ANALYZING ECONOMIC DISCRIMINATION AGAINST BLACKS AND WOMEN WITH THE PUBLIC USE SAMPLES*

by Robert P. Strauss and Francis W. Horvath,

University of North Carolina

ABSTRACT

Economists have studied economic differences between blacks and whites and men and women in terms of competing theories of pervasive discrimination versus "crowding", that is, a situation where minorities face barriers to entering higher paying occupations but within homogeneous work groups experience no differences in wage rates. Large amounts of micro-data are required to properly account for the multitude of factors other than prejudice which may cause wage rate differentials. The 1960 and 1970 1/1000 Public Use Samples were used to analyze the extent and location of black-white, male-female economic differences. A number of processing problems were encountered which could have been avoided with more thoughtful planning. Initial results show that blacks suffered less than women in terms of the frequency of unequal pay for equal work; however, substantial evidence of "crowding" was found.

BIOGRAPHIES

Robert P. Strauss is Associate Professor of Economics at the University of North Carolina, Chapel Hill. He received his B.A. from the University of Michigan and his M.A. and Ph.D. degrees in Economics from the University of Wisconsin. In 1971-72, he was Special Assistant to the Deputy Secretary of the U.S. Treasury with long range planning and evaluation responsibilities for general revenue sharing. He has published widely in the fields of public finance and labor economics.

Francis W. Horvath received his B.A. degree from the College of the Holy Cross, Worcester, Massachusetts. He is currently pursuing a Ph.D. in Economics at the University of North Carolina, Chapel Hill, with special interests in labor and population economics.

INTRODUCTION

Economic differences between blacks and whites and men and women are well documented in the literature. Economists have focused on income differences and to a lesser extent on occupational attainment differences in contrast to sociologists who have focused on differences in social status more broadly defined. Virtually all of the recent research has relied on Decennial Census data, and while the published data by race and sex, especially the 1970 *Detailed Characteristics,* Vol. I, Chapter D, and Vol. II *Subject Reports,* are quite rich, it inevitably is the case that researchers desire tabulations of variables by race or sex which simply are not available in published form. Two avenues are open to solve this problem: one can either fund special tabulations by the Bureau to meet unique needs, or one may manipulate publicly available data on tape to generate the necessary analysis.

The recent availability of the 1960 and 1970 1/1000 Public Use Samples (PUS) in comparable format will no doubt greatly increase the ability of researchers to perform large scale micro-analysis. The 1/100 1970 and 1960 samples will increase this research ability that much more so; however, the more compact form and hence smaller associated computer costs may make the 1/1000 samples ultimately the more attractive.

Our purpose in this paper is severalfold. First, we shall describe the range of problems we entertained from an analytical point of view, with our goal being to obtain conclusions on the extent and pattern of wage rate and occupational discrimination against blacks and women. Secondly, we shall go into detail about the

*Financial support from the National Science Foundation, Grant No. GS3254, is gratefully acknowledged. Also, we wish to thank the Social Science Data Library, University of North Carolina, Chapel Hill, for access to and advice on using the 1960 and 1970 Public Use Samples.

procedures we used in manipulating the 1960 and 1970 (5%) Public Use Samples to achieve our substantive research. Finally, we shall summarize our empirical results to date and suggest additional pertinent research in areas of economic discrimination that may be accomplished with the two sample data bases.

ANALYZING ECONOMIC DIFFERENCES BETWEEN BLACKS AND WHITES, MEN AND WOMEN

The theoretical literature on economic differences between the races has emphasized the importance of differential wage rates for blacks and whites who do the same job. For Becker, the demand for labor services gets depressed by the employer's "taste" for discrimination.[1] In effect, this theory says that blacks must accept lower wage rates than whites to induce the employer to hire them. When this taste is operative, then, fewer blacks at lower wage rates will be hired; for employed whites, a benefit in the form of higher (than if there were no discrimination) wages obtains. A corollary of this theory is that discriminating firms must incur monetary costs as a result of bidding up the wages of the numerical majority of whites. Krueger has shown that under certain conditions a gain may be expected,[2] a point that Thurow has made more recently.[3] Quite clearly, if discrimination contributes positively to an employer's sense of well being or utility, and he maximizes his utility rather than his money profits, then it follows that the employer may on balance benefit from the discrimination since the higher monetary costs of production may be more than offset by the utility gain.

