

IZA DP No. 877

Decomposing Differences in the First Moment

Myeong-Su Yun

September 2003

Decomposing Differences in the First Moment

Myeong-Su Yun

Tulane University and IZA Bonn

Discussion Paper No. 877
September 2003

IZA

P.O. Box 7240
D-53072 Bonn
Germany

Tel.: +49-228-3894-0
Fax: +49-228-3894-210
Email: iza@iza.org

This Discussion Paper is issued within the framework of IZA's research area *Welfare State and Labor Market*. Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent, nonprofit limited liability company (Gesellschaft mit beschränkter Haftung) supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public. The current research program deals with (1) mobility and flexibility of labor, (2) internationalization of labor markets, (3) welfare state and labor market, (4) labor markets in transition countries, (5) the future of labor, (6) evaluation of labor market policies and projects and (7) general labor economics.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available on the IZA website (www.iza.org) or directly from the author.

ABSTRACT

Decomposing Differences in the First Moment*

We propose a simple methodology for decomposing differences in the first moment into characteristics and coefficients effects. This methodology provides a way to apply the Blinder-Oaxaca decomposition to a non-linear function for both aggregate and detailed decompositions.

JEL Classification: C20, J70

Keywords: decomposition analysis, characteristics effect, coefficients effect

Myeong-Su Yun
Department of Economics
Tulane University
206 Tilton Hall
New Orleans, LA 70118
USA
Tel.: +1 504 862 8352
Fax: +1 504 865 5869
Email: msyun@tulane.edu

* The author thanks Ira Gang and an anonymous referee for comments.

I. Introduction

The Blinder-Oaxaca decomposition of wage differentials has played a significant role in studying wage discrimination (Blinder (1973) and Oaxaca (1973)). The basic idea of the Blinder-Oaxaca decomposition is that differences in wages can be explained by the differences in characteristics and by the differences in OLS coefficients. The application potential of this decomposition analysis is not restricted to studying continuous variables such as wage differentials. The variable whose differences in the first moment we are interested in can be either continuous or discrete.

The simple idea of the Blinder-Oaxaca decomposition has been applied whenever we need explain the differences between two comparison groups. To do this, researchers usually construct an auxiliary equation based on one group's characteristics and the estimated coefficients of the other group's equation. The calculation of the overall or aggregate "characteristics" and "coefficients" effects is straightforward and now standard. The contribution of the differences in characteristics and coefficients of individual variables (i.e., the detailed decomposition) can be easily found when linear equations are used (e.g., wage decomposition using Blinder-Oaxaca decomposition), but not when non-linear equations are used. Detailed decomposition is useful and instructive since it can answer questions like "How much of the differences in labor market participation rates between two comparison groups can be explained by differences in educational attainment (characteristics effect) and how much by differences in behavioral response associated with educational attainment (coefficients effect) ?".

This paper proposes a general and systematic methodology for the detailed decomposition

of differences in the first moment, i.e., differences in the mean value of the variable of interest.¹ The proposed detailed decomposition methodology does not depend on the functional form as long as the dependent variable is a function of a linear combination of independent variables and the function is once differentiable. Therefore, the standard Blinder-Oaxaca decomposition methodology is included as a special case of the methodology proposed in this paper.

II. Decomposition Methodology

Suppose that a dependent variable is a function of a linear combination of independent variables, though the function (F) itself may or may not be linear. That is,

$$Y = F(X\beta),$$

where Y , X , and β are, respectively, an $N \times 1$ vector, an $N \times K$ matrix of independent variables, and a $K \times 1$ vector of coefficients; F is a mapping of a linear combination of X ($X\beta$) to Y , and the function F is any once differentiable function.

