Yun, Myeong-Su

Working Paper
Wage differentials discrimination and inequality: a cautionary note on the Juhn, Murphy and Pierce decomposition method

IZA Discussion Papers, No. 2937

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: Yun, Myeong-Su (2007) : Wage differentials discrimination and inequality: a cautionary note on the Juhn, Murphy and Pierce decomposition method, IZA Discussion Papers, No. 2937, Institute for the Study of Labor (IZA), Bonn

This Version is available at:
http://hdl.handle.net/10419/34739

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Wage Differentials, Discrimination and Inequality: A Cautionary Note on the Juhn, Murphy and Pierce Decomposition Method

Myeong-Su Yun

July 2007
ABSTRACT

Wage Differentials, Discrimination and Inequality: A Cautionary Note on the Juhn, Murphy and Pierce Decomposition Method

This paper shows how difficult it is to study the roles of discrimination and unobserved skills when studying changes in racial and gender wage gaps over time by examining merits and shortcomings of a popular decomposition method by Juhn, Murphy and Pierce (1991). The JMP method shows that wage dispersion can offer a compelling explanation of the wage gap. However, JMP have to rely on a few strong assumptions in order to derive their decomposition equation which introduces wage inequality as the price of unobserved skills (the standard deviation of the residuals) into their decomposition equation.

JEL Classification: J70, D30

Keywords: decomposition analysis, wage differentials, discrimination, inequality

Corresponding author:

Myeong-Su Yun
Department of Economics
Tulane University
206 Tilton Hall
New Orleans, 70118 LA
USA
E-mail: msyun@tulane.edu

* The author wishes to thank Jim Davies and Ira Gang for their comments.
1. Introduction

The Oaxaca decomposition method explains wage differentials in terms of differences in individual characteristics (characteristics effect), differences in the coefficients of wage equations (coefficients effect) and differences in residuals (residuals effect). This method has been widely used to understand racial and gender wage differentials (Oaxaca, 1973). The coefficients effects is often interpreted as a measure of discrimination.

The existence and degree of discrimination has been a controversial issue. One of the main sources of the controversy is that the wage equation cannot include all relevant variables measuring skills and individual productivity, hence observationally equivalent people based on the characteristics in the wage equation may not be truly equivalent (so-called omitted variable problem). In this case the Oaxaca decomposition would over-estimate the degree of discrimination, as the coefficients effect is now the sum of discrimination and differences in unobserved skills. The main attraction of the Juhn, Murphy and Pierce (1991, below denoted as JMP) decomposition equation for changes in wage differentials is that it seems to provide a way to show the effect of unobserved skills on racial and gender wage gap explicitly.

The JMP method explains wage differentials in terms of differences in characteristics

---

1. On the other hand, those who believe that there is prevalent discrimination and/or the magnitude of discrimination is bigger than the coefficients effect itself argue that even differences in qualifications and credentials may be the result of pre-market discrimination (so-called included variable problem).

2. Several extensions and modifications of the Oaxaca decomposition method are now in general use. For example, Juhn, Murphy and Pierce (1993) and Yun (2006) apply the decomposition methodology to the second moment of income distribution, and Even and Macpherson (1990) and Yun (2004) extend its use to nonlinear equations. However, JMP (1991) is unique, offering an alternative decomposition formula which includes wage inequality.
(predicted gap) and in terms of differences in residuals (residual gap). The residual gap is further specified in terms of the standard deviation of the residuals and standardized residuals. The standard deviation of the residuals of the wage equation is considered as both within group wage inequality and the price of unobserved skills. Using the JMP method, researchers explain changes in racial and gender wage gaps partly by changes in wage inequality (see Altonji and Blank, 1999, pp. 3226-3230 for details). In short, the JMP method has been perceived as an innovation that changed the course of studies on wage differentials and has enjoyed great popularity and acceptance (e.g., Blau and Kahn, 1997).³

Of course, the JMP method has not been immune to criticism. Suen (1997) focuses on its approach to decomposing the residuals into the standard deviation of residuals and standardized residuals since the two measures are not necessarily independent. Even if some degree of arbitrariness in defining the price of unobserved skills is unavoidable, and using the standard deviation of the residuals is acceptable as the “standard,” the question of how to properly introduce the standard deviation of the stochastic term (wage inequality) into decomposition analysis deserves close examination. Focusing on the strong assumptions required in order to introduce wage inequality into the decomposition of wage differentials, this paper examines the merits and pitfalls of the JMP method relative to the standard Oaxaca decomposition, and shows how difficult it is to identify the roles of discrimination and unobserved skills when studying changes in racial and gender wage gaps over time. This paper also provides an extension of the standard Oaxaca decomposition in order to incorporate the price of unobserved skills (standard deviation of the

³ As Welch (2000, p. 199) puts it, JMP moved the discussion of wage differentials “from first to second derivatives,” that is, from the mean to the dispersion (standard deviation) of wages.
residuals) for a given time period, as well as over time.