Others, notably Baran and Sweezy, have argued by contrast that employers gain as a result of discrimination in monetary as well as psychic terms.[4] They argue that the use of labor-intensive methods of production, which are made possible by the presence of a segregated (black) labor force, provide extraordinary profits. Building on Pigou,[5] Robinson made the same sort of monopsonistic argument (one buyer) in the case of women.[6] Since there is only one buyer or a tacit understanding among potential buyers of labor services to make them act as one for all practical purposes, the minority (women, blacks) may have to accept lower wages than the majority but without bidding up the wages of the majority.

A third group of theories of economic discrimination, due to Edgeworth,[7] more recently Bergmann,[8] and Strauss,[9] have argued that the observed economic differences between blacks and whites and men and women are due to industrial and occupational "crowding". That is, minorities face certain barriers to entry which prevent them from getting into higher paying industries and occupations. However, within homogeneous work groups, there is no difference between black and white wage rates; whether or not the male-female wage rate differences occur is usually said to be an empirical matter.

This brief overview of the more important points of view on economic discrimination strongly suggests a need for micro-data sets to do the requisite hypothesis testing. In terms of ascertaining the causes of income differences by race and sex from an economist's point of view, we can sort out the various theories described above in terms of their predictions about wage rate differences by race and/or sex. Thus, both the Becker theory of discrimination and the Baren-Sweezy-Robinson view of discrimination suggest that different (i.e., discriminatory) wage rates for the same work cause the well known income differentials. The "crowding" theorists by contrast expect there to be no differences in wage rates for comparable work groups. To test whether or not there are wage rate differences then requires data on wages by race, sex, job classification and firm, and experience. The Public Use Samples contain data of these kinds if we entertain that occupation proxies for job classification, and industry proxies for the firm or employer. The third assumption we need to entertain is that reported earnings last year, weeks worked, and hours worked information are sufficient to make inferences about wages.[10] We turn now to the statistical problems of making these assumptions.

STATISTICAL CONSIDERATIONS

We have two types of analysis that we wish to perform. First, we wish to make inferences about wage rate discrimination (unequal pay for equal work) against blacks and women. Second, we wish to understand the occupational and industrial exclusion process that apparently prevents blacks and women from obtaining certain types of jobs.

To make statements about wage rate discrimination, let us first inquire what information is available to us from the Public Use Samples. The figure closest to wage rate is the "earnings from wage and salary" (W) earned in 1959 and 1969. It is useful to write out, in an accounting sense, what this figure should represent (e.g., were there no response error): (see equation 1)

$$(1)\quad W = \sum_{mn}\sum (hr_{mn}w_{mn} + k_{mn}q_{mn} + w^*_{mn}hr^*_{mn} + w^{**}_{mn} + hr^{**}_{mn})$$

where:

w is the basic hourly wage rate
hr is the number of hours actually worked
k is the piece work wage rate
* denotes overtime during the week
** denotes overtime during the weekend
n'th day
m'th week
q is the piecework output

Of course this complicated sum, over n days in the week, and m weeks in the working year, is exclusive of additional monetary benefits that the employer may provide through contributions to pension plans as well as the more difficult-to-measure nonmonetary psychic benefits from having a clean job, nice surroundings and so forth.

What (1) alerts us to is that a wide variety of factors for one worker responding to the question about total wages and salary last year could cause his W to differ from others, even though their hourly wage rate, w, was the same. Important intervening factors include different number of hours worked in any week, different number of weeks worked in a year, differential treatment in terms of access to overtime during the week and weekends. Moreover, in a year's period, workers change jobs, get promoted to higher paying ones, experience unemployment and illness; all of these factors can cause annual earnings to vary, even though basic hourly wage rates might be the same.

