ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Yun, Myeong-Su

Working Paper

Decomposition Analysis for a Binary Choice Model

Working Paper, No. 2000-01

Provided in Cooperation with:
Department of Economics, Rutgers University

Suggested Citation: Yun, Myeong-Su (2000) : Decomposition Analysis for a Binary Choice Model, Working Paper, No. 2000-01, Rutgers University, Department of Economics, New Brunswick, NJ

This Version is available at:
https://hdl.handle.net/10419/94314

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

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.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Decomposition Analysis for a Binary Choice Model

Myeong-Su Yun
Department of Economics
Rutgers University
75 Hamilton Street
New Brunswick, NJ 08901-1248

yun@economics.rutgers.edu

April 10, 2000

Abstract

This paper introduces a new and simple decomposition method for a binary choice model that is equivalent to the Blinder-Oaxaca decomposition analysis for wage differentials. The decomposition method is first developed for a single probit model and later generalized to a simultaneous equations model. Using Taylor expansion, we approximate the differences in the probabilities of choosing option 1 over option 0 between two groups in order to find the effects of the differences in "each" individual characteristic and the differences in "each" coefficient. We implement this decomposition analysis studying the racial gap in female labor market participation rates. The racial gap of participation rates among women can be almost exclusively explained by the differences in the coefficients.

JEL codes: J21, C25, C35

Keywords: decomposition analysis, binary choice, probit, selection bias, maximum likelihood estimation, labor market participation

The author wishes to thank Ira Gang, Mark Killingsworth, Roger Klein, and Hiroki Tsurumi.

1 Introduction

This paper introduces a new and simple decomposition method for a binary choice model which is equivalent to that of Blinder (1973) and Oaxaca (1973) for wage differentials.[1] In the Blinder-Oaxaca decomposition analysis for wage differentials, the observed (log) wage gap between two groups is decomposed into a part explained by differences in the average individual characteristics, and a part explained by differences in coefficients (discrimination).[2] In our decomposition method for a binary choice model, the binary choice equation is estimated by probit. Based on the probit analysis, the observed differences in the probabilities of choosing option 1 over option 0 between two groups may be explained by differences in individual characteristics and differences in probit coefficients (behavioral response or discrimination).[3] This method is helpful in answering numerous questions. For example, "Why are the participation rates in labor market different by race or gender?", or "Why have the participation rates in labor market been changed over time?", or "Who chooses a union job?", or "Who receives a job offer from a firm?", and so on.

Our contributions to the decomposition analysis for a probit model are twofold. First, we are able to find the effects of "each" individual characteristic and "each"

[1] The decomposition analysis developed in this paper is based on probit analysis. However, it can be extended to a logit model.

[2] See Becker (1971), Cain (1986) and chapter 2 of Joshi and Paci (1998) for discussions of the concept of discrimination.

[3] We may interpret the differences in coefficients between two groups as differences in the behavioral response to individual characteristics if the choice is made by her own will, or as discrimination if the choice is made by others.

coefficient using an approximation method (a first order Taylor expansion) for the differences between two groups in the probability of choosing option 1. This approximation method is introduced because the probabilities are estimated using nonlinear function (standard normal cumulative distribution function (CDF)).

Second, we extend the decomposition analysis to the case where the choice equation is estimated jointly with other equations. We measure differences in the probabilities caused by the differences in the observed individual characteristics and their coefficients when the stochastic component (unobserved individual characteristics) of the binary choice equations for both groups has the same shape (i.e., same normal distribution with mean zero and variance one), and differences in the probabilities caused by the differences in the distribution of the unobserved individual characteristics.[4]

In next section, we develop a decomposition method for a binary choice model. At first, we consider a case when probit equation is estimated independently (single probit model). Later, we extend the decomposition analysis to a case when a binary choice equation is jointly estimated with other equations (simultaneous equations model). In section 3, we illustrate the implementation of the decomposition method to racial differences in labor market participation rates using the female sample from the current population survey (CPS). The final section

[4]Yun (1999) introduces a similar approach for decomposing wage differentials when there are selection issues, the "generalized selection bias (GSB) approach" to decomposition analysis for wage differentials. It divides the wage differentials into two parts. The first are differentials predicted by the observed individual characteristics and their coefficients in the wage equation assuming the mean value of the stochastic component of wages to be zero. The second are remaining differentials representing the effects of differences in unobserved individual characteristics (selection effects).

concludes the paper.

2 Decomposition Analysis for a Binary Choice Model

First, we discuss the decomposition of a binary choice model when the choice equation is estimated independently (single probit model). Later, we further develop our decomposition analysis when the choice equation is estimated jointly with other equations (simultaneous equations model).

2.1 Basic Model: Single Probit Model

Let us assume that we observe a binary choice of individual n in group g (S_{gn}) which has value of one if a latent variable (S^*_{gn}) is positive and zero otherwise. The latent variable typically takes the following form,

$$S^*_{gn} = Z_{gn}\,\gamma_g + v_{gn} \qquad (n = 1, \ldots, n_g), \tag{1}$$

where S^*_{gn}, Z_{gn}, and v_{gn} are a latent variable, $1 \times K_S$ socio-economic characteristics, and a stochastic component of individual n in group g (a and b), respectively; γ_g is a $K_S \times 1$ vector of parameters; $v_{gn} \sim N(0, \sigma^2_{v_g})$, where $\sigma^2_{v_g} = 1$ for identification purposes.

If P_{gn} is the probability that $S_{gn} = 1$ and $(1 - P_{gn})$ is the probability that $S_{gn} = 0$, then $\mathrm{E}(S_{gn}) = P_{gn} = \Phi(Z_{gn}\gamma_g)$, where $\mathrm{E}(.)$ and Φ are expectations and a standard normal CDF. The coefficients of the binary choice equation (γ_g) are estimated by probit analysis. We can asymptotically establish the following relationship between the observed rate of choosing option 1 ($\overline{S}_g$) and the average of the computed probability of choosing option 1 using a standard normal CDF. That is,

$$\overline{S}_g = \overline{\widehat{P}}_g = \overline{\Phi(Z_g\,\widehat{\gamma}_g)}, \tag{2}$$

where $\widehat{\gamma}_g$ is a $K_S \times 1$ vector of probit coefficients, $\overline{S}_g = \sum_{n=1}^{n_g} S_{gn}/n_g$, $\overline{\widehat{P}}_g = \sum_{n=1}^{n_g} \widehat{P}_{gn}/n_g$, $\widehat{P}_{gn} = \Phi(Z_{gn}\,\widehat{\gamma}_g)$, and $\overline{\Phi(Z_g\,\widehat{\gamma}_g)} = \sum_{n=1}^{n_g} \Phi(Z_{gn}\,\widehat{\gamma}_g)/n_g$.

