WORKING PAPER #312
INDUSTRIAL RELATIONS SECTION
PRINCETON UNIVERSITY
FEBRUARY 1993

Wage Dispersion, Returns to Skill, and Black-White Wage Differentials

David Card
Princeton University

and

Thomas Lemieux
University of Montreal

February 1993

*We are grateful to Jim Powell for many helpful discussions, and to participants in the Princeton Labor Economics Lunch Seminar for comments and suggestions.

Wage Dispersion, Returns to Skill, and
Black-White Wage Differentials

ABSTRACT

During the 1980s wage differentials between younger and older workers and between more and less educated workers expanded rapidly. Wage dispersion among individuals with the same age and education also rose. A simple explanation for both sets of facts is that earnings represent a return to a one-dimensional index of skill, and that the rate of return to skill rose over the decade.

We explore a simple method for estimating and testing 'single index' models of wages. Our approach integrates 3 dimensions of skill: age, education, and unobserved ability. We find that a one-dimensional skill model gives a relatively successful account of changes in the structure of wages for white men and women between 1979 and 1989. We then use the estimated models for whites to analyze recent changes in the relative wages of black men and women.

David Card
Department of Economics
Princeton University

Thomas Lemieux
Department of Economics
University of Montreal

It is now a well-established fact that wage inequality grew over the 1980s (see for example Tilly, Bluestone, and Harrison (1987), Murphy and Welch (1992), Juhn, Murphy, and Pierce (1992), Bound and Johnson (1992)). Wage differentials between younger and older workers and between more and less educated workers expanded from the late 1970s to the late 1980s. Wage dispersion among men and women with the same age and education also rose. A unified explanation for all these changes is suggested by the hypothesis that labor market earnings represent a return to a one-dimensional bundle of "human capital" or "skill". Changes over time in the rate of return to skill would be expected to increase the wage gaps between different age and education groups, and increase wage dispersion within narrowly defined age/education cells. As noted by Juhn, Murphy and Pierce (1991) a generic rise in the return to skill also has implications for other measured wage gaps, including black-white and male-female differentials. To the extent that unobservable components of skill differ by race or sex, a rise in the return to skill would be expected to widen the gap between black and white or male and female workers.[1]

In this paper we propose a simple technique for estimating and testing a "one-dimensional skill" model of changes in the structure of wages. The method is based on comparing means and quantiles of wages for narrowly-defined age and education cells over time. This approach integrates three alternative dimensions of "skill": education, age (or labor market experience), and unobserved ability within age/education categories. We

[1]It should be noted at the outset that although wage differentials within the male and female populations grew over the 1980s, the male-female gap in average hourly earnings closed dramatically: from 38% in 1979 to 28% in 1989 (see Blau and Kahn (1992) for a recent analysis). A simple one-dimensional skill model cannot reconcile this change with other changes over the 1980s.

fit a series of single-skill models to the wage structures of white men and women in 1973-74, 1979, and 1989. We then use these models to analyze and interpret changes in black-white relative wages over the 1980s.

A one-dimensional skill model provides a relatively accurate account of changes in the structure of white female wages from 1979 to 1989. Over the 1980s we estimate that the return to skill for white women increased by 40 percent. Similar models are less successful in describing changes in the structure of wages among white men. In particular, the rise in relative earnings of young college-educated men is too large, even granting the 25 percent overall rise in the return to skill for white men during the decade.

Our integration of observable and unobservable skill components suggests that changes in within-cell wage dispersion follow the same pattern as changes in the structure of returns to age and education. Patterns of wage growth for male college graduates again pose the greatest difficulty for a one-dimension skill model. Results for men and women suggest that 40-45 percent of residual wage variation within age and education cells is attributable to unobserved abilities whose market valuation rose during the 1980s.

Comparisons of the wage gains achieved by black men and women during the 1980s with the predictions generated by models of the white wage structure lead to two sets of conclusions. On the one hand, changes in the white wage structure provide a surprisingly good forecast of average wage growth for blacks. Black men's wages grew 0.7% faster than predicted by the pattern of white male wage growth, while black women's wages fell 1.8% short of the prediction based on white female wage growth. On the other hand, there were sizeable relative gains and losses within the black labor force. Wages of older blacks rose faster than predicted while wages of younger blacks lagged behind. College-educated black women suffered significant losses relative to predictions based on the wage growth of white women.

I. Single Index Models of Wages

This section outlines the conceptual framework used throughout this paper to model changes in the structure of wages. We begin by considering the special case in which observed (log) earnings are a linear function of a one-dimensional bundle of skill. Let k_i represent the skill index of individual i and assume that the log wage of i in period t is a linear function of k_i: say $\beta_t k_i$. The observed log wage of individual i is w_{it}, where

$$(1) \qquad w_{it} = \beta_t k_i + \varepsilon_{it}$$

and ε_{it} can be interpreted as measurement error.[2] If $\beta_1/\beta_0 > 1$, then we say that the return to skill has increased between periods 0 and 1.

This simple model can be implemented empirically by assuming that

$$(2) \qquad k_i = x_i\theta + a_i,$$

where x_i is a vector of observable characteristics (education, age, etc.) and a_i is an unobservable component of skill. Equations (1) and (2) imply a series of linear regression models with time-dependent coefficients:

$$(3) \qquad w_{it} = x_i\alpha_t + e_{it},$$

where $\alpha_t = \beta_t \bullet \theta$ and $e_{it} = a_i + \varepsilon_{it}$. In this framework, an increase in the

---

[2]Alternatively, ε_{it} can represent the result of randomness, luck, or mistakes in the labor market.

return to skill implies a uniform re-scaling of the regression coefficients associated with observed skill attributes (education, age, etc.).[3] An increase in the return to skill also raises the residual standard deviation of wages, although the increase will be smaller than percentage increase in the α's if the measurement error variance is constant over time. In particular, the cross-sectional variance of e_{it} is

$$s_t^2 = \beta_t^2 \, \sigma_a^2 + \sigma_\varepsilon^2,$$

where σ_a^2 is the cross-sectional variance of unobserved ability and σ_ε^2 is the variance of ε_{it}.

Nonlinear Models

A more general version of the single index model assumes that earnings in period t are a monotonically increasing function of skill, plus measurement error:

(4) $$w_{it} = f_t(k_i) + \varepsilon_{it},$$

where without loss of generality $f_0(k) = k$. In this framework we would say that the return to skill rose between periods 0 and 1 if $f_1'(k) > 1$ for all k: in other words, if f_1 is everywhere steeper than f_0. Equation (4) implies

(5) $$w_{i1} = f_1 \left(w_{it} - \varepsilon_{it} \right) + \varepsilon_{i1}.$$

Thus we can evaluate changes in the return to skill by estimating the transformation between w_{i0} and w_{i1} and asking whether its slope is greater than unity. Nonlinearities in f_0 permit wage differentials at different

[3]Notice that with two periods of data only the ratio β_1/β_0 is identified.

points in the wage distribution to expand more or less rapidly without abandoning the hypothesis of a single index of skill. For example, a more rapid expansion of wage differentials among highly skilled workers implies that f_0 is convex.

In principle it is possible to estimate equation (5) using panel data on the same individuals over time. An alternative procedure that we pursue in this paper is to consider repeated cross-sectional observations on groups of individuals with the same observable skill characteristics. In particular, suppose that individuals can be stratified into J cells (based on single years of age and education in the analysis below). Let k_{ij} represent the skill index of person i in cell j, where

$$k_{ij} = k_j + a_{ij}, \quad \text{with } E(a_{ij}) = 0.$$

The term a_{ij} is interpreted as the unobserved component of skill of person i, relative to mean skill for cell j. Finally, assume that log wages of person i in cell j in period t are generated by

$$(6) \qquad w_{ijt} = f_t(\, k_j + a_{ij} \,) + \varepsilon_{ijt},$$

where (as before) ε_{ijt} is interpreted as measurement error or some random component of wages, and $f_0(k) = k$.

The mean log wage for cell j in period 0 is w_{j0}, where

$$w_{j0} = E\,(\, f_0(\, k_j + a_{ij} \,)\,) = k_j,$$

the mean level of skill for cell j.[4] The mean log wage of cell j in period 1 is:

[4]Our normalization $f_0(k) = k$ implies that "skill" is measured by wages in period 0.

$$w_{j1} = E\left(f_1(k_j + a_{ij} \right) \approx f_1(k_j) + 1/2 \ \mathrm{var}[a_{ij}] \ f_1''(k_j).$$

Mean cell wages in period 1 are therefore related to mean cell wages in period 0 by

$$(7) \quad w_{j1} = f_1(w_{j0}) + r_j,$$

where the "remainder term" r_j is 0 if f_1 is linear or if the variance of unobserved skills is negligible. Otherwise,

$$r_j \approx 1/2 \ \mathrm{var}[a_{ij}] \ f_1''(k_j),$$

which will be constant across cells if the within-cell variance of unobserved ability is constant across cells and if the change in the structure of wages is not "too far" from a quadratic transformation.

Equation (7) suggests a simple and intuitively appealing method for estimating the degree of change in the structure of wages: one simply finds a suitable approximation to the mapping between mean cell wages in different periods. In the empirical analysis below we consider polynomial approximations to f_1, although more general functions could be easily used. In principle, panel data are not required, so long as individuals in a given cell in one period are viewed as exchangeable with individuals in the same cell in a different period. This exchangeability condition will <u>fail</u> if individuals from different cohorts have different mean levels of unobservable skill, or if the relation between skill and the cell classifications changes between cohorts.[5]

Equation (7) also suggests a simple procedure for <u>testing</u> a one-dimensional skill model. Apart from sampling errors (and errors in the

---

[5]For example, women of a given age from earlier cohorts may have lower actual labor market experience than women of the same age from later cohorts.

approximation of f_1) mean cell wages in period 1 are a function of mean cell wages in period 0. Given a choice of the approximation function, this type of restriction can be readily tested by conventional goodness-of-fit tests.

Models of Unobservable Skill

Under a set of simplifying assumptions the preceding framework can be extended to model changes in the overall distribution of wages in different cells. A one-dimensional skill model suggests a parsimonious structure for both mean cell wages and the quantiles of the within-cell wage distribution. Following the notation of the last section, the wage of individual i in cell j and period 0 (the base period used to define "skill") is

$$w_{ij0} = w_{j0} + a_{ij} + \varepsilon_{ij0},$$

where w_{j0} is the mean log wage in the cell, a_{ij} represents unobserved ability, and ε_{ij0} represents measurement error. Assume that a_{ij} and ε_{ijt} are normally distributed with variances σ_j^2 and σ_ε^2, respectively, and let $e_{ij0} = a_{ij} + \varepsilon_{ij0}$. The qth percentile of wages in the jth cell in period 0 is

$$w_{j0}^q = w_{j0} + s_{j0}\, z^q,$$

where $s_{j0}^2 = \sigma_j^2 + \sigma_\varepsilon^2$ is the variance of e_{ij0} and z^q is the qth percentile of the standard normal distribution.

Wages in period 1 are determined by

$$w_{ij1} = f_1\,(k_j + a_{ij}) + \varepsilon_{ij1}.$$

Assume that the transformation of wages for individuals in cell j is locally linear with intercept γ_j and slope β_j. Then

$$w_{ij1} \approx \gamma_j + \beta_j w_{j0} + \beta_j a_{ij} + \epsilon_{ij1}.$$

The mean wage for cell j in period 1 is

$$(8) \quad w_{j1} = \gamma_j + \beta_j w_{j0},$$

while the variance of wages within the jth cell in period 1 is

$$s_{j1}^2 = \beta_j^2 \sigma_j^2 + \sigma_\epsilon^2.$$

Finally, the qth percentile of wages in period 1 is

$$w_{j1}^q = w_{j1} + s_{j1} z^q.$$

Let R_j denote the fraction of within-cell variance attributable to measurement error (or random wage factors) for cell j in period 0. Then

$$s_{j1}^2 = s_{j0}^2 \left(\beta_j^2(1-R_j) + R_j \right).$$

Combining the last two expressions with equation (8) we obtain

$$(9) \quad w_{j1}^q = \gamma_j + \beta_j w_{j0}^q + s_{j0} \bullet z^q \bullet \delta_j,$$

where

$$\delta_j = \left(\beta_j^2 (1 - R_j) + R_j \right)^{1/2} - \beta_j.$$

Notice that for the median wage z^q=0, implying that changes in mean and median cell wages are identical (as must be true under the normality assumption). If $\beta_j > 1$ (i.e. the return to skill has increased) and R_j>0 (i.e., some fraction of within-cell variation is noise) the expression δ_j is negative. In this case the lower quantiles of wages increase by more than the mean or median, whereas the higher quantiles increase by less. This compression reflects the fact that an increase in the return to skill increases the within-cell standard deviation of wages less than proportionately whenever some fraction of wage dispersion is attributable to noise rather than "skill".

Equation (9) is derived under the assumption that the transformation f_0 is "locally linear". If f_0 is linear, then (9) implies that the cell quantiles in period 1 are linearly related to the cell quantiles in period 0, with quantile-specific intercepts.[6] More generally, assume that γ_j and β_j, the intercept and slope of f_0 for wage observations in cell j, are approximately linear functions of w_{j1} (in other words, that f_0 is approximately quadratic). Then

$$(10) \quad w_{j1}^q \approx \nu_0 + \nu_1 w_{j0}^q + \nu_2 (w_{j0}^q)^2 + s_{j0} \bullet z^q \bullet \delta_j,$$

for some constant coefficients (ν_0, ν_1, ν_2). In this case the cell quantiles in period 1 are (approximately) a quadratic function of the corresponding cell quantiles in period 0, with quantile-specific intercepts.

