INTERNATIONAL ECONOMIC REVIEW
Vol. 14, No. 3, October, 1973

MALE-FEMALE WAGE DIFFERENTIALS IN URBAN LABOR MARKETS*

BY RONALD OAXACA[1]

CULTURE, TRADITION, AND OVERT DISCRIMINATION tend to make restrictive the terms by which women may participate in the labor force. These influences combine to generate an unfavorable occupational distribution of female workers vis-à-vis male workers and to create pay differences between males and females within the same occupation. The result is a chronic earnings gap between male and female full-time, year-round workers. Unfortunately, explanations at this level of generality are mainly descriptive. It is the purpose of this paper to estimate the average extent of discrimination against female workers in the United States and to provide a quantitative assessment of the sources of male-female wage differentials.

1. INTRODUCTION

In his study of sex differentials from the 1950 U. S. Census, Sanborn [**13**] found the female/male annual income ratio to be .58 which implies a male-female wage differential (as a proportion of the female wage) of .72. Sanborn's stated objective was to consider discrimination only in the context of equal pay for equal work and not to deal with discrimination stemming from occupational barriers. Using both male and female adjustment weights, Sanborn adjusted the income ratio for occupational distribution, annual hours of work, education, urbanness, race, turnover, absenteeism, and work experience. In his attempt to approximate equal work, Sanborn controlled for 262 detailed occupations. These adjustments brought the income ratio up to .87—.88. The residual difference was therefore .13 and thus about 18% of the original differential. As an estimate of the upper limits of the effects of discrimination it is rather low, but this is not surprising in view of the aspect of discrimination under study.

Using data from the 1960 U. S. Census, Fuchs [**6**] calculated the hourly earnings of females relative to males to be .60 which implies a male-female wage differential of .66. Fuchs does not believe that inherent differences in physical strength account for the sizeable pay difference simply because jobs requiring heavy labor are in a minority in the present day occupational structure. He contended that role differentiation stemming from social attitudes and discrimination affects the determinants of earnings. The earnings ratio was raised to .66

* Manuscript received May 1, 1972; revised November 13, 1972.

1 The author wishes to thank Orley Ashenfelter, Daniel Hamermesh, Albert Rees, Sherwin Rosen, and a referee for comments on an earlier version of this paper. Naturally, the responsibility of any remaining errors lies with the author. Financial assistance for this study was provided by the U.S. Department of Labor, Manpower Administration, under the provisions of Title 1 of the Manpower Development and Training Act, Public Law 87-415, as amended.

after adjustments for color, schooling, age, city size, marital status, class of worker, and length of trip to work. This adjusted ratio implies a residual differential of .51, which is 77% of the original differential. From the results of his regressions of hourly earnings across occupations, Fuchs concluded that nearly all of the wage differential could be explained away if one chooses sufficiently narrow occupational categories. As Fuchs observed, this merely casts the problem interms of why occupational distributions are so different between males and females.

In Cohen's study of sex differentials [**5**], an annual difference of $5,000 for full-time employed males and females was calculated from a 1969 working conditions survey. Cohen adjusted the difference in pay by excluding those under the age of 22 and over 64, the self-employed, persons without a steady job, and professionals. He also adjusted for differences in annual hours worked, fringe benefits, absenteeism, seniority, education, and unionization. The income difference was reduced to $2,550 or 51% of the original difference. Cohen attributed this large residual to the concentration of women in lower paying jobs.

The Malkiels' study [**9**] of professional workers revealed male-female wage differentials that ranged from .48 to .51 for the years 1966, 1969–1971. Adjustment for differences in schooling and experience yielded residual differentials varying between 37% to 49% of the original wage differentials. For one of the years the unexplained residual was reduced to only 3% of the original wage difference by adjusting for differences in schooling, experience, job level, critical area of study, and publications. This clearly indicates that unequal pay for equal work is not a significant source of discrimination. The problem of discrimination centers around the assignment of job levels. The Malkiels found that 53% of the male-female difference in job levels could not be explained by differences in personal characteristics.

2. A MEASURE OF DISCRIMINATION

Discrimination against females can be said to exist whenever the relative wage of males exceeds the relative wage that would have prevailed if males and females were paid according to the same criteria. We can formalize this notion by proposing the concept of a discrimination coefficient (D) as a measure of discrimination:

$$D = \frac{W_m/W_f - (W_m/W_f)^0}{(W_m/W_f)^0}\,; \tag{1}$$

where

$$(W_m/W_f) = \text{the observed male-female wage ratio;}$$

and

$$(W_m/W_f)^0 = \text{the male-female wage ratio in the absence of discrimination.}$$

An equivalent expression in natural logarithms is

$$\ln(D+1) = \ln(W_m/W_f) - \ln(W_m/W_f)^0 . \tag{2}$$

Assuming that employers in a nondiscriminating labor market adhere to the principle of cost minimization, we have

$$\left(\frac{W_m}{W_f}\right)^0 = \frac{MP_m}{MP_f} ;$$

where MP_m and MP_f are the marginal products of males and females, respectively.

In [2] Becker defined the market discrimination coefficient as the percentage wage differential between two types of perfectly substitutable labor. For those cases in which the two factors were not necessarily perfect substitutes, Becker defined the discrimination coefficient as the simple difference between the observed wage ratio and the wage ratio in the absence of discrimination. The discrimination coefficient defined by (1) is simply Becker's generalized measure divided by the wage ratio in the absence of discrimination. The generalized measures admit perfect substitutes as a special case and thus afford more flexibility for empirical work.