The difference in Y at the first moment, i.e., the mean difference of Y between groups A and B can be decomposed as

$$\bar{Y}_A - \bar{Y}_B = [\overline{F(X_A\beta_A)} - \overline{F(X_B\beta_A)}] + [\overline{F(X_B\beta_A)} - \overline{F(X_B\beta_B)}], \quad (1)$$

¹ Indeed, the idea of the Blinder-Oaxaca decomposition method has also been applied to differences related to the second moment of a variable in the inequality literature (e.g., Juhn, Murphy and Pierce (1993) and Yun (2002)). Yun (2002) shows a way to decompose differences in variance at both aggregate and individual variable levels. However, it seems that a general and systematic decomposition methodology for differences related to the second moment has not yet been developed.

where “over bar” represents the value of the sample’s average.²

The above decomposition is done at the aggregate or overall level; this is widely accepted as a way to decompose the differences in the first moment in terms of differences in characteristics $[\overline{F(X_A\beta_A)} - \overline{F(X_B\beta_A)}]$ and in terms of differences in coefficients $[\overline{F(X_B\beta_A)} - \overline{F(X_B\beta_B)}]$. The next step is to find the contribution of each variable to the total difference (detailed decomposition). While this has been studied in a linear regression framework, only a few attempts have been made to examine the detailed decomposition beyond using linear regression.³ The key question is how to properly weight the contribution of each variable to the characteristics and coefficients effects. In order to obtain a proper weight, we use two types of approximation; first, we evaluate the value of the function using mean characteristics;⁴ second, we use a first order Taylor expansion to linearize the characteristics and coefficients effects around $\bar{X}_A\beta_A$ and $\bar{X}_B\beta_B$, respectively.

² The decomposition equation (1) is not unique. The discussion below can be easily applied to the other decompositions with alternative parameterizations.

³ If the function F is a linear regression function, then obtaining the detailed decomposition equation is trivial, and is noted in the example below. There are a few papers on decomposing probit or logit functions (e.g., Even and Macpherson (1990), Nielsen (1998) and Yun (2000)). This paper provides a systematic and general decomposition methodology.

⁴ It is not necessarily required to use the mean characteristics to derive the final weights, $W_{\Delta X}^i$ and $W_{\Delta\beta}^i$. That is, it is possible to take a Taylor expansion of the characteristics and coefficients effects in the equation (1) without the first approximation with mean characteristics. Nonetheless we keep the approximation using mean characteristics since it makes the derivation of $W_{\Delta X}^i$ and $W_{\Delta\beta}^i$ easy to explain. The author thanks an anonymous referee for pointing this out.

$$\begin{aligned}
\bar{Y}_A - \bar{Y}_B &= [F(\bar{X}_A \beta_A) - F(\bar{X}_B \beta_A)] + [F(\bar{X}_B \beta_A) - F(\bar{X}_B \beta_B)] + R_M \\
&= [(\bar{X}_A - \bar{X}_B) \beta_A] f(\bar{X}_A \beta_A) + \bar{X}_B (\beta_A - \beta_B) f(\bar{X}_B \beta_B) + R_M + R_T,
\end{aligned} \tag{2}$$

where

$$\begin{aligned}
R_M &= [\overline{F(X_A \beta_A)} - \overline{F(X_B \beta_A)}] + [\overline{F(X_B \beta_A)} - \overline{F(X_B \beta_B)}] \\
&\quad - [F(\bar{X}_A \beta_A) - F(\bar{X}_B \beta_A)] - [F(\bar{X}_B \beta_A) - F(\bar{X}_B \beta_B)],
\end{aligned}$$

$$\begin{aligned}
R_T &= [F(\bar{X}_A \beta_A) - F(\bar{X}_B \beta_A)] + [F(\bar{X}_B \beta_A) - F(\bar{X}_B \beta_B)] \\
&\quad - [(\bar{X}_A - \bar{X}_B) \beta_A] f(\bar{X}_A \beta_A) - \bar{X}_B (\beta_A - \beta_B) f(\bar{X}_B \beta_B),
\end{aligned}$$

and $f(\bar{X}_j \beta_j) = \frac{dF(\bar{X}_j \beta_j)}{d(\bar{X}_j \beta_j)}$, i.e., $f(\cdot)$ is the first order derivative of the function $F(\cdot)$, where $j = A$

and B . All three, $(R_M, R_T$ and $f(\bar{X}_j \beta_j)$), are scalars. R_M and R_T are approximation residuals resulting from evaluating the function $F(\cdot)$ at the mean values and by using the first order Taylor expansion, respectively.

