DISCUSSION PAPER SERIES

IZA DP No. 1822

Normalized Equation and Decomposition Analysis: Computation and Inference

Myeong-Su Yun

October 2005

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Normalized Equation and Decomposition Analysis: Computation and Inference

Myeong-Su Yun
Tulane University
and IZA Bonn

Discussion Paper No. 1822
October 2005

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
Email: iza@iza.org

IZA Discussion Paper No. 1822
October 2005

ABSTRACT

Normalized Equation and Decomposition Analysis: Computation and Inference*

This paper joins discussions on normalized regression and decomposition equations in devising a simple and general algorithm for obtaining the normalized regression and applying it to the Oaxaca decomposition. This resolves the invariance problem in the detailed Oaxaca decomposition. An algorithm to calculate an asymptotic covariance matrix for estimates in the normalized regression for hypothesis testing is also derived. We extend these algorithms to non-linear equations where the underlying equation is linear and decompose differences in the first moment.

JEL Classification: C20, J70

Keywords: detailed decomposition, invariance, identification, characteristics effect, coefficients effect, normalized regression

Corresponding author:

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

*The author wishes to thank Ira Gang for comments and providing insights on the averaging approach.

I. Introduction

"Normalized" regressions are a useful estimation tool for separately identifying the intercept and coefficients of sets of dummy variables. The coefficients of a set of dummy variables and the intercept are identified in a regression framework by imposing the restriction that the coefficients of a set of dummy variables sum to zero. However, with few exceptions (e.g., Suits, 1984, Greene and Seaks, 1991, Gardeazabal and Ugidos, 2004, and Yun, 2005) the normalized regression has not penetrated everyday estimation practices despite its merits, with the exception of studies on industrial wage differentials (Krueger and Summers, 1988, Edin and Zetterberg, 1992, and Haisken-DeNew and Schmidt, 1997).

Recently it has been found that the normalized regression is very useful in resolving the invariance or identification problem in detailed decompositions of wage differentials (Yun, 2005). Oaxaca and Ransom (1999) show that the detailed Oaxaca decomposition of wage differentials is not invariant to the choice of reference group when a set of dummy variables is used.[1] That is, if we use dummy variable(s), then the detailed coefficients effects attributed to individual variables are not invariant to the choice of the omitted group(s). This invariance or identification problem is well-known to labor economists and has plagued decomposition analysis for a long time.[2]

This paper devises a simple and general algorithm for obtaining the normalized regression and applying to the Oaxaca decomposition. The simple algorithm derives coefficients of the

---

[1] Decomposition analysis, since the paper by Oaxaca (1973), explains wage differentials in terms of differences in individual characteristics (characteristics effect) and differences in the OLS coefficients of wage equations (coefficients effect).

[2] Some have stopped doing detailed decompositions of the coefficients effect (e.g., Ham, Svenjnar and Terrell, 1998).

normalized regression and their covariance matrix using estimates and their covariance matrix from the usual regression equation where reference groups are left out. Once the normalized regression is derived, we construct the Oaxaca decomposition equation in order to solve the invariance problem. We discuss the inference of Oaxaca decomposition equation when the normalized regression is used and point out that incorporating the normalized regression into the Oaxaca decomposition does not change the inference of the decomposition equation for the overall characteristics and coefficients effects.

We extend the discussion to generalized decomposition analysis for differences in the first moment, that is, the differences in the mean value of the variable of interest. We show that the algorithms used for linear regression to derive the normalized regression equation can be used for deriving a normalized equation for a non-linear equation. We also show that the Oaxaca-type decomposition for the differences in the first moment can be easily combined with the normalized equation. For illustration purposes, we study changes in the labor market participation rate of white women between 1980 and 2001 using the 1980 and 2001 waves of Panel Study of Income Dynamics.

II. Deriving Normalized Regression

For illustration purposes, suppose that we have two sets of dummy variables in addition to continuous variables in the regression equation.[3] The regression equation, suppressing individual subscripts, is

$$y = \alpha + \left[\sum_{j=2}^{J} d_j \gamma_j + \sum_{k=2}^{K} q_k \theta_k\right] + \sum_{l=1}^{L} z_l \delta_l + e, \tag{1}$$

[3] The extension of the discussion to incorporate more sets of dummy variables is trivial.

where there are two sets of categorical variables (d's and q's) and L continuous variables (z's); the first and second sets of dummy variables (d's and q's) have J and K categories and J - 1 and K - 1 dummy variables in the equation, respectively; without loss of generality, the reference group is the first category for each set of dummy variables. We refer to equation (1) as the usual regression equation.

Our question is how to derive coefficients for the normalized regression which does not omit the reference groups in the estimating equation. The proto-type normalized regression looks like

$$y = \alpha^* + \left[\sum_{j=1}^{J} d_j \gamma_j^* + \sum_{k=1}^{K} q_k \theta_k^*\right] + \sum_{l=1}^{L} z_l \delta_l + e .$$

In order to construct the normalized regression, Suits (1984) proposes the restrictions $\sum_{j=1}^{J} \gamma_j^* = 0$ and $\sum_{k=1}^{K} \theta_k^* = 0$. Since these restrictions do not have unique solutions, he specifies the coefficients of the normalized regression as $\gamma_j^* = \gamma_j + m_\gamma$ and $\theta_k^* = \theta_k + m_\theta$, and refines the problem of deriving the normalized regression as finding values of m_γ and m_θ. It turns out that their values are $m_\gamma = -\sum_{j=1}^{J} \gamma_j / J$ and $m_\theta = -\sum_{k=1}^{K} \theta_k / K$, where $\gamma_1 = \theta_1 = 0$. Since the publication of Suits (1984), there have been several additional developments for deriving the normalized regression.

Greene and Seaks (1991) derive the normalized regression by obtaining expressions for the restricted least square estimator and its covariance matrix in the classical regression model when the matrix of exogenous variables is not necessarily of full rank. Gardeazabal and Ugidos (2004) run a regression after transforming the dummy variables to a deviation from the reference group, such as $d_j - d_1$ and $q_k - q_1$. In order to transform the dummy variables, they use the same restrictions as Suits (1984), that is, $\sum_{j=1}^{J} \gamma_j^* = 0$ and $\sum_{k=1}^{K} \theta_j^* = 0$. However, the methods developed by Greene and Seaks (1991) and Gardeazabal and Ugidos (2004) may not be flexible enough to handle more complicated econometric models (e.g., the selection model).