We decompose the differences in the probabilities of choosing option 1 over option 0 between group a and b using a simple identity to find the effects of the differences in individual characteristics and the effects of the differences in the coefficients (behavioral response or discrimination). Formally, the decomposition of the differences in probabilities between group a and b ($\overline{\widehat{P}}_a - \overline{\widehat{P}}_b$) can be shown as follows,

$$\overline{S}_a - \overline{S}_b = \left(\overline{\Phi(Z_a\,\widehat{\gamma}_a)} - \overline{\Phi(Z_b\,\widehat{\gamma}_a)}\right) + \left(\overline{\Phi(Z_b\,\widehat{\gamma}_a)} - \overline{\Phi(Z_b\,\widehat{\gamma}_b)}\right), \text{ or} \tag{3.1}$$

$$\overline{S}_a - \overline{S}_b = \left(\overline{\Phi(Z_a\,\widehat{\gamma}_b)} - \overline{\Phi(Z_b\,\widehat{\gamma}_b)}\right) + \left(\overline{\Phi(Z_a\,\widehat{\gamma}_a)} - \overline{\Phi(Z_a\,\widehat{\gamma}_b)}\right). \tag{3.2}$$

The first decomposition equation (3.1) uses an auxiliary equation $\overline{\Phi(Z_b\,\widehat{\gamma}_a)}$

while the second decomposition equation (3.2) uses $\overline{\Phi(Z_a\,\widehat{\gamma}_b)}$ in order to divide the differences in probabilities into two parts; first, differences in standard normal CDFs which represent the effects of different individual characteristics between two groups, and, second, differences in standard normal CDFs which represent the effects of different probit coefficients between two groups.

It is worth noting that what the decomposition equations, (3.1) and (3.2), provide us are aggregate measures of the effects of differences in "every" individual characteristic and the effects of differences in "every" coefficient between two groups. Because the standard normal CDF is a nonlinear function, there is no unique way to find the effects of differences in "each" individual characteristic and the effects of differences in "each" coefficient separately from equations (3.1) and (3.2).[5] We offer a modification of decomposition equations (3.1) and (3.2) which allows us to find the effects of differences in "each" individual characteristic and the effects of differences in "each" coefficient.

The modification consists of a two step approximation. In the first step, we approximate the sample average of the standard normal CDF with the standard normal CDF at sample average of the individual characteristics, i.e., $\overline{\Phi(Z_g\widehat{\gamma}_g)} \approx \Phi(\overline{Z}_g\widehat{\gamma}_g)$, where $\overline{Z}_g = \sum_{n=1}^{n_g} Z_{gn}/n_g$.[6] The equations after first step approximation

[5] See Even and Macpherson (1990) for alternative approach.

[6] The fact that $\overline{\Phi(Z_g\widehat{\gamma}_g)} \neq \Phi(\overline{Z}_g\widehat{\gamma}_g)$ comes from Jensen's inequality.

are,

$$(4.1) \qquad \overline{S}_a - \overline{S}_b = \big(\Phi(\overline{Z}_a\,\widehat{\gamma}_a) - \Phi(\overline{Z}_b\,\widehat{\gamma}_a)\big) + \big(\Phi(\overline{Z}_b\,\widehat{\gamma}_a) - \Phi(\overline{Z}_b\,\widehat{\gamma}_b)\big) + \widehat{R}_M, \text{ or}$$

$$(4.2) \qquad \overline{S}_a - \overline{S}_b = \big(\Phi(\overline{Z}_a\,\widehat{\gamma}_b) - \Phi(\overline{Z}_b\,\widehat{\gamma}_b)\big) + \big(\Phi(\overline{Z}_a\,\widehat{\gamma}_a) - \Phi(\overline{Z}_a\,\widehat{\gamma}_b)\big) + \widehat{R}_M,$$

where $\widehat{R}_M = (\overline{\Phi(Z_a\,\widehat{\gamma}_a)} - \overline{\Phi(Z_b\,\widehat{\gamma}_b)}) - (\Phi(\overline{Z}_a\,\widehat{\gamma}_a) - \Phi(\overline{Z}_b\,\widehat{\gamma}_b))$.

In the second step, we approximate the differences between two standard normal CDFs in equations (4.1) and (4.2) by a first-order Taylor expansion about $\overline{Z}_g\,\widehat{\gamma_{g'}} = \overline{Z}_g\,\widehat{\gamma}_g$, where $g \neq g'$. The final decomposition equations after the Taylor expansion are,

$$(5.1) \quad \overline{S}_a - \overline{S}_b = (\overline{Z}_a - \overline{Z}_b)\,\widehat{\gamma}_a\,\phi(\overline{Z}_a\widehat{\gamma}_a) + \overline{Z}_b\,(\widehat{\gamma}_a - \widehat{\gamma}_b)\,\phi(\overline{Z}_b\widehat{\gamma}_b) + \widehat{R}_M + \widehat{R}_{T1}, \text{ or}$$

$$(5.2) \quad \overline{S}_a - \overline{S}_b = (\overline{Z}_a - \overline{Z}_b)\,\widehat{\gamma}_b\,\phi(\overline{Z}_b\widehat{\gamma}_b) + \overline{Z}_a\,(\widehat{\gamma}_a - \widehat{\gamma}_b)\,\phi(\overline{Z}_a\widehat{\gamma}_a) + \widehat{R}_M + \widehat{R}_{T2},$$