II. Econometric Issues

This section briefly outlines the econometric methods used in estimation and testing of the models proposed in the previous section. A more complete development is presented in the Appendix.

The models describe the relationship between cell-specific means or quantiles of wages in two different periods. According to equation (7) the mean log wage for cell j in period 1 is a simple function of the mean wage for the same cell in period 0, plus an approximation error which we take to be constant across cells. Equation (10) implies a similar relation between cell quantiles in different periods, with quantile-specific intercepts.

---

[6]Strictly speaking this also requires that the within-cell standard deviation of wages is constant across cells.

There are two main problems in estimation: choice of functional form, and the presence of sampling errors in the observed cell data. Our choice of functional form was determined by plotting mean cell wages (and wage quantiles) in one year against the corresponding means (and quantiles) in other years. As noted in more detail below, these plots suggest a smooth function with only limited curvature. In light of this evidence we have restricted our attention to linear and quadratic functional forms. For convenience we refer to these as linear and quadratic single index models.

Given a particular functional form, the presence of sampling errors in the observed means or cell quantiles in the base period induces a measurement error problem in the estimation. Unlike many applications, however, estimates of the variances of the measurement errors are readily available from the sampling errors of the base period data. Following Fuller and Hidiroglou (1978), these can be used to construct "measurement error corrected" least squares estimates.

For concreteness, consider estimation of the quadratic single index model for mean cell wages. The true model is

$$w_{j1} = a + b\, w_{j0} + c\, w_{j0}^2.$$

Let $\hat{w}_{j0}$ represent the estimated mean wage for cell j in period 0, and let $\hat{s}_{j0}$ represent the estimated standard deviation of wages. An estimate of the sampling variance of $(\hat{w}_{j0} - w_{j0})$ is $\hat{s}_{j0}^2/N_j$, where N_j is the number of observations in cell j in period 0. Note that $\hat{w}_{j0}^2$ is <u>not</u> an unbiased estimator of w_{j0}^2 (although it is consistent as the overall sample size tends to infinity). Rather, we use $\hat{w}_{j0}^2 - \hat{s}_{j0}^2/N_j$ as an unbiased estimator of the squared cell mean wage. Thus our statistical model is:

(11) $\hat{w}_{j1} = a + b\, \hat{w}_{j0} + c\, \{ \hat{w}_{j0}^2 - \hat{s}_{j0}^2/N_j \} + \eta_j,$

where η_j includes three terms:

$$\eta_j = (\hat{w}_{j1} - w_{j1}) - b\,(\hat{w}_{j0} - w_{j0})$$
$$- c\,\{\ (\hat{w}_{j0}^2 - \hat{s}_{j0}^2/N_j\ ) - w_{j0}^2\ \}.$$

The first term is the sampling error in the dependent variable, and poses no particular problem for estimation. The second and third terms, however, are functions of the sampling errors in the independent variables, creating a bias in ordinary least squares estimates of the coefficients (a,b,c).

Fuller and Hidiroglou (1978) propose a measurement-error corrected estimator that makes use of a priori information on the covariance matrix of the measurement errors of the independent variables. In obvious notation, write the true model as

$$y_j = x_j\pi, \qquad j=1,\ldots J,$$

and denote the observed data by $(\hat{y}_j,\ \hat{x}_j)$. Let

$$\hat{y}_j - y_j = \varepsilon_j, \quad \text{and}$$
$$\hat{x}_j - x_j = u_j.$$

Suppose that an estimate $\hat{\Sigma}$ of $E(u_j u_j')$ is available. Let $\hat{M}_{xx}$ denote the corrected second moments matrix of $\hat{x}_j$, and let $\hat{M}_{xy}$ denote the corrected cross-products of $\hat{x}_j$ and $\hat{y}_j$. Fuller and Hidiroglou (1978) propose the "measurement-error corrected least squares" estimator

$$\hat{\pi} = (\ \hat{M}_{xx} - \hat{\Sigma}\ )^{-1}\ \hat{M}_{xy}.$$

Under standard conditions, this estimator is consistent and asymptotically normally distributed with a readily computed covariance matrix (see the Appendix).

In our application to cell means, this estimator requires information on the joint sampling covariance matrix of $\hat{w}_{j0}$ and $\hat{w}_{j0}^2 - \hat{s}_{j0}^2/N_j$. Given the sampling covariance matrix of $\hat{w}_{j0}$ and $\hat{s}_{j0}^2$, we use the delta method to construct the required sampling variance matrix. In our application to cell quantiles, we follow a similar approach, making use of the assumption of normality to compute the sampling covariance matrix of the various quantiles and their squares (see the Appendix). Our tests for the goodness-of-fit of the single index model derive directly from equation (11), making use of estimates of the sampling errors of the dependent variable as well as the independent variables to construct an appropriate test statistic.

III. Single Index Models of the Wage Structure for White Men and Women

Data Description

This section summarizes our findings on the use of single index models to characterize changes in the structure of wages for white men and women. Our analysis is based on data from the 1973, 1974, 1979, 1984, and 1989 Current Population Surveys (CPS). Since 1979 the CPS survey has collected earnings information from one-quarter of the sample. Combined data from the 12 monthly surveys yield over 150,000 wage observations per year. Prior to 1979 comparable data were only collected in the May surveys. To increase the available sample size, we have pooled responses from the May 1973 and May 1974 surveys, yielding a sample of 70,000 wage observations from the mid-1970s.[7]

---

[7]Our samples exclude individuals with allocated hourly or weekly earnings data, as well as individuals whose reported or constructed hourly wage is below $2.01 or above $60.00 (in constant 1989 dollars). We also adjusted 1984 earnings observations for individuals whose weekly earnings are censored at $999 per week. Based on weekly wage patterns in the 1989 sample, we allocated an hourly wage of $26.89 to these observations.

In addition to their generous sample sizes these data sets have another advantage for studying the structure of wages. The earnings information refers to hourly or weekly earnings for the respondent's main job[8], rather than to total earnings on all jobs in the previous year, as in the Decennial Census or the March CPS. Thus the wage measure is closer in spirit to a point-in-time "price" of labor, and is unaffected by measurement error in the report of weeks worked. A potential disadvantage is the sample frame of individuals who held a job in the previous week. Individuals with lower employment probabilities will be under-represented in this frame relative to the population of individuals who held a job in the previous year (the sample frame for earnings data in the Census or March CPS).

To investigate the differences associated with alternative sample frames we compared average hourly earnings from our 1979 and 1989 samples to average hourly earnings constructed from retrospective earnings and hours data in the March 1980 and March 1990 CPS files. In addition to constructing an average hourly wage for individuals who held a job last year in the March data sets, we constructed two other wage measures: a weighted average hourly wage rate with individual weights based on the number of weeks of employment last year[9]; and an average hourly wage rate

---

Finally, May 1974 wage observations were deflated by 8.05 percent before being pooled with May 1973 observations.

[8]Individuals who are paid by the hour report an hourly wage rate. Others report usual weekly earnings and usual weekly hours, which we use to construct an hourly rate.

[9]In principle, weighting by weeks worked last year should adjust the March CPS data to a sample frame of individuals who were employed last week.

for "full-time full-year" workers.[10] The results of our investigation are summarized in Appendix Table 1. Average log hourly wage rates from our samples and the March CPS samples are surprisingly close. Contrary to our expectations, average hourly wage rates in the March CPS tend to be as high or higher than hourly wage rates in our samples. The weeks-weighted average and the average for full-time full-year workers are higher still. Black-white wage gaps are comparable across the alternative data sets, and are very similar for the three wage measures derived from the March CPS.

Tables 1a and 1b begin our data analysis by presenting some simple evidence on recent changes in wage differentials among white men (Table 1a) and white women (Table 1b). Rows 1a-1c show estimated wage differences between 46-55 and 26-35 year old workers at three different levels of education. Among both women and men age differentials for less educated workers expanded sharply in the 1980s. For college-educated workers, however, age differentials have been relatively stable. Rows 2a-2d show wage gaps between similarly-aged workers with different levels of education. These expanded at a roughly uniform rate for women. For men, however, the college-high school wage gap expanded more for young men and less for older men. Finally, rows 3a-3d show estimated standard deviations of log wages for 4 narrow age/education cells. These contracted slightly from 1973-74 to 1979 but then expanded during the 1980s -- with a generally greater increase among women.

It is clear from these two tables that age and education-based wage differentials have not expanded at a uniform rate over the 1980s (nor did

[10]Many recent studies of wage dispersion concentrate on full-time full-year workers e.g. Pierce and Welch (1992). In part, this choice is dictated by the absence of accurate annual hours information in March CPS surveys before 1976.

they change uniformly from 1973-74 to 1979). Young college educated white men made significant relative wage gains over the 1980s --leading to an expansion of the college-high school premium for young men and a reduction in the age differential for college-educated men. Among women the growth in wage differentials is more uniform, although the collapse of the age premium for college educated women is a notable exception.[11]

Single Index Results

To implement the estimation methods described in sections I and II we divided wage earners between the ages of 16 and 65 into 225 individual age and education cells. The cells are based on single years of education (with ≤ 8 years in the lowest cell and ≥ 18 years in the highest cell) and 1, 2 or 3 year age ranges (single year age ranges for ages up to 23, 2 year age ranges for ages 24 to 43, and 3 year age ranges for ages 44 and older).[12] We then computed the mean, median, 25th percentile and 75th percentile of log wages in each cell.

Figure 1 gives an overview of the relation between mean cell wages from men and women in 1973-74, 1979, and 1989. For reference, we have also plotted in each panel the line representing constant real wages between the base year and the ending year. All four panels of the figure indicate a strong correlation between mean cell wages in different years.[13] Only one

---

[11]Comparisons of the pattern of age profiles for college-educated women in different years suggests that there may be important cohort effects biasing down the cross-sectional age profiles.

[12]Our sample excludes individuals whose age is less than 6 plus their years of completed schooling.

[13]The correlations between mean cell wages in any two years range from 0.97 to 0.99.

of the four panels -- the panel showing men's wages in 1979 and 1989 -- shows a noticeable degree of curvature. The graphs of 1979 wages against 1973-74 wages show that real wages grew at about the rate of inflation over the late 1970s, although there was a tendency for higher-wage workers to lose ground (particularly among women). As shown in the lower panels, however, real wages of many workers fell sharply over the 1980s. Women with above-average wages enjoyed modest real wage gains, while most men had real wage losses.

Table 2 presents coefficient estimates and goodness-of-fit statistics for various single index models of male and female wages. All the models are estimated by the measurement-error-corrected least squares procedure described in section II.[14] Columns 1 and 4 present simple linear models while columns 2 and 5 present quadratic models. As suggested by the absence of curvature in the plots in Figure 1, the quadratic model does about as well as the linear model between 1973-74 and 1979. The same is true between 1979 and 1989 for women, but not for men.

The goodness-of-fit statistics for the single index models are well above conventional critical values.[15] The fit is relatively poorer for men than women -- particularly between 1979 and 1989. To gain some insights into the causes of failure of the single index model, we plotted fitted and actual mean cell wages in 1989 against mean cell wages in 1979,

---

[14]In fact, OLS estimates of the linear models are not too different than the estimates reported in the table. OLS estimates of the quadratic model, however, are slightly different and generally show a smaller quadratic term.

[15]A 1 percent critical value for the fit statistics in the tables is approximately 275.

using different indicators for cells with different levels of education.[16] The results are presented in Figure 2.

The plot in the upper panel of Figure 2 shows that college-educated men near the middle of the 1979 wage distribution had much faster wage growth than predicted by a single index model. These cells are composed of younger college graduates. On the other hand the wage growth of older college-educated men (those near the top of the 1979 wage distribution) was consistent with patterns for other education groups. Cells of college-educated workers also stand out in the lower panel of Figure 2. The plot suggests that wages of female college graduates near the top of the 1979 wage distribution grew "too slowly" over the 1980s.

A more formal way to test the single index specification is to add regressors to the model for mean cell wages (representing the levels of age or education in the cell). If the single index hypothesis is correct, mean wages in the base year are a "sufficient statistic" for wages in the ending year, and age or education should not help predict end-period wages. This idea is pursued in columns 3 and 6 of Table 2, where we have added the mean years of education in the cell as an additional predictor of wage growth. As is suggested by the simple wage gaps in Table 1 (and other previous research on the returns to age and education) the models in column 3 suggest that cells with higher levels of education had relatively lower wages in 1979 than would be predicted on the basis of their 1973-74 wages. The models for wage growth between 1979 and 1989, however, differ between men and women. For women, education has no significant effect on 1989 wages, controlling for 1979 wages. For men, cells with higher education

---

[16]We use the quadratic models in column 5 of Table 2 to form the predictions.

had significantly higher wages in 1989, controlling for wages in 1979.

Further evidence on the fit of single index models for mean cell wages is presented below. Before turning to this evidence, however, we discuss the results of fitting similar models to the wage quartiles of men and women between 1979 and 1989. Following equation (10), we assume that the 25th percentile, median, or 75th percentile of wages for a particular cell in 1989 is a linear or quadratic function of the corresponding wage quantile in 1979. Thus we fit models for 675 cell quantiles (3 quantiles for each of 225 cells). Our modified least squares estimation procedure makes no allowance for possible correlations between the 3 observed quantiles from each cell, although our estimated standard errors and goodness-of-fit statistics do take account of these correlations (see the Appendix).