3. ESTIMATION PROCEDURES

Since $(W_m/W_f)^0$ is unknown, the estimation of D is equivalent to estimating $(W_m/W_f)^0$. On the basis of either of two assumptions, we can estimate the male-female wage ratio that would exist in the absence of discrimination: If there were no discrimination, 1) the wage structure currently faced by females would also apply to males; or 2) the wage structure currently faced by males would also apply to females. Assumption one (two) says that females (males) would on average receive in the absence of discrimination the same wages as they presently receive, but that discrimination takes the form of males (females) receiving more (less) than a nondiscriminating labor market would award them.

Ordinary least squares estimation of a wage equation for any given group of workers provides an estimate of the wage structure applicable to that group. The wage equation to be estimated separately for each race-sex group has the semi-log functional form

$$\ln(W_i) = Z_i'\beta + u_i , \qquad i = 1, \ldots , n \tag{3}$$

where

W_i = the hourly wage rate of the i-th worker,
Z_i' = a vector of individual characteristics,
β = a vector of coefficients,
u_i = a disturbance term.

When the male-female wage differential is expressed in natural logarithms, the formulation of the discrimination coefficient in (2) and our alternative assumptions about which wage structure would prevail in the absence of discrimination together imply that the wage differential can be decomposed into the effects of discrimination and the effects of differences in individual characteristics.

Let

$$G = \frac{\overline{W}_m - \overline{W}_f}{\overline{W}_f},$$

then

$$\ln(G+1) = \ln(\overline{W}_m) - \ln(\overline{W}_f) \tag{4}$$

where $\overline{W}_m$ and $\overline{W}_f$ are the average hourly wages for males and females, respectively.[2] From the properties of ordinary least squares estimation, we have

$$\ln(\overline{W}_m) = \overline{Z}_m{}'\hat{\beta}_m \tag{5}$$

$$\ln(\overline{W}_f) = \overline{Z}_f{}'\hat{\beta}_f \tag{6}$$

where

$\overline{Z}_m{}'$ and $\overline{Z}_f{}'$ = the vectors of mean values of the regressors for males and females, respectively.

$\hat{\beta}_m$ and $\hat{\beta}_f$ = the corresponding vectors of estimated coefficients.

Upon substitution of (5) and (6) into (4), we obtain

$$\ln(G+1) = \overline{Z}_m{}'\hat{\beta}_m - \overline{Z}_f{}'\hat{\beta}_f. \tag{7}$$

If we let

$$\Delta\overline{Z}' = \overline{Z}_m{}' - \overline{Z}_f{}' \tag{8}$$

$$\Delta\hat{\beta} = \hat{\beta}_f - \hat{\beta}_m \tag{9}$$

and substitute $\hat{\beta}_m = \hat{\beta}_f - \Delta\hat{\beta}$ in (7), then the male-female wage differential can be written as

$$\ln(G+1) = \Delta\overline{Z}'\hat{\beta}_f - \overline{Z}_m{}'\Delta\hat{\beta}. \tag{10}$$

On the basis of equation (2) and the assumption that the current female wage structure would apply to both males and females in a nondiscriminating labor market, it can be shown that

$$\ln\widehat{\left(\frac{W_m}{W_f}\right)^0} = \Delta\overline{Z}'\hat{\beta}_f \tag{11}$$

$$\ln(\widehat{D+1}) = -\overline{Z}_m{}'\Delta\hat{\beta}. \tag{12}$$

Thus expressions (11) and (12) represent the decomposition of the wage differential into the estimated effects of differences in individual characteristics and the estimated effects of discrimination, respectively.

An alternative decomposition of the wage differential is obtained by substi-

[2] These wage figures are computed as geometric means, i.e.,

$$\overline{W} = \exp\left\{\left[\sum_{i=1}^{n} \ln(W_i)\right]\Big/ n\right\}$$

tuting $\hat{\beta}_f = \Delta\hat{\beta} + \hat{\beta}_m$ in (7):

$$\ln(G+1) = \Delta\bar{Z}'\hat{\beta}_m - \bar{Z}_f{}'\Delta\hat{\beta}\,. \tag{13}$$

On the basis of (2) and the assumption that the current male wage structure would apply to both males and females in the absence of discrimination, it can be shown that

$$\ln\left(\widehat{\frac{W_m}{W_f}}\right)^0 = \Delta\bar{Z}'\hat{\beta}_m \tag{14}$$

$$\ln(\widehat{D+1}) = -\bar{Z}_f{}'\Delta\hat{\beta}\,. \tag{15}$$

Our method of estimating the effects of discrimination involves the familiar index number problem. Therefore, the separate estimates obtained from using both the male and female regression weights establish a range of possible values.

4. SPECIFICATION OF THE CONTROL VARIABLES

In accordance with the post-schooling investment model of human capital formation as developed in [**8**] and [**10**], a quadratic experience variable is included in the wage equations. The corresponding coefficients measure the combined effects of the average rate of return to on-the-job training, the initial proportion of time allocated to OJT, and the length of the investment horizon.

Since data on the actual number of years of work experience for a large sample of workers are generally unavailable, we define a proxy for actual work experience:

$$X_i = A_i - E_i - 6 \tag{16}$$

where

X_i = potential experience,

A_i = the age of the i-th individual,

E_i = the number of years of schooling completed by the i-th individual.