where

$$\widehat{R}_{T1} = (\Phi(\overline{Z}_a\,\widehat{\gamma}_a) - \Phi(\overline{Z}_b\,\widehat{\gamma}_b)) - [(\overline{Z}_a - \overline{Z}_b)\,\widehat{\gamma}_a\,\phi(\overline{Z}_a\widehat{\gamma}_a) + \overline{Z}_b\,(\widehat{\gamma}_a - \widehat{\gamma}_b)\,\phi(\overline{Z}_b\widehat{\gamma}_b)], \text{ or}$$

$$\widehat{R}_{T2} = (\Phi(\overline{Z}_a\,\widehat{\gamma}_a) - \Phi(\overline{Z}_b\,\widehat{\gamma}_b)) - [(\overline{Z}_a - \overline{Z}_b)\,\widehat{\gamma}_b\,\phi(\overline{Z}_b\widehat{\gamma}_b) + \overline{Z}_a\,(\widehat{\gamma}_a - \widehat{\gamma}_b)\,\phi(\overline{Z}_a\widehat{\gamma}_a)].$$

The difference between the Blinder-Oaxaca decomposition equation for wage differentials and decomposition equations (5.1) and (5.2) for a probit model, is that, except for the residual components due to the approximation, the coefficients in equations (5.1) and (5.2) are multiplied by a standard normal probability

density function (PDF) evaluated at average characteristics.

Since we have used a two step approximation, the share explained by the differences in individual characteristics and the share explained by the differences in coefficients in equations (5.1) and (5.2) may differ from the share in equations (3.1) and (3.2). We may rescale $(\overline{Z}_a - \overline{Z}_b)\,\widehat{\gamma}_g\,\phi(\overline{Z}_g\widehat{\gamma}_g)$ and $\overline{Z}_g\,(\widehat{\gamma}_a - \widehat{\gamma}_b)\,\phi(\overline{Z}_g\widehat{\gamma}_g)$ in equations (5.1) and (5.2) in order to recover the proportions in equations (3.1) and (3.2).[7]

2.2 Simultaneous Equations Model

In this section, we extend the decomposition analysis for a binary choice model to the case where the binary choice equation is estimated with other equations. For illustration purposes, we explain the decomposition analysis using a two equation model. The multivariate extension of the two equation model is straightforward. Formally, we have following two equation model. For each group a and b, equations for individual n are,

$$(1) \qquad S^*_{gn} = Z_{gn}\,\gamma_g + v_{gn}$$

$$(6) \qquad Y^*_{gn} = X_{gn}\beta_g + e_{gn} \qquad (n = 1, \ldots, n_g),$$

[7]For example, $(\overline{Z}_a - \overline{Z}_b)\,\widehat{\gamma_a}\,\phi(\overline{Z}_a\widehat{\gamma_a})$ and $\overline{Z}_b\,(\widehat{\gamma_a} - \widehat{\gamma_b})\,\phi(\overline{Z}_b\widehat{\gamma_b})$ in equation (5.1) may be rescaled to $\left(\overline{\Phi(Z_a\,\widehat{\gamma_a})} - \overline{\Phi(Z_b\,\widehat{\gamma_a})}\right)$ and $\left(\overline{\Phi(Z_b\,\widehat{\gamma_a})} - \overline{\Phi(Z_b\,\widehat{\gamma_b})}\right)$ in equation (3.1) by multiplying $\left(\overline{\Phi(Z_a\,\widehat{\gamma_a})} - \overline{\Phi(Z_b\,\widehat{\gamma_a})}\right)/(\overline{Z}_a - \overline{Z}_b)\,\widehat{\gamma_a}\,\phi(\overline{Z}_a\widehat{\gamma_a})$ and $\left(\overline{\Phi(Z_b\,\widehat{\gamma_a})} - \overline{\Phi(Z_b\,\widehat{\gamma_b})}\right)/\overline{Z}_b\,(\widehat{\gamma_a} - \widehat{\gamma_b})\,\phi(\overline{Z}_b\widehat{\gamma_b})$, respectively.

where Z_{gn} and X_{gn} are respectively $1 \times K_S$ and $1 \times K_Y$ vectors of socio-economic characteristics of individual n in group g (a and b); coefficients γ_g and β_g are $K_S \times 1$ and $K_Y \times 1$ vectors of parameters, respectively; $\mathrm{E}(v_{gn}) = 0$, $\mathrm{E}(e_{gn}) = 0$, $\mathrm{E}(v_{gn}^2) = \sigma_{v_g}^2 = 1$, $\mathrm{E}(e_{gn}^2) = \sigma_{e_g}^2$, $\mathrm{E}(v_{gn}\, e_{gn'}) = \sigma_{e_g v_g}$ if $n = n'$ and zero if $n \neq n'$.

We observe a binary variable S_{gn} of individual n in group g which has a value of one if the latent variable S_{gn}^* is positive and zero otherwise. Y_{gn}^* is another latent variable. According to the model specification, an observed variable Y_{gn} could be a continuous or a binary variable. Since we have continuous latent variables in the model, we can conceptually derive the conditional expectations and conditional variance of error (v_{gn}) given the value of Y_{gn}. Let $\Lambda_{gn} = \mathrm{E}(v_{gn}|Y_{gn})$ and $\sigma_{cg}^2 = \mathrm{Var}(v_{gn}|Y_{gn})$ be, respectively, the conditional expectations and the conditional variance of v_{gn} given the value of Y_{gn}.[8]

Let us assume that we have consistet estimators, denoted by tilde (˜).[9] The binary choice equation (1) given the value of other equations may be written using the consistent estimators as,

$$S_{gn}^* = Z_{gn}\, \widetilde{\gamma}_g + \widetilde{\Lambda}_{gn} + \widetilde{u}_{gn}, \tag{7}$$

where $\widetilde{\gamma}_g$ is a $K_S \times 1$ vector of the consistent estimators, $\widetilde{v}_{gn} = S_{gn}^* - Z_{gn}\, \widetilde{\gamma}_g$, $\widetilde{\Lambda}_{gn} =$

[8] Λ_{gn} is called the "generalized residuals" (Gourieroux, Monfort, Renault, and Trognon (1987)).