We can through regression analysis estimate the effect of race, sex, education, and labor force experience on earnings last year. However, the estimated coefficients on race and sex may be misleading if we do not account for at least some of the above considerations. Also, we need to compare blacks and whites and men and women of comparable job responsibilities; that is, we need to hold constant industry and occupation as well. A possible regression model to account for these factors is as shown in equation (2):

$$(2)\quad W = \beta_1 + \beta_2\ \text{Race} + \beta_3\ \text{Sex} + \beta_4\ \text{Experience} + \beta_5\ \text{Education} + \sum_{j=6}^{k-1}\beta_j\ \text{Ind} + \sum_{j=k}^{k+m}\beta_j\ \text{Occ} + e$$

That is, we run a multiple regression of earnings last year with regressors of race, sex, experience, education, and dummy variables for industry and occupation. In particular, we construct dummy variables for two-digit census occupations and industries. The coefficients on race and sex tell us if blacks earn less and if women earn less, given comparable levels of experience, education, etc.

A second set of problems arises when we inquire about the likely correlation between race and sex with the industry and occupation dummy variables. We feel *a priori* that blacks and women are excluded from certain jobs. Collinearity is likely then to be a problem. A solution to this is to stratify the sample into separate industry and occupation subsamples and perform the regression analysis separately within each industry-occupation group.

Since we wish to disaggregate temporally as well as by industry and occupation, we need to combine information available about time spent working. Two pieces of information are available from the Public Use Samples; first, information on weeks worked, and information on hours worked in the reference week.

The temporal disaggregation can be achieved by estimating average weekly earnings for individuals by dividing their annual earnings figure by weeks worked. Alternatively, we may perform two separate regressions and use the information from each to make inferences about racial and sex differences in average weekly salary. Table 1 shows the nine combinations of coefficients on race that can occur and the kind of inferences we make about "salary" discrimination. We could perform analogous regressions on average weekly salary and hours worked, and make analogous inferences about hourly wage rate discrimination.

The second set of analysis involves the prediction of industry and occupation of blacks and women under certain conditions. Given educational attainment, and labor market experience, we seek to find out where industrially and occupationally we expect blacks and women to be. Of particular interest is whether or not the chances of being in certain industries and occupations have changed over the period 1960 to 1970. As a preliminary step in this analysis, we suggest the following regression models (3) and (4):

$$(3)\quad Occ_i = \beta_{1i} + \beta_{2i}\,\text{Race} + \beta_{3i}\text{Sex} + \beta_{4i}\text{Experience} + \beta_{5i}\text{Education} + e_3$$

$$(4)\quad Ind_j = \beta_{1j} + \beta_{2j}\text{Race} + \beta_{3j}\text{Sex} + \beta_{4j}\text{Experience} + \beta_{5j}\text{Education} + e_4$$

That is, we regress the demographic characteristics on the dummy variables for industry and occupation. The estimated coefficients are then the conditional probability that one will be in a particular industry or occupation.

DATA BASE AND COMPUTATIONAL CONSIDERATIONS

The above statistical design was made with the availability of the Public Use Samples in mind. However, the peculiarities of the documentation and limitations of computer size and speed also affected the research design, and it is to these matters that we now turn.

Since the analysis is directed at the measurement of wage-rate discrimination, we are only interested in the employed with non-zero earnings. Moreover, since it is generally known that part-time wage institutions differ considerably from full time, it was decided to select a sample for 1960 and 1970 which contained non-self-employed, full-time workers. To this end we generated FORTRAN extract routines for each year to create the basic working tapes from the 1960 and 1970 (5%) 1/1000 samples.[11] Also, it should be noted that we worked only with person records since housing data and family structure were not thought to substantially affect wage rates.