Estimation results are presented in Table 3. In all models we include dummy variables for the 50th and 75th percentile observations (with the 25th percentile as a base). In the models for women we also include dummy variables indicating whether the 25th percentile of wages in either 1979 or 1989 is at or below the minimum wage for the particular year. These dummies were added after a visual inspection of the data (see below) showed the importance of the minimum wage in attenuating wage dispersion at the lower tail of the female wage distribution.

Under the assumptions underlying equation (11) (including normality of the within-cell wage distribution) the quantiles of wages should follow the same model as the mean of wages, with quantile-specific intercepts. Furthermore, the intercepts should be higher for lower quantiles. Both of these predictions are confirmed by the estimates in Table 3. Linear and quadratic single index models for the quantiles of male and female wages are very similar to the corresponding models for mean wages in Table 2. And the estimated 50th and 75th percentile dummies (in rows 5 and 6 of Table 3) show slower growth for the higher quantiles, controlling for the initial value of wages.

Figure 3 presents plots of the 25th and 75th percentiles of wages in 1989 against the corresponding quantiles in 1979. For reference, we have also plotted the fitted quadratic models for the mean of wages. The plots illustrate the basic conclusions from Table 3. Higher and lower quantiles of wages follow roughly parallel models, with more rapid wage growth for lower quantiles. Furthermore, models based on mean wages are relatively good predictors of wage growth for different quantiles of wages.

The data in the lower panel of Figure 3 also illustrate the effect of the relatively high minimum wage in 1979 on the dispersion of wages for younger and less-educated women. The logarithm of the 1979 minimum (1.06) acts as a lower bound for the 25th percentile of wages for all but 4 cells.[17] Over the 1980s the real value of the minimum wage eroded significantly (45 percent): only a few cells had 25 percent or more of workers at or below the minimum in 1989.

Under the assumptions of linearity and normality, equation (11) offers a simple interpretation of the quantile-specific intercepts in Table 3. Consider a linear single-index model with a constant within-cell standard deviation of wages s. Then the predicted coefficient of the 50th percentile dummy is

$$d_{50} = -.6745\ s\ \{\ (\beta^2\ (\ 1\ -\ R)\ +\ R)^{\ 1/2}\ -\ \beta\ \},$$

where β is the slope coefficient of the single index model and 1-R is the

---

[17]37 cells have the 25th percentile equal to the minimum wage.

fraction of within-cell variation attributable to ability. The predicted coefficient of the 75th percentile dummy is $2d_{50}$. Using an estimate of $s \approx 0.40$ and $\beta = 1.2$, the estimated coefficients in column (1) of table 3 imply $R \approx 0.57$ for men. Using an estimate of $s \approx 0.35$ and $\beta = 1.36$, the estimated coefficients in column (4) of table 3 imply $R \approx 0.55$ for women. Similar implications follow from the quadratic models in columns (2) and (5). Relative changes in higher and lower quantiles of wages suggest that 40-50 percent of within-cell wage variation is attributable to unobserved skill.

Table 4 provides a summary of the ability of single index models to describe changes in mean wages and the quantiles of wages for white men and women over the 1980s. The entries in the table are mean prediction errors of 1989 wages for the age/education groups shown in the row headings. These means are weighted averages of cell-specific prediction errors (over the subset of relevant cells) from the quadratic single index models in Tables 2 and 3.

Examination of the prediction errors for various age groups suggests that single index models are relatively successful in modelling changes in age-related wage gaps. Among men there is some over-prediction of wages for 36-46 year olds and under-prediction of wages for 56-65 year olds. Wages for 36-46 year old women are also under-predicted. Average prediction errors for mean wages and the quartiles of wages tend to be very similar for the different age groups.

Examination of the prediction errors for different education groups reveals a 2.5-3.0% over-prediction of wages for male high school graduates and a 5-10% under-prediction for male college graduates. By comparison, prediction errors for different education groups of women are smaller and unsystematic.

A closer examination of college graduates by age (in the bottom panel of the table) confirms the visual impression in Figure 2. A single-index model under-predicts the wage gains of young college educated men over the 1980s. The model does much better describing changes in mean wages for older college graduates, but cannot account for the relative closing of the inter-quartile range of wages among "prime-age" male college graduates. Although a single-index model does very well in describing the wage growth of white female college graduates as a whole, within narrow age ranges the model does less well. As suggested by the evidence in Table 1b, younger female college graduates gained while older ones lost.

One possible explanation for the failure of the single index model to explain wage growth within narrow age ranges is that wages depend on several, as opposed to only one, dimensions of human capital. For instance, Murphy and Welch (1992) find that the structure of wages in the U.S. from 1963 to 1989 is better described by a linear two-index model than by a linear single-index model. The estimated effect of education for men in columns 3 and 6 of Table 2 also suggests that adding another index would improve the fit of the model. We formally test the linear single-index specification against a more general linear two-index model by fitting the following equation by measurement error corrected least squares:

$$\hat{w}_{j89} = a + b\hat{w}_{j79} + c(\hat{w}_{j79} - w^{p}_{j79})$$

where w^{p}_{j79} is the wage for cell j in 1979 linearly predicted on the basis of $\hat{w}_{j73}$.[18] We show in Appendix 4 that a t-test on the estimated value of the coefficient c is a specification test of the linear single-index model

[18]The prediction equation is obtained by fitting a linear equation of mean cell wages in 1979 on mean cell wages in 1973 by measurement error corrected least squares.

against a linear two-index model. The estimated value of c is equal to .692 with an estimated standard error of .376 for men, and to 1.885 with an estimated standard error of 1.414 for women. This suggests that adding a second index does not improve the fit of the single-index model at conventional significance levels although it comes close in the case of men. Interestingly, the results reported in column 5 of Table 2 suggests that the fit of the models improves more by adding a quadratic term to the linear single-index model than by adding a second linear index.

Our conclusions from Table 4, the goodness-of-fit tests in Tables 2 and 3, and the tests for linear two-skills models are mixed. A single index framework provides a parsimonious and relatively accurate description of overall changes in the wage structure. Nevertheless, a one-dimensional skill model is clearly inadequate to fully capture some specific features of the changing wage distribution -- particularly recent changes among the college graduate labor force. In our view the evidence suggests that a single index framework is a valuable starting point for any descriptive analysis of changes in the wage structure. It is particularly helpful in identifying "unusual" changes in the wage structure in an environment of rapidly changing wage inequality, and in unifying the analysis of "observed" and "unobserved" skill.

IV. Changes in Black-White Wage Differentials

We turn to the second objective of this paper, which is to analyze changes in wages for black men and women over the 1980s in light of the changing structure of wages for whites. As a point of departure we present in Table 5 a set of "conventional" estimates of the black-white wage gap, using our 1973/74, 1979, 1984, and 1989 CPS samples. These are derived from OLS regression models that include a linear education term, a quartic expression in potential experience, 8 region dummies, and an indicator for Hispanic ethnicity, as well as a black race indicator or interactions of a race dummy with indicators for different age/education classes.[19]

Row 1 of Table 5 presents unadjusted differences in mean log wages for black and white workers over our 15 year sample period. As previous researchers have noted (see Bound and Freeman (1992), for example), the black-white wage gap for men closed slightly between the mid- and late-1970s, then re-opened in the 1980s. The black-white wage gap for women followed a parallel course. Time series patterns of regression-adjusted wage gaps (in row 2) are roughly similar although the adjusted gaps are smaller in magnitude.

Comparisons of levels and changes in the wage gaps by age and education show considerable diversity within the black labor force. Wage gaps for black men and women aged 26-35 expanded significantly over the 1980s (growing by 7% for men and 10% for women) while gaps for older men and women were stable. Wage gaps for better-educated blacks also grew more while the gaps for male and female dropouts were stable. The trend in the wage gap for college-educated black women is notable: these women had wages well above their white counterparts in the mid-1970s but saw sharp relative declines over the late 1970s and early 1980s.

How do changes in the <u>overall</u> black relative wage gap during the 1980s compare with predictions based on the changing structure of white wages? The answer is presented in Table 6, where we report a simple decomposition

---

[19]We include as blacks only those individuals who report their race as "black". Results for models that pool "blacks" and "other races" are very similar.

of black and white average wage growth from 1979 to 1989. Let $\bar{w}_t$ denote the mean log wage of one race/sex group in period t (t=0 for 1979, t=1 for 1989), let w_{jt} represent the mean log wage for the particular group in age/education cell j in period t, and let s_{jt} represent the fraction of the group in cell j in period t. Finally, let w^p_{j0} represent the predicted mean log wage for cell j in 1989 based on the quadratic single index model for whites and 1979 wages in cell j. Then

$$\bar{w}_1 - \bar{w}_0 = \Sigma_j\ s_{j0}\{\ (w^p_{j1} - w_{j0}) + (w_{j1} - w^p_{j1})\ \} + (s_{j1} - s_{j0})\ w_{j1}.$$

The first term in this decomposition represents an average of cell-specific predicted growth rates based on the single index model for whites. The second is a weighted average of cell-specific prediction errors. Finally, the third term is a distributional effect reflecting changes in the relative fractions of workers in specific age/education cells between 1979 and 1989.

The decomposition in Table 6 leads to slightly different conclusions for men and women. As noted by Juhn, Murphy, and Pierce (1992), changes in the distribution of wages among white men imply that the relative wages of black men would fall over the 1980s. Our estimate (in row 1 of Table 6) is that increases in the return to skill led to a 5.3% fall in the relative wages of black men. Relative changes in demographic structure (including the retirement of older cohorts of less-educated blacks) and slightly better than expected wage growth within narrow age/education cells moderated this relative wage decline.

For black women, the increase in return to skill over the 1980s led to a much smaller relative decline in wages (-2.0%). Even though returns to skill rose more for white women than white men, black women's wages are less concentrated in the lower tail of the white female wage distribution. Thus widening wage inequality had a smaller net impact on their relative position. Within narrow age and education cells, however, black female wages grew more slowly than predicted by the white female wage structure. On net, then, black men and women had similar relative wage losses over the 1980s.

This overall assessment masks substantial relative gains and losses within the black labor force. Columns (1) and (3) of Table 7 present mean prediction errors of black wages in 1989 by age and education group. Columns (2) and (4) compute the relative prediction errors of blacks and whites in the same subgroups. If the single index model provided a "perfect fit" to the white wage distribution, the white prediction errors would be negligible and the relative prediction errors would simply equal the black prediction errors (as is the case for all workers in row 1). Since the single index model is imperfect, some fraction of the relative prediction error in specific age or education categories arises from the under- or over-prediction of white wages.

Examination of the patterns of relative and race-specific prediction errors by age suggests that older black workers enjoyed substantial gains over the 1980s while younger black workers lost ground. This is an important conclusion because some of the conventional explanations for black relative wage gains in the 1960s and 1970s (such as improved school quality) imply continued gains in the 1980s for the oldest groups of workers. Evidence in Card and Krueger (1992) suggests that black relative school quality improved more or less continuously from 1900 to the early 1950s. This improvement should have led to "unexplained" wage gains for older blacks during the 1980s, as individuals born before 1935 retired and were replaced by younger cohorts. Inspection of the positive prediction errors for men and women over age 46 lends some support to this story.

Analysis of the prediction errors by education reveals that poorly-educated blacks did better than expected over the 1980s, given the pattern of white wage changes. Wage growth for better-educated black men was about equal to predicted growth given the white wage structure. Nevertheless, the positive prediction error for white male college graduates (see Table 4) implies that the relative prediction error for college-educated black men is negative. Closer examination of the college subgroup by age shows a 19% relative loss for young male college graduates, equally attributable to the over-predication of black wages and the under-prediction of white wages.[20] Older black male college graduates, by comparison, did surprisingly well relative to predictions based on the white wage structure. This is not the case for older black female college graduates. Indeed, wages of high school- and college-educated black women are 5-10 percent below their predicted values, given white women's wage changes over the 1980s.

A detailed analysis of the prediction errors by age and education categories reveals that with the exception of women college graduates, blacks over age 46 of all education levels did better than expected in the 1980's. The corresponding set of prediction errors for young blacks is not as uniform. Table 8 indicates that young black women of all education levels did worst than predicted, both in race-specific and in relative terms. There is no other systematic pattern among black women under age 36. By contrast, less educated young black men did systematically better

[20]Note that college graduates age 22-25 are included in the overall college group but not shown separately by age.

than expected relative to more educated young black men. This pattern is even more accentuated relative to whites, since young white men with a college degree did much better than expected in the 1980's.

It is thus hard to find a unifying explanation for the relative wages changes of young blacks and young whites over the 1980s. Although arguments based on changed in the quality of education explain relatively well changes in the relative wages of workers over age 46, they can hardly explain why young black college graduates experienced large relative wages losses over that period.[21] A worsening in the quality of education of young black men should mostly affect high school graduates. The results in Table 8 rather indicate that young black men with a high school diploma did just as well as young whites with a high school diploma.

We have also computed prediction errors for the 25th, 50th, and 75th percentiles of black wages, and relative prediction errors between black and white workers at various quantiles of wages. For the most part, the patterns of the black prediction errors and the black-white relative prediction errors are similar to the patterns for mean wages. The most obvious differences emerge for college-educated men. Compared to the relative prediction error for mean wages of male college graduates (-5.9%) the relative error for the 25th percentile is more negative (-12.9%) while the relative error for the 75th percentile is less negative (-1.5%). This "tilting" (which also appears within age subgroups of the male college graduate population) suggests that the dispersion in wages for black college graduates widened substantially.