2. Coefficients Effect and Residual Gap

The JMP method is almost exclusively used for studying changes in wage differentials over time. The idiosyncracy of the JMP method can be, however, more easily laid out by studying the basic decomposition equation for wage differentials in a given time period; this is what we do first. The caveats to the JMP method discussed in this section have a direct bearing on interpreting and explaining changes in wage differentials over time, which we treat in the next section.

Suppose that we are interested in comparing the wages of groups A and B. We can write the regression equation,

\[ \bar{Y}_j = \bar{X}_j \beta_j + \epsilon_j, \tag{1} \]

where \( \bar{Y}_j \) is an \( N_j \times 1 \) vector of (log) wages in group \( j \) (A and B); \( \bar{X}_j \) is an \( N_j \times K \) matrix of exogenous variables; \( \beta_j \) is a \( K \times 1 \) vector of consistent estimates of the wage equation; \( \epsilon_j \) is an \( N_j \times 1 \) vector of stochastic component distributed with a mean of zero and a variance of \( \sigma_{\epsilon_j}^2 \).

Following the Oaxaca decomposition analysis, we decompose the (mean) wage differentials between groups A and B, as follows,

\[ \bar{Y}_A - \bar{Y}_B = (\bar{X}_A - \bar{X}_B) \beta_A + \bar{X}_B (\beta_A - \beta_B) + (\bar{\epsilon}_A - \bar{\epsilon}_B), \tag{2} \]

where the first, the second, and the third components represent the characteristics effect, the coefficients effect, and the residuals effect, respectively, and “over bar” represents the value of the sample average. In the Oaxaca decomposition, the coefficients effect is frequently interpreted as discrimination and the residuals effect as differences in the distribution of unobserved skills.

As in JMP, suppose we estimate using OLS. The residuals effect disappears because \( \bar{\epsilon}_j = 0 \).
Since the residual is considered to represent unobserved skills, decomposition analysis cannot address the effect of unobserved skills when OLS is used to estimate the wage equation. JMP propose an alternative decomposition equation which enables us to study the effect of unobserved skills even if the OLS is used for estimation.\(^4\)

JMP assume that the returns to individual characteristics are the same for both groups \(A\) and \(B\) (i.e., \(\beta_B = \beta_A\) ) and construct an auxiliary wage function for group \(B\),
\[
Y_B = X_B\beta_A + U_B.
\]

Using this auxiliary wage equation for the group \(B\), JMP propose the decomposition equation,
\[
\bar{Y}_A - \bar{Y}_B = (\bar{X}_A - \bar{X}_B)\beta_A - \bar{U}_B = (\bar{X}_A - \bar{X}_B)\beta_A - \bar{U}_B \bar{\sigma}_A,
\]
where \(\sigma_A\) is the standard deviation of the residual \((e_A)\), and \(\bar{U}_B = \bar{U}_B / \sigma_A\).\(^5\)

In JMP terminology, the first and second terms on the right hand side of equation (4) represent the predicted gap and the residual gap, respectively. In the JMP decomposition, the residual gap is supposed to be related to the distribution of unobserved skills. Obviously the residual gap is identical to the coefficients effect in the standard Oaxaca decomposition when OLS is used.\(^6\)

\(^4\) The ability to make statements about wage inequality when OLS is used for estimation may have led to the method’s great popularity. Virtually all papers using the JMP method employ OLS for estimating the wage equation for group \(A\). Note that the JMP method does not estimate the wage equation for group \(B\).

\(^5\) It is interesting to note that Althauser and Wigler (1972) criticize the decomposition equation used in Duncan (1968) which is identical to what JMP use, \(\bar{Y}_A - \bar{Y}_B = (\bar{X}_A - \bar{X}_B)\beta_A - \bar{U}_B\), for not recognizing the importance of discrimination.

