ON DECOMPOSING CHANGES IN MALE-FEMALE WAGE GAP

by

Sang-Hyop Lee

Working Paper No. 00-12
May 2000

On Decomposing Changes in Male-Female Wage Gap

by

Sang-Hyop Lee*

Working Paper No. 00-12

March 2000

*Assistant Professor of Economics at University of Hawaii at Manoa. Address: 2424 Maile Way, SSB 542, Honolulu, Hawaii 96822. E-mail address: leesang@hawaii.edu

Abstract

Several researchers have used trend decomposition techniques to decompose the change in the wage gap between two groups. In contrast to the previous decomposition techniques which are flawed on both conceptual and technical grounds, this paper provides alternative decomposition methods which have clearer interpretations. The alternative decomposition is then applied to the May CPS from 1983 and 1993. The results from the empirical application in this two-period model show that the previous decomposition methods yield substantially lower estimates of the portion due to changes in characteristics, and therefore higher estimates of the portion due to changes in coefficients. This implies the conclusions drawn from previous methods might overstate the change in the wage gap attributable to decline in discrimination.

1. Introduction

Several researchers have used trend decomposition techniques to decompose the *change* in the wage gap between two parts. These analyses are important, since they show how the changes in the means and the coefficients of the explanatory variables combine to affect the change in the wage gap over time. The previous results from these analyses suggest that, all else equal, the proportion of the male-female wage gap attributable to discrimination declined during 1970's (Blau and Beller (1988)). This is also interpreted as evidence that government policy play a role in declining wage gap, due to social discrimination. Stronger evidence of the effect of anti-discrimination policies has also been obtained for many other countries.[1]

However, since no specification seems to be clearly better than the other, the choice of the decomposition technique has been arbitrary.[2] This paper re-examines the previous decomposition techniques, and argues that the decomposition methods adopted by Blau and Beller (1988), Wellington (1992), and O'Neill and Polachek (1993) are flawed on both conceptual and technical grounds. In contrast, this paper suggests an alternative decomposition method which might avoid the shortcomings of interpretation found in previous treatments. The alternative decomposition is then applied to the May CPS from 1983 and 1993, and the results are compared to the results obtained using the previous methods.

The results from the empirical application show that the previous decomposition methods yield substantially lower estimates of the portion due to changes in characteristics, and therefore higher estimates of the portion due to changes in coefficients. This implies the conclu-

[1] See Blau and Kahn (1995) for a discussion.
[2] See Wellington (1992) for a discussion.

sions drawn from previous methods may overstate the change in the wage gap attributable to decline in discrimination. In section 2, the two-period decomposition method is derived from single-period decomposition. Its implications are also discussed. Section 3 presents an empirical application. Section 4 summarizes the paper.

2. Decomposition of the Change in the Wage Gap

A Critique of the Previous Decompositions

The most common forms of the decomposition are developed by Blau and Beller (1988), Wellington (1992), and O'Neil and Polachek (1993). Let $\overline{\ln(w_{mt})}$ and $\overline{\ln(w_{ft})}$ be the means of the log of male (m) and the log of female (f) wages. If the wage model is estimated separately by sex, then the means of the log wage gap can be expressed as the following form

$$(1) \qquad \overline{\ln(w_{mt})} - \overline{\ln(w_{ft})} = \overline{X}_{mt}\,\hat{\beta}_{mt} - \overline{X}_{ft}\,\hat{\beta}_{ft}$$

where $\overline{X}_{mt}$ and $\overline{X}_{ft}$ are vectors containing the means of the variables, and $\hat{\beta}_{mt}$ and $\hat{\beta}_{ft}$ are the estimated coefficients. The subscript t represents the time at which the variables are measured. Let the time increment be measured as, $\Delta_t\overline{\ln(w)} = \overline{\ln(w_t)} - \overline{\ln(w_{t-1})}$, $\Delta_t\overline{X} = \overline{X}_t - \overline{X}_{t-1}$, and $\Delta_t\hat{\beta} = \hat{\beta}_t - \hat{\beta}_{t-1}$. Given the equation (1), the change in the wage gap, $\Delta_t\overline{\ln(w_m)} - \Delta_t\overline{\ln(w_f)}$ has taken the following forms

(2) Blau and Beller (1988): $(\hat{\beta}_{mt-1}\Delta_t\overline{X}_m - \hat{\beta}_{ft-1}\Delta_t\overline{X}_f) + (\overline{X}_{mt-1}\Delta_t\hat{\beta}_m - \overline{X}_{ft-1}\Delta_t\hat{\beta}_f) + \alpha$

