Wage Rate Differences by Race and Sex in the US Labour Market: 1960–1970

ROBERT P. STRAUSS and FRANCIS W. HORVATH

(*Reprinted from 'Economica', August 1976*)

Economica, **43**, 287–298.

Wage Rate Differences by Race and Sex in the US Labour Market: 1960–1970

By ROBERT P. STRAUSS and FRANCIS W. HORVATH
University of North Carolina

Important new evidence on racial employment discrimination in the United States has been provided by Comanor (1973) who demonstrated that racial discrimination in hiring was a pervasive part of the US labour market in the mid-1960s. While Comanor showed significant employment effects of discrimination, he did not directly examine the related issue of racial wage rate differentials. In view of recent interest in the stability of US black–white *income* differences (e.g. Ashenfelter, 1970; Freeman, 1973a, b; Wohlstetter and Coleman, 1972), the issue of wage rate discrimination is of added interest, for these income differentials may result just from discriminatory hiring practices (blacks are forced to work in low-wage industries), or from a combination of wage rate discrimination and a disadvantageous employment distribution.

We seek in this paper to test statistically for the existence of wage rate differences between comparably employed blacks and whites and men and women. Because we have individual level information for 1960, 1967 and 1970, we are able to examine the question of wage rate discrimination more closely than in other studies. The plan of the paper is as follows. Section I discusses the measurement problems posed by actually testing for wage rate discrimination with Census data; Section II discusses the problems of inferring discrimination with various statistical models; Section III presents and interprets the empirical results for 1960, 1967 and 1970, and discusses the problem of sample size fragmentation; Section IV concludes.

I. THE DATA BASE

The central difficulty in comparing wage rates of various employees is to establish in fact that the workers (blacks and whites, males and females) are performing the same jobs. Optimally, we would like plant or firm level data that would allow us to hold constant various other, non-discriminatory factors which might produce observed wage rate differentials; regional location, marginal output per worker or differing productivity of capital are several of the more obvious candidates for "control" variables. However, such disaggregated data are generally unavailable on a nationwide basis, and we must accordingly turn to the traditional, Census sources of economy-wide information on employment and earnings of persons by race and sex. (See Dewey, 1952; Marshall, 1965; and Marshall and Briggs, 1967 for results of surveys of local labour markets. Of interest here is Dewey's earlier findings that blacks and whites doing the same jobs in the plants he surveyed received the same hourly wage rates.)

It is instructive to spell out what the response to a Current Population Survey or decennial Census questionnaire entails in terms of earnings as opposed to desired wage rate data. First, Census earnings data are deficient in regard to fringe benefits because employer contributions to retirement programmes are not included. Moreover, one cannot obtain from a Census source direct observations on hourly wage rates. Rather, one can obtain earnings over an entire year. The reported annual earnings figure, W, is exclusive of additional employee monetary benefits that may be available (special discounts for employees, for example) as well as the more difficult to measure but none the less important non-monetary benefits that accrue from having a clean job, congenial surroundings and likeable co-workers.

The response to the Census inquiry about annual earnings contains then a variety of factors that may mislead us about wage rate discrimination. For example, earnings might vary across race or sex not because of differing hourly wage rates but because of differing amounts of weeks worked, hours worked per week or simply differing productivity levels that affect workers in piece-work. Inference about wage rate discrimination from earnings is complicated further when we entertain the possibility that a worker may have "moonlighted," changed jobs, moved up a job classification during the year's period, been unemployed for some weeks, been sick and received partial compensation or not been willing or able to work the overtime his counterpart in the same firm did.

To reduce the confounding effect of such temporal aggregation, we stratify the three cross-sections—1960, 1967 and 1970—to exclude the self-employed and those working part-time, and also those over sixty-five years of age and non-Negro non-whites. (Perhaps the most serious measurement difficulty that does not readily admit of solution involves the undercount of blacks in the 1960 and 1970 Census; Siegel (1968, 1973) gives the official Census assessment of the problem. With regard to the changes in three-digit occupation and industry definitions between 1960 and 1970, see Greene and Priebe (1972).) The 1960 and 1970 cross-sections are 1/1000 representative samples of the United States; the 1967 data include both the representative and overweighted portions of the Survey of Economic Opportunity (SEO). (See Strauss and Horvath (1973) for a more complete discussion of the data.) The results for 1967 must therefore be interpreted with caution as the entire sample of full-time workers is overweighted for the poor. Of course, if the statistical model specified below is without error, then a biased sample will not lead to erroneous inferences about the effects of, say, race or sex.