Our analysis of black wage growth patterns over the 1980s points to

---

[21]Juhn, Murphy and Pierce (1991) argue that school quality explains an important fraction of the relative wage changes of young black men.

three conclusions. First, relative to predictions based on the white wage structure, older black men and women enjoyed 8-10% relative wage gains. These are similar in magnitude to the relative wage gains of black men in the 1960s and 1970s (see Smith and Welch (1989), Card and Krueger (1992)). Second, younger black men and women, particularly the better-educated, suffered wage losses relative to predictions from the white wage structure. The gains for older workers and the losses for younger workers add up to small net changes overall. Third, young college graduate black males and college graduate black females of all ages suffered the largest unexpected losses.

Conclusions

We have proposed a simple technique for estimating and testing a one-dimensional skills model of the wage structure. The method compares means and/or quantiles of wages within specific age-education cells over time. A single skill model provides a parsimonious and relatively accurate model of changes in the structure of log hourly earnings for white men and women from the mid-1970s to the late 1980s. Within this framework, we find that the return to skill for women rose by 40 percent over the 1980s. For men, the rise was smaller -- approximately 25 percent -- and somewhat greater in the upper tail of the wage distribution than in the lower tail. We also find that 40-50 percent of residual wage variation (around race/sex and age/education means) can be attributed to unobserved ability whose market value rose in the 1980s.

We use the estimated model of changes in the structure of white wages to analyze changes in black-white relative wages from 1979 to 1989. The widening of the white wage distribution would have been expected to lower black men's relative wages by some 5 percentage points during the 1980s. However, changes in the relative demographic distribution of blacks and a small net gain in black wages relative to the white benchmark moderated this loss. The widening wage distribution of white women would have been expected to lead to a 2 percentage point loss in relative wages for black women over the 1980s. Unlike men, black women's relative wages fell short of white benchmark, accentuating the relative decline in their earnings.

There were also significant relative losses and gains within the black labor force. Our estimates suggest that the wages of older black men and women grow 8-12 percent relative to whites during the 1980s. On the other hand, young college-educated black men and college-educated black women in all age groups had wage declines of 5-10 percent relative to single-skill models fit to whites.

APPENDIX 1: CONSISTENT ESTIMATION OF REGRESSION MODELS WITH MEASUREMENT ERROR OF A KNOWN (ESTIMATED) FORM

Consider a true regression model:

(1) $y_j = x_j'\pi, \quad j=1,2,\ldots,K,$

and denote the observed data by $(\hat{y}_j, \hat{x}_j)$. Let

(2) $\hat{y}_j - y_j = \epsilon_j$, and

(3) $\hat{x}_j - x_j = u_j$.

For example, x_j is the following vector in the quadratic regression model for cell means:

$$\hat{x}_j = \begin{bmatrix} 1 \\ \hat{w}_{jo} \\ \hat{w}_{jo}^2 - \frac{\hat{s}_{jo}^2}{N_{jo}} \end{bmatrix}$$

Substituting equations (2) and (3) into equation (1) yields

(4) $\hat{y}_j = \hat{x}_j'\pi + \epsilon_j - u_j'\pi.$

OLS estimates of this regression equation are inconsistent since the component u_j' of the error term is correlated with the regressors $\hat{x}_j'$. Assume that

$$E(u_j u_j') = V_j,$$

$E(u_j \epsilon_j) = 0$, and that

$$\text{plim } \frac{1}{K}\sum_{j=1}^{K} x_j x_j' = M_{xx}.$$

The following sample moments can be constructed from the available data $\hat{y}_j$, $\hat{x}_j$, and an unbiased estimate $\hat{V}_j$ of V_j:

$$\hat{\Sigma} = (1/K) \textstyle\sum_j \hat{V}_j,$$

$$\hat{M}_{xx} = (1/K) \textstyle\sum_j \hat{x}_j\hat{x}_j',$$

$$\hat{M}_{xy} = (1/K) \textstyle\sum_j \hat{x}_j\hat{y}_j'.$$

Although $\hat{M}_{xx}$ does not converge to the true cross-product moment M_{xx} of x_j, it is easily shown that

$$\text{plim } (\hat{M}_{xx} - \hat{\Sigma}) = M_{xx}.$$

Given a consistent estimate $\hat{\Sigma}$ of the variance of the measurement error, π can be consistently estimated by correcting $\hat{M}_{xx}$ for measurement error. The proposed estimator $\hat{\pi}$ is given by:

$$(5) \qquad \hat{\pi} = (\hat{M}_{xx} - \hat{\Sigma})^{-1} \hat{M}_{xy}.$$

To establish consistency of this estimator, note that:

$$\begin{aligned} \hat{M}_{xy} &= \frac{1}{K}\sum_{j=1}^{K} \hat{x}_j(\hat{x}_j'\pi + e_j - u_j'\pi) \\ &= \hat{M}_{xx}\pi + \frac{1}{K}\sum_{j=1}^{K} \hat{x}_j e_j - \frac{1}{K}\sum_{j=1}^{K} \hat{x}_j u_j'\pi \\ &= (\hat{M}_{xx} - \hat{\Sigma})\pi + \frac{1}{K}\sum_{j=1}^{K} \hat{x}_j e_j - \frac{1}{K}\sum_{j=1}^{K} (\hat{x}_j u_j' - \hat{V}_j), \end{aligned}$$

where we have used the definition of $\hat{\Sigma}$ to obtain the last expression. Substituting this last expression into equation (3) yields:

$$\hat{\pi} = \pi + (\hat{M}_{xx} - \hat{\Sigma})^{-1}[\frac{1}{K}\sum_{j=1}^{K} \hat{x}_j e_j - (\hat{x}_j u_j' - \hat{V}_j)\pi)].$$

Define $\eta_j = \epsilon_j - u_j\pi$. $\hat{\pi}$ can be rewritten as:

$$\begin{aligned} \hat{\pi} - \pi &= (\hat{M}_{xx} - \hat{\Sigma})^{-1}[\frac{1}{K}\sum_{j=1}^{K} (\hat{x}_j\eta_j - \hat{V}_j\pi)] \\ &= (\hat{M}_{xx} - \hat{\Sigma})^{-1}[\frac{1}{K}\sum_{j=1}^{K} (\hat{x}_j\eta_j - V_j\pi) - (\frac{1}{K}\sum_{j=1}^{K} (\hat{V}_j - V_j)\pi]. \end{aligned}$$

Since $E(\hat{x}_j\eta_j) = - V_j\pi$ and $E(\hat{V}_j) = V_j$, it follows that

$$\text{plim } (\hat{\pi} - \pi) = 0.$$

This establishes the consistency of the estimator $\hat{\pi}$. To calculate the asymptotic covariance matrix of $\hat{\pi}$, consider the case where we ignore that $\hat{V}_j$ is estimated by simply setting $\hat{V}_j$ equal to V_j. Under standard regularity conditions, it is easily shown that:

$$(\hat{\pi}-\pi) \rightarrow N(0, \frac{1}{K}M_{xx}^{-1}WM_{xx}^{-1})$$

where $W = E[(\hat{x}_j\eta_j - V_j\pi)(\hat{x}_j\eta_j - V_j\pi)']$

A consistent estimate of W is obtained using the method of White (1980):

$$\hat{W} = \frac{1}{K}\sum_{j=1}^{K} [(\hat{x}_j\hat{\eta}_j - \hat{V}_j\hat{\pi})(\hat{x}_j\hat{\eta}_j - \hat{V}_j\hat{\pi})']$$

where $\hat{\eta}_j = \hat{y}_j - \hat{x}_j'\hat{\pi}$. Note that a consistent estimate of W can also be obtained when the error term η_j is correlated across observations. This situation occurs, for example, when several wage quantiles from the same age-education cell are used in the analysis. The consistent estimation of W in this special case is discussed in detail in Appendix 2. Given a consistent estimate $\hat{W}$ of W, a consistent estimate of the covariance matrix of $\hat{\pi}$ is given by the following expression:

$$\text{cov}(\hat{\pi}-\pi) = \frac{1}{K}(\hat{M}_{xx}-\hat{\Sigma})^{-1}\hat{W}(\hat{M}_{xx}-\hat{\Sigma})^{-1}.$$

Finally, an additional variance component could be included to take account of the sampling variability of the estimate $\hat{V}_j$ of the covariance matrix V_j.

GOODNESS-OF-FIT STATISTIC

Under the null hypothesis that equation (1) is correct, the error term $\epsilon_j - u_j'\pi$ in the equation relating $\hat{y}_j$ and $\hat{x}_j$ (equation (4)) consists entirely of sampling error. If y_j and x_j refer to cell moments or quantiles, it is possible to obtain estimates of the variances of the corresponding sampling errors (see Appendix 2). Assume that estimates of $var(\epsilon_j)$ and $var(u_j)$ are available, and assume that $cov(\epsilon_j, u_j) = 0$. Rewrite the regression residual $\hat{\eta}_j$ as:

$$\begin{aligned}\hat{\eta}_j &= \hat{y}_j - \hat{x}_j'\hat{\pi} \\ &= \epsilon_j - u_j'\pi - x_j'(\hat{\pi}-\pi).\end{aligned}$$

Under the null hypothesis that model (1) is correct, the variance of $\hat{\eta}_j$ is given by

$$\begin{aligned}(5)\qquad var(\hat{\eta}_j) &= var(\epsilon_j) + var(u_j'\pi) + var[\hat{x}_j'(\hat{\pi}-\pi)] - 2cov(\epsilon_j, u_j'\beta) \\ &\quad - 2cov[\epsilon_j, \hat{x}_j(\hat{\pi}-\pi)] + 2cov[u_j'\pi, \hat{x}_j'(\hat{\pi}-\pi)].\end{aligned}$$

The first covariance term in equation (5) is equal to zero since (by assumption) ϵ_j and u_j are uncorrelated. We also ignore the two other covariance terms in calculating the variance of the residuals.[1] The variance of $x_j(\hat{\pi}-\pi)$ can be estimated using the delta method

$$var[\hat{x}_j(\hat{\pi}-\pi)] \approx \hat{x}_j var(\hat{\pi}-\pi)\hat{x}_j'.$$

The variance of $\hat{\eta}_j$ can thus be rewritten as

$$var(\hat{\eta}_j) = var(\epsilon_j) + \hat{\pi}\hat{V}_1\hat{\pi}' + \hat{x}_j var(\hat{\pi}-\pi)\hat{x}_j',$$

Similarly, the covariance between $\hat{\eta}_j$ and $\hat{\eta}_k$ is given by:

$$cov(\hat{\eta}_j, \hat{\eta}_k) = \hat{x}_j \bullet var(\hat{\pi} - \pi) \bullet \hat{x}_j'.$$

---

[1] Empirically, incorporating the variance of $\hat{\pi}$ and its covariance with the error components ϵ_j and u_j in the calculations of the goodness-of-fit statistics has a negligible effect for the samples of the size we are using in this study.

Consider the vector of residuals $\hat{\eta}$ and the estimate $\hat{C}$ of its covariance matrix:

$$\hat{\eta} = \begin{bmatrix} \hat{\eta}_1 \\ \hat{\eta}_2 \\ . \\ . \\ \hat{\eta}_K \end{bmatrix}, \quad \hat{C} = \begin{bmatrix} \mathrm{var}(\hat{\eta}_1) & \mathrm{cov}(\hat{\eta}_1, \hat{\eta}_2) & . \\ \mathrm{cov}(\hat{\eta}_1, \hat{\eta}_2) & \mathrm{var}(\hat{\eta}_2) & . \\ . & . & . \end{bmatrix}$$

Under the null hypothesis that model (1) is well-specified, the goodness of fit statistic G is asymptotically distributed as chi-squared with K-3 degrees of freedom:

$$G = \hat{\eta}'\hat{C}^{-}\hat{\eta} \sim \chi^2(K-3)$$

where $\hat{C}^{-}$ is a generalized inverse of the estimated covariance matrix $\hat{C}$.

APPENDIX 2: ESTIMATION OF THE MEASUREMENT ERROR VARIANCE

In this appendix, we discuss the estimation of the variances of the sampling errors ϵ_j and u_j for models of cell means or quantiles.

A2.1 Model for Cell Means.

Consider K age/education cells. Let w_{ijt} represent the wage observation for individual i in cell j at time t, and let w_{jt} represent the true mean log wage in cell j in period t (t=0,1). Changes in the structure of wages are summarized by the following quadratic model:

$$w_{j1} = a + b\, w_{j0} + c\, w_{j0}^2 , \quad j=1,2,\ldots,K$$

For each cell, we observe:

$\hat{w}_{jt}$: mean wage in cell j in period t.

$\hat{s}_{jt}^2$: variance of wages in cell j in period t.

N_{jt}: Number of observations in cell j in period t.