An intuitive method to derive the normalized regression uses an averaging approach (Yun, 2005). The normalized regression is the outcome of averaging all possible estimates with permuting reference groups. There is no need to run large numbers of regression equation in order to exhaust all possible specifications of reference groups since any one regression equation can provide all necessary information to obtain the normalized regression. For example, suppose that we obtain estimates for dummy variables d_j's using the first category as the reference group as shown in equation (1). Using the obtained regression estimates, we can calculate the estimates when the reference group is changed from the first category to the *r*th category. The estimates made by changing the reference group from the first category to the *r*th category will change from $\gamma_j - \gamma_1$ to $\gamma_j - \gamma_r$, where $\gamma_1 = 0$ and $j = 1, 2, \cdots, J$. And the intercept changes from $\alpha + \gamma_1$ to $\alpha + \gamma_r$ with the change in the reference group. We can do the same for variables (q_k). This averaging approach implies that Suits' constraint can be interpreted as the outcome of averaging of all possible estimates with permuting reference groups. Both the constraint approach by Suits (1984) and averaging approach by Yun (2005) can be used whenever estimates of equation (1) are available.

Once consistent estimates of equation (1) are obtained, we can manipulate these estimates in order to obtain a normalized regression equation.[4] The normalized equation is:

$$y = (\alpha + \bar{\gamma} + \bar{\theta}) + \left[\sum_{j=1}^{J} d_j(\gamma_j - \bar{\gamma}) + \sum_{k=1}^{K} q_k(\theta_k - \bar{\theta})\right] + \sum_{l=1}^{L} z_l \delta_l + e, \tag{1'}$$

where $\bar{\gamma} = \dfrac{\sum_{j=1}^{J} \gamma_j}{J}$, $\bar{\theta} = \dfrac{\sum_{k=1}^{K} \theta_k}{K}$, and $\gamma_1 = \theta_1 = 0$.

To be succinct and derive a systematic solution, we may represent the above in terms of a matrix.[5] Define the matrix of independent variables, $\boldsymbol{X} = (\mathbf{1} \vdots \boldsymbol{D} \vdots \boldsymbol{Q} \vdots \boldsymbol{Z})$, where $\boldsymbol{D} = (\boldsymbol{d}_2 \vdots \cdots \vdots \boldsymbol{d}_J)$, $\boldsymbol{Q} = (\boldsymbol{q}_2 \vdots \cdots \vdots \boldsymbol{q}_K)$, and $\boldsymbol{Z} = (z_1 \vdots \cdots \vdots z_L)$. $\boldsymbol{X}$, $\boldsymbol{D}$, $\boldsymbol{Q}$ and $\boldsymbol{Z}$ are, respectively, $N\times T$, $N\times(J-1)$, $N\times(K-1)$ and $N\times L$ matrices where T=$(J+K+L-1)$, and $\boldsymbol{D}$ and $\boldsymbol{Q}$ are matrices of two sets of dummy variables, and $\boldsymbol{Z}$ is a matrix of continuous variables; $\mathbf{1}$ is a vector of ones ($N\times 1$); $\boldsymbol{Y}$ is the vector of the dependent variable ($N\times 1$); $\boldsymbol{\beta} = (\alpha, \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\delta})$ is a coefficient vector ($T\times 1$), where $\boldsymbol{\gamma} = (\gamma_2, \cdots, \gamma_J)$, $\boldsymbol{\theta} = (\theta_2, \cdots, \theta_K)$, and $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_L)$.

A matrix representation of equation (1) is, $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}$. In order to obtain the normalized equation (1') it is useful to rewrite the equation as $\boldsymbol{Y} = \boldsymbol{X}^* \boldsymbol{\beta}_0 + \boldsymbol{e}$, where $\boldsymbol{X}^* = (\mathbf{1} \vdots \boldsymbol{d}_1 \vdots \boldsymbol{D} \vdots \boldsymbol{q}_1 \vdots \boldsymbol{Q} \vdots \boldsymbol{Z})$ and $\boldsymbol{\beta}_0 = (\alpha, 0, \boldsymbol{\gamma}, 0, \boldsymbol{\theta}, \boldsymbol{\delta})$. The normalized regression, $\boldsymbol{Y} = \boldsymbol{X}^* \boldsymbol{\beta}^* + \boldsymbol{e}$, is obtained by transforming

---

[4] This paper assumes simple average of coefficients of dummy variables is used in order to derive the normalized regression equation. Though it is easy to derive a normalized regression using the average of the dummy variables' coefficients weighted by the share of each group, it has the implication that the sum of the product of the dummy variables and their coefficients should be zero, which is not attractive for the Oaxaca decomposition (Yun, 2005).

[5] The derivation of the normalized regression equation is developed by extending a method employed by Haisken-DeNew and Schmidt (1997).

$\boldsymbol{\beta}_0$ to $\boldsymbol{\beta}^*$ using a weight matrix, $\boldsymbol{W}$, that is, $\boldsymbol{\beta}^* = \boldsymbol{W}\boldsymbol{\beta}_0$, which yields $(T^* \times 1)$ vector of $\left((\alpha + \bar{\gamma} + \bar{\theta}), (\gamma_1 - \bar{\gamma}), (\boldsymbol{\gamma} - \bar{\gamma}), (\theta_1 - \bar{\theta}), (\boldsymbol{\theta} - \bar{\theta}), \boldsymbol{\delta}\right)$, where $T^* = T+2 = J+K+L+1$. The weight matrix $\boldsymbol{W}$ is defined as

$$\boldsymbol{W} = \begin{bmatrix} 1 & (1/J)\cdot\mathbf{1}_{1\times J} & (1/K)\cdot\mathbf{1}_{1\times K} & \mathbf{0}_{1\times L} \\ \mathbf{0}_{J\times 1} & \boldsymbol{M}_{J\times J} & \mathbf{0}_{J\times K} & \mathbf{0}_{J\times L} \\ \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times J} & \boldsymbol{M}_{K\times K} & \mathbf{0}_{K\times L} \\ \mathbf{0}_{L\times 1} & \mathbf{0}_{L\times J} & \mathbf{0}_{L\times K} & \mathbf{I}_{L\times L} \end{bmatrix}, \tag{2}$$

where $\boldsymbol{M}_{P\times P} = \mathbf{I}_{P\times P} - (1/P)\cdot\mathbf{1}_{P\times P}$, and $\mathbf{0}$, $\mathbf{1}$, and $\mathbf{I}$ are a matrix of zeros, a matrix of ones and an identity matrix.[6]

