DISCUSSION PAPER SERIES

IZA DP No. 69

Generalized Selection Bias and The Decomposition of Wage Differentials

Myeong-Su Yun

November 1999

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Generalized Selection Bias and The Decomposition of Wage Differentials

Myeong-Su Yun
Rutgers University, New Brunswick, NJ, USA

Discussion Paper No. 69
November 1999

IZA

P.O. Box 7240
D-53072 Bonn
Germany

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

IZA Discussion Paper No. 69
November 1999

ABSTRACT

Generalized Selection Bias and The Decomposition of Wage Differentials*

The major contribution of this paper is ending a new and flexible way to measure the effects of selection on log-wages. In this context, we offer a general approach to performing decomposition analysis when selection effects are present. We call the difference between unconditional and conditional expectations of the log-wages a generalized selection bias (GSB) when the two expectations are measured using the estimates from the joint estimation of the whole model (log-wages and selection equations) by the MLE method. The unconditional and conditional expectations are, respectively, the deterministic component of log-wages, and the deterministic component plus the conditional expectation of the stochastic component of log-wages, where the deterministic component is computed using the estimates from the joint estimation. That is, the GSB is the expectation of the residuals estimated from the joint estimation. It is appropriate to apply the Blinder-Oaxaca decomposition method to the wage differentials adjusted for the GSB. The GSB approach to decomposition analysis is not only easy to implement and flexible enough to apply to virtually any kind of selection issue, but also efficient because it uses full information. We illustrate the GSB approach by applying it to the racial wage differentials among women using data from the Current Population Survey. We discuss the possibility of using semi-parametric or Bayesian sampling method for the joint estimation and related modifications of decomposition analysis.

JEL Classification: J71, J31, C34

Keywords: Decomposition analysis of wage differentials, discrimination, generalized selection bias, maximum likelihood estimation, Heckman's two-step method, semi-parametric, Bayesian sampling

Myeong-Su Yun
Department of Economics
Rutgers University
75 Hamilton Street
New Brunswick, NJ 08901-1248
Email: msyun@rci.rutgers.edu

* The author wishes to thank Mark Killingsworth, Roger Klein, Hiroki Tsurumi, and Frank Vella. Special thanks to Ira Gang for his generous assistance and encouragement in writing this paper.

1. INTRODUCTION

The decomposition method introduced by Blinder (1973) and Oaxaca (1973) has been widely adopted in the analysis of wage differentials.[1] In this approach, log-wages are regressed on various socio-economic characteristics. Based on the regression analysis, the observed log-wage gap of average workers is decomposed into a part explained by difference in the average characteristics, and a part explained by difference in coefficients, traditionally labeled "discrimination."[2]

It is well recognized that sample selection causes bias in the OLS coefficients of log-wages.[3] A decomposition analysis which does not take account of sample selection, therefore, could over- or underestimate "true" discrimination. Studies on wage differentials and discrimination have adopted the well-known Heckman's two-step method (selection bias correction method) to obtain consistent estimates of log-wage parameters and applied the Blinder-Oaxaca decomposition method to the log-wage differentials adjusted for the selection bias.[4] We call the decomposition analysis which relies on Heckman's two-step method a "selection bias correction (SBC) approach."

However, previous studies which have adopted the SBC approach are limited

[1]Though we discuss the decomposition method in terms of (hourly) wages in this paper, our argument still holds when wage is replaced with (monthly) earnings. We also follow the convention of analyzing log-wages rather than level wages. Our argument can be applied to level wages with minor modification.

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

[3]We interpret sample selection in broad sense; it includes not only the self-selection issue but also any kind endogeneity caused by censoring, truncation, etc.