(3) Wellington (1992): $(\hat{\beta}_{mt}\Delta_t\overline{X}_m - \hat{\beta}_{ft}\Delta_t\overline{X}_f) + (\overline{X}_{mt-1}\Delta_t\hat{\beta}_m - \overline{X}_{ft-1}\Delta_t\hat{\beta}_f)$

(4) O'Neill and Polachek (1993): $(\bar{\hat{\beta}}_m\Delta_t\overline{X}_m - \bar{\hat{\beta}}_f\Delta_t\overline{X}_f) + (\overline{\overline{X}}_m\Delta_t\hat{\beta}_m - \overline{\overline{X}}_f\Delta_t\hat{\beta}_f) + \alpha'$

where $\overline{\overline{X}}_m$ and $\overline{\overline{X}}_f$ are vectors containing the means of the variables pooling two periods for males and females, while $\bar{\hat{\beta}}_m$ and $\bar{\hat{\beta}}_f$ are the means of the estimated coefficients pooling two periods for males and females. In each of these, the first term has been interpreted as the change in the wage gap due to a change in characteristics, while the second term has been interpreted as the change in the wage gap due to a change in coefficients (discrimination). Notice that the above equations look very similar to each other. The only difference between (2) and (3) is that the first term of (2) is evaluated at base year coefficients, while that of (3) is evaluated at current year coefficients. Similarly, the difference between (3) and (4) also results from the different time at which the variables and coefficients are measured. One common problem in both (2) and (4) is that the sum of the first term and second term is not equal to the total change in the wage gap. The last term (α and α') in their decompositions has no clear interpretation. Although the α term does not appear in equation (3), the (3) has different (and probably more serious) problem. It does not answer why the changes in the characteristics are evaluated at *current year* coefficients, while the changes in the coefficients are evaluated at *base year* characteristics. In addition, there seem to be more important flaws in these decompositions both on conceptual and technical grounds.

First, although Wellington (1990) argues that she employs these kinds of decomposition in the spirit of Oaxaca's (1973) decomposition, these decompositions are far from the spirit of Oaxaca's decomposition. Let's consider Oaxaca's single-period decomposition model. In a single period earnings function, Oaxaca shows that we can decompose the wage gap between two groups into differences in the means and differences in the coefficients including the constant term. Given equation (1), the means of the log wage gap can be decomposed in two ways. That is

$$(5) \qquad \overline{\ln(w_m)} - \overline{\ln(w_f)} = \hat{\beta}_m \Delta_g \overline{X} + \overline{X}_f \Delta_g \hat{\beta}$$

or

$$(6) \qquad \overline{\ln(w_m)} - \overline{\ln(w_f)} = \hat{\beta}_f \Delta_g \overline{X} + \overline{X}_m \Delta_g \hat{\beta}$$

where $\Delta_g \overline{X} = \overline{X}_m - \overline{X}_f$, $\Delta_g \hat{\beta} = \hat{\beta}_m - \hat{\beta}_f$. The first term of either (5) or (6) is the part of the wage gap due to the different characteristics of males and females, and the second term is the part of the gap due to different coefficients. If in the absence of discrimination males and females receive identical returns for the same characteristics, and differences in wages would therefore be due only to differences in characteristics, then this second term can be interpreted as the wage gap due to discrimination. For the time being, assume that in the absence of discrimination the male wage structure would prevail at both time t and t-1.[3] It is the assumption made in using (5). Oaxaca's one-period decomposition then can be calculated at both time t and t-1, which have the following forms

$$(7) \qquad \overline{\ln(w_{mt})} - \overline{\ln(w_{ft})} = \hat{\beta}_{mt} \Delta_g \overline{X}_t + \overline{X}_{ft} \Delta_g \hat{\beta}_t$$

$$(8) \qquad \overline{\ln(w_{mt-1})} - \overline{\ln(w_{ft-1})} = \hat{\beta}_{mt-1} \Delta_g \overline{X}_{t-1} + \overline{X}_{ft-1} \Delta_g \hat{\beta}_{t-1}$$

where subscripts t and t-1 are the times at which the variables are measured. Now notice that we never get the previous decomposition forms by using Oaxaca's decomposition method, since neither $\hat{\beta}_{ft}$ nor $\hat{\beta}_{ft-1}$ appears in the first term of the right hand sides in both (7) and (8). Similarly, we would get the same result if we started with $\hat{\beta}_{ft}$ and $\hat{\beta}_{ft-1}$ as non-discriminatory wage structure.