The covariance matrix of estimates in the normalized regression equation ($\boldsymbol{\beta}^*$) can be also easily obtained. Suppose that the covariance matrix of estimates in the usual regression equation ($\boldsymbol{\beta}$) is obtained and is defined as

$$\boldsymbol{\Sigma}_{\boldsymbol{\beta}} = \begin{bmatrix} \sigma^2_{\alpha} & \boldsymbol{\Sigma}_{\alpha,\boldsymbol{\gamma}'} & \boldsymbol{\Sigma}_{\alpha,\boldsymbol{\theta}'} & \boldsymbol{\Sigma}_{\alpha,\boldsymbol{\delta}'} \\ \boldsymbol{\Sigma}_{\boldsymbol{\gamma},\alpha} & \boldsymbol{\Sigma}_{\boldsymbol{\gamma},\boldsymbol{\gamma}'} & \boldsymbol{\Sigma}_{\boldsymbol{\gamma},\boldsymbol{\theta}'} & \boldsymbol{\Sigma}_{\boldsymbol{\gamma},\boldsymbol{\delta}'} \\ \boldsymbol{\Sigma}_{\boldsymbol{\theta},\alpha} & \boldsymbol{\Sigma}_{\boldsymbol{\theta},\boldsymbol{\gamma}'} & \boldsymbol{\Sigma}_{\boldsymbol{\theta},\boldsymbol{\theta}'} & \boldsymbol{\Sigma}_{\boldsymbol{\theta},\boldsymbol{\delta}'} \\ \boldsymbol{\Sigma}_{\boldsymbol{\delta},\alpha} & \boldsymbol{\Sigma}_{\boldsymbol{\delta},\boldsymbol{\gamma}'} & \boldsymbol{\Sigma}_{\boldsymbol{\delta},\boldsymbol{\theta}'} & \boldsymbol{\Sigma}_{\boldsymbol{\delta},\boldsymbol{\delta}'} \end{bmatrix}.$$

---

[6] A normalized regression with weighted average of coefficients of dummy variables can be easily obtained by changing the weight matrix, $\boldsymbol{W}$. Define $\boldsymbol{W_D} = (W_{\boldsymbol{d}_1} \vdots \cdots \vdots W_{\boldsymbol{d}_J})$ and $\boldsymbol{W_Q} = (W_{\boldsymbol{q}_1} \vdots \cdots \vdots W_{\boldsymbol{q}_K})$ be vectors of shares of dummy variables, $\boldsymbol{D}$ and $\boldsymbol{Q}$. In order to find a weight matrix for obtaining a normalized regression with weighted averages, replace $(1/J)\cdot\mathbf{1}_{1\times J}$ and $(1/K)\cdot\mathbf{1}_{1\times K}$ with $\boldsymbol{W_D}$ and $\boldsymbol{W_Q}$, and replace $\boldsymbol{M}_{P\times P} = \mathbf{I}_{P\times P} - (1/P)\cdot\mathbf{1}_{P\times P}$ with $\boldsymbol{M}_{J\times J} = \mathbf{I}_{J\times J} - \boldsymbol{W}^+_{\boldsymbol{D}}$ and $\boldsymbol{M}_{K\times K} = \mathbf{I}_{K\times K} - \boldsymbol{W}^+_{\boldsymbol{Q}}$ when $P = J$ and K, respectively. $\boldsymbol{W}^+_{\boldsymbol{D}} = (\boldsymbol{W_D}, \cdots, \boldsymbol{W_D})$ and $\boldsymbol{W}^+_{\boldsymbol{Q}} = (\boldsymbol{W_Q}, \cdots, \boldsymbol{W_Q})$, which are $J\times J$ and $K\times K$ matrices, respectively.

Similar to the derivation of $\boldsymbol{\beta}_0$, the covariance matrix for $\boldsymbol{\beta}_0$ can be obtained by adding zero vectors to the covariance matrix of $\boldsymbol{\beta}$ as

$$\boldsymbol{\Sigma}_{\boldsymbol{\beta}_0} = \begin{bmatrix} \sigma^2_{\alpha} & 0 & \Sigma_{\alpha,\gamma'} & 0 & \Sigma_{\alpha,\theta'} & \Sigma_{\alpha,\delta'} \\ 0 & 0 & \mathbf{0}_{1\times(J-1)} & 0 & \mathbf{0}_{1\times(K-1)} & \mathbf{0}_{1\times L} \\ \Sigma_{\gamma,\alpha} & \mathbf{0}_{(J-1)\times 1} & \Sigma_{\gamma,\gamma'} & \mathbf{0}_{(J-1)\times 1} & \Sigma_{\gamma,\theta'} & \Sigma_{\gamma,\delta'} \\ 0 & 0 & \mathbf{0}_{1\times(J-1)} & 0 & \mathbf{0}_{1\times(K-1)} & \mathbf{0}_{1\times L} \\ \Sigma_{\theta,\alpha} & \mathbf{0}_{(K-1)\times 1} & \Sigma_{\theta,\gamma'} & \mathbf{0}_{(K-1)\times 1} & \Sigma_{\theta,\theta'} & \Sigma_{\theta,\delta'} \\ \Sigma_{\delta,\alpha} & \mathbf{0}_{L\times 1} & \Sigma_{\delta,\gamma'} & \mathbf{0}_{L\times 1} & \Sigma_{\delta,\theta'} & \Sigma_{\delta,\delta'} \end{bmatrix}. \tag{3}$$

Finally the covariance matrix of estimates for the normalized regression equation ($\boldsymbol{\beta}^*$) is computed as $\boldsymbol{\Sigma}_{\boldsymbol{\beta}^*} = \boldsymbol{W}\,\boldsymbol{\Sigma}_{\boldsymbol{\beta}_0}\,\boldsymbol{W}'$.

III. Oaxaca Decomposition with Normalized Regression: Computation and Inference

The Oaxaca decomposition explains wage differentials due to differences in mean characteristics and due to differences in returns to characteristics. The decomposition equation using the usual regression equation is as follows;

$$\overline{\boldsymbol{Y}}_A - \overline{\boldsymbol{Y}}_B = (\overline{\boldsymbol{X}}_A - \overline{\boldsymbol{X}}_B)\boldsymbol{\beta}_A + \overline{\boldsymbol{X}}_B(\boldsymbol{\beta}_A - \boldsymbol{\beta}_B) + \overline{\boldsymbol{e}}_A - \overline{\boldsymbol{e}}_B, \tag{4}$$

where $\overline{\boldsymbol{X}}_g = (1 \vdots \overline{\boldsymbol{D}}_g \vdots \overline{\boldsymbol{Q}}_g \vdots \overline{\boldsymbol{Z}}_g)$ is an $1 \times T$ vector of mean values of exogenous variables for group g (A or B), $\boldsymbol{\beta}_g = (\alpha_g, \boldsymbol{\gamma}_g, \boldsymbol{\theta}_g, \boldsymbol{\delta}_g)$, and $\overline{\boldsymbol{e}}_g$ is the average of the residuals for group g whose value is zero when OLS is used.

