Wage discrimination: The case for reverse regression

Kapsalis, Constantine

Data Probe Economic Consulting

8 December 2010

DATA PROBE ECONOMIC CONSULTING INC.

9 Maki Place, Ottawa, Canada K2H 9R5
www.dataprobeinc.ca *dataprobeinc@hotmail.com*

WAGE DISCRIMINATION: THE CASE FOR REVERSE REGRESSION

Constantine Kapsalis, Ph.D.

Working Papers Series: No. 2010-2

December 8, 2010

ABSTRACT
The reverse regression method of measuring wage discrimination is the main challenge to the dominant direct regression method based on the Oaxaca/Blinder approach. In this article, it is argued that the choice between the two methods is fundamentally a choice of assumptions regarding the nature of the wage determination process and the nature of the unexplained regression residual of the wage regression equation. In particular, this article concludes that the reverse regression method is more likely to produce the correct wage discrimination measure if *any* of the following three assumptions is correct:
(a) qualifications do not determine how much individuals earn (as the direct regression method assumes) but, instead, determine which candidates are selected for existing jobs with fixed wages; (b) errors in the measurement of qualifications are larger than errors in the measurement of wages, in which case the direct regression method would understate the importance of differences in qualifications; and (c) differences in unobserved qualifications (e.g., importance of job flexibility; relevance of past work experience) between two groups are not zero (as the direct regression method assumes) but tend to favour the group with the better observed qualifications. Finally, this article shows that application of the reverse regression technique simply requires the augmentation of the qualification component of the direct regression method by dividing it by the R^2 coefficient.

The author would appreciate receiving comments. Please email your comments to: kapsalis@sympatico.ca

I. Introduction

Wage discrimination is an important public policy issue, both from the equity and the economic efficiency point of view. From the equity point of view, wage discrimination unfairly undervalues the work of part of the labour force. From the economic efficiency point of view, wage discrimination and its associated job discrimination lead to an under-utilization of available human resources.

The topic of wage discrimination has attracted considerable attention in the economic literature.[1] Much of the focus in the literature has been on the decomposition of wage differences between various groups (e.g., men and women, visible minorities and the rest of the work force, immigrants and non-immigrants) into a component due to differences in qualifications (e.g., differences in education and past work experience) and a residual component that is typically attributed to discrimination. The focus of this article is on gender discrimination, although the conclusions are applicable to any comparison groups.

There are two main discrimination concepts in the literature. They are referred to here as *within-qualifications discrimination* and *within-wages discrimination*:

a) <u>Within-qualifications discrimination</u>: According to this concept, which is the dominant one in the economic literature, women are discriminated against if they earn less than men with similar qualifications; and
b) <u>Within-wages discrimination</u>: According to this concept, women are discriminated against if they have higher qualifications than men with similar wages.

As it will be explained below, the first discrimination concept requires a direct regression of male wages as a function of male qualifications, while the second discrimination concept requires a reverse regression of male qualifications as a function of male wages.

In what follows, Section II describes the use of direct regression for estimating within-qualifications discrimination, while Section III describes the use of reverse regression for estimating within-wages discrimination. Section IV compares the two methods and makes a case in favour of the reverse regression technique. Section V summarizes the main conclusion of this article. Appendix A provides a mathematical proof of the relation between the two concepts of wage discrimination.

II. Within-Qualifications Discrimination

Most wage discrimination studies are based on the concept of within-qualifications discrimination – i.e., wage discrimination exists if men and women with similar qualifications receive different wages. The estimation method associated with this concept of discrimination, originally developed by Oaxaca (1973) and Blinder (1973), involves the estimation of a direct regression of male wages as a function of male qualifications. The male wage regression is used to estimate what the average expected female wage would have been if female qualifications were rewarded in the same way as male qualifications were. Finally, the gap between the average actual male wage rate and the average expected female wage rate is attributed to differences in qualifications, while

---

[1] The literature is too vast to review here. For a recent survey see Weichselbaumer and Winter-Ebmer, 2005.

the gap between the overall male-female wage difference and the difference in qualifications is attributed to discrimination.[2]

The male wage regression is specified as follows:

(1) $$Y_m = a_m + b_m X_m + u_m$$

where subscript *m* refers to male employees; *Y* is the wage rate (typically defined in a natural log form); *a* is the constant term; *b* is the vector of regression coefficients; *X* is the matrix of explanatory variables; and *u* is the error term, assumed to satisfy the standard assumptions for a best, linear, unbiased estimator.