Second, a calculation (interpretation) problem with the previous decomposition can be demonstrated using a relatively simple example. Suppose the change in characteristics over time is same for both males and females, but that there is an initial difference in the level of characteristics between males and females (that is $\Delta_t \overline{X}_m = \Delta_t \overline{X}_f$, but $\overline{X}_{mt-1} \neq \overline{X}_{ft-1}$). In addition, assume that the change in coefficients over time is the same for both males and females, and that there is no difference in the level of coefficient between males and females (that is $\Delta_t \hat{\beta}_m = \Delta_t \hat{\beta}_f$, and $\hat{\beta}_{mt-1} = \hat{\beta}_{ft-1}$, and therefore $\hat{\beta}_{mt} = \hat{\beta}_{ft}$). In this case, the previous decomposition methods suggest that the change in the wage gap is totally due to change in coefficients (discrimination), when this is clearly not what has occurred. It does not answer why an initial difference in the level of characteristics leads to the change in the wage gap totally due to change in coefficients, but not due to change in characteristics.[4] In the next part, I consider alternative decomposition methods which have a clearer interpretation.

[3] Wage structure describes the array of prices set for various labor market skills.
[4] We would get an exact same result if we instead assumed that there is an initial difference in the coefficient, but not in the characteristics.

Two-Period Model of Oaxaca's Decomposition

Let's subtract (8) from (7) side by side. Then, we get

$$(9) \quad \Delta_t \overline{\ln(w_m)} - \Delta_t \overline{\ln(w_f)} = [\hat{\beta}_{mt} \Delta_g \overline{X}_t - \hat{\beta}_{mt-1} \Delta_g \overline{X}_{t-1}] + [\overline{X}_{ft} \Delta_g \hat{\beta}_t - \overline{X}_{ft-1} \Delta_g \hat{\beta}_{t-1}]$$

The right hand side of equation (9) can be transformed into the following form.

$$(10) \quad [\hat{\beta}_{mt-1} (\Delta_t \overline{X}_m - \Delta_t \overline{X}_f) + \Delta_g \overline{X}_t \Delta_t \hat{\beta}_m] + [\overline{X}_{ft-1} (\Delta_t \hat{\beta}_m - \Delta_t \hat{\beta}_f) + \Delta_g \hat{\beta}_t \Delta_t \overline{X}_f]$$

or

$$(11) \quad [\underbrace{\hat{\beta}_{mt} (\Delta_t \overline{X}_m - \Delta_t \overline{X}_f)}_{(a)} + \underbrace{\Delta_g \overline{X}_{t-1} \Delta_t \hat{\beta}_m}_{(b)}] + [\underbrace{\overline{X}_{ft} (\Delta_t \hat{\beta}_m - \Delta_t \hat{\beta}_f)}_{(c)} + \underbrace{\Delta_g \hat{\beta}_{t-1} \Delta_t \overline{X}_f}_{(d)}]$$

These forms clearly show how the change in the wage gap over time can be decomposed into four parts. Consider (11) only. The term (a) represents the change in the wage gap due to a change in the characteristics evaluated at males' current year coefficient. It is evaluated at *males'* coefficient, $\hat{\beta}_{mt}$, since we temporarily assumed that in the absence of discrimination the male wage structure would prevail at both time t and t-1. This form implies that if characteristics of males increase faster than those of females, then the wage gap increases due to a change in characteristics as long as the coefficients attached to characteristics are positive. The term (b) is an adjustment term, which implies that if there is a difference in the level of characteristics between males and females, then there exists a change in the wage gap due to a difference in the level of characteristics, even if the change in the female and male wage structure is the same. Notice that neither $\hat{\beta}_{ft}$ nor $\hat{\beta}_{ft-1}$ appears in both terms (a) and (b), since we assumed that in the absence of discrimination the male wage would prevail at both time t and t-1. Similarly, the term (c) represents the change in the wage gap due to a change in coefficients. The term (d) is again an adjustment term, which implies that if there is a difference in the level of coefficients between males and females, then there exists a change in the wage gap due to a difference in the level of coefficients, even though men and women experience the same change in characteristics.