The 1960 and 1970 sources provide annual earnings, weeks worked and hours worked in the reference week. The 1967 SEO also provides actual weekly earnings in the reference week. As a consequence, we may examine the effect of temporal aggregation in 1967 by comparing the effects of our explanatory variables on weekly earnings, E, and on our estimates of average weekly earnings, E^* (annual earnings $\div$ weeks worked). We expect that using the estimate will involve some spurious correlation with the regressors because the denominator may covary with them.

Finally, we use two-digit occupation as a proxy for job classification, and two-digit industry as a proxy for type of firm. While the two-digit classifications correspond in a distant fashion to our desired measures of job classification and firm or plant, and in fact are aggregations across such classifications as well as regions, the impact of this aggregation is to bias our results towards the conclusion that wage differences exist. We expect this bias because failure to hold constant such other "control" factors, such as region, which covaries with race, will allow the race regression coefficient to pick up these other effects.

II. Regression Model and Problems of Inference of Discrimination

To examine wage differences with the three cross-sections, we use a least-squares regression model. Following recent work (e.g. Weiss, 1970) we might specify our regression equation as:

$$W=\beta_{21}+\beta_{22}R+\beta_{23}S+\beta_{24}X+\beta_{25}T+\sum_{j=6}^{n}\beta_{2j}M_j + \sum_{j=n+1}^{p}\beta_{2j}I_j+\sum_{j=p+1}^{q}\beta_{2j}O_j+e \tag{1}$$

where W is total annual wages, R is a zero-one dummy for race ($R=1$ for whites, $R=0$ for blacks), S is a zero-one dummy for sex ($S=1$ for males, $S=0$ for females) (to the extent women are attached to the labour market only periodically because of child-rearing responsibilities, the experience measure as constructed may be inaccurate), X is a proxy for labour market experience measured as age minus years of schooling minus 5, T is years of schooling, M_j other socio-economic regressors, I and O are vectors of dummy variables for industry and occupation of employment, and e is a disturbance term with the usual properties of normality and homogeneity.

Two kinds of problems occur with (1). The first involves the issue of whether or not β_{22} and β_{23} are plausible estimates of the effect of race or sex on earnings; the second involves the problem of temporal aggregation identified earlier.

Equation (1) may be a reduced form equation of a recursive model of the labour market in which employment or quantity effects are conditioned by such exogenous variables as race and sex, as a consequence of taste on the part of workers in their job search or crowding by employers. If it is unreasonable to suppose that industry and occupation are independent of each other, then we may not obtain sound inferences about wage differences by race and sex. For example, we know that inter-occupation wage differences vary with industry. Moreover, it is reasonable to suppose that the racial "crowding" Comanor (1973) observed depresses wage levels in certain industries. Similar industry–occupation employment results for sex have been obtained by Schmidt and Strauss (1975).

One solution to this problem of industry–occupation interaction (and race–sex industry–occupation interaction) is to estimate a complete ("saturated") analysis of variance model which would contain a full set of interaction terms. As there are 11 occupations and 17 industries, this entails 176 interaction coefficients for just the industry–occupation interaction. Consideration of further sex–race interactions further adds to the computational burden and requires, even in the fully saturated model, the assumption that the variance of wages is constant across industry and occupation.

A simple solution to this first set of statistical problems is to stratify our three cross-sections by industry and occupation. That is, for the ith industry and jth occupational group in each cross-section, we perform regressions across individuals of the form:

$$W=\gamma_{11}+\gamma_{12}R+\gamma_{13}S+\gamma_{14}X+\gamma_{15}T+e_1. \tag{2}$$

clusion that wage differences exist. We expect this bias because failure to hold constant such other "control" factors, such as region, which covaries with race, will allow the race regression coefficient to pick up these other effects.

II. Regression Model and Problems of Inference of Discrimination

To examine wage differences with the three cross-sections, we use a least-squares regression model. Following recent work (e.g. Weiss, 1970) we might specify our regression equation as:

$$(1) \qquad W = \beta_{21} + \beta_{22}R + \beta_{23}S + \beta_{24}X + \beta_{25}T + \sum_{j=6}^{n} \beta_{2j}M_j + \sum_{j=n+1}^{p} \beta_{2j}I_j + \sum_{j=p+1}^{q} \beta_{2j}O_j + e$$

where W is total annual wages, R is a zero-one dummy for race ($R=1$ for whites, $R=0$ for blacks), S is a zero-one dummy for sex ($S=1$ for males, $S=0$ for females) (to the extent women are attached to the labour market only periodically because of child-rearing responsibilities, the experience measure as constructed may be inaccurate), X is a proxy for labour market experience measured as age minus years of schooling minus 5, T is years of schooling, M_j other socio-economic regressors, I and O are vectors of dummy variables for industry and occupation of employment, and e is a disturbance term with the usual properties of normality and homogeneity.