The difference in qualifications between men and women is estimated by summing up the products of the differences in qualifications, times the respective b_m coefficients. Thus, the qualification component is given by the following equation:

(2) $$b_m(\overline{X}_m - \overline{X}_f)$$

where vectors $\overline{X}_m$ and $\overline{X}_f$ represent respectively the average value of the male and female qualifications. Finally, the remaining wage differential is attributed to wage discrimination:

(3) $$(\overline{Y}_m - \overline{Y}_f) - b_m(\overline{X}_m - \overline{X}_f)$$

III. Within-Wages Discrimination

The concept of within-wages discrimination – originally proposed by Kapsalis (1979, 1982), Dempsters (1979), and Roberts (1980) – is the main challenge to the conventional approach of measuring wage discrimination. The estimation method associated with this concept of discrimination involves the estimation of a reverse regression of male qualifications as a function of male wages. The reverse regression has the same specification as the direct regression, except that the independent variables (e.g., years of education) become the dependent variables and the dependent variable (wage rate) becomes the independent variable.

For the sake of simplicity, the discussion here assumes that there is only one relevant qualification (years of education), in which case reversing the two variables is straightforward. In the more realistic case of several qualifications, one of the approaches that have been employed in the literature is to use the direct regression results to create an index of qualification and then use that index as the single independent variable. However, as it will be shown later, this step is not necessary since *the reverse regression results can be easily obtained from the direct regression results by simply dividing the qualifications component by the R^2 coefficient*.

---

[2] Of course, one could also use a female wage regression instead of a male wage regression or, as suggested by some, produce two estimates of discrimination and, using a male and a female wage regression, "establish a range of possible values" (Oaxaca, 1973, p. 697). This discussion relates to the familiar index problem, and the three approaches can be viewed as corresponding to the Laspeyres, Paasche and Fisher indices. This point applies equally to both concepts of discrimination.

The primary rationale for estimating a reverse regression of male years of education as a function of wages is to estimate the expected average male years of education that corresponds to a wage level that is equal to the average female wage level. According to the within-wages discrimination concept, if women are discriminated, then within the same wage bracket the expected average male years of education should be lower than the actual average female years of education. The final step is converting the female-male educational gap into a wage gap, which represents the within-wages discrimination.
In the case of a single qualification (years of education) the male qualifications regression is specified as follows:

(4) $$X_m = a'_m + b'_m Y_m + u'_m$$

where *X* is now the years of education and the rest of the symbols have a similar interpretation to the corresponding symbols in equation 1.

Consequently, the expected average male years of education at a wage level equal to the actual average female wage rate is as follows:

(5) $$(X_m \mid Y_m = \overline{Y}_f) = a'_m + b'_m \overline{Y}_f$$

Based on the above equation, the difference in years of education at a level of wages equal to the average female wage rate is:

(6) $$\overline{X}_f - (a'_m + b'_m \overline{Y}_f)$$

Equation 6 provides a measure of the extent of discrimination expressed in years of education. This estimate can be translated into wage discrimination by multiplying it by the implicit estimate of the effect of education on wages that can be derived from equation 4:

(7) $$\Delta Y_m / \Delta X_m = 1/b'_m$$

Consequently, the extent of discrimination according to the within-wages discrimination concept is:

(8) $$(\overline{X}_f - (a'_m + b'_m \overline{Y}_f))(1/b'_m)$$

By manipulating equation 8 it can be converted as follows (see Appendix A for a detailed proof):

(9) $$(\overline{X}_f - (a'_m + b'_m \overline{Y}_f))(1/b'_m) = (\overline{Y}_m - \overline{Y}_f) - (b_m / R^2)(\overline{X}_m - \overline{X}_f)$$

where the left part of the equation represents the within-wages discrimination (same as equation 8); and the right part of the equation represents the within-qualifications discrimination (same as equation 3, except the *b* coefficients are divided by the R^2 coefficient).

Equation 9 applies equally to the general case of more than one qualification, except that the various qualifications are replaced with a single index of qualifications (e.g., weighting the various qualifications by the *b* coefficients of the male direct wage regression).