[9] How one obtains the estimators is not crucial to our decomposition method as long as the estimators are consistent. In fact, what this paper is asking is "how can we use the already obtained consistent estimators of the binary choice equation for decomposition analysis?"

$\mathrm{E}(\widetilde{v}_{gn}|Y_{gn})$, $\widetilde{u}_{gn} = \widetilde{v}_{gn} - \widetilde{\Lambda}_{gn}$, $\mathrm{E}(\widetilde{u}_{gn}|Z_{gn}, \widetilde{\Lambda}_{gn}, Y_{gn}) = 0$, and $\mathrm{Var}(\widetilde{u}_{gn}) = \widetilde{\sigma^2}_{cg}$.[10]

The probability that $S_{gn} = 1$ with given Y_{gn} is $\widetilde{P}_{gn|Y_{gn}} = \Phi((Z_{gn}\,\widetilde{\gamma}_g + \widetilde{\Lambda}_{gn})/\widetilde{\sigma}_{cg})$, and $\mathrm{E}(S_{gn}|Y_{gn}) = \widetilde{P}_{gn|Y_{gn}}$. We can asymptotically establish following relationship,

$$\overline{S}_g = \overline{\widetilde{P}}_{g|Y_g} = \overline{\Phi\left(\frac{Z_g\,\widetilde{\gamma}_g + \widetilde{\Lambda}_g}{\widetilde{\sigma}_{cg}}\right)}, \tag{8}$$

where $\overline{S}_g = \sum_{n=1}^{n_g} S_{gn}/n_g$, $\overline{\widetilde{P}}_{g|Y_g} = \sum_{n=1}^{n_g} \widetilde{P}_{gn|Y_{gn}}/n_g$, and $\overline{\Phi((Z_g\,\widetilde{\gamma}_g + \widetilde{\Lambda}_g)/\widetilde{\sigma}_{cg})} = (1/n_g)\sum_{n=1}^{n_g} \Phi((Z_{gn}\,\widetilde{\gamma}_g + \widetilde{\Lambda}_{gn})/\widetilde{\sigma}_{cg})$.

We may divide the observed probability of choosing option 1 ($\overline{S}_g$), i.e., the average conditional expectations of S_{gn} ($\overline{\widetilde{P}}_{g|Y_g}$), into two parts; one ($\overline{\widetilde{P}}_g$) representing the average unconditional expectations of S_{gn}, and the other ($\overline{\widetilde{P}}_{g\,\widetilde{\Lambda}_g}$) representing the average effects of other equation (Y_{gn}) on the choice of S_{gn}. Asymptotically, the division is,

$$\begin{aligned} \overline{S}_g =& \overline{\widetilde{P}}_{g|Y_g} = \overline{\widetilde{P}}_g + \overline{\widetilde{P}}_{g\,\widetilde{\Lambda}_g} \\ =& \overline{\Phi(Z_g\,\widetilde{\gamma}_g)} + [\ \overline{S}_g - \overline{\Phi(Z_g\,\widetilde{\gamma}_g)}\ ]. \end{aligned} \tag{9}$$

$\overline{\widetilde{P}}_g$ can be easily computed by evaluating a univariate standard normal CDF and $\overline{\widetilde{P}}_{g\,\widetilde{\Lambda}_g}$ can be computed as the average differences between the observed probability of choosing option 1 and the unconditional expectations of S_{gn}.

The differences in the observed probability of choosing option 1 between two groups may be divided into the differences in $\overline{\widetilde{P}}_g$ and the differences in $\overline{\widetilde{P}}_{g\,\widetilde{\Lambda}_g}$. The

[10] The exact expressions of $\widetilde{\Lambda}_{gn}$ and $\widetilde{\sigma^2}_{cg}$ depend on the specification.

differences in $\overline{\widetilde{P}}_g$ (unconditional expectations of S_{gn}) represent the differences in the probability of choosing option 1 when the stochastic component, i.e., the unobserved individual characteristics, of the binary choice equation for both groups has the same distribution (normal distribution with mean zero and variance one), and the differences in $\overline{\widetilde{P}}_{g\,\widetilde{\Lambda}_g}$ represent the differences in the probability of choosing option 1 because of the difference in the distribution of the stochastic components between two groups (different means and variances).

Furthermore, the differences in unconditional expectations of S_{gn} between two groups ($\overline{\widetilde{P}}_a - \overline{\widetilde{P}}_b$) may be explained by the differences in individual characteristics observed in the binary choice equation and differences in coefficients by applying the decomposition analysis developed for the single probit model.[11] The decomposition equations for the differences in probabilities ($\overline{\widetilde{P}}_{a|Y_a} - \overline{\widetilde{P}}_{b|Y_b}$) are,

$$(10.1)\quad \overline{S}_a - \overline{S}_b = \left(\overline{\Phi(Z_a\,\widetilde{\gamma}_a)} - \overline{\Phi(Z_b\,\widetilde{\gamma}_a)}\right) + \left(\overline{\Phi(Z_b\,\widetilde{\gamma}_a)} - \overline{\Phi(Z_b\,\widetilde{\gamma}_b)}\right) + \Delta\left(\overline{\widetilde{P}}_{g\,\widetilde{\Lambda}_g}\right), \text{ or}$$

$$(10.2)\quad \overline{S}_a - \overline{S}_b = \left(\overline{\Phi(Z_a\,\widetilde{\gamma}_b)} - \overline{\Phi(Z_b\,\widetilde{\gamma}_b)}\right) + \left(\overline{\Phi(Z_a\,\widetilde{\gamma}_a)} - \overline{\Phi(Z_a\,\widetilde{\gamma}_b)}\right) + \Delta\left(\overline{\widetilde{P}}_{g\,\widetilde{\Lambda}_g}\right),$$

where $\Delta(\overline{\widetilde{P}}_{g\,\widetilde{\Lambda}_g}) = \overline{\widetilde{P}}_{a\,\widetilde{\Lambda}_a} - \overline{\widetilde{P}}_{b\,\widetilde{\Lambda}_b}$. The first differences in standard normal CDFs represent the effects of differences in individual characteristics, and the second differences in standard normal CDFs represent the effects of differences in coefficients of the binary choice equations between the two groups.