Table 1 demonstrates several examples of this decomposition, which show the importance of these adjustment terms in this analysis.[5] Case I is the assumption which we made to show a calculation problem in the previous decompositions. Notice that how a difference in both change and level of characteristics (coefficient) leads to the change in the wage gap due to differences in both change and level of characteristics (coefficient).

Now suppose that in the absence of discrimination the female wage would prevail at both time t and t-1. It is the assumption made in equation (6). A corresponding decomposition is of the form

$$(12) \quad [\hat{\beta}_{ft}(\Delta_t \overline{X}_m - \Delta_t \overline{X}_f) + \Delta_g \overline{X}_{t-1} \Delta_t \hat{\beta}_f] + [\overline{X}_{mt}(\Delta_t \hat{\beta}_m - \Delta_t \hat{\beta}_f) + \Delta_g \hat{\beta}_{t-1} \Delta_t \overline{X}_m]$$

[5] By using similar examples, Blau and Kahn (1995) demonstrate the importance of wage structure in explaining the international differences in the male-female wage gap.

which has a similar interpretation as equation (11). Again, it is noticeable that only coefficients of females appear in the left hand side, since we assumed that, in the absence of discrimination, the female wage would prevail at both time t and t-1.

However, the validity of these new decomposition also hinge on the rationale of their assumptions: that is, in the absence of discrimination, the male (female) wage structure would prevail at *both* time t and t-1.

The two decomposition methods above do not guarantee the same result, because different wage structure assumptions are used in the alternative decomposition methods. Neumark (1988) argues that the non-discriminatory wage structure should be derived from a theoretical model of discriminatory behavior, and shows how different assumptions about employers' discriminatory tastes lead to Oaxaca's estimators. A corresponding decomposition based on his argument is of the form

$$(13) \quad \underbrace{[\hat{\beta}_t(\Delta_t \overline{X}_m - \Delta_t \overline{X}_f)}_{(a)'} + \underbrace{\Delta_g \overline{X}_{t-1} \Delta_t \hat{\beta}]}_{(b)'} + \underbrace{[\{\overline{X}_{mt}(\Delta_t \hat{\beta}_m - \Delta_t \hat{\beta}) + \overline{X}_{ft}(\Delta_t \hat{\beta} - \Delta_t \hat{\beta}_f)\}}_{(c)'} +$$

$$\underbrace{\{(\hat{\beta}_{mt-1} - \hat{\beta}_{t-1})\Delta_t \overline{X}_m + (\hat{\beta}_{t-1} - \hat{\beta}_{ft-1})\Delta_t \overline{X}_f\}]}_{(d)'}$$

In this decomposition, terms (a)' and (b)' represents a change in the wage gap due to characteristics evaluated at the current year non-discriminatory wage structure. Similarly, terms (c)' and (d)' can be interpreted as the part due to coefficients. If it is assumed that in the absence of discrimination the current male wage structure would prevail at both t and t-1, then

$\hat{\beta}_{t-1} = \hat{\beta}_{mt-1}$, $\hat{\beta}_t = \hat{\beta}_{mt}$, and (13) reduces to (11). If instead it is assumed that in the absence of discrimination the current female wage structure would prevail at both t and t-1, then $\hat{\beta}_{t-1} = \hat{\beta}_{ft-1}$, $\hat{\beta}_t = \hat{\beta}_{ft}$, and (13) reduces to (12). Thus, (11) and (12) are two special cases of (13), and the critical issue is the choice of $\hat{\beta}_t$ and $\hat{\beta}_{t-1}$, the non-discriminatory wage structure at each time period, and therefore $\Delta_t\hat{\beta}$. Also notice that both (a)' and (b)' does not depend on either coefficients of males or coefficients of females. Neumark (1988) proposes that this estimator of the non-discriminatory wage structure can be implemented simply, as the coefficients estimated from the log wage regression for the whole sample, using predicted wages from the log wage regression as the dependent variable.[6]

Notice that equation (13) also show how a difference in both changes and levels of characteristics (coefficients) leads to the change in the wage gap due to a difference in both changes and levels of characteristics (coefficients). In the next part, I examine how different decomposition methods lead to different results.