A detailed decomposition equation of (1) may be written using the information in (2) as follows:

$$\bar{Y}_A - \bar{Y}_B = \sum_{i=1}^{i=K} W_{\Delta X}^i [\overline{F(X_A \beta_A)} - \overline{F(X_B \beta_A)}] + \sum_{i=1}^{i=K} W_{\Delta \beta}^i [\overline{F(X_B \beta_A)} - \overline{F(X_B \beta_B)}], \tag{3}$$

where

$$W_{\Delta X}^i = \frac{(\bar{X}_A^i - \bar{X}_B^i) \beta_A^i f(\bar{X}_A \beta_A)}{(\bar{X}_A - \bar{X}_B) \beta_A f(\bar{X}_A \beta_A)} = \frac{(\bar{X}_A^i - \bar{X}_B^i) \beta_A^i}{(\bar{X}_A - \bar{X}_B) \beta_A},$$

$$W_{\Delta\beta}^i = \frac{\bar{X}_B^i (\beta_A^i - \beta_B^i) f(\bar{X}_B \beta_B)}{\bar{X}_B (\beta_A - \beta_B) f(\bar{X}_B \beta_B)} = \frac{\bar{X}_B^i (\beta_A^i - \beta_B^i)}{\bar{X}_B (\beta_A - \beta_B)}, \text{ and } \sum_{i=1}^{i=K} W_{\Delta X}^i = \sum_{i=1}^{i=K} W_{\Delta\beta}^i = 1.$$

As equation (3) shows, it turns out that the weight is very simple and easy to implement. As long as the coefficient estimates are available, the weight can be easily calculated using the mean values of characteristics and their coefficients.

II.1 Two Examples

As the function F is required to be only once differentiable, the above detailed decomposition equation (3) may be considered as a general decomposition equation for differences in the first moment. In this section, we provide two examples; the standard Blinder-Oaxaca decomposition equation when the function F is a linear regression equation, and the decomposition of a binary choice variable when the function F is a probit equation.

Example A. Blinder-Oaxaca Decomposition

It is simple to show that equation (3) reduces to the standard Blinder-Oaxaca decomposition equation when $F(X\beta) = X\beta + e$, where e is a stochastic term with mean of zero ($\bar{e} = 0$).

Obviously, $\overline{F(X\beta)} = \bar{X}\beta$, and decomposition equation (3) reduces to

$$\bar{Y}_A - \bar{Y}_B = (\bar{X}_A - \bar{X}_B)\beta_A + \bar{X}_B(\beta_A - \beta_B),$$

which is identical to the equation (2) since $f(\bar{X}\beta) = 1$ and the two residual terms (R_M and R_T) are zero.

Example B. Probit Decomposition

Suppose that Y is a binary choice variable, and $\text{prob}(Y = 1) = M(X\beta)$, where M is a standard normal cumulative distribution function. Then decomposition equation (3) equals

$$\bar{Y}_A - \bar{Y}_B = \sum_{i=1}^{i=K} W_{\Delta X}^i [(\overline{\Phi(X_A \beta_A)} - \overline{\Phi(X_B \beta_A)})] + \sum_{i=1}^{i=K} W_{\Delta \beta}^i [\overline{\Phi(X_B \beta_A)} - \overline{\Phi(X_B \beta_B)}] .$$

This probit decomposition has been used in various studies, e.g., labor market participation (Yun (2000)), anti-foreigner attitudes (Gang, Rivera-Batiz, and Yun (2002)), and poverty rates (head-count ratio, Gang, Sen and Yun (2002)).

II.2. Comparison to Previous Methodologies

Contributions of a single variable can be calculated by replacing the value of group A with that of group B sequentially one by one. Using this sequential replacement, the characteristics and coefficients effects of individual variable can be calculated no matter what kind of functional form is employed. However, this sequential replacement method is not only tedious but also problematic since it may be sensitive to the order of switching (see Ham, Svejnar and Terrell (1998), p. 1137 for a discussion of path-dependency). The detailed decomposition proposed in this paper is free from path-dependency.