Two kinds of problems occur with (1). The first involves the issue of whether or not β_{22} and β_{23} are plausible estimates of the effect of race or sex on earnings; the second involves the problem of temporal aggregation identified earlier.

Equation (1) may be a reduced form equation of a recursive model of the labour market in which employment or quantity effects are conditioned by such exogenous variables as race and sex, as a consequence of taste on the part of workers in their job search or crowding by employers. If it is unreasonable to suppose that industry and occupation are independent of each other, then we may not obtain sound inferences about wage differences by race and sex. For example, we know that inter-occupation wage differences vary with industry. Moreover, it is reasonable to suppose that the racial "crowding" Comanor (1973) observed depresses wage levels in certain industries. Similar industry–occupation employment results for sex have been obtained by Schmidt and Strauss (1975).

One solution to this problem of industry–occupation interaction (and race–sex industry–occupation interaction) is to estimate a complete ("saturated") analysis of variance model which would contain a full set of interaction terms. As there are 11 occupations and 17 industries, this entails 176 interaction coefficients for just the industry–occupation interaction. Consideration of further sex–race interactions further adds to the computational burden and requires, even in the fully saturated model, the assumption that the variance of wages is constant across industry and occupation.

A simple solution to this first set of statistical problems is to stratify our three cross-sections by industry and occupation. That is, for the ith industry and jth occupational group in each cross-section, we perform regressions across individuals of the form:

$$(2) \qquad W = \gamma_{11} + \gamma_{12}R + \gamma_{13}S + \gamma_{14}X + \gamma_{15}T + e_1.$$

We allow variances to vary across industry and occupation and provide an analysis that is straightforward in interpretation. As there are 11 two-digit occupations and 16 two-digit industries, we obtain 176 $\hat{\gamma}_{12}$ and $\hat{\gamma}_{22}$s to make inferences about earnings differences by race not attributable to productivity differences as held constant by experience and education, and 176 $\hat{\gamma}_{13}$ and $\hat{\gamma}_{23}$s to make inferences about earnings differences by sex. (Of course, one could attempt to construct appropriate interaction terms in (1); however, the multiple analysis of covariance approach in (2) is much more readily interpreted and more certain to provide homogeneous comparisons by race and sex.)

We turn now to the second set of problems involving temporal aggregation. Our goal is to make inferences about hourly wage rate differences by race and sex, but our basic data are annual earnings and weeks worked. To obtain a proxy for weekly earnings, we divide annual earnings by weeks worked, Wk, to get average weekly earnings, E^*. We may then insert E^* in (2) and see if E^* varies systematically by race and sex. However, if Wk is not independent of R and S, the new $\hat{\gamma}_{12}$ and $\hat{\gamma}_{13}$s may be biased towards finding significant differences by race and sex. That is, they would be reflecting the effects of R and S on Wk but not on the true wage rate which is unobserved. We may isolate this effect of R and E on Wk by estimating

$$(3) \qquad \begin{aligned} W &= \gamma_{11}+\gamma_{12}R+\gamma_{13}S+\gamma_{14}X+\gamma_{15}T+e_1 \\ Wk &= \gamma_{21}+\gamma_{22}R+\gamma_{23}S+\gamma_{24}X+\gamma_{25}T+e_2. \end{aligned}$$

Clearly, if $\hat{\gamma}_{22}$ and $\hat{\gamma}_{23}$ are not significantly different from zero, then $\hat{\gamma}_{12}$ and $\hat{\gamma}_{13}$ reflect true weekly earnings effects. Also, if whites (men) earned more in a year and worked fewer weeks than did blacks (women) in the same industry and occupation, given experience and educational attainment which proxy for productivity, we may conclude that "weekly earnings discrimination" occurred. If $\hat{\gamma}_{12}$ and $\hat{\gamma}_{22}$ are equal to zero, then we conclude that discrimination has not occurred. If whites (men) earned more (less) but also worked more (fewer) weeks than blacks (women), in the same industry and occupation, we cannot readily infer that discrimination has occurred since we need to know by how much more (less) they earned as compared to how much longer (shorter) they worked. Finally, we may have a situation in which blacks (women) earn more than whites (men), given comparable weeks worked, i.e. "reverse discrimination".