3. An Empirical Application

In this section, alternative decomposition methods are applied to the May Current Population Survey (CPS) samples from 1983 and 1993. To simplify the discussion, the sample is restricted to white, full time, year round, private sector workers. Individuals in agriculture, forestry and fishery, and personal service industries are dropped. Hourly wages are used as the wage variable in order to control for the change in hours worked between genders over peri-

[6] See Neumark (1998) pp. 283-89 for a detailed procedure.

ods. Table 2 presents descriptive statistics for the data set in 1983 and 1993. The log wage gap in 1983 was 0.355, implying that female wages are 70% of the male wages in 1983. The wage gap falls to 0.251 by 1993, implying that female wages are 78% of the male wages in 1993. The mean value of schooling is higher for males in both periods, which is partly due to males being older on average.

In order to compare the alternative decomposition with the previous decomposition, I construct the following specification. A basic wage equation of the form

$$(14) \qquad \ln(w_i) = X_i\beta + \varepsilon_i$$

is estimated without industry or occupation dummy variables, where X is a vector of workers' characteristics.

Table 3 reports the OLS estimates for the basic form of the log wage equation. The coefficients in these estimations are used in the calculation of decomposition. It is noticeable that the coefficients for the schooling variable increase over time. However, they increase faster for females than males, suggesting that this variable might play a role in decreasing the wage gap through the change in the coefficient effect. The coefficients for the marital status dummy variable increase for females, but decrease for males, suggesting that this variable also might play a role in decreasing the wage gap through the change in the coefficient effect.

Table 4 presents a comparison between the previous decomposition and the alternative decomposition methods for the basic specification. At the bottom of the table, the change in the wage gap attributable to each variable is added up, in order to summarize the result. Results show that the new decomposition produces more or less lower estimates of the percentage of the wage gap due to coefficients (therefore higher estimates of the percentage of the wage gap due to a change in characteristics). Using $\hat{\beta}_{mt}$ as the male wage structure, it is estimated that 39% of the change in the wage gap is due to characteristics, while using Blau and Beller's (similarly Wellington's) method leads to an estimate of 22%. The adjustment terms are in parentheses. They markedly vary by assumptions on wage structure, implying a potentially important role for wage structure at each time period in decomposing the change in the wage gap.

Table 5 presents results for alternative specifications which consider changes in the employment distributions of males and females across industries and occupations. The first specification in Table 5 includes the portion due to changes in the employment distributions of men and women across industry. Based on the argument by Macpherson and Hirsch (1995), the second specification considers the effect of sex segregation due to a change of differences in gender density in specific occupation and industry. The last specification includes the portion due to changes in the employment distributions of men and women across industry and occupation. Occupation and industry dummy variables are for one digit 1980 code.

Some results are easily noticeable from Table 5. First, when these variables are added as additional control variables, the portion due to characteristics rises substantially in both previous decompositions and alternative methods. This might reflect the shift in industry and occupation structure and relative demand for labor over time. Second, the rise in the portion due to changes in characteristics is substantially higher when we use the alternative methods. This implies one cannot arbitrarily choose a decomposition method without considering its differences from other alternatives, since they lead to quite different interpretation; the conclusions drawn from previous methods may overstate the change in the wage gap due to decline in discrimination. Third, as Neumark (1988) points out in a single-period model, the estimates using (11) (male wage structure) and (12) (female wage structure) do not provide any range for the non-discriminatory wage structure in the two-period model, either. However, unlike Neumark's single period model, the estimates based on non-discriminatory wage structure are not necessarily more sensitive than the estimates based on (11) and (12) to differences in the distribution of characteristics across men and women.

4. Conclusion

Although several researchers have used trend decomposition techniques to decompose the change in the wage gap between two groups, their decomposition methods are flawed on both conceptual and technical grounds. In contrast, this paper suggests an alternative decomposition method which might avoid the shortcomings of interpretation found in previous treatments. The alternative decomposition is then applied to the May CPS from 1983 and 1993, and the results are compared to the results obtained using the previous methods.

The results from the empirical application in this two-period model show that the previous decomposition methods yield substantially lower estimates of the portion due to changes in characteristics, and therefore higher estimates of the portion due to changes in coefficients. This implies the conclusions drawn from previous methods may overstate the change in the wage gap attributable to decline in discrimination.

Reference

Blau, Francine D. and Lawrence M. Kahn (1995), The Gender Earnings Gap: Some International Evidence, in Richard B. Freeman & Lawrence F. Katz (eds.), *Differences & Changes in Wage Structures,* NBER Comparative Labor Markets Series, The University of Chicago Press.

Blau, Francine D. and Andrea H. Beller (1988), Trends in Earning Differentials by Gender, 1971-1981, *Industrial and Labor Relations Review*, Vol. 41, No. 4, pp. 513-529.