Table 1

Inferences on Salary Discrimination against Blacks from Regressions of the Form per Industry-Occupation Cell a/

$$W = \beta_{11} + \beta_{12}\,\text{Race} + \beta_{13}\text{Sex} + \beta_{14}\text{Experience} + \beta_{15}\,\text{Education} + e_1$$

$$Wks = \beta_{21} + \beta_{22}\,\text{Race} + \beta_{23}\text{Sex} + \beta_{24}\text{Experience} + \beta_{25}\,\text{Education} + e_2$$

W \ Wks →	$\hat{\beta}_{22} < 0$	$\hat{\beta}_{22} = 0$	$\hat{\beta}_{22} > 0$
$\hat{\beta}_{12} < 0$	Discrimination b/ / Reverse c/ Discrimination	Reverse Discrimination	Reverse Discrimination
$\hat{\beta}_{12} = 0$	Discrimination	No Discrimination	Reverse Discrimination
$\hat{\beta}_{12} > 0$	Discrimination	Discrimination	Discrimination b/ / Reverse c/ Discrimination

a/ Note: Race = 1 for whites, 0 for blacks; Sex = 1 for men, 0 for women; W = Earnings last year; Wks = Average weekly salary

b/ If whites earn relatively more (less) and work relatively less (more) than blacks. then salary discrimination is inferred.

c/ If whites earn relatively more (less) and work relatively more (less) than blacks, then "reverse discrimination" in terms of weekly salary may be inferred, although whites still earn more in a year than blacks.

To pass the 1970 PUS (5%), we performed several test runs to insure that the extraction routine was working properly. In particular, we retrieved on the following variables:

1. Race, excluded all those except whites and Negroes
2. Age, excluded those under 14 years of age
3. Hours Worked, excluded those who did not work in the reference week
4. Weeks Worked, excluded those who did not work last year
5. Job Worked in the Reference Week, excluded those who never worked, persons in the armed forces, persons not in the labor force who did not report the last year worked, and persons not in the labor force for whom the last year worked was 1959 or earlier
6. Class of Worker, excluded those working without pay and not applicable.

In total, 23 variables were extracted from the above universe. Starting with 277,755 records, the above retrieval reduced the file to 71,662 adult records of 69 characters each. In addition to selecting on the above six variables, further stratification was performed by our statistical software when we performed the regression analysis. Running on the Triangle University Computation Center's IBM 370/165, the above extraction was accomplished at a cost of $42.23, requiring 7.31 minutes of CPU and 3.35 minutes of charge for passing the entire tape file.[12] In addition to extracting the basic 23 variables that we needed for subsequent analysis, we also transformed some of the more important variables so that they would be directly available for statistical analysis: hours worked, earnings last year, and weeks worked were converted from coded values into midpoints of the intervals provided in the technical documentation.

Using essentially the same variables listed above, we performed an analogous extraction on the 1960 Public Use Sample. Passing the 237,990 records, we created a working tape of 60,596 working adults. This file contained 17 variables of total length 47 characters. Costs here were more modest; total dollar charge was $33.69 using 5.39 minutes of CPU and 2.56 minutes for tape. We should note that while the 1960 and 1970 PUS's are essentially the same for our purposes, there are several significant differences that affected our extraction procedure. The most important difference for us between 1970 and 1960 is in the coded intervals for earnings. The 1960 intervals are in $100's through $9,999 and then in $1,000 intervals to $25,000; the 1970 intervals are by $100's to $50,000.

Since our statistical design was reasonably well thought out before we began our extractions, we generated relatively short and efficient data tapes. Working in 1600 BPI with the 1970 tape blocked at 6900 and the 1960 tape blocked at 7050, we experienced subsequent statistical production tape reads of only 29.3 seconds for the 1970 tape and 16.6 seconds for the 1960 tape on the Model 165.

While extraction was relatively trouble-free, subsequent use for statistical analysis ultimately became a burden. The most important problem we faced in actual analysis involved definitional questions that were not resolved by quick reference to the documentation. The most persistent problem we had with regard to definitions involved the mapping of three-digit industry and occupation codes into major two-digit categories (11 major occupation groups and 16 major industry groups—see Table 5).