When work experience is acquired without interruption after the completion of formal schooling, potential and actual experience coincide. Potential experience is a reasonable proxy for actual experience in the case of males since males on average exhibit a strong attachment to the labor force. However, potential experience overstates the actual years of work experience of females to the extent that many female workers have left the labor force for some period in the past due to their household and childbearing activities. The difficulty this presents for estimation is of course the errors in variables problem. It is not clear what the net effect would be on our estimates of discrimination. It seems reasonable to suppose that the potential experience-wage profile would be flatter than the actual experience-wage profile. If the estimator of the coefficient on the linear experience term were biased downward for females, then $-\Delta\hat{\beta}$ would be biased upward in this instance. Consequently, there would be a bias toward finding discrimination.[3]

As a rough attempt to handle the problem of lost experience, we have controlled for the number of children (C) born to the female. The linear children variable reflects the cost of lost experience due to child care, including the costs from the depreciation of skills during the periods of absence from the labor force. Accordingly, we expect the estimated coefficient ($\hat{\beta}_c$) to have a negative sign. There is some difficulty associated with the introduction of this variable since it is obviously correlated with the proxy experience variable; however, in the absence of data on actual work experience, the use of the children variable seems justified.[4]

The remaining control variables are briefly described:

Education: years of schooling completed (linear and quadratic terms);

Class of Worker: dummy variables for union membership (privately employed wage and salary worker), government employed, and self-employed with non-union private wage and salary workers as the reference group;

Industry: dummy variables for U.S. Census two digit industries with retail trade as the reference group;

Occupation: dummy variables for U.S. Census two digit occupations with sales workers as the reference group;

Health Problems: dummy variable = 1 if the individual reports health problems that affect the kind or amount of work he or she can perform, and '0' otherwise;

Part-Time: dummy variable = 1 if the individual works less than thirty-five hours a week, and '0' otherwise;

Migration: a) dummy variable = 1 if the individual has maintained a residence more than fifty miles from his or her current address since the age of seventeen, and '0' otherwise, b) *YRSM*: number of years since the individual last migrated (linear and quadratic terms);

Marital Status: dummy variables for spouse present, spouse absent, widowed, and divorced (or separated) with never married individuals as the reference group;

Size of Urban Area: dummy variables for residence in Standard Metropolitan Statistical Areas less than 250,000 (SMSA$<$250), greater than or equal to 250,000 but less than 500,000 (SMSA 250–500), greater than or equal to 500,000 but less than 750,000 (SMSA 500–750), and greater than or equal to 750,000 (SMSA 750+) with urban, non-SMSA's as the reference group; and

Region: dummy variables for U.S. Census regions North East, North Central, and West with South as the reference group.[5]

One difficulty with the present formulation of the wage equation is that it controls for what many would consider to be major sources of discrimination

[3] These considerations also apply to wage models that specify an age variable in lieu of experience.

[4] The consequences of omitting this variable are discussed in footnote 6.

[5] The rationale for including most of these variables in a wage regression is fairly well known; nevertheless a more detailed discussion is given in [**11**].

against women. By controlling for broadly defined occupation, we eliminate some of the effects of occupational barriers as sources of discrimination. As a result, we are likely to underestimate the effects of discrimination. Therefore, we estimate another set of equations that do not control for occupation, industry, and class of worker. We shall refer to this set of regressions as the personal characteristics wage regressions, and to the original set as the full-scale wage regressions.

It is clear that the magnitude of the estimated effects of discrimination crucially depends upon the choice of control variables for the wage regressions. A researcher's choice of control variables implicitly reveals his or her attitude toward what constitutes discrimination in the labor market. If it were possible to control for virtually all sources of variation in wages, one could pretty well eliminate labor market discrimination as a significant factor in determining wage differentials by sex (or race). This is because $\Delta\hat{\beta}$, and therefore $\ln(D+1)$, would be very small. The result is that whatever the wage differential observed, it is completely justified on the grounds of alleged productivity differences. The other extreme is to control for virtually nothing and thereby minimize the role of productivity differences, $\Delta\bar{Z}$. This is tantamount to declaring at the outset that the two labor inputs are near perfect substitutes and therefore attributing virtually all of the observed wage differential to labor market discrimination, i.e., relatively high values of $\Delta\hat{\beta}$.

In reference to the type of regression approach we have adopted, it is impractical to control for detailed occupation. When we control for broadly defined two digit categories, we are not assuming that the conditions of equal work are met. Although the full-scale wage regressions eliminate male-female differences in broad occupational attachment as possible sources of discrimination, they can still reflect job and pay discrimination within each two digit category. The personal characteristics wage equations were specified with the intention of examining the issue of equal pay for roughly similar personal characteristics and not just for equal work.

5. EMPIRICAL RESULTS

The data for the study are from the 1967 Survey of Economic Opportunity. The particular subsample used for this study consists of the intersection of the following sets: those individuals who show an hourly wage for the week preceding the survey; adults sixteen years or older; those who live in urban areas; and those who report their race as either White or Negro.

The regression coefficients corresponding to the full-scale and personal characteristics wage equations are presented in Tables 1 and 2, respectively. Coefficients were not estimated in the following cases: 1) the particular characteristic served as the base group; 2) there were no observations in a particular cell: 3) the same observations were found in another cell; or 4) the regressor was left out on the basis of poor results from earlier regressions. The joint tests of significance for $\Delta\hat{\beta}$ in both the full-scale and personal characteristics wage regressions

TABLE 1

FULL-SCALE WAGE REGRESSIONS[a]