Although it has not attracted much attention in the literature, equation 9 is particularly useful because it provides a *simple way of estimating the extent of within-wages* *discrimination by simply dividing the qualification component of the within-qualifications discrimination by the R^2 coefficient*.

Equation 9 also shows that *when the R^2 coefficient of the direct male wage regression equals one, then the two methods lead to identical results*. This observation makes it clear that at the heart of the difference between the two methods is the difference in the nature of the assumptions about the unexplained regression residual.

IV. Assessment of the Two Methods

This article argues that if *any* of the following assumptions are valid, then reverse regression is preferable over the direct regression method.

A. The Nature of Wage Determination

By regressing wages as a function of qualifications, the direct regression method assumes that employers pay employees according to their qualifications. For example, a department store hiring sales persons, a school hiring teachers, or a high tech company hiring programmers will pay more to those with better qualifications.

However, in most situations a reverse direction of causality may be more realistic. In particular, according to the alternative process "an employer reviews a group of candidates for a particular salary level and selects the candidate who appears to have the best qualifications for that salary. In this case, the salary level is fixed, while the qualification level varies across the eligible candidates" (White and Piette, 1998, pp. 127). Consequently, according to this process, it is the nature of jobs and their pay structure that determines the likely qualifications of the successful applicants, rather than the other way around.

In reality, both processes are likely taking place at the same time – i.e., applicants with better qualifications are more likely to be hired while, at the same time, those with better qualifications are also more likely to get promoted and/or receive pay increases. However, to the extent that the alternative process is more prevalent, a reverse regression (of qualifications as a function of wages) may be more appropriate than a direct regression (of wages as a function of qualifications).

B. Errors in Independent Variables

The conventional argument for reverse regression is the presence of errors in independent variables (Goldberger, 1984). The argument is that qualifications are more likely to be subject to a measurement error than wages– either because key qualifications are often excluded (e.g., field of study, importance of job flexibility) or because they are represented by crude proxies (e.g., years in the labour force being used as a proxy for years of relevant experience). This problem is referred to in the literature as "errors in independent variables" and its consequence is biasing the b coefficients toward zero (and, as a result, understating the importance in differences in qualifications between the two genders).

C. Nature of the Unexplained Residual

An important limitation of the direct regression method that has long been recognized is that "the job qualifications actually available typically comprise a very incomplete listing of pertinent qualifications for any job" (Conway and Roberts, 1983, p. 75). For example, typically no account is taken of male-female differences with respect to the relative importance of wages and job flexibility. However, there is evidence that "men and women value job flexibility differently" (Bender, Donohue, and Heywood, 2005, p. 479). This difference may be one of the possible explanations for the apparent paradox that, despite their lower wages, women tend to report higher job satisfaction than men.

Let's assume that the "true" wage regression consists of two sets of qualification variables: observed (*X*) and unobserved (*Z*). Moreover, let's assume that if both sets of qualifications were known to us and were included in the direct male wage regression, then the R^2 coefficient would have been 1. In this case, the qualifications component of the direct regression method can be written as follows:

(10) $$b_m(\overline{X}_m - \overline{X}_f) + k_m(\overline{Z}_m - \overline{Z}_f)$$

As demonstrated earlier, the reverse regression method is equivalent to the direct regression method, except that the qualifications component is augmented by dividing it by the R^2 coefficient. One interpretation of this result is that the reverse regression method implicitly assumes that the qualifications component with respect to unobserved variables is not zero, but "proportional" to the observed endowment – i.e.,

(12) $$k_m(\overline{Z}_m - \overline{Z}_f) / b_m(\overline{X}_m - \overline{X}_f) = (1 - R^2) / R^2$$

Although there is no way of knowing how omitted variables compare between the two genders, the assumption of "proportionality" is as reasonable as the assumption of no difference.

V. Conclusion

This article has provided evidence that the choice between the direct regression method and reverse regression method is fundamentally a choice of assumptions regarding the nature of the wage determination process and the nature of the unexplained regression residual of the wage regression equation. In particular, this article has concluded that the reverse regression method is more likely to produce the correct wage discrimination measure if *any* of the following three assumptions is correct:

(a) Qualifications, for the most part, do not determine how much individuals earn (as the direct regression method assumes) but, instead, determine who is selected for existing jobs with fixed wages.

(b) Errors in the measurement of qualifications are likely to be larger than errors in the measurement of wages and, as a result, the direction regression method likely understates the importance of differences in qualifications.