The sampling variance of $\hat{w}_{jt}$ is given by:

$$\mathrm{Var}(\hat{w}_{jt}) = (1/N_{jt})\hat{s}_{jt}^2$$

Since $E(\hat{w}_{jt}^2) = E(w_{jt}) + s_{jt}^2$, unbiased estimates of $\hat{w}_{j0}$ and $\hat{w}_{j0}^2$ are $\hat{w}_{j0}$ and $\hat{w}_{j0}^2 - (1/N_{j0})\hat{s}_{j0}^2$. Let:

$$\theta_j = \begin{bmatrix} \hat{w}_{j0} \\ \hat{s}_{j0}^2 \end{bmatrix}, \text{ and } \hat{x}_j = \begin{bmatrix} 1 & \hat{w}_{j0} & \hat{w}_{j0}^2 - \frac{\hat{s}_{j0}^2}{N_{j0}} \end{bmatrix}'.$$

We can write $\hat{w}_j = f(\theta_j)$. Using the delta method, we find that

$$\mathrm{cov}(\hat{x}_j) \approx F_j A_j F_j',$$

where $F_j = \partial f(x_j)/\partial x_j$, and where $A_j = \mathrm{cov}(\theta_j)$. The Jacobian matrix F_j is given by:

$$F_j = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 2\hat{w}_{j0} & -\frac{1}{N_{j0}} \end{bmatrix},$$

while

$$A_j = \begin{bmatrix} \frac{c_{2j0}}{N_{j0}} & \frac{c_{3j0}}{N_{j0}} \\ \frac{c_{3j0}}{N_{j0}} & \frac{c_{4j0} - c_{2j0}^2}{N_{j0}} \end{bmatrix}.$$

where c_{kjt} is the kth central moment of the distribution of wages in cell j at time t:

$$c_{1jt} = \sum_{i=1}^{N_{jt}} w_{ijt},$$

$$\text{and } c_{kjt} = \frac{1}{N_{jt}} \sum_{i=1}^{N_{jt}} (w_{ijt} - c_{1jt}) \text{ for } k \geq 2.$$

A consistent estimate of $\mathrm{cov}(x_j)$ is thus given by:

$$\hat{V}_j = \hat{F}_j\hat{A}_j\hat{F}_j,$$

where $\hat{F}_j$ and $\hat{A}_j$ are the sample analogs of F_j and A_j.

A2.2 Model for Wage Quantiles

The model for wage percentiles is the following:

$$w_{j1}^{25} = a^{25} + bw_{j0}^{25} + c(w_{j0}^{25})^2,$$
$$w_{j1}^{50} = a^{50} + bw_{j0}^{50} + c(w_{j0}^{50})^2,$$
$$w_{j1}^{75} = a^{75} + bw_{j0}^{75} + c(w_{j0}^{75})^2.$$

These three equations can be combined in a single equation:

$$y_j^q = x_j^{q\prime}\pi \text{ , for } q=.25,\ .50, \text{ and } .75,$$

$$\text{where } x_j^q = \begin{bmatrix} 1 \\ D_{j50}^q \\ D_{j75}^q \\ w_{j0}^q \\ (w_{j0}^q)^2 \end{bmatrix} \text{ and } \pi = \begin{bmatrix} a^{25} \\ a^{50}-a^{25} \\ a^{75}-a^{25} \\ b \\ c \end{bmatrix}.$$

Note that D_{j50}^q is an indicator variable that is equal to one when $q=.50$ while D_{j75}^q is an indicator variable that is equal to one when $q=.75$. As in the case of the model for cell means, π can be consistently estimated by replacing x_j^q by an unbiased estimate $\hat{x}_j^q$ and adjusting the cross products of $\hat{x}_j^q$ for measurement error. To simplify the calculations, assume that the (log) wage observations w_{ijt} are drawn from a normal distribution with mean w_{jt} and variance $s_{jt}{}^2$. The sampling variance of the q^{th} estimated quantile ($q = .25, .5, .75$) is given by

$$\text{var}(\hat{w}_{jt}^q - w_{jt}^q) = \frac{1}{N_{jt}} k^q s_{jt}^2,$$

where $k^q = q(1-q)/\phi(z^q)^2$ and where z^q is the qth quantile of the standard normal distribution ($\phi(.)$ is density of the standard normal distribution). As discussed in section 2.1 of this appendix, the sampling variance of the estimate $\hat{s}_{jt}^2$ of s_{jt}^2 is given by:

$$\mathrm{var}(\hat{s}_{jt}^2 - s_{jt}^2) = \frac{1}{N_{jt}}[c_{4jt} - c_{2jt}^2].$$

It can also be shown that[2]

$$\mathrm{cov}(\hat{w}_{jt}^q, \hat{s}_{jt}^2) = -z^q s_{jt}^3.$$

An unbiased estimate of w_{j0}^q is the estimated quantile $\hat{w}_{j0}^q$, while an unbiased estimate of $(w_{j0}^q)^2$ is $(\hat{w}_{j0}^q)^2 - (1/N_{j0})\bullet k^q \bullet \hat{s}_{j0}^q$. Let:

$$\theta_j^q = \begin{bmatrix} \hat{w}_{j0}^q \\ \hat{s}_{io}^2 \end{bmatrix}.$$

We can write $x_j = f(\theta_j)$. Using the delta method, we find that:

$$\mathrm{cov}(x_j^q) \approx F_j^q A_j^q F_j^{q\prime}.$$

where $F_j^q = \partial f(x_j^q)/\partial x_j^q$ and where $A_j^q = \mathrm{cov}(\theta_j^q)$. The Jacobian matrix F_j^q is now given by:

$$F_j^q = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 2w_{j0}^q & -\frac{k^q}{N_{j0}} \end{bmatrix},$$

while

---

[2] The proof of this result is contained in an appendix available on request.

$$A_j^q = \begin{bmatrix} \frac{1}{N_{j0}} k^q s_{j0}^2 & \frac{-1}{N_{j0}} z^q s_{j0}^3 \\ \frac{-1}{N_{j0}} z^q s_{j0}^3 & \frac{1}{N_{j0}} (c_{4j0} - s_{j0}^4) \end{bmatrix}.$$

A consistent estimate of $\mathrm{cov}(\hat{x}_j^q)$ is thus given by:

$$\hat{V}_j^q = \hat{F}_j^q \hat{A}_j^q \hat{F}_j^{q\prime},$$

where $\hat{F}_j^q$ and $\hat{A}_j^q$ are the sample analogs of F_j^q and A_j^q. A consistent estimate of π is given by:

$$\hat{\pi} = (\hat{M}_{xx} - \hat{\Sigma})^{-1} \hat{M}_{xy},$$

where the average cross products are now averaged over both quantiles q and cells j

$$\hat{M}_{xx} = \frac{1}{3K} \sum_{j=1}^{K} \sum_q \hat{x}_j^q \hat{x}_j^{q\prime},$$

$$\hat{M}_{xy} = \frac{1}{3K} \sum_{j=1}^{K} \sum_q \hat{x}_j^q \hat{y}_j^q,$$

$$\hat{\Sigma} = \frac{1}{3K} \sum_{j=1}^{K} \sum_q \hat{V}_j^q.$$

To calculate the asymptotic covariance matrix of $\hat{\pi}$, consider again the case where we ignore that $\hat{V}_j$ is estimated by simply setting $\hat{V}_j$ equal to V_j. Under standard regularity conditions, it is easily shown that:

$$(\hat{\pi} - \pi) \to N(0, \frac{1}{3K} M_{xx}^{-1} W M_{xx}^{-1}),$$

$$\text{where } W = E[\sum_q (\hat{x}_j^q \eta_j^q - V_j^q \pi) \sum_q (\hat{x}_j^q \eta_j^q - V_j^q \pi)'],$$

and where $\hat{\eta}_j^q = \hat{y}_j^q - \hat{x}_j^{q\prime} \pi$. Note that the covariance matrix W takes account of the correlation between the residuals of the three wage quantiles in a given cell. A consistent estimate of W is obtained using the suggestion of White (1980):

$$\hat{W} = \frac{1}{3K}\sum_{j=1}^{K}\left[\sum_{q}(\hat{x}_j^q\hat{\eta}_j^q-\hat{v}_j^q\hat{\pi})\sum_{q}(\hat{x}_j^q\hat{\eta}_j^q-\hat{v}_j^q\hat{\pi})'\right].$$

Finally, the goodness-of-fit test of the model is similar to the test for the cell means model.

APPENDIX 3: DECOMPOSITION OF CHANGES IN WAGES

In this appendix, we decompose the change in the average wage of a given group of workers into three components: 1) the change predicted by the estimated model holding the skill composition constant, 2) the change not predicted by the estimated model holding the skill composition constant, and 3) the change due to changes in the distribution of skills. We perform this decomposition for blacks, whites, and for black-white wage differentials. We also calculate the standard errors for each component of the decomposition.

Consider $\bar{w}_{bt}$ and $\bar{w}_{wt}$, the sample mean (log) wages for blacks and whites respectively in period t,

$$\overline{w}_{bt} = \sum_{j=1}^{K} m_{bjt}\hat{w}_{bjt} \quad \text{and} \quad \overline{w}_{wt} = \sum_{j=1}^{K} m_{wjt}\hat{w}_{wjt},$$

where m_{bjt} (m_{wjt}) is the weighted share of the black (white) population in cell j at time t, and $\bar{w}_{bjt}$ ($\bar{w}_{wjt}$) is the sample average log wage of blacks (whites) in cell j at time t.

The predicted wage for race group r (r=b,w) in cell j in period 1 is

$$w_{rj1}^{p} = \hat{x}_{rj1}'\hat{\pi}\ , \text{ for } r=b,w,$$

where $\hat{\pi}$ is the vector of parameter estimates based on the wages of whites and $\hat{x}_{rj1}$ is a race- and cell-specific vector of regressors. Consider the following decomposition of the change in average wages of blacks $\bar{w}_{b1} - \bar{w}_{b0}$

$$\overline{w}_{b1}-\overline{w}_{b0} = \sum_{j=1}^{K} m_{bjo}(\hat{w}_{bj1}-\hat{w}_{bjo}) + \sum_{j=1}^{K} \hat{w}_{bj1}(m_{bj1}-m_{bjo}) .$$

The first component of this equation can be further decomposed into a predicted and an unpredicted component:

$$\sum_{j=1}^{K} m_{bjo}(\overline{w}_{bj1}-\overline{w}_{bjo}) = \sum_{j=1}^{K} m_{bjo}(w^{p}_{bj1}-\hat{w}_{bjo}) + \sum_{j=1}^{K} m_{bjo}(\hat{w}_{bj1}-w^{p}_{bj1}) .$$

Overall, the change in average wages of blacks, $\bar{w}_{b1} - \bar{w}_{b0}$, can thus be decomposed in three components:

1) Δ^{b}_{pred}, the predicted change in wages of blacks:

$$\Delta^{b}_{pred} = \sum_{j=1}^{K} m_{bjo}(w^{p}_{bj1}-\hat{w}_{bjo}) .$$

2) Δ^{b}_{res}, the unpredicted or residual change in wages of blacks:

$$\Delta^{b}_{res} = \sum_{j=1}^{K} m_{bjo}(\hat{w}^{b}_{bj1}-w^{p}_{bj1}) .$$

3) Δ^{b}_{dist}, the distributional effect due to the change in the skill composition of the black workforce:

$$\Delta^{b}_{dist} = \sum_{j=1}^{K} \hat{w}_{bj1}(m_{bj1}-m_{bjo}) .$$

Similarly, the change in average wages of whites can be decomposed in three components Δ^{w}_{pred}, Δ^{w}_{res}, and Δ^{w}_{dist}. Finally, the change in the wage gap Δ^{bw} can also be decomposed into three components Δ^{bw}_{pred}, Δ^{bw}_{res}, and Δ^{bw}_{dist} representing the differences between blacks and whites of the three components.

To find the sampling variance of the three components of changes in wages first consider the first order approximation of w^{p}_{rj1}:

$$w^p_{bj1} = \hat{x}'_{bjo}\hat{\pi} \approx x'_{bjo}\pi + x'_{bjo}(\hat{\pi}-\pi) + (\hat{x}_{bjo}-x_{bjo})'\pi$$

For blacks, it follows that

$$\operatorname{var}(w^p_{bj1}) \approx x'_{bjo}\operatorname{var}(\hat{\pi})x_{bjo} + \pi'V_{bj}\pi,$$
$$\operatorname{cov}(w^p_{bj1}, w^p_{bk1}) \approx x'_{bjo}\operatorname{var}(\hat{\pi})x_{bko},$$

and that

$$\operatorname{var}(w^p_{bj1}-\hat{w}_{bjo}) \approx x'_{bjo}\operatorname{var}(\hat{\pi})x_{bjo} + \delta'V_{bj}\delta,$$
$$\operatorname{cov}(w^p_{bj1}-\hat{w}_{bjo}, w^p_{bk1}-\hat{w}_{bko}) \approx x'_{bjo}\operatorname{var}(\hat{\pi})x_{bko},$$

where δ is such that

$$\hat{x}'_{bj1}\delta = \hat{x}'_{bj1}\pi - \hat{w}_{bjo}.$$

The same formulas can also be used for whites, although they ignore a negligible covariance term between the estimated parameters $\hat{\pi}$ and the regressors $\hat{w}_{wj0}$.