Variable	Whites Male	Whites Female	Whites $\Delta\beta$	Blacks Male	Blacks Female	Blacks $\Delta\beta$
Constant	.0365 (.77)	−.1024 (−1.34)	−.1389 (−1.60)	.0953 (1.71)	−.3851* (−6.35)	−.4804* (−5.85)
Experience						
Experience	.0176* (13.89)	.0138* (8.19)	−.0038 (−1.88)	.0117* (7.73)	.0067* (4.38)	−.0050* (−2.29)
Experience**2	−.000288* (−12.22)	−.000248* (−7.31)	.000040 (.98)	−.000204* (−7.59)	−.000122* (−4.33)	.000082* (2.11)
Education						
Education	.0082 (1.27)	−.0118 (−.98)	−.0200 (−1.53)	−.0308* (−4.60)	−.0175* (−1.98)	.0133 (1.19)
Education**2	.00169* (5.92)	.00194* (3.53)	.00025 (.42)	.00300* (8.23)	.00245* (5.26)	−.00055 (−.93)
Class of Worker						
Union	.1113* (9.39)	.1500* (6.70)	.0387 (1.59)	.2129* (14.15)	.0719* (3.11)	−.1410* (−5.14)
Nonunion Private Wage and Salary	—	—	—	—	—	—
Government	.0646* (3.15)	.1445* (5.89)	.0799* (2.54)	.1328* (5.44)	.1263* (5.19)	−.0065 (−.19)
Self-Employed	−.1290* (3.51)	.1137 (1.22)	.2427* (2.54)	−.0128 (−.15)	−.3437* (−2.67)	−.3309* (−2.15)
Industry						
Agriculture	.1285 (1.81)	.2847 (1.09)	.1562 (.61)	−.0067 (−.08)	−.0190 (−.21)	−.0123 (−1.18)
Mining	.3604* (6.83)	.4112* (2.02)	.0508 (.26)	.0697 (.40)	—	—
Construction	.2997* (13.72)	.2444* (3.80)	−.0553 (−.86)	.2729* (10.54)	.0395 (.22)	−.2334 (−1.30)
Manufacturing-Durable	.2398* (13.76)	.2562* (8.39)	.0164 (.48)	.2101* (9.15)	.2590* (6.46)	.0489 (1.06)
Manufacturing-Non Durable	.2086* (11.03)	.1968* (6.60)	−0.118 (−.35)	.1679* (6.85)	.2305* (6.46)	.0626 (1.45)
Transportation	.2332* (9.81)	.3154* (5.54)	.0822 (1.40)	.2182* (7.39)	.5463* (5.73)	.3281* (3.32)
Communications	.2370* (5.62)	.2290* (4.56)	−.0080 (−.12)	.1555 (1.78)	.2657* (3.71)	.1102 (.98)
Utilities	.2414* (7.32)	.2451* (2.83)	.0087 (.04)	.1433* (3.45)	.7026* (2.76)	.5593* (2.19)
Wholesale Trade	.2039* (8.45)	.1979* (4.74)	−.0060 (−.13)	.1204* (3.76)	.3065* (4.34)	.1861* (2.41)
Retail Trade	—	—	—	—	—	—
Finance	.2224* (8.25)	.1761* (5.65)	−.0463 (−1.14)	.0184 (.47)	.1593* (3.22)	.1409* (2.25)

(*Continued on next page*)

TABLE 1
(CONTINUED)

Variable	Whites			Blacks		
	Male	Female	$\Delta\beta$	Male	Female	$\Delta\beta$
Business and Repair Services	.1385* (4.44)	.1525* (3.24)	.0140 (.26)	.0766* (2.10)	.1326* (2.31)	.0560 (.83)
Personal Services	−.0618 (−1.71)	−.0183 (−.50)	.0435 (.85)	−.1055* (−3.22)	.0118 (.40)	.1173* (2.65)
Recreation	.0488 (.97)	.1527* (1.97)	.1039 (1.16)	.0020 (.04)	.1019 (1.29)	.0999 (1.05)
Professional Services	−.0629* (−2.53)	.0528* (2.01)	.1157* (3.24)	.0633* (2.13)	.1181* (4.45)	.0548 (1.38)
Public Administration	.1970* (6.58)	.2165* (4.86)	.0195 (.37)	.2374* (6.75)	.2170* (5.61)	−.0204 (−.39)
Occupation						
Professional Workers	.1563* (6.62)	.3736* (10.25)	.2173* (5.16)	.2144* (4.62)	.4631* (10.80)	.2487* (3.43)
Managers	.1822* (8.27)	.2759* (6.85)	.0937* (2.12)	.0810 (1.49)	.2792* (3.53)	.1982* (2.07)
Clerical Workers	−.0639* (−2.68)	.1665* (6.03)	.2304* (6.41)	.0208 (.54)	.1509* (4.50)	.1301* (2.55)
Sales Workers	—	—	—	—	—	—
Craftsmen	.0275 (1.28)	.0932 (1.31)	.0657 (.93)	.0733* (1.99)	.1297* (1.97)	.0564 (.75)
Operatives	−.1064* (−4.92)	.0128 (.37)	.1192* (3.00)	−.0271 (−.77)	.0236 (.62)	.0507 (.98)
Private Household Workers	−.1900 (−1.03)	−.3060* (−5.46)	−.1160 (−.58)	−.0458 (−.28)	−.1432* (−3.58)	−.0974 (−.58)
Service Workers	−.1358* (−5.19)	−.0219 (−.72)	.1139* (2.89)	−.0998 (−2.84)	−.0164 (−.53)	.0834 (1.78)
Farm Laborers	−.4570* (−5.38)	.1579 (.43)	.6149 (1.71)	−.1421 (−1.36)	—	—
Laborers	−.1540* (−5.59)	−.0166 (−.15)	.1374 (1.29)	−.0637 (−1.77)	.0317 (.37)	.0954 (1.03)
Health Problems	−.1001* (−6.08)	−.0710* (−2.70)	.0291 (.97)	−.0811* (−3.79)	−.0270 (−1.31)	.0541 (1.82)
Part-Time	−.1874* (−9.14)	−.0445* (−2.64)	.1429* (5.37)	−.1117* (−4.80)	.0034 (.21)	.1151* (4.04)
Migration						
Migration	−.0356* (−2.48)	−.1073* (−5.03)	−.0717* (−2.87)	.0052 (.44)	−.0361 (−1.94)	−.0413 (−1.88)
YRSM	.0072* (4.22)	.0087* (3.33)	.0015 (.48)	—	.0025* (2.73)	—
YRSM**2	−.000140* (−3.08)	−.000147* (−2.14)	−.000007 (−.10)	—	—	—
Marital Status						
Spouse Present	.1841* (11.88)	.0883* (4.51)	−.0958* (−3.91)	.1211* (6.43)	.0995* (5.13)	−.0216 (−.80)