\(^6\) The residual in the JMP method, \(U_B = Y_B - X_B\beta_A\), is identical to \(-X_B(\beta_A - \beta_B) + e_B\). When OLS can be applied to group \(B\), then \(-\bar{U}_B \bar{\sigma}_A = \bar{X}_B(\beta_A - \beta_B)\) since \(e_B = 0\). Gupta, Oaxaca and Smith (2000) suggest using the coefficients from a pooled sample to evaluate the effects on the gender wage gap of both wage dispersion and the changing relative ranking of women in the pooled wage distribution.
Their modified decomposition equation is 

\[ \bar{Y}_A - \bar{Y}_B = (\bar{X}_A - \bar{X}_B) \beta_P + (\bar{\omega}_A - \bar{\omega}_B) \sigma_P, \]  

(4')

where \( \beta_P \) is a \( K \times 1 \) vector of consistent estimates of the wage equation for the pooled sample; \( \sigma_P \) is the standard deviation of wage residuals from the pooled sample; \( \bar{\omega}_j = \bar{Y}_j - \bar{X}_j \beta_P \). Unlike the JMP method, the residual means for groups \( A \) and \( B \) will not be zero if unobserved skills are unevenly distributed between groups \( A \) and \( B \). However, this method assumes that coefficients of the two groups are identical. One might argue that the modified JMP decomposition equation (4') has the same shortcomings as the JMP decomposition equation (4). That is, \( \bar{\omega}_A \sigma_P = \bar{X}_A (\beta_A - \beta_P) \) and \( \bar{\omega}_B \sigma_P = \bar{X}_B (\beta_B - \beta_P) \), hence they represent a premium to a favored group \( A \) and disadvantage for the discriminated against group \( B \), respectively. This kind of shortcoming cannot be avoided due to \( \bar{\sigma}_j = 0 \).

The issue is how the residual gap can be related to unobserved skills since the coefficients effect is related only to the deterministic component of the regression equation. The JMP method should reconcile the fact that the stochastic term in regression analysis is supposed to reflect the unobserved skills with the fact that the residual gap is equivalent to the coefficients effect related only to the deterministic component of the regression.

The computational procedure to generate a non-zero residual gap is simple as shown above. By constructing the counterfactual wage equation without estimating it for group \( B \), JMP are able to introduce the residual gap into the decomposition equation even though average value of the stochastic components in the wage equations for group \( A \) and \( B \) are zero. The logic behind renaming the coefficients effect of the standard Oaxaca decomposition as the residual gap in the JMP method, and interpreting it as reflecting differences in unobserved skills seems to follow this reasoning: (1) Every measure in labor market is not perfect and we cannot have all the possible relevant variables. (2) Observed characteristics might be correlated with unobserved skills leading to biased estimates. (3) Therefore, differences in unobserved skills are packaged in the coefficients effect which is
showing the wage differentials due to differences in “biased” estimates.⁷

A problem with this argument is that the biased estimates are also used in characteristics effects. Hence, the characteristics effect may also capture a portion of differences in unobserved skills. Only when the estimates of group A are not biased, the characteristics effect does not capture the differences in unobserved skills. However, it is arbitrary to say that estimates of group A are unbiased while the estimates of the other group, B, are biased.

Suppose that the assumption of the JMP method is correct, that is, the OLS estimates of group A are unbiased, though the OLS estimates of group B are biased. In this case, the residual gap is, as JMP argue, sum of the coefficients effect and the residuals effect in the Oaxaca decomposition equation (2). Even when accurate, it may be more desirable to extend the Oaxaca decomposition equation (2), rather than lumping the coefficients and residuals effects into the residual gap as JMP do. In general, we can modify Oaxaca decomposition equation (2) to allow non-zero mean residuals as follows;

\[ \bar{Y}_A - \bar{Y}_B = (\bar{X}_A - \bar{X}_B)\beta_A + \bar{X}_B(\beta_A - \beta_B) + (\theta_A - \theta_B)\sigma_A + \theta_B(\sigma_A - \sigma_B), \]

where \( \sigma_j \) is standard deviation of the residuals \( (e_j) \) and \( \theta_j = e_j / \sigma_j \). To follow the JMP assumption, the (2’) can be easily revised by imposing \( \theta_A = 0 \).

In this section, we show that, based on the assumption that OLS estimates of group A are not biased while those of group B are biased, JMP revise the standard Oaxaca decomposition equation, no longer investigating bias nor separately identifying the coefficients and residuals effects. The

---

⁷ Of course, it is not JMP who first ask whether measured discrimination (coefficients effect) is just an artifact of a mis-specified model. See Ashenfelter and Oaxaca (1987) for a discussion of the court debate and the Supreme Court decision on this issue.
next section shows how JMP use their decomposition for wage differentials at a given time to study changes in wage differentials over time.