Based on the asymptotic variance for the characteristics and coefficients effects, the hypothesis test that the overall characteristics (coefficients) effect is significantly different from zero can be derived. The t-test for the characteristics and coefficients effects can be constructed as $t_{\Delta X} = \dfrac{(\overline{\boldsymbol{X}}_A - \overline{\boldsymbol{X}}_B)\,\boldsymbol{\beta}_A}{\sigma_{\Delta X}}$ and $t_{\Delta\beta} = \dfrac{\overline{\boldsymbol{X}}_A(\boldsymbol{\beta}_A - \boldsymbol{\beta}_B)}{\sigma_{\Delta\beta}}$, where $\sigma^2_{\Delta X} = (\overline{\boldsymbol{X}}_A - \overline{\boldsymbol{X}}_B)\,\Sigma_{\beta_A}(\overline{\boldsymbol{X}}_A - \overline{\boldsymbol{X}}_B)'$ and

$\sigma^2_{\Delta\beta} = \overline{\boldsymbol{X}}_B(\Sigma_{\beta_A} + \Sigma_{\beta_B})\overline{\boldsymbol{X}}_B'$ (Oaxaca and Ransom, 1998).

The invariance problem in the detailed Oaxaca decomposition is that the sum of the coefficients effects of dummy variables of $\boldsymbol{D}$, i.e., $\overline{\boldsymbol{D}}_B(\boldsymbol{\gamma}_A - \boldsymbol{\gamma}_B) = \sum_{j=2}^{J} \overline{d}_{jB}(\gamma_{jA} - \gamma_{jB})$, is not invariant when the reference group is changed. The same is true of the coefficients effects of dummy variables of $\boldsymbol{Q}$ (Oaxaca and Ransom, 1999). The normalized equation is ideal for solving the invariance problem in the detailed Oaxaca decomposition (Yun, 2005). The decomposition equation with the normalized regression equation is

$$\overline{\boldsymbol{Y}}_A - \overline{\boldsymbol{Y}}_B = (\overline{\boldsymbol{X}}^*_A - \overline{\boldsymbol{X}}^*_B)\boldsymbol{\beta}^*_A + \overline{\boldsymbol{X}}^*_B(\boldsymbol{\beta}^*_A - \boldsymbol{\beta}^*_B) + \overline{\boldsymbol{e}}_A - \overline{\boldsymbol{e}}_B, \tag{4'}$$

where $\overline{\boldsymbol{X}}^*_g = (1 \vdots \overline{d}_{1g} \vdots \overline{\boldsymbol{D}}_g \vdots \overline{q}_{1g} \vdots \overline{\boldsymbol{Q}}_g \vdots \overline{\boldsymbol{Z}}_g)$, and

$\boldsymbol{\beta}^*_g = \left((\alpha_g + \overline{\gamma}_g + \overline{\theta}_g), (\gamma_{1g} - \overline{\gamma}_g), (\boldsymbol{\gamma}_g - \overline{\gamma}_g), (\theta_{1g} - \overline{\theta}_g), (\boldsymbol{\theta}_g - \overline{\theta}_g), \boldsymbol{\delta}_g\right)$ for group g (A or B).

The t-test for the characteristics and coefficients effects can be constructed as

$t_{\Delta X^*} = \dfrac{(\overline{\boldsymbol{X}}_A^* - \overline{\boldsymbol{X}}_B^*)\,\boldsymbol{\beta}_A^*}{\sigma_{\Delta X^*}}$ and $t_{\Delta\beta^*} = \dfrac{\overline{\boldsymbol{X}}_A^*\,(\boldsymbol{\beta}_A^* - \boldsymbol{\beta}_B^*)}{\sigma_{\Delta\beta^*}}$, where $\sigma^2_{\Delta X^*} = (\overline{\boldsymbol{X}}_A^* - \overline{\boldsymbol{X}}_B^*)\,\Sigma_{\boldsymbol{\beta}_A^*}\,(\overline{\boldsymbol{X}}_A^* - \overline{\boldsymbol{X}}_B^*)'$ and

$\sigma^2_{\Delta\beta^*} = \overline{\boldsymbol{X}}_B^*\,(\Sigma_{\boldsymbol{\beta}_A^*} + \Sigma_{\boldsymbol{\beta}_B^*})\,\overline{\boldsymbol{X}}_B^{*\prime}$.

One may wonder whether utilizing the normalized regression equation for computing the decomposition equation changes the size and inference of the overall characteristics and coefficients effects, and therefore, the outcome of the t-tests. It can be easily shown that both the size and the asymptotic variances of the overall characteristics and coefficients effects do not change. To show these, first, note that $\overline{\boldsymbol{X}}\boldsymbol{\beta} = \overline{\boldsymbol{X}}^*\,\boldsymbol{\beta}_0$ and $\overline{\boldsymbol{X}}\,\Sigma_{\boldsymbol{\beta}}\,\overline{\boldsymbol{X}}' = \overline{\boldsymbol{X}}^*\,\Sigma_{\boldsymbol{\beta}_0}\,\overline{\boldsymbol{X}}^{*\prime}$. The next step is showing that $\overline{\boldsymbol{X}}^*\,\boldsymbol{\beta}_0 = \overline{\boldsymbol{X}}^*\,\boldsymbol{\beta}^*$ and $\overline{\boldsymbol{X}}^*\,\Sigma_{\boldsymbol{\beta}_0}\,\overline{\boldsymbol{X}}^{*\prime} = \overline{\boldsymbol{X}}^*\,\Sigma_{\boldsymbol{\beta}^*}\,\overline{\boldsymbol{X}}^{*\prime}$. This step can be verified from the equality, $\boldsymbol{X}^*\,\boldsymbol{W} = \boldsymbol{X}^*$. Since $\boldsymbol{\beta}^* = \boldsymbol{W}\boldsymbol{\beta}_0$ and $\Sigma_{\boldsymbol{\beta}^*} = \boldsymbol{W}\,\Sigma_{\boldsymbol{\beta}_0}\,\boldsymbol{W}'$, the equality on the size and asymptotic variances can be established as $\overline{\boldsymbol{X}}\boldsymbol{\beta} = \overline{\boldsymbol{X}}^*\,\boldsymbol{\beta}^*$ and $\overline{\boldsymbol{X}}\,\Sigma_{\boldsymbol{\beta}}\,\overline{\boldsymbol{X}}' = \overline{\boldsymbol{X}}^*\,\Sigma_{\boldsymbol{\beta}^*}\,\overline{\boldsymbol{X}}^{*\prime}$.[7] Therefore, the hypothesis test for the overall characteristics and coefficients effects is identical whether the usual or the normalized regression equation is used for the decomposition equation. Obviously, the size and variance of the detailed decomposition for continuous variables (***Z***) do not change, but the size and variance of the detailed decomposition for the sets of dummy variables (***D*** and ***Q***) and the intercept change.

[7] Even if the normalized equation is constructed using weighted averages of sets of dummy variables, the conclusion on the inference is still valid.