Blinder, Alan S. (1973), Wage Discrimination: Reduced Form and Structural Estimates, *The Journal of Human Resources*, Vol. 8, No. 4, pp. 436-455.

Butler, Richard (1982), Estimating Wage Discrimination in the Labor Market, *The Journal of Human Resources*, Vol. 17, No. 4, pp. 606-621.

Cotton, Jeremiah (1988), On the Decomposition of Wage Differentials, *The Review of Economics and Statistics*, pp. 236-243.

Fields, Judith and Edward N. Wolff (1995), Interindustry Wage Differentials and the Gender Wage Gap, *Industrial and Labor Relations Review*, Vol. 49, No. 1, pp. 105-120.

Jones, F. L. (1983), On Decomposing the Wage Gaps: A Critical Comment on Blinder's Method, *The Journal of Human Resources*, Vol. 18, No. 1, pp. 126-130.

Kuhn, Peter (1987), Sex Discrimination in Labor Markets: The Role of Statistical Evidence, *American Economic Review*, Vol. 77, No. 4, pp. 567-583.

Macpherson, David A. and Barry T. Hirsch (1995), Wages and Gender Composition: Why Do Women's Job Pay Less?, *Journal of Labor Economics*, Vol. 13, No. 3, pp. 426-471.

Neumark, David (1988), Employers' Discriminatory Behavior and the Estimation of Wage Discrimination, *The Journal of Human Resources*, Vol. 23, No. 3, pp. 279-295.

Oaxaca, Ronald (1973), Male-Female Wage Differentials in Urban Labor Markets, *International Economics Review*, Vol. 14, No. 3, pp. 693-709.

Oaxaca, Ronald and Michael R. Ransom (1994), On Discrimination and the

Decomposition of Wage Differentials, *Journal of Econometrics*, Vol. 61, March, pp. 5-21.

O'Neil, June and Solomon Polachek (1993), Why the Gender Gap in Wages Narrowed in the 1980s, *Journal of Labor Economics*, Vol. 11, No. 1, pp. 205-225.

Wellington, Alison J. (1993), Changes in the Male/Female Wage Gap, 1976-85, *The Journal of Human Resources,* Vol. 28, No. 2, pp. 383-411.

Table 1. Examples of Decomposition

Case	Assumptions on Characteristics: Assumptions	Assumptions on Characteristics: Results	Assumptions on Coefficients: Assumptions	Assumptions on Coefficients: Results
I	Change: same ($\Delta_t \overline{X}_m = \Delta_t \overline{X}_f$) Level: different ($\overline{X}_{mt-1} \neq \overline{X}_{ft-1}$)	due to a change in characteristics = 0 difference in levels of characteristic ≠ 0	Change: same ($\Delta_t \hat{\beta}_m = \Delta_t \hat{\beta}_f$) Level: same ($\hat{\beta}_{mt-1} = \hat{\beta}_{ft-1}$)	due to a change in coefficient = 0 difference in levels of coefficient = 0
II	Change: different ($\Delta_t \overline{X}_m \neq \Delta_t \overline{X}_f$) Level: same ($\overline{X}_{mt-1} = \overline{X}_{ft-1}$)	due to a change in characteristics ≠ 0 difference in levels of characteristics = 0	Change: same ($\Delta_t \hat{\beta}_m = \Delta_t \hat{\beta}_f$) Level: same ($\hat{\beta}_{mt-1} = \hat{\beta}_{ft-1}$)	due to a change in coefficient = 0 difference in levels of coefficient = 0
III	Change: different ($\Delta_t \overline{X}_m \neq \Delta_t \overline{X}_f$) Level: same ($\overline{X}_{mt-1} = \overline{X}_{ft-1}$)	due to a change in characteristics ≠ 0 difference in levels of characteristics = 0	Change: same ($\Delta_t \hat{\beta}_m = \Delta_t \hat{\beta}_f$) Level: different ($\hat{\beta}_{mt-1} \neq \hat{\beta}_{ft-1}$)	due to a change in co-efficient = 0 difference in levels of coefficient ≠ 0
IV	Change: same ($\Delta_t \overline{X}_m = \Delta_t \overline{X}_f$) Level: different ($\overline{X}_{mt-1} \neq \overline{X}_{ft-1}$)	due to a change in characteristics = 0 difference in levels of characteristics ≠ 0	Change: different ($\Delta_t \hat{\beta}_m \neq \Delta_t \hat{\beta}_f$) Level: same ($\hat{\beta}_{mt-1} = \hat{\beta}_{ft-1}$)	due to a change in co-efficient ≠ 0 difference in levels of coefficient = 0
V	Change: same ($\Delta_t \overline{X}_m = \Delta_t \overline{X}_f$) Level: different ($\overline{X}_{mt-1} \neq \overline{X}_{ft-1}$)	due to a change in characteristics = 0 difference in levels of characteristics ≠ 0	Change: different ($\Delta_t \hat{\beta}_m \neq \Delta_t \hat{\beta}_f$) Level: different ($\hat{\beta}_{mt-1} \neq \hat{\beta}_{ft-1}$)	due to a change in co-efficient ≠ 0 difference in levels of coefficient ≠ 0