Basically, the industry-occupation code problem had two dimensions. First, the range of numerical values of three-digit codes for the two years are different, so that to group into major categories, one must first make comparable lists of codes. Second, the 1970 classification scheme is much more detailed, and as a result, there are many more codes to contend with. The two problems interacted in the case of accountants who have an occupation code of 000 in 1960 but a nonzero code in 1970. We would urge for other potential users a thorough reading of Census Working Paper 26 which details the changes in definition and allocation procedures.[13] Correspondence of percent distribution of major occupations and industries for 1960 and 1970 from our PUS extracts with published figures is reasonably strong, even though the published tables combine part-time with full-time workers in 1960 and 1970. Of interest is that we begin with working samples of well over 60,000 persons to analyze differences in wage rates and occupational and industrial achievement.

Because our statistical design required regressions performed within industry and occupation specific groups, we had to generate a stratified (by industry and occupation) data base to do the regressions. Of course, neither Public

Use Sample is ordered in this manner. Fortunately, a rather efficient statistical system is available at the Triangle University Computation Center and supported by the Duke University Computation Center which readily solved the problem. The Tele-Storage and Retrieval System (TSAR) performs sorting and merging operations as the first step of any statistical analysis by creating storage areas in core within which it accumulates sums of squares as it reads or processes each logical record. Hence, to create our matrix of all 11 occupations by 16 industries or 176 cells, only one pass through each data tape was necessary. This contrasts with the better known Statistical Package for the Social Sciences (SPSS) solution to the sort-merge problem which essentially entails multiple passes of an internally generated file. While the TSAR routine requires only one pass of the data, it required a minimum of 300K to do the job. By far, this part of the analysis was the most costly. For example, to perform the sorting into 176 groups, perform 4 regressions per group, and do regression analysis on the entire universe used 13.35 minutes of CPU on the 370/165 at a total dollar cost of $81.29. Costs to analyze the 1960 PUS were proportionately lower because of the smaller number of observations; however, the overall cost was considerable.

In retrospect, had we known that we would be passing the tapes as many times as we did to do regression analysis, it would have been more efficient to have initially created a sorted output tape which TSAR would have done as part of one run. With a sorted tape, we could have performed our analysis with other subroutine regression packages. We should note, however, that TSAR is written virtually entirely in double precision and in good measure in assembler language. It is not clear then that going to, say, one of the Scientific Subroutines, would have yielded us faster and more accurate results.

One matter that we did anticipate was the desire for additional variables, a desire that was effectively resisted. Once we specified our regression analysis, we decided to extract the smallest set of data, so that we would not be led down the primrose path of "one more run to see how variable x affects earnings." In fact, the multi-passing that contributed so heavily to the final computer time bill was due to our errors in getting the industry and occupation codes sorted into the two-digit categories and to properly sorting the universe beyond that performed with the FORTRAN extract routine. Perhaps the best advice we can give to the user who wants to save both computer time and through-put time in the long run is to assemble the various stages of his analysis to the point of setting up each run before actually performing any of them. Of course, many errors can not be caught until one sees results and interacts with various types of control check totals. Another bit of advice is to generate as much in the way of optional printout in each pass as possible. For example, due to an oversight, we neglected to get means, correlations, and cross-products for each regression run on the sorted data. This was simply an oversight, but required a separate pass (i.e., $80.00) to get these results which were ultimately needed.

OVERVIEW OF EMPIRICAL RESULTS

We can not provide here a detailed display and interpretation of the empirical results generated by the above procedures as they are the subject of several separate papers. Several points will, however, be highlighted.

First, when we performed a "traditional" statistical analysis on the 1960 and 1970 samples by running dummy variable (for industry and occupation) regressions on earnings last year, we found important differences between earnings of blacks and whites and men and women. Table 2 displays the regression coefficients for the two years and indicates that in 1960 blacks earned $623.57 less than whites and women earned $1954.56 less than men. By 1970, the difference in annual earnings for blacks had risen to $812.98 and for women to $3194.94. The "traditional" analysis then suggests a pervasive pattern of annual earnings differences which may be due to wage rate discrimination.