IV. Decomposing Differences in the First Moment and Normalized Equation

We will generalize what we have discussed using a linear equation in previous sections. 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,

$$\boldsymbol{Y} = F(\boldsymbol{X}\boldsymbol{\beta}),$$

where matrices $\boldsymbol{Y}$, $\boldsymbol{X} = (\mathbf{1} \vdots \boldsymbol{D} \vdots \boldsymbol{Q} \vdots \boldsymbol{Z})$, and $\boldsymbol{\beta} = (\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\delta})$ are defined same as in the section **II**; F is a mapping of a linear combination of $\boldsymbol{X}$ ($\boldsymbol{X}\boldsymbol{\beta}$) to $\boldsymbol{Y}$, and the function F is any once differentiable function.

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

$$\overline{\boldsymbol{Y}}_A - \overline{\boldsymbol{Y}}_B = \left[\overline{F(\boldsymbol{X}_A\boldsymbol{\beta}_A)} - \overline{F(\boldsymbol{X}_B\boldsymbol{\beta}_A)}\right] + \left[\overline{F(\boldsymbol{X}_B\boldsymbol{\beta}_A)} - \overline{F(\boldsymbol{X}_B\boldsymbol{\beta}_B)}\right], \tag{5}$$

where "over bar" represents the value of the sample's average.

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, $\overline{F(\boldsymbol{X}_A\boldsymbol{\beta}_A)} - \overline{F(\boldsymbol{X}_B\boldsymbol{\beta}_A)}$, and in terms of differences in coefficients, $\overline{F(\boldsymbol{X}_B\boldsymbol{\beta}_A)} - \overline{F(\boldsymbol{X}_B\boldsymbol{\beta}_B)}$. The next step is to find the contribution of each variable to the total difference (detailed decomposition). Yun (2004) proposes the following detailed decomposition equation,[8]

---

[8] 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; second, we use a first order Taylor expansion to linearize the characteristics and coefficients effects around $\overline{\boldsymbol{X}}_A\boldsymbol{\beta}_A$ and $\overline{\boldsymbol{X}}_B\boldsymbol{\beta}_B$, respectively.

$$\overline{\boldsymbol{Y}}_A - \overline{\boldsymbol{Y}}_B = \sum_{i=1}^{T} W_{\Delta X}^i \left[\left(\overline{F(\boldsymbol{X}_A \boldsymbol{\beta}_A)} - \overline{F(\boldsymbol{X}_B \boldsymbol{\beta}_A)}\right) \right] + \sum_{i=1}^{T} W_{\Delta\beta}^i \left[\overline{F(\boldsymbol{X}_B \boldsymbol{\beta}_A)} - \overline{F(\boldsymbol{X}_B \boldsymbol{\beta}_B)} \right], \quad (6)$$

where

$$W_{\Delta X}^i = \frac{(\overline{\boldsymbol{X}}_A^i - \overline{\boldsymbol{X}}_B^i)\ \boldsymbol{\beta}_A^i}{(\overline{\boldsymbol{X}}_A - \overline{\boldsymbol{X}}_B)\ \boldsymbol{\beta}_A} \ , \ W_{\Delta\beta}^i = \frac{\overline{\boldsymbol{X}}_B^i\ (\boldsymbol{\beta}_A^i - \boldsymbol{\beta}_B^i)}{\overline{\boldsymbol{X}}_B\ (\boldsymbol{\beta}_A - \boldsymbol{\beta}_B)} \ , \text{ and } \sum_{i=1}^{T} W_{\Delta X}^i = \sum_{i=1}^{T} W_{\Delta\beta}^i = 1 .$$

This decomposition methodology is free from path dependency, unlike a sequential replacement approach that computes the contribution of an individual variable or its coefficient to the differences in the first moment by switching values of one group with those of a comparison group, such as in the method proposed by Fairlie (2003).[9] The sequential replacement approach is sensitive to the order of switching (see Ham, Svejnar and Terrell, 1998, p. 1137 for a discussion of path-dependency). The detailed decomposition equation (2) is a generalization of what Even and Macpherson (1990, 1993) propose for only the characteristics effect when probit is used. The generalization by Yun (2004) is to both characteristics and coefficients effects and to when any non-linear equation is used.[10]

Suppose that we have obtained consistent estimates of $\boldsymbol{\beta}$ and its covariance matrix, $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$. The normalized equation for $\boldsymbol{Y} = F(\boldsymbol{X}\boldsymbol{\beta})$ can be obtained following the procedures described in the section **II**. That is, the normalized equation is $\boldsymbol{Y} = F(\boldsymbol{X}^* \boldsymbol{\beta}^*)$, where $\boldsymbol{X}^* = (\mathbf{1} \vdots \boldsymbol{d}_1 \vdots \boldsymbol{D} \vdots \boldsymbol{q}_1 \vdots \boldsymbol{Q} \vdots \boldsymbol{Z})$

---

[9] The sequential replacement approach has its roots in simulation methods which switch one group's coefficient with other's. It is usually done one by one in order to see the changes incurred due to the substitution. See Abowd and Killingsworth (1984).

[10] See Doriron and Riddell (1994) for another approach to generalizing Even and Macpherson's methodology.

and $\boldsymbol{\beta}^* = \boldsymbol{W}\boldsymbol{\beta}_0$, where $\boldsymbol{\beta}_0 = (\boldsymbol{\alpha}, 0, \boldsymbol{\gamma}, 0, \boldsymbol{\theta}, \boldsymbol{\delta})$ and the matrix $\boldsymbol{W}$ is defined same as in equation (2). The asymptotic covariance of $\boldsymbol{\beta}^*$ is computed as $\boldsymbol{\Sigma}_{\boldsymbol{\beta}^*} = \boldsymbol{W}\,\boldsymbol{\Sigma}_{\boldsymbol{\beta}_0}\,\boldsymbol{W}'$, where $\boldsymbol{\Sigma}_{\boldsymbol{\beta}_0}$ is defined same as in equation (3).

The decomposition equation using the normalized equation is

$$\overline{\boldsymbol{Y}}_A - \overline{\boldsymbol{Y}}_B = \left[\overline{F(\boldsymbol{X}_A^*\boldsymbol{\beta}_A^*)} - \overline{F(\boldsymbol{X}_B^*\boldsymbol{\beta}_A^*)}\right] + \left[\overline{F(\boldsymbol{X}_B^*\boldsymbol{\beta}_A^*)} - \overline{F(\boldsymbol{X}_B^*\boldsymbol{\beta}_B^*)}\right], \tag{5'}$$

and the detailed decomposition equation is

$$\overline{\boldsymbol{Y}}_A - \overline{\boldsymbol{Y}}_B = \sum_{i=1}^{T^*} W_{\Delta X^*}^i \left[\overline{(F(\boldsymbol{X}_A^*\boldsymbol{\beta}_A^*)} - \overline{F(\boldsymbol{X}_B^*\boldsymbol{\beta}_A^*)}\right] + \sum_{i=1}^{T^*} W_{\Delta\boldsymbol{\beta}^*}^i \left[\overline{F(\boldsymbol{X}_B^*\boldsymbol{\beta}_A^*)} - \overline{F(\boldsymbol{X}_B^*\boldsymbol{\beta}_B^*)}\right], \tag{6'}$$