Only one variable case is considered.

Table 2. Means and Standard Deviations of Variables

	1983			1993		
	Total	Male	Female	Total	Male	Female
log (wage)	6.415	6.551	6.196	6.772	6.876	6.625
	(.497)	(.495)	(.414)	(.528)	(.538)	(.476)
Schooling	12.92	12.98	12.81	13.12	13.14	13.10
	(2.56)	(2.74)	(2.25)	(2.40)	(2.54)	(2.20)
Age	36.91	37.45	36.05	37.73	37.95	37.41
	(12.25)	(12.18)	(12.32)	(11.21)	(11.28)	(11.09)
Experience	19.00	19.47	18.24	19.60	19.81	19.31
	(12.83)	(12.73)	(12.96)	(11.50)	(11.46)	(11.55)
Married	.655	.719	.551	.633	.681	.565
Union	.209	.259	.129	.130	.171	.071
Central City	.206	.195	.224	.189	.185	.195
In SMSA	.373	.385	.354	.395	.396	.393
# of observation	7165	4421	2744	7569	4448	3121

Means are reported with standard deviations in the parentheses. Age is included only for comparison between groups. Experience is calculated by a formula, age - schooling - 5. This variable construction implicitly assumes that all years since school were spent in the labor force, which is not necessarily true. This is a common problem we face when we use the CPS data.

Table 3. OLS Estimates (Basic Specification)

	1983			1993		
	Total	Male	Female	Total	Male	Female
Constant	4.90	5.02	4.99	5.03	5.12	5.00
	(.012)	(.042)	(.055)	(.010)	(.048)	(.060)
Schooling	.074	.070	.068	.098	.094	.098
	(.001)	(.002)	(.003)	(.001)	(.003)	(.004)
Experience	.031	.037	.024	.029	.032	.026
	(.001)	(.002)	(.002)	(.0004)	(.002)	(.002)
Experience x 10^{-2}	-.048	-.058	-.039	-.044	-.046	-.046
	(.001)	(.004)	(.005)	(.001)	(.004)	(.005)
Married	.144	.157	.011	.128	.146	.033
	(.004)	(.015)	(.015)	(.003)	(.016)	(.015)
Union	.223	.158	.181	.197	.147	.125
	(.004)	(.015)	(.022)	(.004)	(.018)	(.029)
Central City	.051	.023	.097	.071	.043	.118
	(.005)	(.018)	(.019)	(.004)	(.019)	(.021)
In SMSA	.125	.120	.112	.113	.094	.149
	(.005)	(.015)	(.017)	(.003)	(.015)	(.017)
R-squared	.7701	.3157	.2055	.8597	.3319	.2782
# of observation	7165	4421	2744	7569	4448	3121

Standard errors are in the parentheses. In the log wage regression for the whole sample, the dependent variable is fitted log wages from the separate wage regression. Other variables include 8 regional dummy (state division categories in the CPS) variables. Occupation and industry dummy variables are not included.