[11] One might want to decompose ($\overline{\widetilde{P}}_{a|Y_a} - \overline{\widetilde{P}}_{b|Y_b}$), not just ($\overline{\widetilde{P}}_a - \overline{\widetilde{P}}_b$), in terms of differences in individual characteristics and differences in coefficients. However, it is not clear how we can decompose $\widetilde{\Lambda}_g$ into individual characteristics and coefficients. Even we are able to do that, it is also problematic whether we can equally treat individual characteristics of $\widetilde{\Lambda}_g$ with observed individual characteristics.

The same two step approximation can be adopted for the first two components of equations (10.1) and (10.2) in order to compute the effects of differences in "each" individual characteristic in Z vector and the effects of differences in "each" coefficient in $\widetilde{\gamma}$. The final equations are,

$$\begin{aligned}
(11.1) \qquad \overline{S}_a - \overline{S}_b =& (\overline{Z}_a - \overline{Z}_b)\, \widetilde{\gamma}_a\, \phi(\overline{Z}_a \widetilde{\gamma}_a) + \overline{Z}_b\, (\widetilde{\gamma}_a - \widetilde{\gamma}_b)\, \phi(\overline{Z}_b \widetilde{\gamma}_b) \\
&+ \widetilde{R}_M + \widetilde{R}_{T1} + \Delta\left(\overline{\widetilde{P}}_{g\,\widetilde{\Lambda}_g}\right), \text{ or} \\
(11.2) \qquad \overline{S}_a - \overline{S}_b =& (\overline{Z}_a - \overline{Z}_b)\, \widetilde{\gamma}_b\, \phi(\overline{Z}_b \widetilde{\gamma}_b) + \overline{Z}_a\, (\widetilde{\gamma}_a - \widetilde{\gamma}_b)\, \phi(\overline{Z}_a \widetilde{\gamma}_a) \\
&+ \widetilde{R}_M + \widetilde{R}_{T2} + \Delta\left(\overline{\widetilde{P}}_{g\,\widetilde{\Lambda}_g}\right),
\end{aligned}$$

where

$$\begin{aligned}
\widetilde{R}_M =& (\overline{\Phi(Z_a\, \widetilde{\gamma}_a)} - \overline{\Phi(Z_b\, \widetilde{\gamma}_b)}) - (\Phi(\overline{Z}_a\, \widetilde{\gamma}_a) - \Phi(\overline{Z}_b\, \widetilde{\gamma}_b)), \\
\widetilde{R}_{T1} =& (\Phi(\overline{Z}_a\, \widetilde{\gamma}_a) - \Phi(\overline{Z}_b\, \widetilde{\gamma}_b)) - [(\overline{Z}_a - \overline{Z}_b)\, \widetilde{\gamma}_a\, \phi(\overline{Z}_a \widetilde{\gamma}_a) + \overline{Z}_b\, (\widetilde{\gamma}_a - \widetilde{\gamma}_b)\, \phi(\overline{Z}_b \widetilde{\gamma}_b)], \text{ and} \\
\widetilde{R}_{T2} =& (\Phi(\overline{Z}_a\, \widetilde{\gamma}_a) - \Phi(\overline{Z}_b\, \widetilde{\gamma}_b)) - [(\overline{Z}_a - \overline{Z}_b)\, \widetilde{\gamma}_b\, \phi(\overline{Z}_b \widetilde{\gamma}_b) + \overline{Z}_a\, (\widetilde{\gamma}_a - \widetilde{\gamma}_b)\, \phi(\overline{Z}_a \widetilde{\gamma}_a)].
\end{aligned}$$

We may rescale $(\overline{Z}_a - \overline{Z}_b)\, \widetilde{\gamma}_g\, \phi(\overline{Z}_g \widetilde{\gamma}_g)$ and $\overline{Z}_g\, (\widehat{\gamma}_a - \widetilde{\gamma}_b)\, \phi(\overline{Z}_g \widetilde{\gamma}_g)$ in equations (11.1) and (11.2) in order to recover the proportions in equations (10.1) and (10.2).

In this section, we have introduced the decomposition method for a binary choice model when the binary choice equation is estimated independently and jointly with other equations.

3 Racial Difference in Female Labor Market Participation

We apply the decomposition method to racial differences in labor market participation rates using the female sample from the March 1995 CPS. At first, we estimate the labor market participation equation (single equation model). Later, we jointly estimate the labor market participation and log-wages equations (simultaneous equations model).

3.1 Data

The female sample used here is drawn from the March 1995 CPS. The data comes from the outgoing rotation group only, and the responses to questions about the survey month are used rather than those for last year.[12] The sample includes females aged between 25 and 60 who were not in school, retired, disabled or self-employed. For married women, we exclude those whose husbands are under 25 years old. We also exclude women whose hourly wage rate is greater than $40, or whose working hours are top-coded (99 hours per week). Table 1 describes the variables used for our study.

Table 2 shows the means and standard deviations of variables used in the decomposition analysis. The characteristics of working women are different from

[12]Information on last year's earnings are used to compute non-labor income. For details, see Table 1.

those of non-working women. Working women are older and have more years of education than non-working women. Non-working women have a higher marriage rate among white women but there is little difference among other race women. Non-working women have more children (for both age under 6 and between age 6 and 18) and larger family size. Non-working women also have a higher non-labor income among white women, but there is not much difference among other race women, which might be related to the marriage rates.

White women are more educated than other race women in both the non-working and working samples. Though white women have a higher rate of marriage, family size of white women is smaller than that of other race women. The number of children is not significantly different between white and other race women. White women, especially non-working white women, have higher non-labor income than other race women do.

Though both white and other race women are working similar hours, their wages (measured both in level and log) are significantly different from each other according to the t-test at 5% (level wage) and 1% level (log-wage), respectively. Yun (1999) explains racial wage differentials in terms of differences in individual characteristics and differences in the coefficients. In this paper, we focus on racial differences in participation rates. The participation rates are significantly different from each other according to the t-test at 1% level. The participation rate of whites is 81.85% and that of others is 75.28%.[13]

[13]Blau (1998) reports that white females have a higher participation rate than do black females in 1995. However, historically black women had a higher participation rate than white women. See Blau (1998) and Chapter 2 in Goldin (1990).