(*Continued on next page*)

TABLE 1

(CONTINUED)

Variable	Whites			Blacks		
	Male	Female	$\Delta\beta$	Male	Female	$\Delta\beta$
Spouse Absent	.1124 (1.72)	.0852 (1.39)	−.0272 (−.30)	.0446 (.79)	.1050* (2.38)	.0604 (.84)
Widowed	.1030* (2.37)	.0687* (2.21)	−.0343 (−.64)	.0920* (2.13)	.0980* (3.47)	.0060 (.12)
Divorced	.0793* (2.74)	.0933* (3.38)	.0140 (.35)	.0396 (1.53)	.0607* (2.72)	.0211 (.62)
Single, Never Married	—	—	—	—	—	—
Children	—	−.0198* (−4.51)	—	—	−.0007 (−.24)	—
Size of Urban Area						
Urban, Non SMSA	—	—	—	—	—	—
SMSA<250	.0332* (1.98)	.0920* (3.86)	.0588* (2.07)	.0523 (1.54)	.1458* (4.19)	.0935 (1.93)
SMSA 250-500	.0727* (3.89)	.0956* (3.65)	.0229 (.73)	.1098* (2.83)	.1833* (4.61)	.0735 (1.32)
SMSA 500-750	.1411* (7.30)	.1524* (5.46)	.0113 (.34)	.1349* (3.55)	.1816* (4.46)	.0467 (.84)
SMSA 750+	.1745* (12.57)	.2186* (11.21)	.0441 (1.89)	.2079* (6.46)	.3643* (10.92)	.1564* (3.38)
Region						
North East	.0738* (5.63)	.0882* (4.69)	.0144 (.64)	.1366* (7.86)	.1724* (9.24)	.0358 (1.41)
North Central	.0749* (5.85)	.0646* (3.52)	−.0103 (−.47)	.1479* (9.37)	.1376* (8.00)	−.0103 (−.44)
South	—	—	—	—	—	—
West	.1200* (8.51)	.1389* (6.83)	.0189 (.79)	.2452* (12.48)	.2612* (12.07)	.0160 (.55)
F Statistic for Joint Test of Significance	128.31*	50.88*	13.28*	71.96*	97.14*	9.93*
R^2	.43	.33		.46	.56	
Standard Error of Estimate	.40	.45		.35	.36	
Number of Observations	8,123	4,962		3,897	3,502	

[a] '*t*' values in parentheses.

* Significant at the 5% level.

reveal that the wage structure for males and females are significantly different with respecct to the regressors common to both groups.

The average logarithms of the hourly wages (and the corresponding geometric mean wages) computed from our sample are as follows: 1.0806 ($2.95) for white males, .6499 ($1.92) for white females, .7721 ($2.16) for black males, and .3732 ($1.45) for black females. The values of the wage differentials in logarithmic terms, ln $(G + 1)$, are .4307 for whites and .3989 for blacks. To facilitate com-

TABLE 2

PERSONAL CHARACTERISTICS WAGE REGRESSIONS[a]

Variable	Whites			Blacks		
	Male	Female	Δβ	Male	Female	Δβ
Constant	−.0681* (−2.03)	.0894 (.94)	.1575 (1.87)	.1472* (2.44)	−.2325* (−3.75)	−.3797* (−4.65)
Experience						
Experience	.0222* (16.51)	.0182* (10.20)	−.0039 (−1.81)	.0195* (12.03)	.0066* (4.02)	−.0129* (−5.53)
Experience**2	−.000354* (−14.19)	−.000349* (−9.69)	.000005 (.13)	−.000340* (−11.69)	−.000133* (−4.26)	.000207* (4.86)
Education						
Education	.0342* (5.24)	−.0394* (−3.30)	−.0736* (−5.63)	−.0434* (−6.16)	−.0660* (−7.12)	−.0226 (−1.95)
Education**2	.00097* (3.51)	.00450* (8.63)	.00353* (6.24)	.00417* (11.46)	.00685* (15.35)	.00268* (4.68)
Health Problems	−.1325* (−7.57)	−.1097* (−3.89)	.0228 (.71)	−.1275* (−5.41)	−.0638* (−2.75)	.0637 (1.92)
Part-Time	−.3154* (−14.84)	−.1560* (−9.09)	.1594* (5.81)	−.1908* (−7.57)	−.1139* (−6.73)	.0769* (2.52)
Migration						
Migration	−.0316* (−2.05)	−.1262* (−5.54)	−.0946* (−3.56)	.0125 (.96)	−.0726* (−3.48)	−.0851* (−3.49)
YRSM	.0073* (3.98)	.0107* (3.81)	.0034 (1.05)	—	.0034* (3.24)	—
YRSM**2	−.000140* (−2.91)	−.000169* (−2.29)	−.000029* (−.34)	—	—	—
Marital Status						
Spouse Present	.2514* (15.44)	.1246* (5.96)	−.1268* (−4.90)	.1584* (7.66)	.0986* (4.53)	−.0598* (−1.99)
Spouse Absent	.1189 (1.71)	.0706 (1.07)	−.0483 (−.51)	.0975 (1.56)	.0964 (1.94)	−.0011 (−.01)
Widowed	.1389* (3.00)	.0804* (2.41)	−.0585 (−1.01)	.1648* (3.46)	.0754* (2.38)	−.0894 (−1.55)
Divorced	.1027* (3.34)	.1064* (3.61)	.0037 (.09)	.0511 (1.78)	.0618* (2.46)	.0107 (.28)
Single, Never Married	—	—	—	—	—	—
Children	—	−.0295* (−6.31)	—	—	−.0025 (−.80)	—
Size of Urban Area						
Urban, Non SMSA	—	—	—	—	—	—
SMSA<250	.0412* (2.30)	.1080* (4.23)	.0668* (2.20)	.0667 (1.78)	.1415* (3.60)	.0748 (1.38)
SMSA 250–500	.0845* (4.26)	.1154* (4.11)	.0309 (.92)	.1103* (2.57)	.1879* (4.20)	.0776 (1.25)
SMSA 500–750	.1739* (8.47)	.1721* (5.77)	−.0018 (−.05)	.1821* (4.34)	.1769* (3.86)	−.0052 (−.08)