[4]See, for example, Bloom and Killingsworth (1982) and Joshi and Paci (1998). See Heckman (1979) for his two-step method.

to models with relatively simple selection, usually a single selection (e.g., participation vs. non-participation). There are numerous studies on various selection issues (e.g., double selection model, censored or truncated regression model, switching regression model) which have not used yet but may be used in the decomposition analysis of wage differentials. Some studies have used only Heckman's two-step method or only the maximum likelihood estimation (MLE) method or both to take selection issues into account. In practice, it is often the MLE method, not Heckman's two-step method, which is used for the complicated selection models. Heckman's two-step method might have conceptual or practical difficulties in handling these complicated selection issues for which the MLE method is usually used. By developing a general way to use the MLE method in the context of the decomposition analysis, we substantially broaden the scope of the decomposition analysis. Furthermore, the decomposition analysis using the MLE method will be more efficient than that using Heckman's two-step method due to the efficiency of the MLE method, even when the selection issues can be handled by both Heckman's two-step and the MLE methods.

What is a proper way to apply the MLE method to the decomposition analysis? In this paper, we introduce a new approach to decomposing wage differentials based on joint estimation of log-wages and selection equations using the MLE method. It is well-known that joint estimation using the MLE method gives consistent estimates of log-wage parameters by taking correlation between log-wages and selection equations into account. Obtaining the consistent esti-

mates of log-wage parameters using the MLE method is just a part of the proper decomposition analysis. As a second part of the proper decomposition analysis, we show a simple and general way to compute selection bias which can be used as long as consistent estimates of log-wage parameters are available. The selection bias is the difference between conditional and unconditional expectations of the log-wages (Rosen (1986), p. 654) which is the source of bias of the OLS estimates of log-wages. The conditional expectations of the log-wages are the expectations of the log-wages given that the condition of being selected in the sample is satisfied. By combining consistent estimates of log-wages and selection bias, we can decompose the observed wage differentials properly. The intuition behind our new approach is remarkably simple.

In general, the sample expectations of any random variable should be the same as the population expectations of the variable if the sample is randomly sampled from the population. If there is any discrepancy between them, we suspect non-randomness of the sample, i.e., the existence of selection bias, provided that distribution of the variable is correctly specified to infer the population expectations. In practice, we are interested in a variable which contains deterministic and stochastic components. In this paper, this variable is log-wages.

Since only the stochastic component may have conditional expectations different from unconditional (population) expectations, the selection bias of a random variable is identical to that of the stochastic component. The stochastic component of the random variable can be computed by subtracting the deterministic

component from the variable. Since we do not know the value of the deterministic component of the variable, we have to recover it by estimating a structural equation which is supposed to predict the value of the deterministic component. The estimation of the structural equation itself, using only the selected sample by, for example, OLS may mislead us to a wrong measurement of both deterministic and stochastic components because the conditional expectations of the stochastic component provided being selected in the sample may be different from the unconditional expectations of the stochastic component. That is the reason why we jointly estimate the structural equation with sample selection equations which are supposed to affect the selection of the sample. By taking the correlation among the structural and selection equations, the source of the selection bias, into account, we can obtain consistent estimates of the structural equation parameters and measure the deterministic and stochastic components of the random variable.

Without loss of generality, we may assume the stochastic component of the random variable follows a distribution whose expectations are zero, e.g., the normal distribution. If unconditional expectations are assumed to be zero, the selection bias can be easily computed by taking expectations of the residuals which are equal to the difference between value of the random variable and its deterministic component. We call the expectations of residuals a "generalized selection bias" (GSB) if the residuals are computed using the consistent estimates of the structural equation parameters from the joint estimation.[5] In this paper, the MLE

[5]Computing selection bias via residuals requires only that the estimates of the structural equation parameters are consistent. This implies that the consistent estimates from Heckman's two-step method can be used to compute selection bias via computing residuals. Our definition

method is used for the joint estimation.[6]

We can apply the Blinder-Oaxaca decomposition method to the log-wage differentials adjusted for the GSB, i.e., the log-wage differentials measured using the structural equations for two groups representing only the differentials of the population (unconditional) expectations. We call the decomposition method which uses the joint estimation of the structural and selection equations and measures selection bias from the residuals computed using the consistent estimates from the joint estimation a "generalized selection bias (GSB) approach."