3. Decomposing Changes in Wage Differentials over Time

To study changes in wage differentials over time, the Oaxaca decomposition (equation 2 or 2') can be applied twice (Smith and Welch, 1989). The (double) decomposition equation for studying the changes in wage differentials based on equation (2') is,

\[
D_{t' t} - D_t = (\Delta X_{t'} - \Delta X_t) \beta_{A t} + \Delta X_{t'} (\beta_{A t'} - \beta_{A t}) \\
+ (\bar{X}_{B t'} - \bar{X}_{B t}) \Delta \beta_A + \bar{X}_{B t'} (\Delta \beta_{A t'} - \Delta \beta_A) \\
+ (\Delta \theta_{t'} - \Delta \theta_t) \sigma_{A t} + \Delta \theta_{t'} (\sigma_{A t'} - \sigma_{A t}) \\
+ (\bar{\theta}_{B t'} - \bar{\theta}_{B t}) \Delta \sigma_A + \bar{\theta}_{B t'} (\Delta \sigma_{A t'} - \Delta \sigma_A),
\]

where \(D_h = \bar{Y}_{A h} - \bar{Y}_{B h}, \Delta X_h = \bar{X}_{A h} - \bar{X}_{B h}, \Delta \beta_h = \beta_{A h} - \beta_{B h}, \Delta \theta_h = \bar{\theta}_{A h} - \bar{\theta}_{B h}\) and \(\Delta \sigma_h = \sigma_{A h} - \sigma_{B h}\) for given time period \(h\) (\(t'\) and \(t\)). When equation (2) is used instead of (2'), the last two lines of equation (5) reduce to \(\Delta \epsilon_{t'} - \Delta \epsilon_t\), where \(\Delta \epsilon_h = \bar{\epsilon}_{A h} - \bar{\epsilon}_{B h}\) for given time period \(h\) (\(t'\) and \(t\)).

JMP devise another (double) decomposition equation, based on their decomposition equation for a single period, equation (4). The changes in wage differentials can be expressed as,

\[
D_{t' t} - D_t = (\Delta X_{t'} - \Delta X_t) \beta_{A t} + \Delta X_{t'} (\beta_{A t'} - \beta_{A t}) \\
+ (\Delta \theta_{t'}^* - \Delta \theta_t^*) \sigma_{A t} + \Delta \theta_{t'}^* (\sigma_{A t'} - \sigma_{A t}),
\]

where \(\Delta \theta_h^* = \bar{\theta}_{A h} - \bar{\theta}_{B h}^*\) for given time period \(h\) (\(t'\) and \(t\)), and when OLS is employed as is the usual practice with the JMP method, \(\Delta \theta_h^* = \bar{\theta}_{A h} - \bar{\theta}_{B h}^* = - \bar{\theta}_{B h}^*\) since \(\bar{\theta}_{A h} = 0\).
As the second line in equation (6) shows, the JMP method considers changes in the residual gap, in terms of changes in prices and quantities, as sources of changes in wage differentials. It is obvious that the last line in equation (6) is equivalent to last three lines in the standard Oaxaca (double) decomposition equation (5); that is, changes in the residual gap are equivalent to changes in the coefficients and the residuals effects over time. As shown in the previous section, this equation is based on the assumption that OLS estimates for group $A$ are unbiased while those for group $B$ are biased in both periods, $t'$ and $t$. If discrimination (coefficients effect) remains stable over time as JMP speculate, that is, the second line of the equation (5) is zero, then changes in the residual gap can be attributed to changes in the price and quantities of unobserved skills. This speculation leads to a convenient decomposition that the first line of (6) is related to changes in observed skills while the second line is related to changes in unobserved skills.

One may argue that wage differentials can be decomposed into different components in numerous ways as long as the decomposition provides insights on the causes of the differentials. It is laudable to aim at explaining changes wage differentials in terms of changes in the prices of unobserved skills, and the JMP decomposition equation is ingenious. However, it should not be overlooked that the JMP method needs quite strong assumptions to derive equation (6). These need justification and close inspection. It is desirable to rigorously study data and estimation issues, and to distinguish wage differentials due to differences in returns to observed characteristics from those due to differences in prices of unobserved skills, as in the equation (5).