(*Continued on next page*)

TABLE 2

(CONTINUED)

Variable	Whites			Blacks		
	Male	Female	$\Delta\beta$	Male	Female	$\Delta\beta$
SMSA 750+	.1972* (13.45)	.2543* (12.27)	.0571* (2.31)	.2452* (6.95)	.3888* (10.39)	.1436* (2.80)
Region						
North East	.0655* (4.73)	.1129* (5.69)	.0474* (2.01)	.1704* (8.97)	.2268* (11.07)	.0564* (2.02)
North Central	.0790* (5.91)	.0685* (3.52)	−.0105 (−.46)	.2255* (13.56)	.1996* (10.57)	−.0259 (−1.03)
South	—	—	—	—	—	—
West	.1111* (7.49)	.1174* (5.43)	.0063 (.25)	.2889* (13.44)	.3027* (12.56)	.0138 (.43)
F Statistic for Joint Test of Significance	213.29*	67.51*	45.05*	103.93*	132.01*	58.14*
R^2	.34	.22		.33	.43	
Standard Error of Estimate	.43	.49		.39	.40	
Number of Observations	8,123	4,962		3,897	3,502	

[a] *t* values in parentheses.

* Significant at the 5% level.

parison of our results with those of other studies, the wage differential *G* is also calculated. The value of *G* is .54 for whites and .49 for blacks.

The effects of discrimination are approximated by the residual left after subtracting the effects of differences in individual characteristics from the overall wage differential. The calculations based on the full-scale wage regressions are presented in Table 3. As a simple average of the two estimates obtained, discrimination accounts for 58.4% of the logarithmic wage differential for whites and 55.6% for blacks. The average value of the discrimination coefficient is .29 for whites and .25 for blacks. Table 4 presents the effects of discrimination calculated from the personal characteristics wage regressions. Predictably, the estimated effects of discrimination are larger than those reported in Table 3: Discrimination accounts for approximately 77.7% of the wage diffferential for whites and 93.6% for blacks. The averaged estimates of the discrimination coefficient are .40 and .45 for whites and blacks, respectively. Under both sets of regressions and for both races, sex differences in the distribution of part-time employment and marital status significantly contributed to a narrowing of the wage differential. It is evident from Tables 1 and 2 that workers with spouse present tend to earn more than others even after controlling for other factors. A smaller proportion of women workers fall into the category of spouse present, and therefore this difference reduces the wage differential due to discrimination. The difference probably reflects the competing activities of production in the home. In the case of whites, the effects of childbearing also narrow

TABLE 3

THE EFFECTS OF DISCRIMINATION ESTIMATED FROM THE FULL-SCALE WAGE REGRESSIONS

Item	Whites				Blacks			
	Female Regression Weights		Male Regression Weights		Female Regression Weights		Male Regression Weights	
	(1)[a]	(2)[b]	(3)[c]	(4)[b]	(5)[a]	(6)[b]	(7)[c]	(8)[b]
Wage differential	.4307	100.0%	.4307	100.0%	.3989	100.0%	.3989	100.0%
Adjustment for sex differences in								
Experience	−.0056	−1.3	−.0074	−1.7	−.0009	−0.2	−.0017	−0.4
Education	−.0051	−1.2	−.0037	−0.9	+.0170	+4.3	+.0140	+3.5
Class of Worker	−.0218	−5.1	−.0144	−3.3	−.0120	−3.0	−.0418	−10.5
Industry	−.0745	−17.3	−.0901	−20.9	−.0995	−24.9	−.1170	−29.3
Occupation	−.0059	−1.4	−.0338	−7.8	−.0451	−11.3	.0090	−2.3
Health Problems	+.0012	+0.3	+.0017	+0.4	−.0006	−0.2	−.0019	−0.5
Part-time	−.0065	−1.5	−.0273	−6.3	+.0006	+0.2	−.0184	−4.6
Migration	+.0030	+0.7	+.0001	0.0	+.0013	+0.3	−.0002	0.0
Marital Status	−.0078	−1.8	−.0271	−6.3	−.0070	−1.8	−.0157	−3.9
Children	−.0309	−7.2	.0000	0.0	−.0015	−0.4	.0000	0.0
Size of Urban Area	−.0015	−0·3	−.0012	−0.3	−.0030	−0.8	−.0022	−0.6
Region	+.0002	0.0	.0000	0.0	−.0045	−1.1	−.0050	−1.3
	$\ln(\widehat{D+1})=.2755$ ($\hat{D}=.32$)	63.9%	$\ln(\widehat{D+1})=.2275$ ($\hat{D}=.25$)	52.9%	$\ln(\widehat{D+1})=.2437$ ($\hat{D}=.27$)	61.1%	$\ln(\widehat{D+1})=.2000$ ($\hat{D}=.22$)	50.1%

[a] The adjustment for the j-th variable using female regression weights is $-\hat{\beta}_{fj}\Delta\bar{Z}_j$, and therefore the sum is $-\Delta\bar{Z}'\hat{\beta}_f$. This implies $\ln(\widehat{D+1}) = \ln(G+1) - \Delta\bar{Z}'\hat{\beta}_f = -\bar{Z}_m'\Delta\hat{\beta}$.