Table 2

Dummy Variable Regressions on Earnings Last Year: 1960, 1970

W =	1960	1970
β_1	-1519.05	-2956.27
β_2 RACE	623.57**	812.98**
	(43.35)	(60.77)
β_3 SEX	1954.56**	3194.94**
	(30.77)	(44.41)
β_4 "AGE"	45.93**	72.28**
	(0.95)	(1.34)
β_5 EDUCT	283.02**	463.66**
	(4.99)	(7.73)

*Significant at 95% level.
**Significant at 99% level.
$H_o : \beta_{2,...5} > 0$.

Note: Figures in parentheses are the standard error of estimate.

Table 3
Patterns of Salary Discrimination against Blacks, Women, Less Experienced and Less Educated in 1960 and 1970, Controlling for Industry and Occupation

Distribution of Significant Coefficients (.01 Level) from Single Equation (1960 Data)

1960 N = 88 (two-digit industry by occupation cells)

	Regressor			
	Race	Sex	Experience	Education
Earnings Last Year	25	49	58	64
Average Weekly Earnings	21	45	47	59
Weeks Worked	2	25	28	9

Distribution of Significant Coefficients (.01 Level) from Single Equation (1970 Data)

1970 N =109 (two-digit industry by occupation cells)

	Regressor			
	Race	Sex	Experience	Education
Earnings Last Year	16	66	66	73
Average Weekly Earnings	12	61	47	63
Weeks Worked	4	37	51	11

Table 4
Patterns of Salary Discrimination against Blacks in 1960 and 1970, Comparing Earnings Last Year and Average Weekly Salary, Controlling for Industry and Occupation

1960 Pattern of Significance at 95% Level for $\hat{\gamma}_{12}$ and $\hat{\gamma}_{22}$ from:

$$W = \gamma_{11} + \gamma_{12}\,\text{Race} + \gamma_{13}\text{Sex} + \gamma_{14}\text{"Age"} + \gamma_{15}\text{Educt} + \mu_1$$

$$WKS = \gamma_{21} + \gamma_{22}\,\text{Race} + \gamma_{23}\text{Sex} + \gamma_{24}\text{"Age"} + \gamma_{25}\,\text{Educt} + \mu_2$$

W ↓ / WKS →	$\hat{\gamma}_{22} < 0$	$\hat{\gamma}_{22} = 0$	$\hat{\gamma}_{22} > 0$
$\hat{\gamma}_{12} < 0$	0	0	0
$\hat{\gamma}_{12} = 0$	2	48	4
$\hat{\gamma}_{12} > 0$	1	30	3 / 0

1970 Pattern of Significance at 95% Level for $\hat{\gamma}_{12}$ and $\hat{\gamma}_{22}$ from:

$$W = \gamma_{11} + \gamma_{12}\,\text{Race} + \gamma_{13}\text{Sex} + \gamma_{14}\text{"Age"} + \gamma_{15}\,\text{Educt} + \mu_1$$

$$WKS = \gamma_{21} + \gamma_{22}\,\text{Race} + \gamma_{23}\text{Sex} + \gamma_{24}\text{"Age"} + \gamma_{25}\,\text{Educt} + \mu_2$$

	$\hat{\gamma}_{12} < 0$	$\hat{\gamma}_{12} = 0$	$\hat{\gamma} > 0$
$\hat{\gamma}_{22} < 0$	0	1	0
$\hat{\gamma}_{22} = 0$	3	68	6
$\hat{\gamma}_{22} > 0$	1	25	5

When we perform regression analysis within the stratified industry-occupation cells and tabulate the pattern of significant coefficients, a rather different picture emerges. Table 3 displays the number of times each of the regressors' co-efficient was significant at the .01 level when regressed on annual earnings, average weekly earnings, and weeks worked. Thus in 1960, out of 88 valid cells (enough blacks, enough women, enough observations), in only 25 did race have a significant effect on total annual earnings, in only 21 did race have a significant effect on average weekly earnings, and in only 2 did race have a significant effect on weeks worked. In 1970 there were 109 valid cells, but in only 16 was race significant in terms of annual earnings, in only 12 in terms of average weekly earnings, and in only 4 in terms of weeks worked.