(c) Differences in omitted qualifications (e.g., job flexibility, relevance of past work experience) between two groups are not necessarily zero (as the direct regression method assumes) but they may very well be "proportional" to the included qualifications.

At the minimum, the reverse regression of wage discrimination should be used to establish a range of the likely magnitude of wage discrimination. The fact that the reverse regression estimate can be easily calculated by dividing the qualification component by R^2 coefficient makes it very simple to complement the traditional direct regression method estimates with alternative estimates based on the reverse regression method.

Appendix A. Mathematical Relation between the Two Discrimination Concepts
Equation A1 represents the male reverse regression of years of education (*X*) as a function of the male wage rate (*Y*). Using equation A1, the average male education is expressed as a function of the average wage rate (equation A2). Then, through a series of manipulations we end up with equation A6, which is the same as equation 9 in the main body of the article. Equation A6 shows that the within-wages discrimination (left side) equals the within-qualifications discrimination with the *b* coefficient adjusted by dividing it by the R^2.

(A1) $X_m = a'_m + b'_m Y_m + u'_m$

(A2) $\overline{X}_m = a'_m + b'_m \overline{Y}_m$

(A3) $\overline{Y}_m = -a'_m / b'_m + 1/b'_m \overline{X}_m$

(A4) $\overline{Y}_m - \overline{X}_m / b'_m + a'_m / b'_m = 0$

(A5) $\overline{Y}_m - \overline{Y}_f + \overline{Y}_f - \overline{X}_m / b'_m + \overline{X}_f / b'_m - \overline{X}_f / b'_m + a'_m / b'_m = 0$

By re-arranging terms, equation *A5* can be written as follows:

(A6) $(\overline{X}_f - (a'_m + b'_m \overline{Y}_f))(1/b'_m) = (\overline{Y}_m - \overline{Y}_f) - (1/b'_m)(\overline{X}_m - \overline{X}_f)$

And, since the *b* coefficients of the direct and indirect regressions are related through the relationship $b_m b'_m = R^2$ (where R^2 is common to both regressions), the term $1/b'_m$ in the right side of equation *A6* can be replaced by b_m / R^2 and, as a result, equation *A6* can be written as follows (which is the same as equation 9 in the main body of the article):

(A7) $(\overline{X}_f - (a'_m + b'_m \overline{Y}_f))(1/b'_m) = (\overline{Y}_m - \overline{Y}_f) - (b_m / R^2)(\overline{X}_m - \overline{X}_f)$

Equation *A7* shows that the within-wages measure of wage discrimination (left side of the equation) equals the within-qualifications wage discrimination (right side of the equation), except that the b_m coefficient is divided by the R^2 coefficient.

References

Bender, A. K, S.M. Donohue, and J.S. Heywood, 2005. "Job satisfaction and gender segregation." *Oxford Economic Papers* 57: 479–496.

Blinder, A., 1973. "Wage discrimination: Reduced form and structural estimates." *Journal of Human Resources* 8 (Fall): 436-455.

Conway, D.A. and H.V. Roberts, 1983. "Reverse regression, fairness, and employment discrimination." *Journal of Business and Economic Statistics*, January: 75-85.

Dempsters, A.P., 1979. *Statistical concepts of discrimination*. Statistics Department, Harvard University, Cambridge, MA.

Goldberger, A.S., 1984. "Reverse regression and salary discrimination." *Journal of Human Resources* XIX, 3.

Kapsalis, C., 1979. *A new method of estimating wage discrimination between men and women.* Economic Council of Canada.

Kapsalis, C., 1982. "A new measure of wage discrimination." *Economics Letters* 9: 287-293.

Oaxaca, R., 1973. "Male-female wage differential in urban labour markets." *International Economic Review* 14 (3): 693-705.

Roberts, H.V., 1980. "Statistical biases in the measurement of employment discrimination." In E.R. Livernash, ed., *Comparable worth: Issues and alternatives.* Equal Employment Advisory Council, Washington, DC.

Weichselbaumer, D. and R. Winter-Ebmer, 2005. "A meta-analysis of the international gender wage gap." *Journal of Economic Surveys* 9 (3): 479-511.

White, F.P. and M.J. Piette, 1998. "The use of 'reverse regression' in employment discrimination analysis." *Journal of Forensic Economics* 11(2): 127-138.