In this paper, we choose maximum likelihood estimation (MLE) method to obtain the consistent estimators. The MLE method for estimation is implemented using both Gauss CML (constrained maximum likelihood) program and the SAS NLP (non linear programming) procedure (SAS Institute, 1997). The likelihood function for the simultaneous equations model (participation and log-wage equations) is omitted here as this model is well known.[14]

3.2 Female Labor Market Participation

In this section, we implement our decomposition method for racial differences in women's labor market participation rates.[15] Women are partitioned into two groups according to their race, whites ($g = w$) and other races ($g = o$). Hence, groups a and b in section 2 are whites (w) and other races (o), respectively. We estimate the labor market participation equation (1) for the single equation model. Later, we jointly estimate the labor market participation (1) and log-wages (6) equations for the simultaneous equations model. Women will participate in the labor market ($S_{gn} = 1$) if S^*_{gN} in equation (1) has a positive value. Log-wages (Y_{gn}) are equal to latent log-wages (Y^*_{gn}) if S^*_{gn} is positive and missing otherwise.

Table 3 shows the estimates of the participation equation.[16] The first and

[14]See Heckman (1974), Mroz (1987), and Zabel (1993) for details.

[15]Participation is usually defined to include both the employed and the unemployed. However most studies of labor supply do not count unemployment in the definition of participation. We treat unemployment as non-participation to keep the analysis simple. Blundell, Ham and Meghir (1987) is a rare exception. They include unemployment in the definition of participation.

[16]The estimates of the log-wage equation are reported in Yun (1999).

second columns of Table 3 report the coefficients of the participation equation (1) when it is estimated independently and jointly with log-wages, respectively. The coefficients of participation show the expected signs; education increases participation, and the presence of children decreases participation. Only the marriage variable has an unexpected positive sign. The estimate of the marriage coefficient is not significant in white women but is significant at 5% in other race women.

Tables 4 and 5 show the decomposition of racial differences in participation choice when single probit model and simultanenous model are used, respectively. According to Tables 4 and 5, the racial differences in participation rates are explained almost exclusively by the differences in coefficients.[17] Especially, the effects of differences in coefficients of the two age variables (age and age^2/100) are very large. We can infer that the difference of participation rates will decrease at least by 0.2 if both race women had the same age-participation profiles. 0.2 is extremely large number, much larger than the observed differentials of participation rates (0.066).

In addition to two age variables, we consider the roles of education, marital status, variables related to family composition (number children under age 6, number of children age 6 -18, and family size), and non-labor income. First, the

[17] In equations (3.1) and (3.1), the single probit model, the differences in individual characteristics explain -0.004 (or 0.003) of total differences 0.066, while the the differences in coefficients explain 0.07 (or 0.063) of total differences 0.066. In equations (10.1) and (10.2), the simultaneous equations model, the differences in individual characteristics explain -0.005 (or 0.001) of total differences 0.066, while the differences in coefficients explain 0.071 (or 0.065) of total differences 0.066.

attainment of higher education levels by white women (about a half year) contributes to the higher participation rate of white women, but the fact that other race women tend to participate more once they acquire education contributes to reducing the differences in participation rates between white and other race women. In sum, education variable contributes to reducing the differences in participation rates between white and other race women.

Second, the coefficients for marital status of both white and other race women are positive, though white women's coefficients are not significant. The higher marriage rate of white women contributes to increasing the racial gap of participation rates, but the magnitude is small. However, the differences in coefficients contribute to reducing the racial gap of participation rates since other race women have larger coefficients. Marital status also contributes to reducing the differences in participation rates between white and other race women.

Third, white women tend to participate labor market less if they have children under age 6. The differences in coefficients contribute to reducing the racial gap in participation rates. The effects of differences in coefficients are the larger effects of the number of children under age 6. The effects of the differences in the number of children under age 6 are negligible, because the number of children under age 6 is not significantly different between two groups as Table 2 shows. The contribution of number of children age 6 to 18 (from the viewpoint of differences in both levels and coefficients) to the differences in participation rates is negligible. Family size also affects the racial gap through the differences in coefficients. The coefficients

on family size by race contribute to reducing the racial gap of participation rates, since white women tend to participate less given the size of the family.

Fourth, the fact that white women have relatively larger non-labor income contributes to reducing the racial differentials in participation rates. In contrast, the larger negative coefficient for non-labor income of other race women relative to that of white women contributes to increasing the racial gap. Non-labor income is the only variable whose contribution from the differences in levels outweighs that from the differences in coefficients.

In summary, the effects of differences in coefficients outweigh those of differences in individual characteristics. In fact, the racial gap in women's participation rates is almost exclusively explained by the effects of differences in coefficients, especially by coefficients of two age variables. Other variables contribute to reducing the gap between two racial groups.

4 Conclusion

This paper introduces a new and simple decomposition method for a binary choice model. We decompose the differences in probabilities of choosing option 1 over option 0 between two groups into a part explained by differences in the individual characteristics and a part explained by differences in coefficients of the binary choice equation (behavioral response or discrimination).

We introduce an approximation method for the differences in probabilities

between two groups in order to find the effects of differences in "each" individual characteristic and the effects of differences in "each" coefficient. The approximation method is introduced because the probabilities are measured using a nonlinear function, the standard normal CDF.

We also extend the decomposition analysis to the case where the choice equation is estimated jointly with other equations. We divide the difference in observed rates of choosing option 1 into the differences in unconditional probabilities and the remaining discrepancies between differences in observed rates of choosing option 1 and differences in unconditional probabilities. Later we apply the approximation method to the differences in unconditional probabilities between two groups in order to find the effects of differences in "each" individual characteristic and the effects of differences in "each" coefficient.

We implement the decomposition analysis to study the racial gap in female participation rates into the labor market using single probit model and simultaneous equations model (participation and log-wage equations). The racial differentials of participation rates among women can be almost exclusively explained by the differences in behavioral response to individual characteristics (i.e., differences in coefficients). The effects of difference in coefficients of age variables are very large, which requires further study.