[b] Each adjustment is expressed as a percentage of the wage differential.

[c] The adjustment for the j-th variable using male regression weights is $-\hat{\beta}_{mj}\Delta\bar{Z}_j$, and therefore the sum is $-\Delta\bar{Z}'\hat{\beta}_m$. This implies $\ln(\widehat{D+1}) = \ln(G+1) - \Delta\bar{Z}'\hat{\beta}_m = -\bar{Z}_f'\Delta\hat{\beta}$.

TABLE 4

THE EFFECTS OF DISCRIMINATION ESTIMATED FROM THE PERSONAL CHARACTERISTICS WAGE REGRESSIONS

Item	Whites				Blacks			
	Female Regression Weights		Male Regression Weights		Female Regression Weights		Male Regression Weights	
	(1)[a]	(2)[b]	(3)[c]	(4)[b]	(1)[a]	(2)[b]	(3)[c]	(4)[b]
Wage differential	.4307	100.0%	.4307	100.0%	.3989	100.0%	.3989	100.0%
Adjusment for sex differences in								
Experience	−.0072	−1.7	−.0094	−2.1	−.0007	−0.2	−.0028	−0.7
Education	−.0122	−2.8	−.0008	−0.2	+.0351	+8.8	+.0190	+4.8
Health Problems	+.0018	+0.4	+.0022	+0.5	−.0015	−0.4	−.0030	−0.8
Part-Time	−.0227	−5.3	−.0459	−10.7	−.0187	−4.7	−.0314	−7.9
Migration	+.0033	+0.8	−.0002	0.0	+.0024	+0.6	−.0004	−0.1
Marital Status	−.0143	−3.3	−.0380	−8.8	−.0086	−2.2	−.0167	−4.2
Children	−.0460	−10.7	.0000	0.0	−.0052	−1.3	.0000	0.0
Size of Urban Area	−.0017	−0.4	−.0012	−0.3	−.0033	−0.8	−.0029	−0.7
Region	+.0003	+0.1	−.0001	0.0	−.0058	−1.5	−.0069	−1.7
	$\ln(\widehat{D+1}) = .3320$ $(\hat{D} = .39)$	77.1%	$\ln(\widehat{D+1}) = .3373$ $(\hat{D} = .40)$	78.4%	$\ln(\widehat{D+1}) = .3913$ $(\hat{D} = .48)$	98.5%	$\ln(\widehat{D+1}) = .3538$ $(\hat{D} = .42)$	88.7%

[a] The adjustment for the j-th variable using female regression weights is $-\hat{\beta}_{fj}\Delta\bar{Z}_j$, and therefore the sum is $-\Delta\bar{Z}'\hat{\beta}_f$. This implies $\ln(\widehat{D+1}) = \ln(G+1) - \Delta\bar{Z}'\hat{\beta}_f = -Z_m'\Delta\hat{\beta}$.

[b] Each adjustment is expressed as a percentage of the wage differential.

[c] The adjustment for the j-th variable using male regression weights is $-\hat{\beta}_{mj}\Delta\bar{Z}_j$, and therefore the sum is $-\Delta\bar{Z}'\hat{\beta}_m$. This implies $\ln(\widehat{D+1}) = \ln(G+1) - \Delta\bar{Z}'\hat{\beta}_m = -\bar{Z}_f'\Delta\hat{\beta}$.

the differential. Differences in the mean years of schooling completed widens the differential for blacks because black females complete on average almost a full year more of schooling than black males. It is clear from Table 3 that sex differences in the distributions by class of worker, industry, and occupation significantly narrow the wage differential even though industry and occupation are represented by highly aggregated categories.

Space limitations permit discussion of only a few selected aspects of the regression coefficients reported in Tables 1 and 2. The estimated experience coefficients under both sets of regressions imply that for the same rate of return to OJT, males invest more initially and for a longer period. If both sexes invest the same initially, then the pattern of differences in the coefficients imply that males earn a higher rate of return and invest for a longer period than females. The coefficients on the children variable indicate that each child lowers the white female wage by 2 or 3% but has a negligible effect on the black female wage.[6] This may suggest that black females do not stay out of the labor force as long as white females for each child born.[7] Our results are consistent with those studies of labor supply, such as [3] and [4], that find that the presence of children inhibits the labor force participation of white females significantly more than for black females. Perhaps some form of extended family arrangement in black communities provides a ready source of child care for working mothers. It may also be that lost experience and acquired skills are not very important in the kinds of jobs black females typically hold. For example, 52% of the black

[6] The personal characteristics wage equations were also estimated without the children variable. Generally, the positive coefficients were reduced in magnitude and the negative coefficients became larger in absolute value. Consequently, $-\Delta\hat{\beta}$ was biased upward, which implies a larger estimated effect of discrimination.