where

$$W_{\Delta X^*}^i = \frac{(\overline{\boldsymbol{X}}_A^{*i} - \overline{\boldsymbol{X}}_B^{*i})\ \boldsymbol{\beta}_A^{*i}}{(\overline{\boldsymbol{X}}_A^{*} - \overline{\boldsymbol{X}}_B^{*})\ \boldsymbol{\beta}_A^{*}}, \quad W_{\Delta\boldsymbol{\beta}^*}^i = \frac{\overline{\boldsymbol{X}}_B^{*i}\ (\boldsymbol{\beta}_A^{*i} - \boldsymbol{\beta}_B^{*i})}{\overline{\boldsymbol{X}}_B^{*}\ (\boldsymbol{\beta}_A^{*} - \boldsymbol{\beta}_B^{*})}, \quad \sum_{i=1}^{T^*} W_{\Delta X^*}^i = \sum_{i=1}^{T^*} W_{\Delta\boldsymbol{\beta}^*}^i = 1, \text{ and}$$

$T^* = T+2 = J+K+L+1$.

Finally, the identity of asymptotic variance of characteristics and coefficients effects between (5) and (5') can be established. Let $\Delta\boldsymbol{X} = \overline{F(\boldsymbol{X}_A\boldsymbol{\beta}_A)} - \overline{F(\boldsymbol{X}_B\boldsymbol{\beta}_A)}$ and $\Delta\boldsymbol{\beta} = \overline{F(\boldsymbol{X}_B\boldsymbol{\beta}_A)} - \overline{F(\boldsymbol{X}_B\boldsymbol{\beta}_B)}$ be characteristics and coefficients effects when the usual equation is used. The asymptotic variances of the $\Delta\boldsymbol{X}$ and $\Delta\boldsymbol{\beta}$ are calculated as

$$\sigma_{\Delta X}^2 = G_{\Delta X}\,\boldsymbol{\Sigma}_{\boldsymbol{\beta}_A} G_{\Delta X}' \text{ and}$$

$$\sigma^2_{\Delta\boldsymbol{\beta}} = G_{\Delta\boldsymbol{\beta}} \begin{bmatrix} \Sigma_{\boldsymbol{\beta}_A} & 0 \\ 0 & \Sigma_{\boldsymbol{\beta}_B} \end{bmatrix} G'_{\Delta\boldsymbol{\beta}} = \left(\frac{\partial \Delta\boldsymbol{\beta}}{\partial \boldsymbol{\beta}'_A}\right) \Sigma_{\boldsymbol{\beta}_A} \left(\frac{\partial \Delta\boldsymbol{\beta}}{\partial \boldsymbol{\beta}'_A}\right)' + \left(\frac{\partial \Delta\boldsymbol{\beta}}{\partial \boldsymbol{\beta}'_B}\right) \Sigma_{\boldsymbol{\beta}_B} \left(\frac{\partial \Delta\boldsymbol{\beta}}{\partial \boldsymbol{\beta}'_B}\right)',$$

where

$$G_{\Delta X} = \frac{\partial \Delta \boldsymbol{X}}{\partial \boldsymbol{\beta}'_A} \text{ and } G_{\Delta\boldsymbol{\beta}} = \left[\frac{\partial \Delta\boldsymbol{\beta}}{\partial \boldsymbol{\beta}'_A} \vdots \frac{\partial \Delta\boldsymbol{\beta}}{\partial \boldsymbol{\beta}'_B}\right],$$

where $G_{\Delta X}$ and $G_{\Delta\boldsymbol{\beta}}$ are $1 \times T$ and $1 \times 2T$ vectors of gradients (see Yun, forthcoming, for details of deriving asymptotic variances).

Similarly we can compute the characteristics and coefficients effects when the normal equation is used. Let $\Delta\boldsymbol{X}^* = \overline{F(\boldsymbol{X}^*_A\boldsymbol{\beta}^*_A)} - \overline{F(\boldsymbol{X}^*_B\boldsymbol{\beta}^*_A)}$ and $\Delta\boldsymbol{\beta}^* = \overline{F(\boldsymbol{X}^*_B\boldsymbol{\beta}^*_A)} - \overline{F(\boldsymbol{X}^*_B\boldsymbol{\beta}^*_B)}$ be characteristics and coefficients effects when the normal equation is used. The asymptotic variances of the $\Delta\boldsymbol{X}^*$ and $\Delta\boldsymbol{\beta}^*$ are calculated as

$$\sigma^2_{\Delta X^*} = G_{\Delta X^*} \Sigma_{\boldsymbol{\beta}^*_A} G'_{\Delta X^*} \text{ and}$$

$$\sigma^2_{\Delta\boldsymbol{\beta}^*} = G_{\Delta\boldsymbol{\beta}^*} \begin{bmatrix} \Sigma_{\boldsymbol{\beta}^*_A} & 0 \\ 0 & \Sigma_{\boldsymbol{\beta}^*_B} \end{bmatrix} G'_{\Delta\boldsymbol{\beta}} = \left(\frac{\partial \Delta\boldsymbol{\beta}^*}{\partial \boldsymbol{\beta}^{*\prime}_A}\right) \Sigma_{\boldsymbol{\beta}^*_A} \left(\frac{\partial \Delta\boldsymbol{\beta}^*}{\partial \boldsymbol{\beta}^{*\prime}_A}\right)' + \left(\frac{\partial \Delta\boldsymbol{\beta}^*}{\partial \boldsymbol{\beta}^{*\prime}_B}\right) \Sigma_{\boldsymbol{\beta}^*_B} \left(\frac{\partial \Delta\boldsymbol{\beta}^*}{\partial \boldsymbol{\beta}^{*\prime}_B}\right)',$$

where

$$G_{\Delta X^*} = \frac{\partial \Delta \boldsymbol{X}^*}{\partial \boldsymbol{\beta}_A^{*\prime}} \text{ and } G_{\Delta \boldsymbol{\beta}^*} = \left[\frac{\partial \Delta \boldsymbol{\beta}^*}{\partial \boldsymbol{\beta}_A^{*\prime}} \vdots \frac{\partial \Delta \boldsymbol{\beta}^*}{\partial \boldsymbol{\beta}_B^{*\prime}}\right],$$

where $G_{\Delta X^*}$ and $G_{\Delta \boldsymbol{\beta}^*}$ are $1 \times T^*$ and $1 \times 2T^*$ vectors of gradients.