Using the two-equation technique increased the number of times we found salary and wage rate discrimination (see Table 4 which follows the form of Table 1), but did not alter the basic finding that unequal pay for equal work for blacks per se is less prevalent than previously thought. Since income differences by race are well documented, this suggests that they are primarily due to crowding into less favorable industries and occupations in terms of salary and wage scales.

We should note that the situation with regard to women is more suggestive of actual discrimination. As shown in Table 3, out of 88 valid cells in 1960, sex had a significant effect on total annual earnings in 49 cells, a significant

Table 5

Predicting Two-Digit Occupational Attainment in 1960 and 1970

1960

Pr (Occupation) =	β_1	β_2 Race	β_3 Sex	β_4 Experience	β_5 Education
1 Professional	-0.439256	-0.028866**	-0.008976**	0.002181**	0.049561**
2 Farmers and Farm Managers	-0.000028	0.000843	0.000919**	-0.000001	-0.000063
3 Managers, Officials, and Proprietors, except Farm	-0.227056**	0.0341145**	0.066050**	0.002271**	0.015212**
4 Clerical and Kindred Workers	0.271354**	0.097422**	-0.291069**	-0.002151**	0.006671**
5 Sales Workers	-0.061432	0.043001**	0.007156**	0.000578**	0.006275**
6 Craftsmen, Foremen, and Kindred Workers	0.017445	0.116329**	0.212512**	0.000564**	-0.010726**
7 Operatives and Kindred Workers	0.653722	0.047614**	0.047146**	-0.003226**	-0.038779**
8 Private Household Workers	0.141978	-0.083199**	-0.049281**	0.000429**	-0.002387**
9 Service Workers except Private Household	0.290261	-0.123058**	-0.063427**	0.000888**	-0.007303**
10 Farm Laborers and Foremen	0.112781	-0.016699**	0.018002**	-0.000674**	-0.00698**
11 Laborers, except Farm and Mine	0.23423	-0.087501**	0.060968**	-0.060859**	-0.011478**

1970

Pr (Occupation) =	β_1	β_2 Race	β_3 Sex	β_4 Experience	β_5 Education
1 Professional	-0.522976	-0.012346**	-0.004679	0.001533**	0.055574**
2 Farmers and Farm Managers	-0.000180	0.000769	0.000862**	0.000013	-0.000042
3 Managers, Officials, and Proprietors, except Farm	-0.244616	0.037583**	0.064917**	0.002126**	0.016861**
4 Clerical and Kindred Workers	0.315433	0.066852**	-0.306019**	-0.001701**	0.004655**
5 Sales Workers	-0.048385	0.037409**	0.013036**	0.000505**	0.004651**
6 Craftsmen, Foremen, and Kindred Workers	0.149764	0.080687**	0.204509**	-0.000143	-0.016581**
7 Operatives and Kindred Workers	0.680961	-0.020464**	0.050078**	-0.002113**	-0.037853**
8 Private Household Workers	0.075822	-0.040249**	-0.023392**	0.000288**	-0.001883**
9 Service Workers except Private Household	0.336672	-0.093699**	-0.062133**	0.000504**	-0.010960**
10 Farm Laborers and Foremen	0.063298	-0.005740**	0.011169**	-0.000240**	-0.004235**
11 Laborers, except Farm and Mine	0.194207	-0.050801**	0.051652**	-0.000831**	-0.010186**

Note: * statistically different from 0 at .05 level
** statistically different from 0 at .01 level

effect on average weekly earnings in 45, and a significant effect on weeks worked in 25. This situation was even more pronounced in 1970.