This method will be useful in studying the differences of behavior between two groups, e.g., racial or gender gap of any binary choice. Especially, the generalization of the decomposition analysis to the simultaneous equations model broadens

the scope of application. This generalization will bear fruit especially when we want to evaluate various aspects of gender or racial gap at the same time.[18]

REFERENCES

Becker, Gary S. (1971): *The Economics of Discrimination*, second edition. Chicago: University of Chicago Press.

Blau, Francine D. (1998): "Trends in the Well-Being of American Women, 1970-1995," *Journal of Economic Literature*, 36, 112-165.

Blinder, Alan S. (1973): "Wage Discrimination: Reduced Form and Structural Estimates," *Journal of Human Resources*, 8, 436-455.

Blundell, Richard, John Ham, and Costas Meghir (1987): "Unemployment and Female Labour Supply," *Economic Journal*, 97, 44-64.

Cain, Glen G. (1986): "The Economic Analysis of Labor Market Discrimination: A Survey," in *Handbook of Labor Economics, Vol. I*, ed. by Orley Ashenfelter and Richard Layard. Amsterdam: Elsevier Science B.V., 693-785.

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

Gang, Ira N. and Myeong-Su Yun (2000): "The Gender Gap during Rapid Transition," manuscript, Rutgers University.

Goldin, Claudia (1990): *Understanding the Gender Gap: An Economics History of American Women*, New York: Oxford University Press.

Gourieroux, Christian, Alain Monfort, Eric Renault, and Alain Trognon (1987): "Generalized Residuals," *Journal of Econometrics*, 34, 5-32.

Heckman, James (1974): "Shadow Prices, Market Wages, and Labor Supply," *Econometrica*, 42, 679-694.

Joshi, Heather and Pierella Paci (1998): *Unequal Pay for Women and Men: Evidence from the British Birth Cohort Studies*, Cambridge: MIT press.

Mroz, Thosmas A. (1987): "The Sensitivity of an Empirical Model of Marred Women's Hours of Work to Economic and Statistical Assumptions," *Econometrica*, 55, 765-799.

[18]See Gang and Yun (2000) for an implement of the decomposition analysis in studying changes in the gender gaps of wages and participation rates in the former German Democratic Republic the unification.

Oaxaca, Roland (1973): "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 14, 693-709.

SAS Institute (1997): *SAS/OR Technical Report: The NLP Procedure*, Cary, NC: SAS Institute Inc.

Yun, Myeong-Su (1999): " Generalized Selection Bias and The Decomposition of Wage Differentials," Working Paper, No. 99-11, Department of Economics, Rutgers University.

Zabel, Jeffrey E. (1993): "The Relationship between Hours of Work and Labor Force Participation in Four Models of Labor Supply Behavior," *Journal of Labor Economics*, 11, 387-416.

Table 1: Variables Used in the Analysis

Variables	Definition and Note
Age	Aged 25 – 60 years.
$\text{Age}^2/100$	Age squared in hundreds.
Education	Number of years of schooling.
Marriage	Married = 1, Single = 0. Married but spouse absent is treated as single.
Children < 6	Number of children under age 6.
Children 6–18	Number of children age 6 – 18.
Family Size	Number of family members.
Non-Labor Inc.	Sum of last year's survivor's income, interest income, dividends income, rent income, child support payment, alimony. If married, husband's annual wage of last year is added. Unit is $1000.
MSA	Metropolitan statistical areas = 1, Else = 0.
West	West region = 1, Else = 0.
South	South region = 1, Else = 0.
Midwest	Midwest region = 1, Else = 0.
Northeast	Reference region.
Wages ($)	Hourly wage rate (level) = usual weekly earnings / usual weekly hours of work.
Hours	Usual weekly hours of work.

Table 2: Mean Characteristics of the Sample

	Whites		Others	
	Mean	(s.d.)	Mean	(s.d.)
Whole Sample				
Age	39.694	(9.371)	39.088	(9.368)
Education	13.290	(2.642)**	12.832	(2.648)
Marriage	0.615	(0.487)**	0.403	(0.491)
Children < 6	0.294	(0.609)	0.318	(0.656)
Children 6–18	0.618	(0.925)*	0.706	(0.967)
Family Size	2.919	(1.439)**	3.125	(1.600)
Non-Labor Inc.	25.154	(27.218)**	13.267	(20.746)
Participation Rate	0.818	(0.385)**	0.753	(0.432)
Sample size	3829		894	
Non-Working Sample				
Age	38.576	(9.434)	37.701	(9.134)
Education	12.326	(2.933)*	11.814	(2.996)
Marriage	0.755	(0.430)**	0.398	(0.491)
Children < 6	0.622	(0.842)	0.534	(0.882)
Children 6–18	0.799	(1.017)	0.928	(1.122)
Family Size	3.555	(1.483)	3.561	(1.743)
Non-Labor Inc.	33.479	(31.286)**	14.142	(23.850)
Sample size	695		221	
Working Sample				
Age	39.942	(9.340)	39.544	(9.406)
Education	13.503	(2.524)**	13.166	(2.435)
Marriage	0.584	(0.493)**	0.404	(0.491)
Children < 6	0.222	(0.517)	0.247	(0.544)
Children 6–18	0.578	(0.899)	0.633	(0.900)
Family Size	2.778	(1.390)**	2.982	(1.524)
Non-Labor Inc.	23.308	(25.876)**	12.979	(19.631)
MSA	0.737	(0.440)**	0.816	(0.388)
West	0.185	(0.388)	0.198	(0.399)
South	0.280	(0.449)**	0.441	(0.497)
Midwest	0.268	(0.443)**	0.160	(0.367)
Wages	11.723	(6.120)*	11.070	(6.229)
Log-wages	2.331	(0.526)**	2.254	(0.576)
Hours	37.594	(9.024)	37.960	(7.482)
Sample size	3134		673	

[a] ** and * imply that the null hypothesis, mean of white women is equal to that of other race women, is rejected at 1% and 5% level of significance, respectively.