The equalities stating that asymptotic variances of the two effects are same whether we use the usual or normal equations, that is, $\boldsymbol{\sigma}^2_{\Delta X} = \boldsymbol{\sigma}^2_{\Delta X^*}$ and $\boldsymbol{\sigma}^2_{\Delta \boldsymbol{\beta}} = \boldsymbol{\sigma}^2_{\Delta \boldsymbol{\beta}^*}$, are easily proven from the equalities $G_{\Delta X^*} W = G_{\Delta X^*}$ and $G_{\Delta \boldsymbol{\beta}^*} W = G_{\Delta \boldsymbol{\beta}^*}$, and $G_{\Delta X^*} \boldsymbol{\Sigma}_{\boldsymbol{\beta}_0} G^{\prime}_{\Delta X^*} = G_{\Delta X} \boldsymbol{\Sigma}_{\boldsymbol{\beta}} G^{\prime}_{\Delta X}$ and $G_{\Delta \boldsymbol{\beta}^*} \boldsymbol{\Sigma}_{\boldsymbol{\beta}_0} G^{\prime}_{\Delta \boldsymbol{\beta}^*} = G_{\Delta \boldsymbol{\beta}} \boldsymbol{\Sigma}_{\boldsymbol{\beta}} G^{\prime}_{\Delta \boldsymbol{\beta}}$. Therefore, the hypotheses for the two effects are same whether the usual or normalized equations are used. Note that if $F(\cdot) = \boldsymbol{X}\boldsymbol{\beta}$, then the above findings are simplified to those expressions in previous two sections, **II** and **III**.

V. Empirical Illustration

We illustrate how to use the normalized equation by studying sources of changes in the labor market participation rates of white women between 1980 and 2001 using 1980 and 2001 waves of Panel Study of Income Dynamics.[11] The participation rate has risen from 69.4% to 81.9%. The first two columns of Table 1 show mean characteristics of white women in 1980 and 2001. We estimate a probit model where the dependent variable has a value of one if woman is participating labor market and zero otherwise. In this probit model we have used two sets of dummy variables, one set

[11] Participation rate is measured as the ratio of the employed to women of aged between 20 and 65.

for marriage (alas one variable, reference is not married) and the other for regions (three variables, reference is the northeast region). The last four columns of the Table 1 show both usual (with reference groups omitted) and normalized estimates of the probit model of labor market participation. As it is clear from comparing the usual and normalized estimates, the intercept and the coefficients on two sets of dummy variables are changed when transforming the usual estimates to normalized estimates.

Table 2 shows a decomposition using the probit estimates. It decomposes the predicted changes in the participation rate of -12.3% into characteristics and coefficients effects. The aggregate characteristics effect shows that 29.8% of the total changes can be attributed to the changes in characteristics between the 1980 and 2001, while 70.2% of total changes can be explain by changes in behavioral response to characteristics (changes in coefficients) between the two years. Table 2 shows that the size and asymptotic variance (or standard error) of the two aggregate effects are identical whether the usual or the normalized estimates are employed. It also shows that decomposition components related to continuous variables are not affected by using the normalized equation. The characteristics effect of sets of dummy variables is not affected by the use of the normalized equation as shown by characteristics columns of Marital Status and Regions. However, coefficients effects of Intercept, Marital Status and Regions, the focus of the identification problem in the detailed decomposition, show changes when the normalized equation is used. By using a normalized equation, we can resolve the long standing issue of the identification or the invariance problem in the detailed decomposition.

VI. Conclusion

The normalized regression can solve the invariance problem in the detailed Oaxaca decomposition. We derive simple algorithms for obtaining the estimates and their asymptotic covariance matrix for the normalized regression equation, provided that estimates and their covariance matrix for the usual regression equation are obtained. A decomposition equation utilizing the normalized regression equation is discussed and its properties are compared with those of the decomposition equation when the usual regression equation is used.

We also extend the discussion to when a non-linear equation is used for calculating the first moment. The findings with the linear regression equation can be generalized to non-linear equations where we study differences in the first moments between comparison groups. As an illustration, we decompose changes in white women's labor market participation rate between 1980 and 2001. The algorithms introduced in this paper for deriving normalized coefficients and their covariance matrix are simple and practical. Adopting these algorithms for deriving a normalized regression equation facilitates a resolution to the identification or invariance problem in the detailed decomposition.

REFERENCES

Abowd, John M. and Mark R. Killingsworth, "Do Minority/White Unemployment Differences Really Exist?," *Journal of Business and Economic Statistics*, 2:1 (1984), 64-72.

Doiron, Denis J. and W. Craig Riddell, "The Impact of Unionization on Male-Female Earnings Differences in Canada," *Journal of Human Resources*, 29:2 (1994), 504-534.

Edin, Per-Anders and Jonny Zetterberg, "Interindustry Wage Differentials: Evidence from Sweden and a Comparison with the United States," *American Economic Review*, 82:5 (1992), 1341-1349.

Even, William E. and David A. Macpherson, "Plant Size and the Decline of Unionism," *Economics Letters* 32:4 (1990), 393-398.

Even, William E. and David A. Macpherson, "The Decline of Private-Sector Unionism and the Gender Wage Gap," *Journal of Human Resources*, 28:2 (1993), 279-296.

Fairlie, Robert W., "An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models," Working paper, Economic Growth Center, Yale University (2003).

Gardeazabal, Javier and Arantza Ugidos, "More on Identification in Detailed Wage Decompositions," *Review of Economics and Statistics*, 86:4 (2004), 1034-1036.

Greene, William H. and Terry G. Seaks, "The Restricted Least Squares Estimator: A Pedagogical Note," *Review of Economics and Statistics*, 73:3 (1991), 563-567.

Haisken-DeNew, John P. and Christoph M. Schmidt, "Interindustry and Interregion Differentials: Mechanics and Interpretation," *Review of Economics and Statistics*, 79:3 (1997), 516-521.

Ham, John C., Jan Svejnar and Katherine Terrell, "Unemployment and the Social Safety Net During Transitions to a Market Economy: Evidence from the Czech and Slovak Republics," *American Economic Review*, 88:5 (1998), 1117-1142.

Krueger, Alan B. and Lawrence H. Summers, "Efficiency Wages and the Inter-Industry Wage Structure," *Econometrica*, 56:2 (1988), 259-293.

Oaxaca, Ronald L., "Male-female Wage Differentials in Urban Labor Markets," *International Economic Review*, 14:3 (1973) 693-709.

Oaxaca, Ronald L. and Michael R. Ransom, "Identification in Detailed Wage Decompositions," *Review of Economics and Statistics*, 81:1 (1999), 154-157.

Oaxaca, Roland L. and Michael Ransom, "Calculation of Approximate Variances for Wage

Decomposition Differentials," *Journal of Economic and Social Measurement*, 24:1 (1998), 55-61.

Suits, Daniel B., "Dummy Variables: Mechanics v. Interpretation," *Review of Economics and Statistics*, 66:1 (1984), 177-180.

