is concerned with improving the earnings of blacks through redistribution across labor markets. These policies flow from observations about the lower schooling levels of blacks and the distinct differences in locational patterns between races, but presume that black rewards to these factors will be sufficient to close substantially the observed earnings differences. The evidence suggests such reductions in earnings will be relatively modest. For example, keeping the current distribution of blacks and whites across areas at the 1970 observed distribution, equating experience and schooling levels of the average black in the high school class would increase relative black–white mean earnings from .65 to .69. Similarly, shuffling the regional distribution of blacks to match that of whites (but keeping average worker characteristics and rewards constant) would increase mean black relative earnings from .65 to .68. On the other hand, holding individual characteristics and geographic location constant but paying blacks according to the white earnings schedules for each SMSA would increase relative earnings from .65 to .93; in other words, 80% of the earnings gap would be closed.

The character of the earnings differences is quite different for blacks with lower levels of schooling and with higher levels. For those with a high school diploma or less, the largest differences between blacks and whites is found in the returns to schooling and experience. For those with at least some college, base earnings differences—those seen regardless of schooling or experience levels—are as important as the different rewards to schooling and experience.

Interpretation of these results must, nevertheless, be made within the

[27] These estimates do seem consistent with some earlier attempts at earnings decompositions, e.g., Duncan (1968) and Seigel (1965). However, because they focus on other features of earnings and employment, direct comparisons are difficult.

context of the very simplified models of individual earnings. While some attempts were made to include effects of differential school quality, they were largely unsuccessful, and the final models describe earnings differences among individuals simply in terms of years of schooling and experience levels of individuals. Because blacks and whites may differ in terms of unmeasured attributes (such as school quality or abilities), the estimated earnings differences would be distorted by these. Therefore, the differences in earnings attributed to differences in reward structures based solely on these measured characteristics cannot be taken as a measure of pure wage discrimination.[28] These estimates do provide some bounds on potential levels of wage discrimination—and the evidence suggests substantial room for discrimination.

Further, in the analysis not all workers or all aspects of employment are considered. Only male workers who are full time, full year are analyzed. (This latter restriction was imposed to focus on longer run differences in wages and to eliminate cyclic phenomena.) To the extent that blacks and whites differ in unemployment probabilities, this analysis will misstate the total differences in income and economic well being.

Finally, all of the estimates refer to earnings in 1969.[29] While these observations do follow some of the major civil rights changes of the mid-sixties, they do not provide any information about changes that might have occurred since then. The evidence about more recent changes is mixed. Smith and Welch (1977) argue that there was steady improvement in the relative earnings of blacks over the decade of the sixties, and this would suggest that the situation may well have improved during the seventies. On the other hand, Lazaer (1979) argues that this improvement may have been illusory—that employers raised current wages in response to governmental pressures through reducing the amount of training provided to blacks. If this were true, one might expect a widening of disparities since 1969. Thus, it seems that extrapolation at this time is difficult.

REFERENCES

Becker, G. S. (1957), The Economics of Discrimination, Univ. of Chicago Press, Chicago.

Cain, G. (1976), "The challenge of segmented labor market theories to orthodox theory: A survey," Journal of Economic Literature, December, 1215–1257.

Coleman, J. S., *et al.*, (1966), Equality of Educational Opportunity, U.S. Govt. Printing Office, Washington, D.C.

[28] Here the definition of discrimination is differences in wages for "identical" individuals of different races. The comparisons only refer to individuals who are the same in terms of the measured characteristics of the earnings functions.

[29] More exactly, they refer to the earnings of full-time, full-year workers. Since the average black is much less likely to be such a worker, the earnings differences for the entire population are understated in this analysis.

Dresch, S. (1975), "Demography, technology, and higher education," Journal of Political Economy **83,** 535–570.

Duncan, O. D. (1968), "Inheritance of poverty or inheritance of race?" *in* On Understanding Poverty (D. P. Moynihan, Ed.), pp. 85–110, Basic Books, New York.

Freeman, R. B. (1974), "Labor market discrimination: Analysis, findings, and problems," *in* Frontiers of Quantitative Economics, Vol. II (M. Intriligator and D. Kendrick, Eds.), North-Holland, Amsterdam.

Freeman, R. B. (1976), The Overeducated American, Academic Press, New York.

Fuchs, V. (1967), Differentials in Hourly Earnings by Region and City Size, Occasional Paper 101, National Bureau of Economic Research, New York.

Greenwood, M. J. (1975), "Research on internal migration in the United States: A survey," Journal of Economic Literature **13,** 392–433.

Griliches, Z. (1977), "Estimating the returns to schooling: Some econometric problems," Econometrica, January.

Griliches, Z., and Mason, W. (1972), "Education, income, and ability," Journal of Political Economy **80,** S74–S103.

Hanoch, G. (1967), "An economic model of earnings and schooling," Journal of Human Resources, Summer.

Hanushek, E. A. (1981), "Alternative models of earnings determination and labor market structure," Journal of Human Resources **16,** 238–259.

Hanushek, E. A., and Quigley, J. M. (1981), "Life-cycle earning capacity and the OJT model" (mimeo).

Johnson, G. (1970), "The demand for labor by educational category," Southern Economic Journal **37,** 379–388.

Lazaer, E. (1979), "The narrowing of black–white wage differentials is illusory," American Economic Review **69,** 553–564.

McCall, J. (1973), Income Mobility, Racial Discrimination, and Economic Growth, Lexington Books, Lexington, Mass.

Mincer, J. (1970), "The distribution of labor income: A survey with special reference to the human capital approach," Journal of Economic Literature **8,** 1–26.

Mincer, J. (1974), Schooling, Experience, and Earnings, National Bureau of Economic Research, New York.

Rosen, S. (1977), "Human capital: A survey of empirical research," *in* Research in Labor Economics (R. G. Ehrenberg, Ed.), Vol. 1, JAI Press, Greenwich, Conn.

Siegel, P. M. (1965), "On the cost of being Negro," Sociological Inquiry **35,** 41–58.

Smith, J. P., and Welch, F. (1977), "Black–white male wage ratios: 1960–1970," American Economic Review **67,** 323–338.

Spence, A. M. (1973), "Job market signalling," Quarterly Journal of Economics **87**(August), 355–374.

Thurow, L. (1969), Poverty and Discrimination, Brookings, Washington, D.C.

Thurow, L. (1975), Generating Inequality, Basic Books, New York.

Weiss, R. D. (1970), "The effect of education on the earnings of blacks and whites," Review of Economics and Statistics **52,** 150–159.

Welch, F. (1973), "Black–white differences in returns to schooling," American Economic Review **63,** 893–907.