Table 3: Participation Choice

	Whites			
	Single		Simultaneous	
	Est.	(s.e.)	Est.	(s.e.)
Constant	-1.626^{**}	(0.525)	-1.646^{**}	(0.525)
Age	0.086^{**}	(0.026)	0.085^{**}	(0.026)
$\text{Age}^2/100$	-0.108^{**}	(0.032)	-0.107^{**}	(0.032)
Education	0.119^{**}	(0.010)	0.120^{**}	(0.010)
Marriage	0.054	(0.077)	0.071	(0.076)
Children < 6	-0.475^{**}	(0.050)	-0.478^{**}	(0.049)
Children 6–18	-0.072	(0.040)	-0.067	(0.040)
Family Size	-0.053	(0.030)	-0.049	(0.030)
Non-Labor Inc.	-0.009^{**}	(0.001)	-0.010^{**}	(0.001)
	Others			
	Single		Simultaneous	
	Est.	(s.e.)	Est.	(s.e.)
Constant	-1.418	(0.945)	-1.411	(0.945)
Age	0.025	(0.046)	0.022	(0.047)
$\text{Age}^2/100$	-0.015	(0.057)	-0.010	(0.057)
Education	0.129^{**}	(0.020)	0.132^{**}	(0.021)
Marriage	0.293^{*}	(0.147)	0.296^{*}	(0.147)
Children < 6	-0.239^{**}	(0.082)	-0.244^{**}	(0.082)
Children 6–18	-0.076	(0.063)	-0.065	(0.065)
Family Size	-0.024	(0.042)	-0.021	(0.042)
Non-Labor Inc.	-0.011^{**}	(0.003)	-0.012^{**}	(0.004)

[a] ** and * mean statistically significant at 1% and 5%, respectively.

Table 4: Decomposition of Participation Choice:

Single Probit Model

	Decomposition 1			
	Diff. in Characteristics		Diff. in Coefficients	
Constant			−0.063	(−95.77)
Age	0.012	(18.55)	0.723	(1091.21)
Age^2/100	−0.012	(−18.44)	−0.458	(−690.82)
Education	0.013	(19.45)	−0.040	(−59.69)
Marriage	0.003	(4.08)	−0.029	(−44.26)
Children < 6	0.003	(3.96)	−0.023	(−34.31)
Children 6–18	0.001	(2.25)	0.001	(1.45)
Family Size	0.003	(3.89)	−0.027	(−40.61)
Non-Labor Inc.	−0.026	(−39.59)	0.007	(10.60)
Sum	−0.004	(−5.85)	0.091	(137.82)
	Decomposition 2			
	Diff. in Characteristics		Diff. in Coefficients	
Constant			−0.049	(−74.60)
Age	0.005	(6.90)	0.572	(863.18)
Age^2/100	−0.002	(−3.21)	−0.367	(−554.06)
Education	0.018	(27.10)	−0.032	(−48.15)
Marriage	0.019	(28.54)	−0.035	(−52.63)
Children < 6	0.002	(2.56)	−0.016	(−24.77)
Children 6–18	0.002	(3.07)	0.001	(0.99)
Family Size	0.002	(2.31)	−0.020	(−29.54)
Non-Labor Inc.	−0.040	(−60.33)	0.010	(15.66)
Sum	0.005	(6.95)	0.064	(96.09)

[a] Decomposition 1 and 2 correspond to equations (5.1) and (5.2), respectively.

[b] Shares of $\overline{\Phi(Z_w\,\widehat{\gamma_w})} - \overline{\Phi(Z_o\,\widehat{\gamma_o})}$ (=0.066) are reported in parentheses in percentage.

[c] $\widehat{R}_M = -0.011$.

[d] $\widehat{R}_{T1} = -0.011$ and $\widehat{R}_{T2} = 0.009$ for decomposition 1 and 2, respectively.

Table 5: Decomposition of Participation Choice:

Simultaneous Equations Model

	Decomposition 1			
	Diff. in Characteristics		Diff. in Coefficients	
Constant			−0.071	(−107.85)
Age	0.012	(18.43)	0.753	(1138.00)
Age^2/100	−0.012	(−18.28)	−0.473	(−715.73)
Education	0.013	(19.74)	−0.045	(−68.75)
Marriage	0.004	(5.36)	−0.028	(−41.77)
Children < 6	0.003	(4.00)	−0.023	(−34.13)
Children 6–18	0.001	(2.11)	−0.0004	(−0.66)
Family Size	0.002	(3.59)	−0.026	(−39.87)
Non-Labor Inc.	−0.028	(−42.18)	0.007	(10.42)
Sum	−0.005	(−7.25)	0.092	(139.66)
	Decomposition 2			
	Diff. in Characteristics		Diff. in Coefficients	
Constant			−0.056	(−84.03)
Age	0.004	(6.01)	0.596	(900.43)
Age^2/100	−0.001	(−2.26)	−0.380	(−574.19)
Education	0.018	(27.79)	−0.037	(−55.48)
Marriage	0.019	(28.88)	−0.033	(−49.68)
Children < 6	0.002	(2.62)	−0.016	(−24.64)
Children 6–18	0.002	(2.62)	−0.0002	(−0.45)
Family Size	0.001	(1.97)	−0.019	(−29.02)
Non-Labor Inc.	−0.042	(−63.47)	0.010	(15.40)
Sum	0.003	(4.15)	0.065	(98.33)

[a] Decomposition 1 and 2 correspond to equations (11.1) and (11.2), respectively.

[b] Shares of $\overline{\Phi(Z_w\,\widetilde{\gamma_w})} - \overline{\Phi(Z_o\,\widetilde{\gamma_o})}$ (=0.066) are reported in parentheses in percentage.

[c] $\overline{\widetilde{P}}_{a\,\widetilde{\Lambda}_a} - \overline{\widetilde{P}}_{b\,\widetilde{\Lambda}_b} = -0.0004$.

[d] $\widetilde{R}_M = -0.011$.

[e] $\widetilde{R}_{T1} = -0.011$ and $\widetilde{R}_{T2} = 0.009$ for decomposition 1 and 2, respectively.