Yun, Myeong-Su, "Hypothesis Tests when Decomposing Differences in the First Moment," *Journal of Economic and Social Measurement* (forthcoming).

Yun, Myeong-Su, "A Simple Solution to the Identification Problem in Detailed Wage Decompositions," *Economic Inquiry*, 43:4 (2005), 766-772.

Yun, Myeong-Su, "Decomposing Differences in the First moment," *Economics Letters*, 82:2 (2004), 275-280.

Table 1. Mean Characteristics and Probit Estimates (Usual and Normalized)

	Mean		Probit Estimates (Usual)		Probit Estimates (Normalized)	
	1980	2001	1980	2001	1980	2001
Intercept			-0.936**	-0.060	-1.027**	-0.221
			(0.425)	(0.841)	(0.426)	(0.835)
Not Married	0.269	0.355			0.142***	0.179***
	(0.444)	(0.478)			(0.043)	(0.060)
Married	0.731	0.645	-0.284***	-0.357***	-0.142***	-0.179***
	(0.444)	(0.478)	(0.085)	(0.121)	(0.043)	(0.060)
Northeast	0.250	0.193			-0.051	-0.018
	(0.433)	(0.395)			(0.052)	(0.085)
Midwest	0.303	0.244	0.098	0.151	0.047	0.133
	(0.460)	(0.429)	(0.081)	(0.138)	(0.049)	(0.082)
South	0.269	0.349	0.022	-0.057	-0.030	-0.074
	(0.443)	(0.477)	(0.083)	(0.124)	(0.051)	(0.069)
West	0.178	0.214	0.086	-0.024	0.034	-0.041
	(0.382)	(0.410)	(0.095)	(0.142)	(0.061)	(0.085)
Age	39.935	45.343	0.059***	0.073**	0.059***	0.073**
	(12.887)	(10.681)	(0.020)	(0.037)	(0.020)	(0.037)

$Age^2/100$	17.609	21.700	-0.105***	-0.136***	-0.105***	-0.136***
	(10.840)	(9.515)	(0.023)	(0.040)	(0.023)	(0.040)
Household Size	3.082	2.638	-0.006	0.064	-0.006	0.064
	(1.489)	(1.306)	(0.045)	(0.073)	(0.045)	(0.073)
Children aged	0.384	0.250	-0.467***	-0.469***	-0.467***	-0.469***
0 - 7	(0.719)	(0.607)	(0.062)	(0.110)	(0.062)	(0.110)
Children aged	0.640	0.400	-0.121**	-0.197**	-0.121**	-0.197**
8 - 18	(0.988)	(0.781)	(0.057)	(0.095)	(0.057)	(0.095)
Education	12.373	13.440	0.134***	0.076***	0.134***	0.076***
	(2.241)	(2.177)	(0.015)	(0.021)	(0.015)	(0.021)
Non-Labor Income	3.459	4.283	-0.052***	-0.011**	-0.052***	-0.011**
($10,000)	(3.562)	(7.945)	(0.009)	(0.005)	(0.009)	(0.005)
Participation Rate	0.694	0.819				
	(0.461)	(0.385)				
Sample Size	2214	1211	2214	1121	2214	1211

Note. a) ***, **, and * denote significance at the 1, 5 and 10 percent respectively. b) Standard deviation and standard error are reported in parenthesis for mean and probit estimates, respectively. c) Observations have been weighted using a weight variable provided by PSID. d) Non-labor income is in 1995 constant dollar.

Table 2. Decomposing Changes in Participation Rates of White Women between 1980 and 2001

	Using Probit Estimates				Using Normalized Estimates			
	Characteristics effect		Coefficients Effect		Characteristics effect		Coefficients Effect	
	Estimate	Share(%)	Estimate	Share(%)	Estimate	Share(%)	Estimate	Share(%)
Aggregate Effect	-0.037***	29.8	-0.087***	70.2	-0.037***	29.8	-0.087***	70.2
	(0.007)		(0.015)		(0.007)		(0.015)	
Intercept			-0.225	182.9			-0.207	168.2
			(0.240)				(0.239)	
Marital Status	-0.008***	6.9	0.012	-9.8	-0.008***	6.9	0.003	-2.2
	(0.002)		(0.024)		(0.002)		(0.005)	
Not Married					-0.004***	3.4	-0.003	2.7
					(0.001)		(0.007)	
Married	-0.008***	6.9	0.012	-9.8	-0.004***	3.4	0.006	-4.9
	(0.002)		(0.024)		(0.001)		(0.012)	
Regions	0.000	-0.3	0.010	-7.9	0.000	-0.3	0.001	-0.9
	(0.002)		(0.027)		(0.002)		(0.003)	
Northeast					-0.001	0.8	-0.002	1.4
					(0.001)		(0.005)	
Midwest	0.002	-1.6	-0.003	2.7	0.001	-0.8	-0.005	4.4
	(0.002)		(0.010)		(0.001)		(0.006)	

South	-0.001	0.5	0.007	-5.7	0.001	-0.7	0.004	-3.2
	(0.002)		(0.013)		(0.001)		(0.008)	
West	-0.001	0.9	0.006	-4.9	-0.000	0.4	0.004	-3.4
	(0.001)		(0.009)		(0.001)		(0.006)	
Age	-0.111***	89.9	-0.170	138.3	-0.111***	89.9	-0.170	138.3
	(0.032)		(0.494)		(0.032)		(0.494)	
$Age^2/100$	0.150***	-121.6	0.171	-138.8	0.150***	-121.6	0.171	-138.8
	(0.028)		(0.258)		(0.028)		(0.258)	
Household size	-0.001	0.8	-0.047	38.5	-0.001	0.8	-0.047	38.5
	(0.007)		(0.058)		(0.007)		(0.058)	
Children aged 0 - 7	-0.022***	17.7	0.000	-0.1	-0.022***	17.7	0.000	-0.1
	(0.004)		(0.008)		(0.004)		(0.008)	
Children aged 8 - 18	-0.010**	8.2	0.008	-6.4	-0.010**	8.2	0.008	-6.4
	(0.005)		(0.011)		(0.005)		(0.011)	
Education	-0.050***	40.3	0.200**	-162.3	-0.050***	40.3	0.200**	-162.3
	(0.006)		(0.095)		(0.006)		(0.095)	
Non-Labor Income	0.015***	-12.0	-0.044***	35.8	0.015***	-12.0	-0.044***	35.8
	(0.003)		(0.012)		(0.003)		(0.012)	

Note. a) ***, **, and * denote significance at the 1, 5 and 10 percent respectively. b) Share is percentage share of predicted changes in the participation rates of -12.3% (= 69.5 - 81.8). The observed changes are -12.5% (= 69.4 - 81.9). c) Standard errors are reported in parentheses.