For exposition purposes, in this paper, we assume that log-wages and selection equations follow a joint normal distribution when we discuss the SBC and GSB approaches. This distributional assumption, joint normality among log-wages and selection equations, is not crucial to our discussion. Our discussion can be easily extended to semi- or nonparametric framework with minor modification.[7]

In next section, we discuss the econometrics of the GSB approach to decomposition (discrimination) analysis in detail. In section 3, we illustrate the implementation of the GSB approach to racial wage differentials using the Current Population Survey. The final section concludes the paper with comments on the possibility of using various joint estimation methods (semiparametric, Bayesian

---

of the GSB, however, also requires the joint estimation of the structural and selection equations. We define the GSB in this way because the selection bias can be directly measured using selection bias correction terms and their coefficients at the second step if Heckman's two-step method is employed.

[6]The choice of estimation technique itself is not crucial as long as the joint estimation can obtain consistent estimates of the structural equation parameters.

[7]See Powell (1994) and Härdle and Linton (1994) for discussions of semi- and nonparametric methods.

sampling) and their implications for decomposition analysis.

2. GSB APPROACH TO DECOMPOSITION ANALYSIS

The foundation of conventional decomposition analysis is a regression of log-wages.[8] We estimate the log-wage function for group g which typically takes the following form,[9]

$$(1) \qquad Y_{gn} = X_{gn}\beta_g + e_{gn} \qquad (n = 1, \ldots, n_g),$$

where Y_{gn}, X_{gn}, and e_{gn} are log-wages, $1 \times K_Y$ socio-economic characteristics, and error of individual n in group g (a and b), respectively; β_g is $K_Y \times 1$ vector of parameters; $\mathrm{E}(e_{gn}) = 0$.

Based on the OLS analysis, the conventional decomposition analysis uses a simple identity to compute the portion of wage differentials between group a and b explained by the difference in average characteristics and by the difference in the coefficients. Formally, the decomposition of the difference in log-wages between

[8]The Blinder-Oaxaca type decomposition analysis without selection issues is called a conventional decomposition analysis in this paper.

[9]We emphasize that this is an equation for the observed log-wages, which may or may not be randomly selected from the population. In the conventional decomposition analysis, the sample is assumed to be randomly sampled from the population.

group a and b can be shown as follows,[10]

$$\overline{Y}_a - \overline{Y}_b = \begin{cases} (\overline{X}_a - \overline{X}_b)\widehat{\beta}_b + \overline{X}_a(\widehat{\beta}_a - \widehat{\beta}_b), & \text{or} \\ (\overline{X}_a - \overline{X}_b)\widehat{\beta}_a + \overline{X}_b(\widehat{\beta}_a - \widehat{\beta}_b), \end{cases} \tag{2}$$

where $\overline{Y}_g$, $\overline{X}_g$, and $\widehat{\beta}_g$ are, for each group g (a and b), the sample average of log-wages ($\sum_{n=1}^{n_g} Y_{gn}/n_g$), $1 \times K_Y$ vector of the average characteristics, and $K_Y \times 1$ vector of OLS coefficients, respectively.

From equation (2), we compute two discrimination coefficients (see Oaxaca (1973) for details), depending on which group's characteristics are used as weights,

$$\mathrm{D}_g = \exp\big(\overline{X}_g(\widehat{\beta}_a - \widehat{\beta}_b)\big) - 1, \tag{3}$$

where $g = a$ or b.

The crucial assumption of the conventional decomposition analysis is that the expectations of e_{gn} are zero, which makes the identity of equation (2) hold. That is, the sample of each group is randomly selected from the population. However, the expectations of e_{gn} might not be zero when the sample is not randomly selected from population. The violation of a mean zero assumption results in biased OLS estimates. This, in turn, implies that the conventional decomposition analysis that does not take account of the bias of the estimates may lead us to a wrong

[10]Another issue in the discrimination literature is that the wage differentials between two groups consists of two parts: the gain above the nondiscriminatory (competitive) wage and the loss below the nondiscriminatory wage. The nondiscriminatory wage is usually estimated using pooled data. See Oaxaca and Ransom (1988, 1994) and Neumark (1988). This issue can be easily incorporated in the GSB approach by estimating log-wages and selection equations jointly using pooled data to compute the nondiscriminatory wage.

measurement of the sources of wage differentials.