The second type of analysis undertaken in our study involved predicting occupation of employment for the two time periods. Table 5 displays the regression coefficients obtained by ordinary least squares. For example, in 1960, every additional year of schooling added .049 to one's odds of being in the Professional occupation (occupation 1). Interestingly, whites were less likely to be Professionals in both 1960 and 1970, holding everything else constant; similarly, women were more likely to be in the Professional occupation than men in both years. Blacks and women seem to experience "crowding" in the Private Household and Service Workers occupations (8 and 9), blacks in the Laborers occupation (11), and women in the Clerical occupation (4). However, the coefficients in 1970 are generally smaller than in 1960. This suggests in turn that the extent of occupational discrimination has lessened, holding constant educational attainment and labor market experience. Of interest is that in 1960 the Operatives and Kindred Workers occupation (7) was more likely to be a white occupation than a black one; however, in 1970 that situation had reversed, so that Operatives was more likely to be a black occupation than a white one.

SUMMARY

Analysis of economic discrimination against blacks and women requires large amounts of micro-data to properly account for the multitude of factors other than prejudice which may cause wage rate differentials. The 1960 and 1970 1/1000 Public Use Samples were successfully used to test a variety of hypotheses about the extent and location of wage rate differences by race and sex in the U.S. economy. While use of the files was reasonably straight forward, more caution in constructing the industry and occupation codes for the two years and more thoughtful planning of the ultimately desired statistical output would have resulted in both computational and through-put savings. In terms of substantive results, it was found that blacks suffered less than women in terms of the frequency of unequal pay for equal work; however, substantial evidence of "crowding" was found. Additional research may explore three-digit industry and occupation patterns of earnings and investigate further the prediction of occupational and industrial attainment.

REFERENCES AND NOTES

1. Gary S. Becker, *The Economics of Discrimination,* 2nd edition (Chicago: University of Chicago Press, 1971).
2. Anne O. Krueger, "The Economics of Discrimination," *Journal of Political Economy,* Vol. 71 (October, 1963), pp. 481-486.
3. Lester C. Thurow, *Poverty and Discrimination* (Washington: Brookings Institution, 1969).
4. Paul A. Baran and Paul M. Sweezy, *Monopoly Capital: An Essay on the American Economic Order* (New York: Monthly Review Press, 1966).
5. A. C. Pigou, *The Economics of Welfare* (London: The Macmillian Co., 1920), Chapter 14.
6. Joan Robinson, *The Economics of Imperfect Competition* (London: The Macmillian Co., 1933), Chapter 26.
7. F. Y. Edgeworth, "Equal Pay to Men and Women for Equal Work," *Economic Journal,* Vol. 32 (December, 1922), pp. 431-457.
8. B. R. Bergmann, "The Effect on White Income of Discrimination in Employment," *Journal of Political Economy,* Vol. 79, No. 2 (March-April, 1971), pp. 294-313.
9. Robert P. Strauss, "Industrial Patterns of Male Negro Employment,"*Journal of Human Resources,* Vol. 7, No. 1 (Winter, 1972), pp. 111-118.
10. We first worked with the 1967 Survey of Economic Opportunity developed by O.E.O. (University of Wisconsin, fixed-format version). Since the sampling procedure was designed to overrepresent poor persons, the number of cases for full-time, non-self-employed workers became too small for a detailed analysis.
11. At that time, the DUALabs PUSH program had just arrived on campus and was not yet operational. Rather than lose through-put time, we coded and executed our own routines. Subsequent time and cost comparisons with PUSH runs here indicate that the PUSH software is twice as fast as our FORTRAN procedures.
12. Cost figures in the text are at overnight rates or 50% of normal charge levels.
13. U.S. Bureau of the Census, *1970 Occupation and Industry Classification Systems in Terms of Their 1960 Occupation and Industry Elements,* Technical Paper No. 26, by John Priebe, Joan Heinkel, and Stanley Greene (Washington: U.S. Government Printing Office, 1972).