---

8 If OLS is used for estimating regression equations for groups $A$ and $B$, then last two lines of the equation (5) disappear, and the last line of the equation (6), i.e., changes in the residual gap, is exactly the second line of the equation (5), i.e., changes in the coefficients effect.
Furthermore, one should not overlook a caveat in interpreting decomposition equation (6), and even (5), related to causation between changes in wage differentials and changes in the price of unobserved skills measured by wage inequality (standard deviation of the residuals). From the standard variance decomposition with within and between-group inequality, we know that increases in wage differentials raise wage inequality, ceteris paribus. However, the reverse is not always true. The increase of within-group inequality ($\sigma_4$ in the JMP method) does not necessarily increase wage differentials. That is, there is no causal relationship between wage differentials and within group inequality.

Suppose within group wage variation arises from only stochastic components, i.e., the deterministic component is identical within each group, and is distributed as follows:

<table>
<thead>
<tr>
<th>Scenario</th>
<th>wage group $A$</th>
<th>wage group $B$</th>
<th>gap</th>
<th>$\sigma_A$</th>
<th>$\sigma_B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>2, 4</td>
<td>1, 3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>II</td>
<td>1, 4</td>
<td>1, 4</td>
<td>0</td>
<td>1.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Note: The numbers in the two wage columns represent the wages earned by the two members of each group. Gap, $\sigma_A$ and $\sigma_B$, mean, respectively, wage differentials, and standard deviation of wages in groups $A$ and $B$.

As the example shows, the increase in within group inequality ($\sigma_A$ and $\sigma_B$) from scenario I to II

---

9 Some may argue that this is not an important issue since the wage inequality, the standard deviation of residuals, is introduced as the price of unobserved skills. However, it is worth noting that the fact that wage inequality is introduced into the study of wage differentials while there has been substantial increases in wage inequality during 1980s has contributed to making the JMP method popular. JMP (1993) explain that residual (or within-group) inequality accounts for most of the growth in overall wage inequality. Combined with increasing returns to human capital variables (typically schooling and labor market experience), JMP argue that the increase in wage inequality are due to skill-biased technological change. For criticism of this argument, see Card and DiNardo (2002) and Lemieux (2006).
However, their explanation of the matching procedure can help in understanding the meaning of the components in equations (5) and (6). This caveat should be borne in mind when the JMP method is applied.

Finally, one interesting computation issue is that there is no reason to compute the second line in the equation (6) using a cumbersome matching process over time (for details of matching, see JMP, 1991; Altonji and Blank, 1999). Unlike the matching process using a cumulative distribution function across years as explained in JMP (1991, pp. 126-7), the last line in equation (6) can be obtained from a simple manipulation because the decomposition equation is based on an identity; given values of $D_h$, $\Delta X_h$, $\beta_{Ah}$, and $\sigma_{Ah}$ in equation (5) or (6), the matching process cannot change the values of remaining component, $\Delta \theta_h$ in equation (5) or $\Delta \theta_h^*$ in equation (6).\(^\text{10}\)

4. Conclusion

Ashenfelter and Oaxaca (1987) proclaim that the Oaxaca decomposition based on Gary Becker’s (1971) discrimination framework has become the standard form by which the litigation of disputes over allegations of race and sex discrimination proceed, in spite of drawbacks springing from using regression analysis. On the other hand, the method proposed in JMP (1991) pays attention to the importance of differences in unobserved skills for studying wage differentials. The role of discrimination and unobserved skills in wage differentials have been the subjects of continuous debates (see e.g. Heckman, 1998 and Riach and Rich, 2002).

JMP devise an alternative decomposition equation in which wage differentials are related to within group inequality or the price of unobserved skills. Because of the ingenuity of the JMP

\(^{10}\) However, their explanation of the matching procedure can help in understanding the meaning of the components in equations (5) and (6).
decomposition equation, researchers are able to explain the changes in wage differentials with changes in the price of unobserved skills, even when OLS is used for estimation. This may be the source of the great success of the JMP method. Researchers may choose between the standard Oaxaca decomposition and JMP decomposition depending on their emphasis on discrimination or the price of unobserved skills. It should be noted that, in order to decompose changes in wage gap over time into changes in the predicted gap and changes in unobserved skills, JMP have to rely on two strong assumptions, first, OLS estimates of one group are not biased while those of the other group are biased, and second, discrimination is stable over time. These assumptions are difficult to verify. The strength of the assumptions needed to construct the JMP decomposition equation must give us pause.
References