The variance of the three components of changes in wages of blacks can then be estimated as follows:

$$\operatorname{var}(\Delta^b_{pred}) = \sum_{j=1}^{K}(m_{bjo})^2\frac{S^2_{bjo}}{N_{bjo}} + \sum_{j=1}^{K}(m_{bjo})^2(\delta'V_{bj}\delta)$$
$$+ (\sum_{j=1}^{K} m_{bjo}\hat{x}_{bjo})'\operatorname{var}(\hat{\pi})(\sum_{j=1}^{K} m_{bjo}\hat{x}_{bjo}),$$

$$\operatorname{var}(\Delta^b_{res}) = \sum_{j=1}^{K}(m_{bjo})^2\frac{S^2_{bj1}}{N_{bj1}} + \sum_{j=1}^{K}(m_{bjo})^2(\hat{\pi}'V_{bj}\hat{\pi})$$
$$+ (\sum_{j=1}^{K} m_{bjo}\hat{x}_{bjo})'\operatorname{var}(\hat{\pi})(\sum_{j=1}^{K} m_{bjo}\hat{x}_{bjo}),$$

$$\operatorname{var}(\Delta^b_{dist}) = \sum_{j=1}^{K}(m_{bj1}-m_{bjo})^2(s^2_{j1}).$$

Similar formulas can be used for whites. The variance of the three components of changes in the black-white wage gap can finally be estimated as follows:

$$\mathrm{var}(\Delta^{bw}_{pred}) = \mathrm{var}(\Delta^{b}_{pred}) + \mathrm{var}(\Delta^{w}_{pred})$$
$$-2\left(\sum_{j=1}^{K} m_{bj0}\hat{x}_{bj0}\right)'\mathrm{var}(\hat{\pi})\left(\sum_{j=1}^{K} m_{wj0}x_{wj0}\right),$$

$$\mathrm{var}(\Delta^{bw}_{res}) = \mathrm{var}(\Delta^{b}_{res}) + \mathrm{var}(\Delta^{w}_{res})$$
$$+\left(\sum_{j=1}^{K} m_{bj0}x_{bj0}\right)'\mathrm{var}(\hat{\pi})\left(\sum_{j=1}^{K} m_{wj0}x_{wj0}\right),$$

$$\mathrm{var}(\Delta^{bw}_{dist}) = \mathrm{var}(\Delta^{b}_{dist}) + \mathrm{var}(\Delta^{w}_{dist}).$$

APPENDIX 4: ESTIMATION OF A TWO-SKILLS MODEL

Assume that wages for cell j in 1973 is the sum of two skills S_{1j} and S_{2j}. Without loss of generality, assume that the two skills are uncorrelated and that they have the same variance $\sigma^2 = \mathrm{var}(w_{j73})/2$

$$w_{j73} = S_{1j} + S_{2j}.$$

Cell wages in 1979 and 1989 are the following linear functions of these two skills

$$w_{j79} = \gamma_{0,79} + \gamma_{1,79}S_{1j} + \gamma_{2,79}S_{2j}, \text{ and}$$

$$w_{j89} = \gamma_{0,89} + \gamma_{1,89}S_{1j} + \gamma_{2,89}S_{2j}.$$

The projection of w_{j79} on w_{j73} yields

$$w^{p}_{j79} = \gamma_{0,79} + \bar{\gamma}_{79}(S_{1j} + S_{2j}),$$

where $\bar{\gamma}_{79} = (\gamma_{1,79} + \gamma_{2,79})/2$. The prediction error $(w_{j79} - w^{p}_{j79})$ is given by

$$(w_{j79} - w^p_{j79}) = (\gamma_{1,79} - \bar{\gamma}_{79})\ S_{1j} + (\gamma_{2,79} - \bar{\gamma}_{79})\ S_{2j}.$$

The linear projection of w_{j89} on w_{j79} and $(w_{j79} - w^p_{j79})$ is thus given by

$$P[w_{j89}\ |w_{j79},\ w_{j79} - w^p_{j79}] = a + bw_{j79} + c(w_{j79} - w^p_{j79}),$$

where

$$a = \gamma_{0,89} - (\bar{\gamma}_{89}/\bar{\gamma}_{79})\gamma_{0,79}$$

$$b = \bar{\gamma}_{89}/\bar{\gamma}_{79}$$

$$c = [\gamma_{1,89}(\gamma_{1,79} - \bar{\gamma}_{79}) + \gamma_{2,89}(\gamma_{2,79} - \bar{\gamma}_{79})]\ /\ [(\gamma_{1,79} - \bar{\gamma}_{79})^2 + (\gamma_{2,79} - \bar{\gamma}_{79})^2].$$

Under the null hypothesis that the single-index model is well-specified, the relative price $\gamma_{1,t}/\gamma_{2,t}$ of the skills must remain constant over time. This condition is necessary and sufficient for the two skills to aggregate in a single skill. Furthermore, the coefficient c is equal to 0 whenever this condition is satisfied. A t-test of the measurement error corrected least squares estimate of c is thus a specification test for the single-skill model against the two-skills model. The measurement error correction used is similar to the procedure described in Appendices 1 and 2.

References

Blau, Francine D. and Lawrence M. Kahn. "The Gender Earnings Gap: Some International Evidence". National Bureau of Economics Research Working Paper Number 4224, December 1992.

Bound, John and Richard B. Freeman. "What Went Wrong? The Erosion of Relative Earnings of and Employment Among Young Black Men in the 1980s". Quarterly Journal of Economics 107 (February 1992): 201-232.

Bound, John and George Johnson. "Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Explanations". American Economic Review 82 (June 1992): 371-392.

Card, David and Alan B. Krueger. " School Quality and Black-White Relative Earnings: A Direct Assessment". Quarterly Journal of Economics 107 (February 1992): 151-200.

Fuller, Wayne A. and Michael A. Hidiroglou. "Regression After Correcting for Attenuation". Journal of the American Statistical Association 73 (March 1978): 99-104.

Juhn, Chinhui, Kevin M. Murphy and Brooks Pierce. "Accounting for the Slowdown in Black-White Wage Convergence". in Marvin Kosters, editor, Workers and Their Wages. Washington DC: American Enterprise Institute Press, 1991

Juhn, Chinhui, Kevin M. Murphy and Brooks Pierce. "Wage Inequality and the Rise in Returns to Skill". Unpublished Working Paper, 1992.

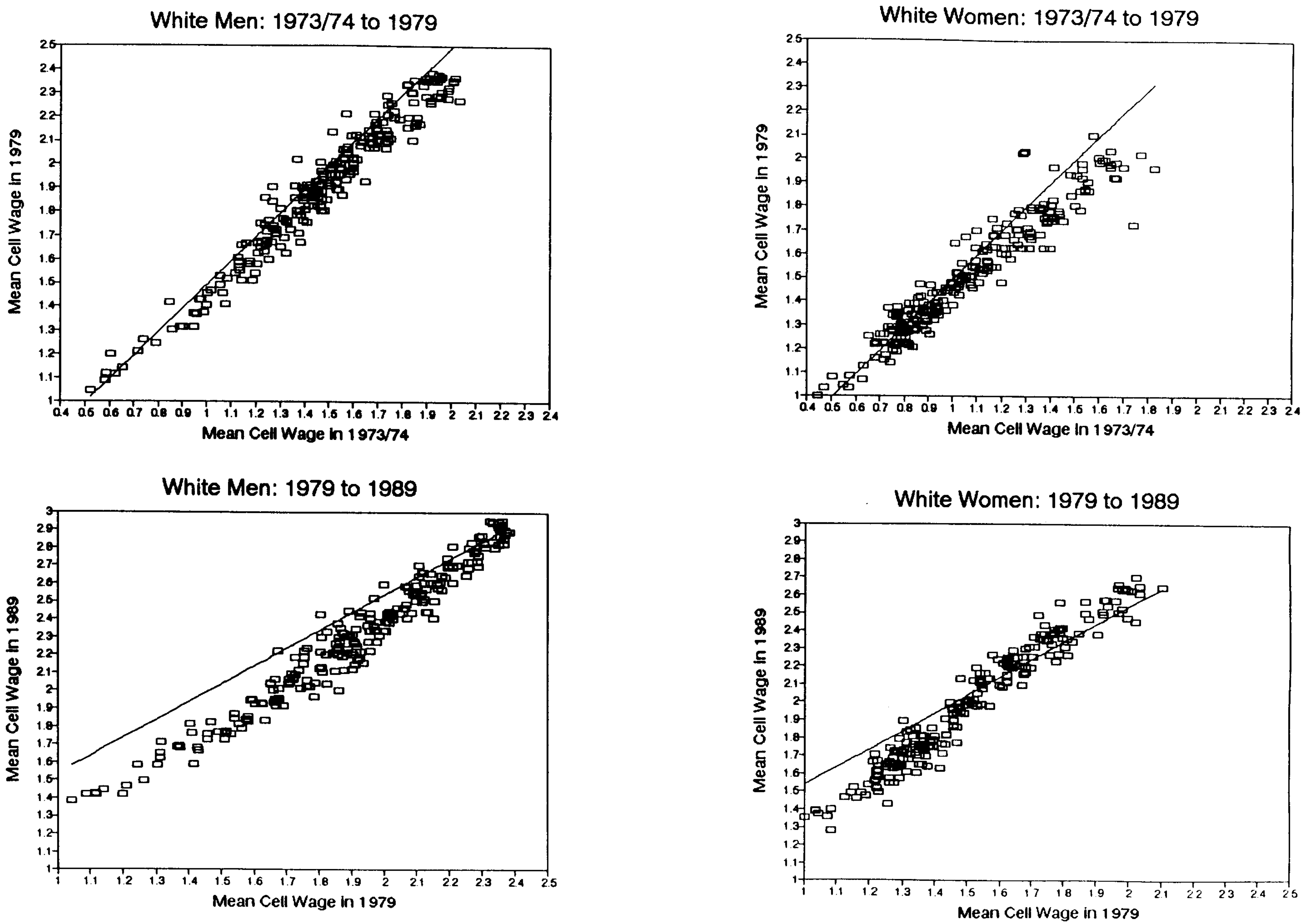
Murphy, Kevin M. and Finis Welch. "The Structure of Wages". Quarterly Journal of Economics 107 (February 1992): 285-326.

Pierce, Brooks and Finis Welch. "Dimensions of Inequality in Labor Income". Unpublished Working Paper, Texas A&M University, 1992.

Smith, James P. and Finis Welch. "Black Economic Progress After Myrdal". Journal of Economic Literature 27 (1989): 519-564.

Tilley, Chris, Barry Bluestone and Bennet Harrison. "What is Making American Wages More Unequal?". Industrial Relations Research Association Proceedings (May 1987).

Figure 1



Note: constant real wage line is shown.

Figure 2

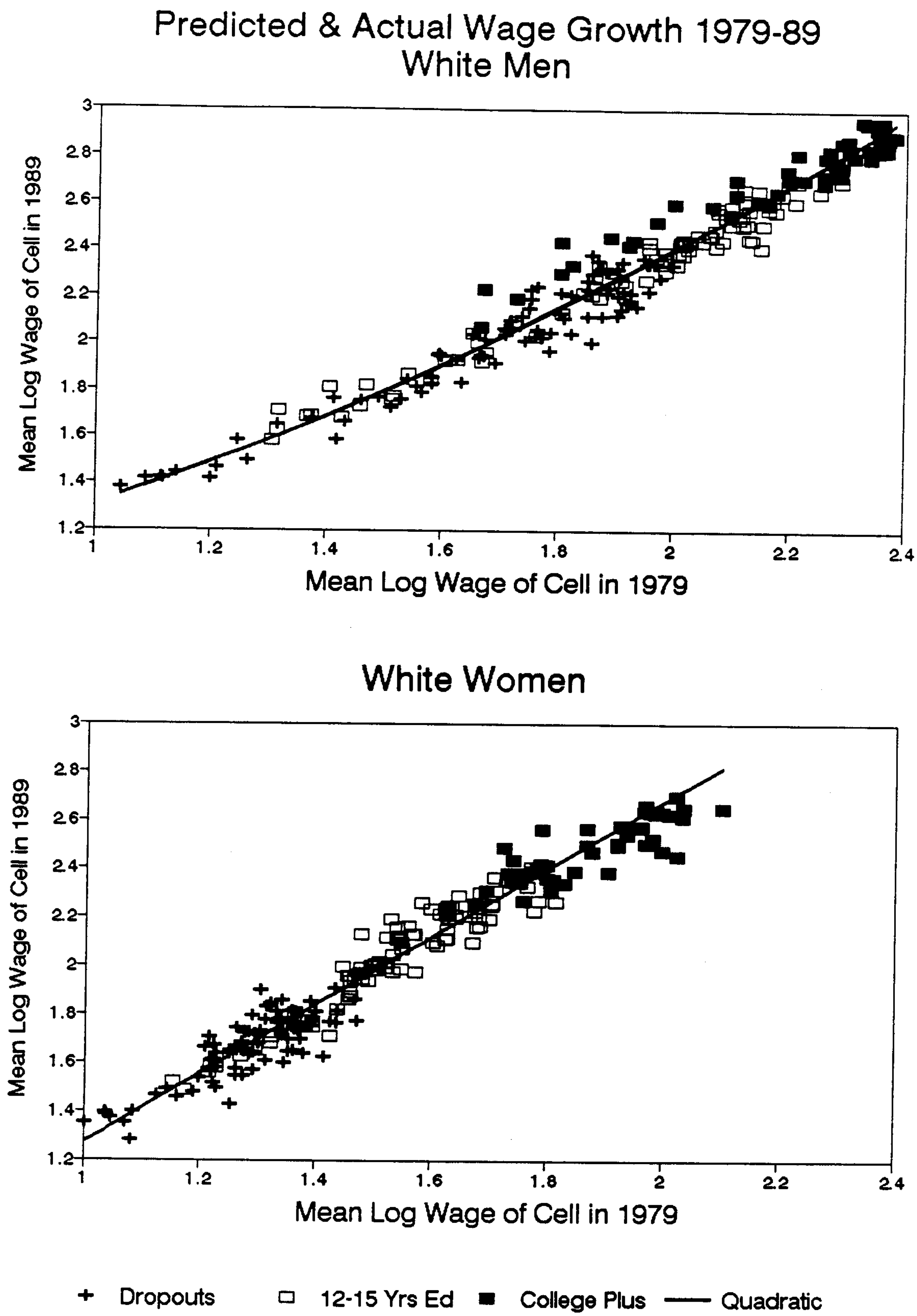
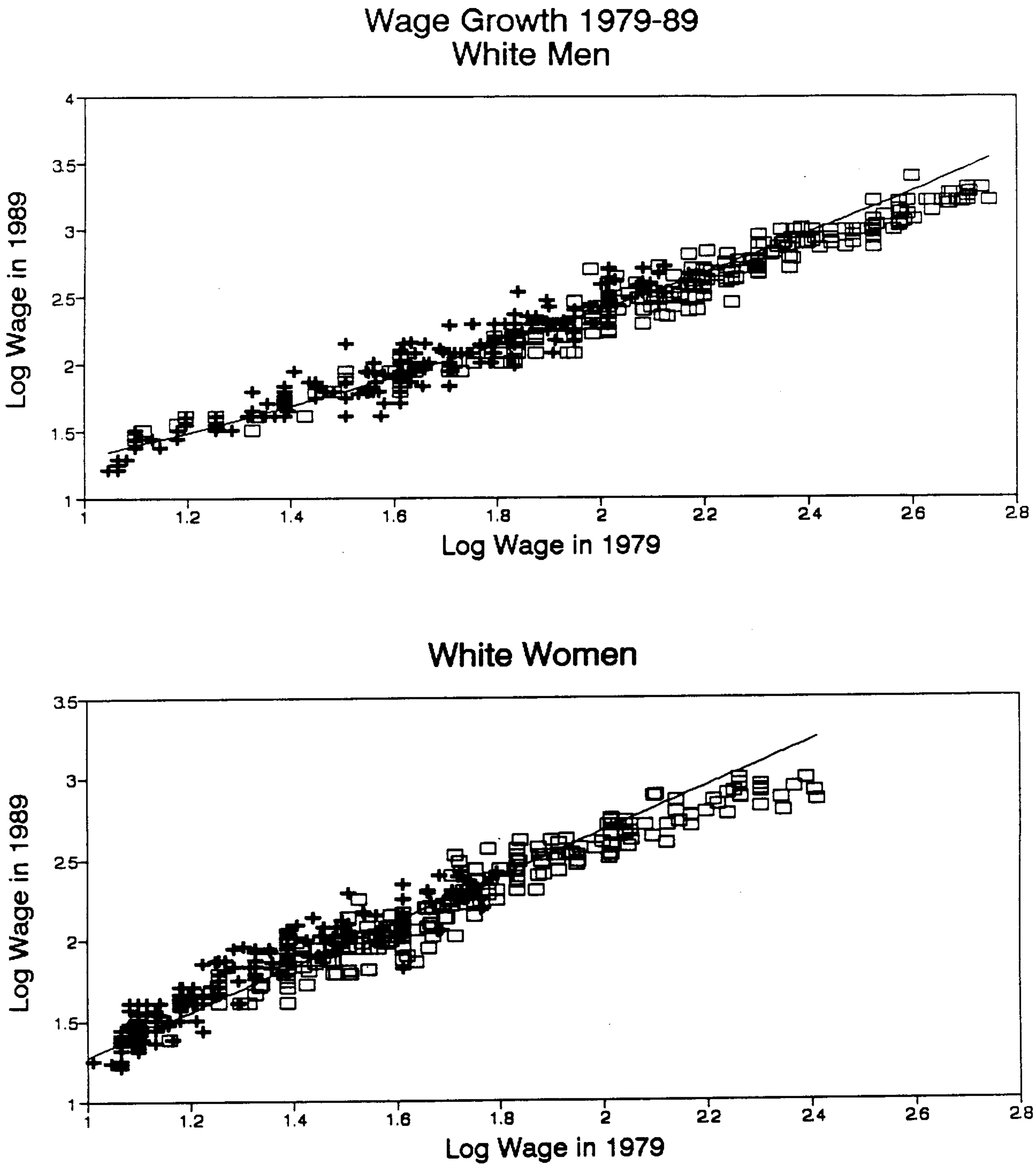