When we move to test for hourly wage rate discrimination as compared with "weekly earnings discrimination", we may repeat the approach in (3) and specify an hours worked (H) equation and an average weekly earnings equation:

$$(4) \qquad \begin{aligned} H &= \gamma_{31}+\gamma_{32}R+\gamma_{33}S+\gamma_{34}X+\gamma_{35}T+e_3 \\ E^* &= \gamma_{41}+\gamma_{42}R+\gamma_{43}S+\gamma_{44}X+\gamma_{45}T+e_4. \end{aligned}$$

As the reader will note, we prefer generally to use separate equations to estimate the impact of temporal aggregation on earnings or wage rates; that is, we prefer to test each coefficient separately. For an alternative point of view, however, see Miller (1966). In the empirical work below, we report results for E^* and W and Wk to ascertain the impact of using ratio or averaged data. Moreover, because the 1967 Survey of Economic Opportunity reports not only total earnings last year, W, but also actual weekly earnings, E, as well as hours worked in the reference week, we can ascertain, at least for 1967, the bias in testing (1) with (4) using E^* rather than the true E. Accordingly, for 1967 we report for (4) results using both E and E^*.

We have then a three-step strategy to isolate wage differences by race and sex. First, we perform what we consider to be a traditional analysis of annual earnings and average weekly earnings with dummy variables for industry and occupation. Second, we correct for statistical dependencies by stratifying by industry and occupation and analysing annual earnings and average weekly earnings within each industry–occupation group. Third, we combine information about earnings and time spent working within the stratified groupings in step two to make inferences about weekly earnings and hourly earnings differences and control for possible spurious correlation that may have affected our results in step two.

III. Empirical Results

Dummy variable results

Table 1 provides the dummy variable regressions on W for the three cross-sections and indicates that, holding constant industry and occupation of employment, blacks suffered an annual earnings difference of over \$600. Surprisingly, the racial disadvantage was lower in 1967 than in 1960, though it grew to \$813 by 1970. This may reflect the non-representative nature of the 1967 sample noted above. In each cross-section, the estimated coefficient is ten times its estimated standard error, indicative of very significant results. The results for women are equally strong and much more sizeable in absolute terms: in 1960 women earned \$1,955 less than men, and in 1970 women earned \$3,195 less than men. Thus, the disadvantage of being female was at least three times that of being black. Not surprising is the fact that additional experience and education yield positive and significant returns to annual earnings.

While we have reasonably strong *a priori* expectations about the effects of the first four regressors, our expectations for the industry and occupation dummy variables are less developed. Accordingly, we perform two tail tests of statistical significance on them. Note that Industry 4 (Manufacturing Durables) and Occupation 7 (Operatives and Kindred Workers) are the suppressed categories. Most of the dummy variable coefficients are significantly different from zero with a negative pattern for the industry vector and a positive pattern for the occupation vector. Of interest is the relative stability of the signs of the coefficients over the three years. However, the relative sizes do not remain constant. Thus, if we compare the sizes of the 1960 and 1970 dummy variable coefficients, we find relative changes of as much as 150 per cent (Utilities and Sanitary Services).

The weeks worked equations (not reported here but available on request) indicate that whites and blacks worked the same number of weeks. Of course, since we are considering only those who are classified as "full-time", these results are not entirely surprising. Men, by contrast, do work significantly more than women in each of the three cross-sections.

The third set of results relates to average weekly earnings (Table 2). In all four equations (there are two for 1967, using E^* and E), whites earn significantly more than blacks (and men significantly more than women). Using the true weekly earnings, E, results in smaller race and sex effects than using E^*, although the difference is less than ten per cent in both instances. The stratification to