We discuss two approaches to decomposition analysis when there are selection issues: the SBC approach based on Heckman's two-step method and the GSB approach based on joint estimation of log-wages and selection equations using the MLE method. The SBC approach has been widely used in the decomposition studies, in fact, it has been virtually the only available approach in the decomposition studies until now. We first compare two approaches using a simple two equation model for exposition purposes, and later argue that the GSB approach is, in general, easy to implement, flexible enough to handle virtually any kind of selection issue, and efficient because it uses full information.

2.1 *Preliminary: A Two Equation Model*

We consider a two equation model to simplify the exposition. For each group a and b, equations for individual N are,

$$Y^*_{gN} = X_{gN}\beta_g + e_{gN} \tag{4}$$

$$S^*_{gN} = Z_{gN}\,\gamma_g + v_{gN} \qquad (N = 1, \ldots, N_g), \tag{5}$$

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

Y^*_{gN} and S^*_{gN} are respectively latent log-wages and selection variables. We observe a binary variable S_{gN} for every individual in group g which has a value of one if $S^*_{gN} > 0$ and zero otherwise. The sample size whose $S_{gN} = 1$ ($S_{gN} = 0$) is n_g ($N_g - n_g$), where $N_g > n_g$. For individuals whose $S_{gN} = 1$, a continuous variable Y_{gN} is observed equal to Y^*_{gN} while, for others whose $S_{gN} = 0$, Y_{gN} is missed.[11] Y^*_{gN} and Y_{gN} may be interpreted as "offered" log-wages and "observed" log-wages (Reimers (1983)).

The population regression function for (4) can be written as

$$\mathrm{E}(Y^*_{gN}|X_{gN}) = X_{gN}\beta_g, \tag{6}$$

since the unconditional expectations of e_{gN} are assumed to be zero, i.e., $\mathrm{E}(e_{gN}) = 0$. The regression function for the non-random sample due to the sample selection determined by equation (5) may be written as

$$\mathrm{E}(Y^*_{gN}|X_{gN}, S_{gN} = 1) = X_{gN}\beta_g + \mathrm{E}(e_{gN}|S_{gN} = 1), \tag{7}$$

where $N = 1, \ldots, n_g$ because only n_g observations have data available on Y^*_{gN}.

The OLS estimates of equation (7) may not be consistent if the conditional expectations of e_{gN}, $\mathrm{E}(e_{gN}|S_{gN} = 1)$, are not zero. It has been well-studied how to obtain the consistent estimates of log-wage parameters from both Heckman's two-step and the MLE methods. Measuring the selection bias, the difference between

[11]The conventional decomposition analysis presumes Y_{gn} is observed without any missing or censoring, that is $n_g = N_g$. Hence it does not consider the second equation (5).

conditional and unconditional expectations of log-wages, is also required in the decomposition analysis to modify the decomposition equation (2) in addition to obtaining consistent estimates. We have to answer two related questions to have a proper decomposition (or discrimination) analysis when the data is not randomly selected from population: "how can consistent estimates of log-wage parameters be obtained when there are selection issues?" and "how can the selection bias of log-wages be measured?".