[7] The cost of children in terms of their effect on the hourly wage can be translated into an equivalent number of years of potential experience; 1) Set the estimated female experience profile minus the children term equal to zero; and 2) Solve the resulting quadratic for the negative root. Let R_1 be the negative root of

$$\hat{\beta}_2X^2 + \hat{\beta}_1X - \hat{\beta}_cC = 0\ .$$

The absolute value $|R_1|$ is an estimate of the equivalent number of years of experience. The mean number of children per white female was 1.6. Using the coefficients from the personal characteristics wage regressions, we evaluated $|R_1|$ at $C = 1.6$. The value of $|R_1|$ was approximately 2.5 years. As an estimate of the number of years of experience lost to child care, 2.5 years seems low. Yet this may be a reasonable estimate when one considers that ours is a sample of employed females. A sample from the total population of adult females would include mothers not in the labor force. The years of actual work experience lost would be higher for such a sample. Since we are interested in the average years of lost experience for all employed females, the estimate of 1.6 children per female was computed using the total number of females. However the sample includes women who have never had chidren. Thus the mean number of children per mother and the estimate of their lost experience would be higher. The average ages of white males and females in our sample were 39.5 and 38.9 years, respectively. The average difference in potential experience was 0.6 years. When 0.6 is added to the estimated 2.5 years. we have a total estimated male-female experience difference of 3.1 years. In the Malkiels' study [**9**, (24)] the difference in actual work experience was 2.9 years in 1971 for males and females whose average ages were 40.8 and 38.9 years, respectively. Another independent source [**14**, (94)] shows that the difference in years of employment covered by Social Security was 3.2 years in 1960 between males and females in the 35-39 age group.

females in our sample held private household and service worker jobs whereas only 16% of the white females held these jobs.

6. CONCLUDING REMARKS

As in other studies we find the sex differential to be quite large. We are in agreement with other researchers that unequal pay for equal work does not account for very much of the male-female wage differential. Rather it is the concentration of women in lower paying jobs that produces such large differentials. Our results suggest that a substantial proportion of the male-female wage differential is attributable to the effects of discrimination.

The effects of discrimination are estimated as the residual left after adjusting the sex differential for differences in various characteristics. This methodological technique is found in other studies as well and may take the form of regression analysis or standardization analysis. There are some difficulties with this general approach which should be mentioned. Could it be that the wage structures for males and females would differ even in the absence of discrimination? For example, male-female differences in the coefficients of the experience variable suggest that the rate of return to OJT may be higher for males, and/or females invest less in OJT. One might argue that even in the absence of discrimination females may plan on a shorter working life and hence invest less than men. The result would be a difference in the parameters of the experience variables, yet these differences contribute to the effects of discrimination under our analysis. In defense of this approach it should be pointed out that occupational barriers against women deny them the opportunities to invest to the same extent as men. Also, the short work life expectancy of women may represent a rational response to anticipated discrimination in the labor market. The issue becomes one of how much of the male-female difference in the coefficients is due to discrimination.

Another difficulty with the residual approach is that it does not take into account the effects of the feedback from labor market discrimination on the male-female differences in the selected individual characteristics. The differences could reflect the adaptation of women to the biases of the labor market; yet under the residual approach all differences in the characteristics contribute to a reduction of the wage differential attributable to discrimination. The problem becomes one of how much of the observed differences in individual characteristics would exist in the absence of discrimination.

These very difficult problems have not been dealt with in this study, but they are clearly important in terms of policy prescriptions for narrowing the male-female wage differential.

University of Massachusetts, Amherst, U. S. A.

REFERENCES

[1] ASHENFELTER, ORLEY, "Racial Discrimination and Trade Unionism," *Journal of Political Economy*, LXXX (May/June, 1972), 435-464.

[2] BECKER, GARY S., *The Economics of Discrimination* (Chicago: University of Chicago Press, 1957).

[3] BOWEN, WILLIAM G., AND T. A. FINEGAN, *The Economics of Labor Force Participation* (Princeton: Princeton University Press, 1969).

[4] CAIN, GLEN G., *Married Women in the Labor Force: An Economic Analysis* (Chicago: University of Chicago Press, 1966).

[5] COHEN, MALCOM S., "Sex Differences in Compensation," *Journal of Human Resources*, VI (Fall, 1971), 434-447.

[6] FUCHS, VICTOR R., "Differences in Hourly Earnings Between Men and Women," *Monthly Labor Review*, XCIV (May, 1971), 9-15.

[7] GWARTNEY, JAMES, "Discrimination and Income Differentials, *American Economic Review*, LX (June, 1970), 396-408.

[8] JOHNSON, THOMAS, "Returns from Investment in Human Capital," *American Economic Review*, LX (September, 1970), 546-559.

[9] MALKIEL, BURTON G., AND JUDITH A. MALKIEL, "Male-Female Pay Differentials in Professional Employment," Working Paper No. 35, Industrial Relations Section, Princeton University.

[10] MINCER, JACOB, "The Distribution of Labor Incomes: A Survey with Special Reference to the Human Capital Approach," *Journal of Economic Literature*, VIII (March, 1970), 1-26.

[11] OAXACA, RONALD L., *Male-Female Wage Differentials in Urban Labor Markets*, unpublished Ph. D. dissertation, Department of Economics, Princeton University, (1971).

[12] REES, ALBERT, AND GEORGE P. SHULTZ, *Workers and Wages in an Urban Labor Market* (Chicago: Chicago University Press, 1970).

[13] SANBORN, HENRY, "Pay Differences Between Men and Women," *Industrial and Labor Relations Review*, XVII (July, 1964), 534-550.

[14] U. S. SOCIAL SECURITY ADMINISTRATION, OFFICE OF RESEARCH AND STATISTICS, *Workers Under Social Security, 1960: Annual and Work History Statistics* (Washington, D. C.: U. S. Government Printing Office, 1968).