TABLE 1

DUMMY VARIABLE REGRESSIONS FOR EARNINGS LAST YEAR: 1960, 1967, 1970

Coefficient	Variable	1960	1967	1970
β_1		−1519·00**	−1383·00**	−2956·00**
		(79·91)	(132·8)	(120·2)
β_2	*Race*	623·60**	602·40**	813·00**
		(43·35)	(50·71)	(60·77)
β_3	*Sex*	1955·00**	2498·00**	3195·00**
		(30·78)	(54·58)	(44·42)
β_4	X	45·93**	47·98**	72·28**
		(0·95)	(1·77)	(1·34)
β_5	T	283·00**	345·50**	463·70**
		(4·99)	(8·92)	(7·73)
β_6	I_1	−1719·00**	−1289·00**	−1836·00**
		(173·0)	(416·7)	(269·6)
β_7	I_2	−204·50**	692·20**	−88·99
		(107·0)	(254·4)	(178·5)
β_8	I_3	−583·10**	−559·30**	−250·80**
		(55·38)	(103·5)	(85·18)
β_9	I_5	−289·30**	−367·50**	−340·80**
		(39·92)	(75·58)	(63·97)
β_{10}	I_6	−163·50**	−20·58	−45·17
		(56·54)	(115·0)	(96·22)
β_{11}	I_7	−69·42	−7·69**	−195·5
		(95·56)	(187·6)	(142·2)
β_{12}	I_8	−336·20**	−250·30	−664·00**
		(89·06)	(157·7)	(129·2)
β_{13}	I_9	−530·70**	−767·70**	−462·70**
		(65·67)	(122·3)	(94·90)
β_{14}	I_{10}	−1615·00**	−1826·00**	−2064·00**
		(45·67)	(89·70)	(68·64)
β_{15}	I_{11}	−738·50**	−834·30**	−654·20**
		(61·76)	(123·2)	(90·85)
β_{16}	I_{12}	−882·80**	−1012·00**	−953·00**
		(83·07)	(164·3)	(114·7)
β_{17}	I_{13}	−1601·00**	−1476·00**	−1695·00**
		(79·90)	(134·6)	(129·6)
β_{18}	I_{14}	−1033·00**	−1350·00**	−1071·00**
		(155·1)	(301·7)	(225·6)
β_{19}	I_{15}	−1612·00**	−1820·00**	−1806·00**
		(50·68)	(90·64)	(69·91)
β_{20}	I_{16}	−796·70**	−519·70**	−397·40**
		(54·73)	(94·05)	(82·56)
β_{21}	O_1	1439·00**	2061·00**	2489·00**
		(52·49)	(97·91)	(75·66)
β_{22}	O_2	1674·00**	203·20	2098·00**
		(475·9)	(917·5)	(648·8)
β_{23}	O_3	3195·00**	3464·00**	4371·00**
		(54·41)	(103·8)	(81·31)
β_{24}	O_4	426·40**	469·80**	463·10**
		(42·94)	(80·24)	(65·33)
β_{25}	O_5	1035·00**	1421·00**	1765·00**
		(58·00)	(130·2)	(90·24)
β_{26}	O_6	764·10**	939·60**	1098·00**
		(39·47)	(76·03)	(62·29)
β_{27}	O_8	85·53	165·8	419·6
		(122·3)	(179·1)	(228·7)
β_{28}	O_9	189·30**	310·10**	82·63
		(52·98)	(89·60)	(77·28)
β_{29}	O_{10}	−471·70**	−1551·00**	−1057·00**
		(192·9)	(448·5)	(319·9)
β_{30}	O_{11}	−418·00**	−513·50**	−881·50**
		(58·53)	(103·8)	(94·81)
		$N=41,257$	$N=17,714$	$N=51,182$
		$R^2=0·3975$	$R^2=0·4278$	$R^2=0·3681$

TABLE 2

DUMMY VARIABLE REGRESSIONS FOR AVERAGE WEEKLY EARNINGS: 1960, 1967, 1970

Coefficient	Variable	1960	1967	1967 (Earnings last week)	1970
β_1		−15·01**	−11·19**	−9·05**	−36·12**
		(1·77)	(2·71)	(2·62)	(2·73)
β_2	*Race*	12·41**	12·49**	11·61**	15·25**
		(0·96)	(1·03)	(1·00)	(1·38)
β_3	*Sex*	34·89**	44·58**	41·15**	57·06**
		(0·68)	(1·11)	(1·08)	(1·01)
β_4	X	0·80**	0·77**	0·65**	1·14**
		(0·02)	(0·04)	(0·04)	(0·03)
β_5	T	5·53**	6·64**	6·53**	8·99**
		(0·11)	(0·18)	(0·17)	(0·18)
β_6	I_1	−32·69**	−21·62*	−22·82**	−32·62**
		(3·83)	(8·49)	(8·23)	(6·13)
β_7	I_2	0·21	13·33*	20·51**	1·26
		(2·37)	(5·18)	(5·02)	(4·06)
β_8	I_3	−8·61**	−5·06*	−6·94**	3·00
		(1·23)	(2·11)	(2·04)	(1·94)
β_9	I_5	−8·59**	−8·31**	−7·16**	−7·78**
		(0·88)	(1·54)	(1·49)	(1·45)
β_{10}	I_6	−4·75**	1·49	2·16	2·13
		(1·25)	(2·34)	(2·27)	(2·19)
β_{11}	I_7	−5·90**	0·18	−1·67	−5·06
		(2·11)	(3·82)	(3·70)	(3·24)
β_{12}	I_8	−12·44**	−8·76**	−3·14	−12·73**
		(1·97)	(3·21)	(3·11)	(2·94)
β_{13}	I_9	−13·22**	−16·51**	−13·87**	−6·14**
		(1·45)	(2·49)	(2·42)	(2·16)
β_{14}	I_{10}	−32·86**	−36·14**	−34·07**	−38·44**
		(1·01)	(1·83)	(1·77)	(1·56)
β_{15}	I_{11}	−16·92**	−17·37**	−12·52**	−11·15**
		(1·37)	(2·51)	(2·43)	(2·07)
β_{16}	I_{12}	−17·23**	−20·93**	−15·62**	−14·00**
		(1·84)	(3·35)	(3·24)	(2·61)
β_{17}	I_{13}	−33·13**	−30·26**	−27·38**	−29·30**
		(1·77)	(2·74)	(2·66)	(2·95)
β_{18}	I_{14}	−20·13**	−18·19**	−25·56**	−15·29**
		(3·43)	(6·15)	(5·96)	(5·13)
β_{19}	I_{15}	−28·92**	−34·81**	−30·43**	−28·19**
		(1·12)	(1·84)	(1·79)	(1·59)
β_{20}	I_{16}	−18·46**	−11·45**	−8·13**	−8·14**
		(1·21)	(1·97)	(1·92)	(1·87)
β_{21}	O_1	28·29**	41·95**	45·35**	51·14**
		(1·16)	(1·99)	(1·93)	(1·72)
β_{22}	O_2	27·05*	−5·59	−6·83	39·11**
		(10·53)	(18·69)	(18·12)	(14·76)
β_{23}	O_3	57·88**	63·58**	64·51**	81·18**
		(1·20)	(2·11)	(2·05)	(1·85)
β_{24}	O_4	4·75**	6·36**	8·72**	5·38**
		(0·95)	(1·63)	(1·58)	(1·49)
β_{25}	O_5	18·43**	26·65**	26·29**	33·81**
		(1·28)	(2·65)	(2·57)	(2·05)
β_{26}	O_6	13·26**	16·74**	17·16**	20·03**
		(0·87)	(1·55)	(1·50)	(1·42)