Predicted & Actual Wage Growth 1979-89
White Men
Mean Log Wage of Cell in 1989
Mean Log Wage of Cell in 1979
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
1.2
1.4
1.6
1.8
2
2.2
2.4
White Women
Mean Log Wage of Cell in 1989
Mean Log Wage of Cell in 1979
Dropouts
12-15 Yrs Ed
College Plus
Quadratic

Figure 3



+ 25th Percentile   □ 75th Percentile   —— Fitted to Cell Mean

Table 1a: Changes in Wage Inequality Among White Men

	1973/4	1979	1984	1989	Ratio 1989/1979
1. Age Differentials:					
Age 46-55 - Age 26-35					
a. 9-11 Years Education	0.088	0.158	0.211	0.235	1.49
	(0.016)	(0.013)	(0.015)	(0.016)	
b. 12 Years Education	0.101	0.146	0.219	0.224	1.53
	(0.010)	(0.007)	(0.008)	(0.008)	
c. 16 Years Education	0.330	0.326	0.318	0.304	0.93
	(0.021)	(0.013)	(0.015)	(0.015)	
2. Education Differentials:					
12 Years - 9-11 Years					
a. Age 26-35	0.100	0.125	0.154	0.177	1.42
	(0.013)	(0.010)	(0.011)	0.011	
b. Age 46-55	0.113	0.114	0.162	0.165	1.45
	(0.013)	(0.010)	(0.013)	(0.014)	
16 Years - 12 Years					
c. Age 26-35	0.176	0.144	0.262	0.324	2.25
	(0.013)	(0.008)	(0.008)	(0.008)	
d. Age 46-55	0.406	0.324	0.361	0.405	1.25
	(0.020)	(0.013)	(0.014)	(0.015)	
3. Within-Cell Standard Deviations:					
12 Years Education					
a. Age 26-27	0.366	0.385	0.408	0.408	1.06
	(0.011)	(0.006)	(0.006)	(0.007)	
b. Age 47-49	0.397	0.388	0.447	0.417	1.07
	(0.013)	(0.006)	(0.009)	(0.008)	
16 Years Education					
a. Age 26-27	0.394	0.409	0.449	0.437	1.07
	(0.017)	(0.010)	(0.011)	(0.012)	
b. Age 47-49	0.452	0.449	0.502	0.528	1.18
	(0.023)	(0.014)	(0.017)	(0.018)	
Overall Standard Deviation of Log Hourly Wage	0.501	0.497	0.552	0.568	1.14

Note: Standard errors in parentheses. Entries are difference in mean log wages between indicated groups, or standard deviations of mean log wages within indicated groups. Samples include individuals age 16-65 who report positive hourly or weekly wages. See text for further details.

Table 1b: Changes in Wage Inequality Among White Women

	1973/4	1979	1984	1989	Ratio 1989/1979
1. Age Differentials:					
Age 46-55 - Age 26-35					
a. 9-11 Years Education	0.065 (0.019)	0.063 (0.014)	0.068 (0.016)	0.107 (0.017)	1.70
b. 12 Years Education	0.047 (0.011)	0.040 (0.007)	0.076 (0.007)	0.078 (0.008)	1.95
c. 16 Years Education	0.068 (0.026)	0.028 (0.019)	0.002 (0.017)	-0.008 (0.016)	--
2. Education Differentials:					
12 Years - 9-11 Years					
a. Age 26-35	0.153 (0.016)	0.151 (0.011)	0.177 (0.012)	0.232 (0.012)	1.54
b. Age 46-55	0.135 (0.015)	0.128 (0.011)	0.184 (0.013)	0.204 (0.014)	1.59
16 Years - 12 Years					
c. Age 26-35	0.344 (0.015)	0.282 (0.009)	0.347 (0.009)	0.441 (0.009)	1.56
d. Age 46-55	0.366 (0.024)	0.270 (0.018)	0.273 (0.017)	0.354 (0.015)	1.31
3. Within-Cell Standard Deviations:					
12 Years Education					
a. Age 26-27	0.368 (0.014)	0.364 (0.008)	0.402 (0.007)	0.421 (0.009)	1.16
b. Age 47-49	0.400 (0.013)	0.375 (0.008)	0.408 (0.007)	0.424 (0.007)	1.13
16 Years Education					
c. Age 26-27	0.332 (0.016)	0.359 (0.010)	0.433 (0.011)	0.428 (0.011)	1.19
d. Age 47-49	0.420 (0.039)	0.454 (0.027)	0.480 (0.022)	0.464 (0.015)	1.02
Overall Standard Deviation of Log Wages	0.437	0.418	0.476	0.514	1.23

Note: Standard errors in parentheses. Entries are difference in mean log wages between indicated groups, or standard deviations of mean log wages within indicated groups. Samples include individuals age 16-65 who report positive hourly or weekly wages. See text for further details.

Table 2: Measurement Error Corrected Estimates of Single Index Model, White Men and Women, 1973/4 to 1979 and 1979 to 1989

	1973/4 to 1979			1979 to 1989		
	(1)	(2)	(3)	(4)	(5)	(6)
I. White Men						
1. Constant	0.507	0.494	0.594	-0.040	0.704	0.404
	(0.016)	(0.042)	(0.052)	(0.033)	(0.095)	(0.123)
2. Mean Cell Wage in Base Year	0.948	0.968	0.886	1.226	0.363	0.624
	(0.011)	(0.067)	(0.073)	(0.017)	(0.114)	(0.132)
3. Mean Cell Wage Squared in Base Year	--	-0.008	0.036	--	0.243	0.153
		(0.026)	(0.030)		(0.033)	(0.039)
4. Mean Years of Education in Cell	--	--	-0.006	--	--	0.011
			(0.002)			(0.003)
5. Goodness-of-Fit (deg. freedom)	419.0	420.2	382.7	933.0	700.3	649.3
	(223)	(222)	(221)	(223)	(222)	(221)
II. White Women						
1. Constant	0.601	0.538	0.610	-0.013	-0.178	-0.173
	(0.017)	(0.048)	(0.070)	(0.032)	(0.171)	(0.191)
2. Mean Cell Wage in Base Year	0.847	0.971	0.954	1.407	1.470	1.466
	(0.018)	(0.098)	(0.111)	(0.022)	(0.234)	(0.251)
3. Mean Cell Wage Squared in Base Year	--	-0.058	-0.020	--	-0.021	-0.019
		(0.049)	(0.062)		(0.080)	(0.089)
4. Mean Years of Education in Cell	--	--	-0.008	--	--	-0.000
			(0.003)			(0.003)
5. Goodness-of-Fit (deg. freedom)	371.1	364.3	325.6	492.0	489.7	488.3
	(223)	(222)	(221)	(223)	(222)	(221)

Note: Dependent variable is mean log wage in age-education cell in final year (1979 in columns 1-3; 1989 in columns 4-6). Cells are weighted by weighted count of workers in age-education cell in 1979. Estimation method is corrected least squares -- see text.

Table 3: Measurement Error Corrected Estimates of Single Index Model, 25th, 50th, and 75th Percentiles of Log Wages for White Men and Women, 1979 to 1989.

	White Men			White Women		
	(1)	(2)	(3)	(4)	(5)	(6)
1. Constant	0.015 (0.020)	0.419 (0.062)	0.173 (0.063)	0.002 (0.020)	-0.430 (0.085)	-0.437 (0.085)
2. Corresponding Wage Percentile of Cell in 1979	1.212 (0.012)	0.746 (0.068)	0.886 (0.064)	1.358 (0.016)	1.938 (0.113)	1.932 (0.112)
3. Corresponding Wage Percentile Squared	--	0.130 (0.019)	0.072 (0.018)	--	-0.185 (0.036)	-0.187 (0.036)
4. Mean Years of Education in Cell	--	--	0.015 (0.002)	--	--	0.001 (0.002)
5. Dummy for 50th Percentile	-0.031 (0.009)	-0.030 (0.009)	-0.010 (0.008)	-0.035 (0.007)	-0.045 (0.008)	-0.026 (0.008)
6. Dummy for 75th Percentile	-0.057 (0.012)	-0.071 (0.011)	-0.025 (0.011)	-0.096 (0.010)	-0.102 (0.010)	-0.097 (0.011)
7. 25th Percentile Below Minimum Wage in 1979[a]	--	--	--	-0.119 (0.012)	-0.093 (0.013)	-0.092 (0.013)
8. 25th Percentile Below Minimum Wage in 1989[b]	--	--	--	0.030 (0.058)	0.052 (0.067)	0.053 (0.067)
9. Goodness-of-Fit (deg. freedom)	3098.0 (671)	2894.6 (670)	2458.5 (669)	1724.6 (669)	1621.1 (668)	1627.5 (667)

Note: Dependent variable is 25th, 50th, or 75th percentile of log wage distribution in age-education cell in 1989. (There are 3 observations for each of 225 age-education cells). Cells are weighted by weighted count of workers in age-education cell in 1979. Estimation method is corrected least squares -- see text.

[a] Dummy variable equal to 1 if 25th percentile of wages in cell in 1979 is less than or equal to minimum wage in 1979.

[b] Dummy variable equal to 1 if 25th percentile of wages in cell in 1989 is less than or equal to minimum wage in 1989.