Table 4. Decomposition Based on Basic Specification

	(1) Blau & Beller, Wellington	(2) O'Neill & Polachek	(3) New Method: Male wage structure	(4) New Method: Female wage structure	(5) New Method: Non-discriminatory wage structure
Schooling (total=-.094)					
Characteristic	-.010	-.012	-.009 (.001)	-.009 (.002)	-.010 (.001)
Coefficient	-.084	-.083	-.085 (-.002)	-.085 (-.002)	-.084 (-.001)
Experience (total=-.134)					
Characteristic	-.015	-.016	-.029 (-.003)	-.017 (.000)	-.024 (-.003)
Coefficient	-.119	-.117	-.105 (.009)	-.117 (.003)	-.110 (.006)
Exp. Square (total=.110)					
Characteristic	.012	.012	.015 (.004)	.008 (-.001)	.012 (.002)
Coefficient	.098	.097	.094 (-.001)	.102 (.005)	.098 (.003)
Married (total=-.026)					
Characteristic	-.006	-.006	-.009 (-.001)	.002 (.004)	-.009 (-.002)
Coefficient	-.020	-.020	-.017 (.002)	-.028 (-.005)	-.017 (.001)
Union (total=-.001)					
Characteristic	-.006	-.006	-.006 (-.003)	-.011 (-.007)	-.009 (-.003)
Coefficient	.005	.004	.005 (.004)	.010 (-.001)	.008 (.002)
Central City (total=.002)					
Characteristic	.003	.003	.000 (-.001)	.002 (-.002)	.001 (-.001)
Coefficient	-.001	-.001	.002 (.002)	.001 (.000)	.001 (.001)
In SMSA (total=-.028)					
Characteristic	-.003	-.004	-.003 (-.001)	-.003 (-.000)	-.004 (-.001)
Coefficient	-.024	-.024	-.024 (-.001)	-.025 (-.001)	-.024 (.001)
Region (total=-.032)					
Characteristic	.003	.000	.000 (-.002)	.001 (-.004)	.000 (-.004)
Coefficient	-.035	-.031	-.032 (.001)	-.032 (.001)	-.032 (.002)
Change due to					
Change in Charact.	..	..	-.038	-.020	-.011
Difference in Charact.	..	..	-.003	-.008	-.032
Total	-.022	-.028	-.041	-.028	-.043 (41.4%)
(%)	(21.5%)	(26.9%)	(39.4%)	(26.9%)	-.075
Change in Coeff.	..	..	-.074	-.075	.014
Difference in Coeff.	..	..	.011	-.001	-.061 (58.6%)
Total	-.081	-.076	-.063	-.076	
(%)	(78.5%)	(73.1%)	(60.6%)	(73.1%)	

The change in the wage differential, $\Delta\overline{\ln(w_m)} - \Delta\overline{\ln(w_f)} = -0.104$ over the period of 1983-1993. Occupation and industry dummy variables are not included.

Table 5. Decomposition Based on Alternative Specification

Decomposition / Specification	(1) Wellington (Blau & Beller)	(2) O'Neill & Polachek	(3) Male wage structure would apply	(4) Female wage structure would apply	(5) Non-discriminatory wage structure
Including industry dummies					
Change in Charact.	..	..	-.032	-.025	-.031
Difference in Charact.	..	..	-.023	-.015	-.026
Total	-.032	-.019	-.056	-.040	-.056
(%)	(30.8%)	(18.3%)	(53.8%)	(38.5%)	(53.8%)
Change in Coeff.	..	..	-.060	-.067	-.060
Difference in Coeff.	..	..	.012	.003	.011
Total	-.072	-.079	-.048	-.064	-.048
(%)	(69.2%)	(76.0%)	(46.2%)	(61.5%)	(46.2%)
Macpherson & Hirsch: (Including FEM)					
Change in Charact.	..	..	-.035	-.022	-.045
Difference in Charact.	..	..	-.043	-.020	-.033
Total	-.027	-.032	-.078	-.042	-.078
(%)	(25.6%)	(30.7%)	(75.2%)	(40.3%)	(75.0%)
Change in Coeff.	..	..	-.033	-.059	-.051
Difference in Coeff.	..	..	.008	-.003	.016
Total	-.077	-.072	-.026	-.062	-.036
(%)	(74.4%)	(69.4%)	(24.8%)	(59.8%)	(25.0%)
Including occupation and industry dummies					
Change in Charact.	..	..	-.046	-.023	-.053
Difference in Charact.	..	..	-.044	-.063	-.032
Total	-.031	-.036	-.090	-.086	-.085
(%)	(30.3%)	(34.8%)	(87.1%)	(83.0%)	(81.7%)
Change in Coeff.	..	..	-.025	-.008	-.037
Difference in Coeff.	..	..	.012	-.010	.020
Total	-.072	-.067	-.013	-.018	-.017
(%)	(69.7%)	(65.2%)	(12.9%)	(17.0%)	(28.3%)

The change in the wage gap, $\Delta\overline{\ln(w_m)} - \Delta\overline{\ln(w_f)}$ = -0.104 during 1983-1993. FEM is the ratio if female to total employment in a worker's occupation and industry.