TABLE 2—*continued*

Coefficient	Variable	1960	1967	1967 (Earnings last week)	1970
β_{27}	O_8	−2·79	−7·98	−7·52	−13·05*
		(2·71)	(3·65)	(3·53)	(5·20)
β_{28}	O_9	0·95	3·64*	6·99**	−0·78
		(1·17)	(1·82)	(1·77)	(1·75)
β_{29}	O_{10}	−12·82**	−36·51**	−29·47**	−21·00**
		(4·26)	(9·14)	(8·86)	(7·28)
β_{30}	O_{11}	−5·25**	−10·13**	−8·12**	−13·86**
		(1·29)	(2·11)	(20·5)	(2·15)
		N=41,257	N=17,714	N=17,714	N=51,182
		R^2=0·3266	R^2=0·3990	R^2=0·3983	R^2=0·2857

* Industry and occupation definitions are as follows:
I_1 = Agriculture, Forestry and Fisheries
I_2 = Mining
I_3 = Construction
I_4 = Manufacturing—Durable Goods
I_5 = Manufacturing—Non-durable Goods
I_6 = Transportation
I_7 = Communications
I_8 = Utilities and Sanitary Services
I_9 = Wholesale Trade
I_{10} = Retail Trade
I_{11} = Finance, Insurance and Real Estate
I_{12} = Business and Repair Services
I_{13} = Personal Services
I_{14} = Entertainment and Recreation Services
I_{15} = Professional and Related Services
I_{16} = Public Administration
O_1 = Professional, Technical and Kindred Workers
O_2 = Farmers and Farm Managers
O_3 = Managers, Officials and Proprietors, except Farm
O_4 = Clerical and Kindred Workers
O_5 = Sales Workers
O_6 = Craftsmen, Foremen and Kindred Workers
O_7 = Operatives and Kindred Workers
O_8 = Private Household Workers
O_9 = Service Workers, Except Private Household
O_{10} = Farm Labourers and Foremen
O_{11} = Labourers, except Farm and Mine

In Table 2, earnings last week is estimated by ratio of total earnings last year divided by the midpoint of the reported weeks-worked interval.

Standard errors are given in parentheses: a single asterisk indicates that a coefficient is significantly different from zero at the 5 per cent level, a double asterisk at the 1 per cent level.

full-time workers then is reasonably productive in eliminating part of the temporal aggregation problems noted above. Finally, when we compare the absolute size of the sex coefficients to the race coefficients, we note again that being female is about three times as disadvantageous as being black.

Stratified results

As is apparent from Table 3, the frequency with which blacks in different industries and occupations of employment have experienced significant wage differences has declined since 1960. In 1960 52 per cent of the industry–occupation groups contained significant differences in racial annual earnings; by 1970 only 36 per cent of the groups contained such differences. Women, by contrast,

TABLE 3

FRACTION OF TOTAL EMPLOYMENT EXPERIENCING SIGNIFICANT RACE AND SEX DIFFERENCES IN W, E^*, E, H AND Wk FOR 1960, 1967, 1970

Dependent variable	1960 Race %	1960 Sex %	1967 Race %	1967 Sex %	1970 Race %	1970 Sex %
W	79·3	93·0	51·5	98·2	68·6	98·2
E^*	77·9	92·5	51·5	97·3	47·6	98·6
E	—	—	49·7	96·6	—	—
H	49·4	75·8	46·4	80·4	49·0	91·9
Wk	10·5	68·7	5·4	67·0	7·9	89·9

experienced a high and stable level of earnings differences. In 1960 78 per cent of the industry–occupation groups contained significant sex differences in earnings; in 1970 the respective figure was 76 per cent. Of interest is the fact that, when actual weekly earnings is used instead of the average, we find essentially the same levels of differences by race and sex in 1967. This suggests that the use of averaged data does not appear to involve excessive amounts of spurious correlation.