Two strands of papers offer related methodologies. One studies only the characteristics effect in a detailed decomposition using a probit equation while the other studies only the coefficients effect in a detailed decomposition using a logit equation. Both methodologies find a weight similar to the weight proposed in this paper. Obviously, these methodologies do not suffer from the path

dependency problem. Even and Macpherson (1990, 1993) suggest a detailed decomposition of the characteristics effect, but not of the coefficients effect, when a probit equation is used to study unionization. Interestingly, their suggested weight in the detailed decomposition for the characteristics effect is identical to what we have derived in this paper, $W_{\Delta X}^i = \frac{(\bar{X}_A^i - \bar{X}_B^i) \beta_A^i}{(\bar{X}_A - \bar{X}_B) \beta_A}$. It

is not shown in their paper how they derive the weight for each variable. On the other hand, Nielsen (1998) suggests a detailed decomposition of the coefficients effect, but not of the characteristics effect, when a logit equation is used to study employment in formal sector in an urban area. In order to derive a weight for the detailed decomposition of the coefficients effect, Nielsen (1998) specifies $\beta_A = \beta_B + \delta$, and applies total differentiation to the coefficients effect, $[F(X_A \beta_A) - F(X_A \beta_B)]$, with respect to δ . Nielsen's proposed weight for the coefficients effect for a variable i is as follows:

$$W_{\delta}^i = \frac{\sum_{A_j} f(X_{A_j} \beta_A) X_{A_j}^i d\delta^i / N_{A_j}}{\sum_{A_j} (X_{A_j} \beta_A) X_{A_j} d\delta / N_{A_j}}, \quad (4)$$

where N_{A_j} is the size of the sample A.

As it is obvious from the weight, W_{δ}^i , the choice of $d\delta$ is left to practitioners. It is quite reasonable to assign the same value to each $d\delta^i$, which further simplifies the weight to

$$W_{\delta}^i = \frac{\sum_{A_j} f(X_{A_j} \beta_A) X_{A_j}^i / N_{A_j}}{\sum_{A_j} (X_{A_j} \beta_A) X_{A_j} / N_{A_j}}.$$

Though using total differentiation is a legitimate method, the implication of the weight proposed by Nielsen (1998) is a little counter-intuitive since the weight for the coefficients effect is not expressed in terms of differences in coefficients, i.e., δ , but in terms of β_A .⁵

In short, the sequential replacement methodology can be theoretically applied to any functional form, but it suffers from path-dependency. While the suggested detailed decompositions by Even and Macpherson (1990, 1993) and Nielsen (1998) are free from path-dependency, they have not envisioned the potential of extending their method to provide a general and systematic detailed decomposition methodology as provided in this paper.

III. Conclusion

We provide a general and systematic decomposition for first moment differences. The differences are decomposed into characteristics and coefficients effects at both an aggregate and individual variable level. This decomposition methodology is applicable to any functional form as long as the function is once differentiable. Since most functional forms we are interested are once differentiable, this decomposition methodology provides a common framework whenever differences

⁵ We may modify the weight, W_δ^i , in the equation (4) a bit in order to make comparison to the weight proposed in this paper clear. If an approximation using mean characteristics is used instead of the sample averaging, then the weight becomes;

$$W_\delta^i = \frac{f(\bar{X}_A \beta_A) \bar{X}_A^i d\delta^i}{f(\bar{X}_A \beta_A) \bar{X}_A d\delta} = \frac{\bar{X}_A^i d\delta^i}{\bar{X}_A d\delta} = \frac{\bar{X}_A^i}{\sum_{i=1}^{i=K} \bar{X}_A^i},$$

where the last equality is obtained when the same value is assigned to each $d\delta^i$.