Table 4: Mean Prediction Errors of 1989 Wages from Single Index Models, White Men and Women

	Men				Women			
	Cell	Cell Percentiles			Cell	Cell Percentiles		
	Mean	25th	50th	75th	Mean	25th	50th	75th
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
By Age:								
16-25 Years	-0.002	-0.025	0.007	0.018	-0.016	-0.030	-0.013	-0.009
	(0.007)	(0.009)	(0.009)	(0.010)	(0.007)	(0.008)	(0.007)	(0.008)
26-35 Years	0.010	0.011	0.002	0.006	-0.005	0.001	-0.010	-0.008
	(0.007)	(0.009)	(0.008)	(0.008)	(0.007)	(0.009)	(0.008)	(0.008)
36-46 Years	-0.020	-0.013	-0.017	-0.025	0.025	0.032	0.029	0.026
	(0.007)	(0.010)	(0.009)	(0.009)	(0.007)	(0.009)	(0.008)	(0.009)
47-55 Years	-0.008	0.016	-0.010	-0.025	0.010	0.010	0.010	0.011
	(0.009)	(0.011)	(0.010)	(0.011)	(0.008)	(0.011)	(0.010)	(0.011)
56-65 Years	0.029	0.036	0.025	0.022	-0.004	0.006	-0.007	-0.019
	(0.010)	(0.013)	(0.012)	(0.013)	(0.010)	(0.013)	(0.012)	(0.013)
By Education:								
9-11 Years Education	-0.007	-0.033	-0.020	0.004	0.003	-0.014	0.004	0.008
	(0.007)	(0.009)	(0.009)	(0.010)	(0.009)	(0.001)	(0.009)	(0.010)
12 Years Education	-0.025	-0.033	-0.031	-0.029	-0.008	-0.003	-0.003	-0.018
	(0.006)	(0.008)	(0.007)	(0.008)	(0.006)	(0.007)	(0.006)	(0.007)
13-15 Years Education	0.005	-0.002	0.006	0.012	0.012	0.018	0.003	0.011
	(0.007)	(0.009)	(0.009)	(0.009)	(0.007)	(0.009)	(0.008)	(0.009)
16+ Years Education	0.049	0.098	0.074	0.037	0.007	0.004	0.002	0.030
(All Ages)	(0.009)	(0.010)	(0.010)	(0.011)	(0.013)	(0.012)	(0.011)	(0.013)
College Graduates, By Age:								
Age 26-35	0.097	0.123	0.112	0.096	0.009	-0.006	-0.003	0.045
	(0.010)	(0.013)	(0.012)	(0.014)	(0.015)	(0.016)	(0.014)	(0.016)
Age 36-46	-0.014	0.064	0.012	-0.036	0.024	0.025	0.015	0.046
	(0.015)	(0.016)	(0.016)	(0.019)	(0.020)	(0.021)	(0.018)	(0.021)
Age 47-55	-0.018	0.088	-0.004	-0.067	-0.054	-0.028	-0.042	-0.041
	(0.021)	(0.023)	(0.022)	(0.026)	(0.026)	(0.030)	(0.026)	(0.029)
Age 56-65	0.000	0.060	0.081	-0.018	-0.040	0.017	-0.052	-0.042
	(0.026)	(0.031)	(0.030)	(0.035)	(0.033)	(0.041)	(0.035)	(0.038)

Notes: Entries are average prediction errors of cell mean log wages (columns 1 and 5) and cell wage percentiles (columns 2-4 and 6-8) in 1989. Predictions of cell mean wages are based on models presented in column 5 of Table 2. Predictions of cell percentiles are based on models presented in columns 2 and 5 of Table 3. Cell-specific prediction errors are weighted by weighted count of workers in cell in 1979. Standard errors in paretheses.

Table 5: Cross-Sectional Black-White Wage Gaps for Men and Women, 1973-1989

	Men				Women			
	1973/74	1979	1984	1989	1973/74	1979	1984	1989
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1. Unadjusted Gap	-0.233	-0.208	-0.226	-0.242	-0.093	-0.049	-0.068	-0.081
	(0.009)	(0.006)	(0.006)	(0.006)	(0.008)	(0.005)	(0.005)	(0.006)
2. Adjusted Gap	-0.146	-0.140	-0.154	-0.178	-0.029	-0.017	-0.033	-0.044
	(0.007)	(0.005)	(0.005)	(0.005)	(0.007)	(0.004)	(0.005)	(0.005)
3. By Age:								
Age 16-25	-0.089	-0.104	-0.130	-0.133	-0.001	-0.028	-0.071	-0.063
	(0.013)	(0.009)	(0.010)	(0.011)	(0.013)	(0.008)	(0.009)	(0.010)
Age 26-35	-0.170	-0.139	-0.160	-0.207	-0.019	0.011	-0.018	-0.087
	(0.013)	(0.008)	(0.008)	(0.008)	(0.013)	(0.008)	(0.008)	(0.008)
Age 36-45	-0.192	-0.177	-0.164	-0.204	-0.029	-0.011	-0.018	-0.012
	(0.015)	(0.010)	(0.010)	(0.010)	(0.014)	(0.009)	(0.009)	(0.009)
Age 46-55	-0.153	-0.155	-0.159	-0.157	-0.056	-0.034	0.000	0.026
	(0.016)	(0.011)	(0.012)	(0.013)	(0.016)	(0.011)	(0.011)	(0.012)
Age 56-65	-0.134	-0.142	-0.169	-0.142	-0.089	-0.060	-0.094	-0.061
	(0.022)	(0.015)	(0.017)	(0.018)	(0.023)	(0.015)	(0.015)	(0.017)
4. By Education:								
Dropout	-0.152	-0.154	-0.163	-0.163	-0.044	-0.040	-0.040	-0.053
	(0.010)	(0.007)	(0.009)	(0.010)	(0.011)	(0.008)	(0.010)	(0.011)
High School	-0.159	-0.150	-0.173	-0.203	-0.074	-0.038	-0.054	-0.081
	(0.012)	(0.007)	(0.007)	(0.007)	(0.011)	(0.007)	(0.007)	(0.007)
Some College	-0.149	-0.133	-0.143	-0.175	-0.024	-0.008	-0.042	-0.038
	(0.020)	(0.011)	(0.011)	(0.010)	(0.018)	(0.010)	(0.009)	(0.009)
College Grad	-0.047	-0.063	-0.091	-0.130	0.200	0.091	0.047	0.052
	(0.026)	(0.014)	(0.014)	(0.013)	(0.022)	(0.012)	(0.011)	(0.011)

Notes: Entries represent estimated differentials in log hourly wages between black and white workers. Adjusted gaps and gaps by age and education represent estimated coefficients of a black indicator variable (or the interaction of a black indicator with age or education indicators) in a linear regression model than includes linear education and quartic experience terms, 8 regional dummies and an Hispanic indicator.

Table 6: Decomposition of Changes in Black Relative Wages, 1979 to 1989

	Male Decomposition: Blacks	Male Decomposition: Whites	Male Decomposition: Blacks-Whites	Female Decomposition: Blacks	Female Decomposition: Whites	Female Decomposition: Blacks-Whites
	(1)	(2)	(3)	(4)	(5)	(6)
<u>Average Across Cells of:</u>						
1. Predicted Within Cell Change [a]	0.328	0.381	-0.053	0.448	0.468	-0.020
	(0.005)	(0.004)	(0.003)	(0.004)	(0.004)	(0.002)
2. Unpredicted Within Cell Change [b]	0.007	0.000	0.007	-0.018	0.000	-0.018
	(0.009)	(0.005)	(0.009)	(0.008)	(0.005)	(0.008)
3. Change in Cell Distribution [c]	0.078	0.066	0.011	0.089	0.083	0.006
	(0.003)	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)
4. Total Change in Mean Log Wages 1979-89	0.413	0.448	-0.035	0.518	0.551	-0.033
	(0.007)	(0.002)	(0.007)	(0.006)	(0.002)	(0.006)

Notes: Standard errors in parentheses. Predictions are based on quadratic single index models fit to whites only.

[a] Weighted average of difference between predicted mean log wage for cell in 1989 and actual mean log wage of cell in 1979.

[b] Weighted average of difference between actual mean log wage for cell in 1989 and predicted mean log wage of cell based on single index model.

c Change in cell weight between 1979 and 1989, weighted by mean log wage of cell in 1989.

Table 7: Unpredicted Changes in Mean Cell Wages for Blacks, 1979-1989

	Men		Women	
	Blacks (1)	Blacks-Whites (2)	Blacks (3)	Blacks-Whites (4)
Overall	0.007 (0.009)	0.007 (0.009)	-0.018 (0.008)	-0.018 (0.008)
By Age:				
16-25 Years	-0.016 (0.014)	-0.014 (0.014)	-0.035 (0.014)	-0.020 (0.014)
26-35 Years	-0.045 (0.016)	-0.055 (0.016)	-0.112 (0.014)	-0.107 (0.014)
36-46 Years	-0.002 (0.017)	0.017 (0.018)	0.023 (0.015)	-0.002 (0.016)
47-55 Years	0.080 (0.024)	0.088 (0.025)	0.086 (0.021)	0.077 (0.023)
56-65 Years	0.152 (0.028)	0.124 (0.029)	0.084 (0.029)	0.087 (0.030)
By Education:				
9-11 Years Education	0.028 (0.014)	0.035 (0.014)	0.036 (0.015)	0.033 (0.015)
12 Years Education	-0.007 (0.134)	0.018 (0.013)	-0.037 (0.012)	-0.028 (0.012)
13-15 Years Education	0.000 (0.019)	-0.005 (0.020)	-0.046 (0.017)	-0.058 (0.018)
16+ Years Education (All Ages)	-0.009 (0.028)	-0.059 (0.029)	-0.046 (0.026)	-0.053 (0.024)
College Graduates, By Age:				
Age 26-35	-0.091 (0.044)	-0.188 (0.044)	-0.076 (0.036)	-0.085 (0.036)
Age 36-46	-0.009 (0.054)	0.006 (0.055)	-0.072 (0.049)	-0.095 (0.047)
Age 47-55	0.083 (0.084)	0.101 (0.086)	-0.069 (0.064)	-0.015 (0.066)
Age 56-65	0.075 (0.110)	0.076 (0.113)	-0.188 (0.111)	-0.147 (0.112)

Notes: Entries are average prediction errors of cell means for blacks (columns 1 and 3) or differences in average prediction errors of cell means between blacks and whites (columns 2 and 4). Standard errors in parentheses.

Table 8: Unpredicted Changes in Mean Cell Wages for Young Blacks, 1979-1989

	Men		Women	
	Blacks (1)	Blacks-Whites (2)	Blacks (3)	Blacks-Whites (4)
Age 16-25, By Education:				
9-11 Years Education	0.012	0.024	-0.033	-0.048
	(0.023)	(0.022)	(0.025)	(0.024)
12 Years Education	-0.046	-0.008	-0.056	-0.013
	(0.021)	(0.021)	(0.020)	(0.020)
13-15 Years Education	-0.044	-0.075	-0.052	-0.035
	(0.032)	(0.033)	(0.028)	(0.029)
Age 26-35, By Education:				
9-11 Years Education	-0.034	0.013	-0.078	-0.055
	(0.032)	(0.034)	(0.031)	(0.034)
12 Years Education	-0.036	0.008	-0.136	-0.128
	(0.023)	(0.023)	(0.020)	(0.021)
13-15 Years Education	-0.045	-0.047	-0.120	-0.112
	(0.031)	(0.032)	(0.029)	(0.030)
16+ Years Education	-0.091	-0.188	-0.076	-0.085
	(0.044)	(0.044)	(0.036)	(0.036)

Notes: Entries are average prediction errors of cell means for blacks (columns 1 and 3) or differences in average prediction errors of cell means between blacks and whites (columns 2 and 4). Standard errors in parentheses.

Appendix: Comparison of Wage Measures in Monthly Earnings Supplement and March CPS

	Outgoing Rotation Files	March CPS All	March CPS Weighted By Weeks[a/]	March CPS FTFY[b/]
Men				
1979 Data				
a. Mean Log Wage-Blacks	1.660	1.634	1.687	1.762
	(0.005)	(0.011)	(0.010)	(0.012)
b. Mean Log Wage-Whites	1.868	1.882	1.935	2.019
	(0.002)	(0.003)	(0.003)	(0.003)
c. Black-White Gap	-0.208	-0.248	-0.248	-0.257
	(0.006)	(0.011)	(0.010)	(0.012)
1989 Data				
a. Mean Log Wage-Blacks	2.074	2.116	2.176	2.246
	(0.006)	(0.011)	(0.011)	(0.012)
b. Mean Log Wage-Whites	2.316	2.361	2.410	2.499
	(0.002)	(0.004)	(0.003)	(0.004)
c. Black-White Gap	-0.242	-0.245	-0.234	-0.252
	(0.006)	(0.012)	(0.011)	(0.013)
Change in Black-White Gap: 1979 to 1989	-0.034	0.003	0.014	0.005
	(0.009)	(0.017)	(0.015)	(0.018)
Women				
1979 Data				
a. Mean Log Wage-Blacks	1.424	1.383	1.428	1.517
	(0.005)	(0.009)	(0.009)	(0.012)
b. Mean Log Wage-Whites	1.474	1.455	1.497	1.583
	(0.002)	(0.003)	(0.003)	(0.004)
c. Black-White Gap	-0.050	-0.072	-0.069	-0.066
	(0.005)	(0.009)	(0.009)	(0.013)
1989 Data				
a. Mean Log Wage-Blacks	1.944	1.957	2.011	2.104
	(0.005)	(0.010)	(0.010)	(0.012)
b. Mean Log Wage-Whites	2.025	2.035	2.079	2.177
	(0.002)	(0.003)	(0.003)	(0.004)
c. Black-White Gap	-0.081	-0.078	-0.068	-0.073
	(0.006)	(0.011)	(0.010)	(0.012)
Change in Black-White Gap: 1979 to 1989	-0.031	-0.006	0.001	-0.007
	(0.008)	(0.014)	(0.014)	(0.018)

Notes: Entries are mean log wages with standard errors in parentheses. Entries in column 1 are from pooled monthly files of individuals in the outgoing rotation groups of the Current Population Survey (CPS) for 1979 and 1989, and are based on reported wages for main job during the survey week. Entries in columns 2-4 are from individuals in the March 1980 and March 1990 CPS, and are based on reported wage and salary earnings from all jobs in the previous calendar year, divided by hours worked last year. Individuals with allocated earnings data and individuals with extreme values for their hourly wage are excluded.

a/Weighted average of log wage rates, using weeks worked last year as a weight.

b/Based on full-time full-year workers only.