Without describing the results here in any detail, we note that the inference of discrimination is less frequent by the two-equation technique. The results are summarized in Table 4 for two-equation techniques (within the stratified industry-occupation cells). A comparison of the two methods of inference indicates that, while the two-equation method shows fewer instances of race and sex differences in average weekly earnings, the orders of magnitude are the same for both techniques. Thus, blacks experienced differences in 1960 48 per cent of the time according to the single-equation technique but 41 per cent of the time according to the two-equation technique.

Complete tabular material on which the results of this section are based is available from the authors.

TABLE 4

RELATIVE FREQUENCY OF SIGNIFICANT RACE AND SEX DIFFERENCES (AT THE 5 PER CENT LEVEL) IN AVERAGE WEEKLY AND HOURLY EARNINGS IN INDUSTRY–OCCUPATION GROUPS BY METHOD OF INFERENCE

Type of comparison	1960 Average weekly earnings %	1960 Hourly earnings %	1967 Average weekly earnings %	1967 Hourly earnings %	1970 Average weekly earnings %	1970 Hourly earnings %
Race differences	40·9	33·0	34·7	28·0 (E^*) 30·6 (E)	31·2	27·5
Sex differences	72·7	67·0	64·0	64·0 (E^*) 60·0 (E)	67·0	67·9

Three caveats: cell size, further stratification and group size

As may be apparent, cross-tabulating even very large samples into 176 groups and performing analysis within is likely to lead to non-empty but none the less small cells, especially in terms of counts of blacks and women. Whether or not our counter-intuitive results are due to small sample size or the underlying postulated behaviour may then be still unanswered. However, size distributions of the numbers in the groups for each of the three years indicates that in 1960 75 per cent of the groups had 90 or more observations; in 1967 75 per cent of the groups had 47 or more observations, and in 1970 75 per cent of the groups had 72 or more observations.

To check for possible aggregation bias, we disaggregate a large 1970 cell within which consistent differences in earnings by race were found. Specifically we disaggregate the Manufacturing–Durables industry at its intersection with the Operatives and Kindred Workers occupation; there are 3,961 persons in the sample, or 3,961,000 in population terms.

If we stratify Manufacturing–Durables by three-digit occupations, only 4 of the 26 cells contain significant race coefficients at the 95 per cent confidence level in terms of annual earnings, 6 in terms of average weekly earnings and 3 in terms of weeks worked. The results for sex are much stronger: significant differences in average weekly earnings occur 14 times and significant differences in weeks worked occur 8 times.

If we stratify the Operatives and Kindred Workers occupation by three-digit industry, the pattern of significance is broadly the same. With 39 cells, blacks experience lower annual earnings 11 times, lower average weekly earnings 7 times, and fewer weeks worked 4 times. Again, differences by sex are more pronounced: we find 30 differences in annual earnings, 24 differences in average weekly earnings and 12 differences in weeks worked.

This fragment of evidence on the relative infrequency of race differences in earnings and time spent working for finer occupations and industries suggests that the overall pattern of results reported above seems plausible. Admittedly, one can go much further in terms of disaggregation; however, these investigations have satisfied us that salary and wage rate discrimination against blacks are not as pervasive as generally thought; we note by contrast the rather unfavourable situation of women.

While we have noted that, among homogeneous work groups, blacks experienced wage rate discrimination in fewer than half the different types of

TABLE 5

FRACTION OF TOTAL EMPLOYMENT EXPERIENCING SIGNIFICANT RACE AND SEX DIFFERENCES IN W, E^*, E, Hrs AND Wks FOR 1960, 1967, 1970

Dependent variable	1960 Race %	1960 Sex %	1967 Race %	1967 Sex %	1970 Race %	1970 Sex %
W	79·3	93·0	51·5	98·2	68·6	98·2
E^*	77·9	92·5	51·5	97·3	47·6	98·6
E	—	—	49·7	96·6	—	—
Hrs	49·4	75·8	46·4	80·4	49·0	91·9
Wks	10·5	68·7	5·4	67·0	7·9	89·9

jobs analysed by 1970, it is of additional interest to analyse how many persons *overall* were adversely affected. Table 5 provides the results at the 95 per cent confidence level and indicates that in 1960, while only 30 to 40 per cent of the different types of jobs exhibited significant differences by race, these jobs where discrimination occurred contained better than 75 per cent of all employed and all black employees. It is of interest that the same 1960–1970 trend found in Table 4 is not apparent in Table 5. The fraction of adversely affected blacks *rose* from 1967 to 1970 in terms of annual earnings, although this appears to have been the result of having worked fewer weeks. Note that the fraction of adversely affected blacks in terms of average weekly earnings was 48 per cent in 1970 as compared with 78 per cent in 1960.