The weight becomes a ratio of mean value of one variable to the sum of mean value of variables. That is, the level of \bar{X}_A^i influences the size of weight, without accounting for potential differences in β 's. This also implies that a rescaling of the X 's could affect the weight substantially.

in the first moment are studied. Obviously, this generalization includes the standard Blinder-Oaxaca decomposition as a special case.

References

Blinder, A. S., 1973, Wage discrimination: reduced form and structural estimates, *Journal of Human Resources* 8, 436-455.

Even, W. E. and D. A. Macpherson, 1990, Plant size and the decline of unionism, *Economics Letters*, 32, 393-398.

Even, W. E. and D. A. Macpherson, 1993, The decline of private-sector unionism and the gender wage gap, *Journal of Human Resources*, 28, 279-296.

Gang, I. N., F. L. Rivera-Batiz and M.-S. Yun, 2002, Economic strain, ethnic concentration and attitudes towards foreigners in the European Union, Working Paper, Department of Economics, Rutgers University.

Gang, I. N., K. Sen and M.-S. Yun, 2002, Caste, ethnicity and poverty in rural India, Working Paper, Department of Economics, Rutgers University.

Ham, J. C., J. Svejnar and K. Terrell, 1998, Unemployment and the social safety net during transitions to a market economy: evidence from the Czech and Slovak Republics, *American Economic Review*, 88, 1117-1142.

Juhn, C., K. M. Murphy and B. Pierce, 1993, Wage inequality and the rise in returns to skill, *Journal of Political Economy*, 101, 410-442.

Nielsen, H. S., 1998, Discrimination and detailed decomposition in a logit model, *Economics Letters*, 61, 115-120.

Oaxaca, R., 1973, Male-female wage differentials in urban labor markets, *International Economic Review* 14, 693-709.

Yun, M.-S., 2002, Earnings inequality in USA, 1961-1999: comparing inequality using earnings equations, Working Paper, Department of Economics, Tulane University.

Yun, M.-S., 2000, Decomposition analysis for a binary choice model, Working Paper, No. 2000-01, Department of Economics, Rutgers University.

IZA Discussion Papers

No.	Author(s)	Title	Area	Date
863	C. U. Chiswick	Immigrant Religious Adjustment: An Economic Approach to Jewish Migrations	1	08/03
864	D. Cobb-Clark C. Ryan R. Breunig	A Couples-Based Approach to the Problem of Workless Families	6	08/03
865	S. W. Polachek	Mincer's Overtaking Point and the Lifecycle Earnings Distribution	5	08/03
866	A. Constant	Immigrant Adjustment in France and Impacts on the Natives	1	08/03
867	T. J. Dohmen B. Kriechel G. A. Pfann	Monkey Bars and Ladders: The Importance of Lateral and Vertical Job Mobility in Internal Labor Market Careers	1	08/03
868	P. Brañas Garza S. Neuman	Analyzing Religiosity Within an Economic Framework: The Case of Spanish Catholics	5	08/03
869	A. B. Krueger A. Mas	Strikes, Scabs and Tread Separations: Labor Strife and the Production of Defective Bridgestone/Firestone Tires	3	09/03
870	S. Klasen A. Launov	Analysis of the Determinants of Fertility Decline in the Czech Republic	4	09/03
871	Y. L'Horty C. Rault	The Impact of Growth, Labour Cost and Working Time on Employment: Lessons from the French Experience	5	09/03
872	J. R. Munch M. Rosholm M. Svarer	Are Home Owners Really More Unemployed?	1	09/03
873	M. Falk B. M. Koebel	The Impact of Office Machinery and Computer Capital on the Demand for Heterogeneous Labour	5	09/03
874	J.-S. Pischke	The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Years	6	09/03
875	C. Grund	Severance Payments for Dismissed Employees in Germany	3	09/03
876	M. Karanassou H. Sala D. J. Snower	The European Phillips Curve: Does the NAIRU Exist?	3	09/03
877	M.-S. Yun	Decomposing Differences in the First Moment	3	09/03

An updated list of IZA Discussion Papers is available on the center's homepage www.iza.org.