IV. Conclusion

Using three extensive cross-sectional data sets, we have tested the proposition that blacks earn less than whites over a year's period, that they earn less in an average week, that they work fewer weeks, and that they experience lower hourly wage rates. We find that, if we estimate the effect of race on annual earnings and average weekly earnings by holding constant industry and occupation of employment with dummy variables in an additive multiple regression, then we accept the proposition in terms of annual earnings and average weekly earnings for blacks and for women. When, by contrast, we more properly stratify by industry and occupation and perform our analysis within each homogeneous group, we find that in the majority of groups, blacks earn the same as whites, work as many weeks and earn the same hourly wage rate. Moreover, this similarity in earnings of comparably employed blacks has increased over the period 1960–1970. Broadly speaking, we found racial differences in average weekly earnings about 40 per cent of the time in 1960 and only 30 per cent of the time in 1970. When we disaggregated more fully in terms of three-digit occupation or three-digit industry, previously significant differences by race disappeared. In terms of hourly wage rate discrimination, we found it 33 per cent of the time in 1960 and 27 per cent of the time in 1970. However, the fraction of adversely affected black employees is higher, although by 1970 only 48 per cent experienced differential weekly earnings.

The relative absence of wage rate differences by race suggests that the observed income differences by race may be due primarily to adverse employment distributions. By contrast, our results for women suggest that wage rate discrimination is still quite prevalent. Indeed, we suggest that women suffer in the United States both from the "crowding" phenomena and from unequal pay for equal work.

ACKNOWLEDGMENTS

We wish to thank the Data and Program Library Service, University of Wisconsin, for providing an efficient version of the 1967 Survey of Economic Opportunity data tape; the Social Science Data Library, University of North Carolina, for access to the 1960 and 1970 Public Use Samples; the National Science Foundation for financial support; and C. A. K. Lovell and Peter Schmidt for valuable comments. Responsibility for errors rests with the authors.

REFERENCES

ASHENFELTER, O. (1970). Change in labor market discrimination over time. *Journal of Human Resources*, **5**, 403–429.

COMANOR, W. W. (1973). Racial discrimination in American industry. *Economica*, **40**, 363–378.

DEWEY, D. (1952). Negro employment in southern industry. *Journal of Political Economy*, **60**, 279–293.

FREEMAN, R. B. (1973a). Changes in the labor market for black Americans, 1948–1972. *Brookings Papers on Economic Activity*, **1**, 67–132.

—— (1973b). Labor market discrimination: analysis, findings, and problems. Harvard Institute of Economic Research Discussion Paper 299.

GREENE, S. and PRIEBE, J. (1972). Changes in the occupation and industry classification systems between the 1960 and 1970 Census of Population. *Proceedings of the Business and Economic Statistics Section, American Statistical Association.*

MARSHALL, F. R. (1965). *The Negro and Organized Labor*. New York: Wiley.

MARSHALL, F. R. and BRIGGS, V. M. (1967). *The Negro Apprenticeship*. Baltimore: Johns Hopkins Press.

MILLER, R. G. (1966). *Simultaneous Statistical Inference*. New York: McGraw-Hill.

SCHMIDT, P. and STRAUSS, R. P. (1975). The prediction of occupation with multiple logit models. *International Economic Review*, **16**, 471–486.

SIEGEL, J. S. (1968). Completeness of coverage of the nonwhite population in the 1960 census and current estimates, and some implications. In *Social Statistics and the City* (David M. Heer, ed.). Cambridge, Mass.: Joint Center for Urban Studies.

—— (1973). Estimates of coverage of the population by sex, race and age in the 1970 Census. U.S. Census Bureau, mimeo.

STRAUSS, R. P. and HORVATH, F. W. (1974). Analyzing economic discrimination against blacks and women with the public use samples. *Review of Public Data Use*, **1**, 10–24.

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

WOHLSTETTER, A. O. and COLEMAN, S. (1972). Racial differences in income. In *Racial Discrimination in Economic Life* (A. H. Pascal, ed.), pp. 3–82. Lexington Books.