2.2 *Selection Bias Correction Approach: A Two Equation Model*

Heckman's two-step method provides one way to answer both questions (see Heckman (1979) for details). It has been a popular choice in studies of selection models ever since its introduction to (empirical) microeconomists.[12]

Heckman's two-step method begins from the analytical formula for the conditional expectations of e_{gN} given the value of $S_{gN} = 1$ ($v_{gN} > -Z_{gN}\gamma_g$), which is

$$\mathrm{E}(e_{gN}|S_{gN} = 1) = \theta_g \lambda_{gN}, \tag{8}$$

[12]See Maddala (1983) for the usefulness of Heckman's two-step method. Heckman's two-step method is implemented in various statistical packages. For example, see Green (1995) for Limdep.

where $\theta_g = \sigma_{e_g v_g}/\sigma_{v_g} = \rho_{e_g v_g}\sigma_{e_g}$, $\rho_{e_g v_g} = \sigma_{e_g v_g}/(\sigma_{e_g}\sigma_{v_g})$, and

$$\lambda_{gN} = \frac{\phi\left(-\frac{Z_{gN}\gamma_g}{\sigma_{v_g}}\right)}{1-\Phi\left(-\frac{Z_{gN}\gamma_g}{\sigma_{v_g}}\right)},$$

where ϕ and Φ are the standard univariate probability density and distribution functions, respectively.

Equation (7) may be written as

$$\mathrm{E}(Y^*_{gN}|X_{gN}, S_{gN}=1) = X_{gN}\beta_g + \theta_g\lambda_{gN},$$

so the log-wages for the selected sample may be written as

$$Y_{gN} = X_{gN}\beta_g + \theta_g\lambda_{gN} + \varepsilon_{gN}, \tag{9}$$

where $\varepsilon_{gN} = e_{gN} - \theta_g\lambda_{gN}$ and $\mathrm{E}(\varepsilon_{gN}|X_{gN}, \lambda_{gN}, S_{gN}=1) = 0$.

The first-step estimates the parameters of the probability that $S_{gN}=1$, i.e., γ_g since σ_{v_g} is normalized to 1 for identification purposes, using probit analysis for the full sample (N_g observations). From the probit estimates of γ_g, we compute a selection bias correction term (inverse Mill's ratio, λ_{gN}) for the selected sample (n_g observations whose Y_{gN} is available). The constructed value of λ_{gN} is used as a regressor in equation (9) in the second step. The OLS estimates $(\widehat{\beta}_g, \widehat{\theta}_g)$ in the second step are consistent because the conditional expectations of ε are zero. This answers the first question, "how can consistent estimates of log-wage

parameters be obtained when there are selection issues?".

The answer to the second question, "how can the selection bias of log-wage be measured?", can be also found from the second step. Let Λ_{gN} be the selection bias of log-wages of individual $N = 1, \ldots, n_g$ in group g (a and b). The Λ_{gN} is the difference between equations (6) and (7) which is identical to equation (8), the conditional expectations of e_{gN} given $S_{gN} = 1$, since the unconditional expectations of e_{gN} are zero. The Λ_{gN} can be evaluated by the selection bias correction term computed using the first-step probit estimates ($\widehat{\lambda}_{gN}$) and its coefficient at the second-step OLS ($\widehat{\theta}_g = \widehat{\rho_{e_g v_g} \sigma_{e_g}}$), that is

$$\widehat{\Lambda}_{gN} = \widehat{\theta}_g \widehat{\lambda}_{gN}, \tag{10}$$

where

$$\widehat{\lambda}_{gN} = \frac{\phi(-Z_{gN}\,\widehat{\gamma}_g)}{1 - \Phi(-Z_{gN}\,\widehat{\gamma}_g)},$$

where $\widehat{\gamma}_g$ is the vector of estimates from the probit analysis.

Once the selection bias is measured in the second-step, we can apply the Blinder-Oaxaca decomposition method to the log-wage differentials adjusted for the selection bias, i.e., the log-wage differentials measured from the population regression equations (6) for two groups using the second step OLS estimates. The decomposition equation (2) is modified as follows using only the selected sample

$(S_{gN} = 1)$,[13]

$$(11) \qquad \overline{Y}_a - \overline{Y}_b = \begin{cases} (\overline{X}_a - \overline{X}_b)\widehat{\beta}_b + \overline{X}_a(\widehat{\beta}_a - \widehat{\beta}_b) + (\widehat{\theta}_a\overline{\widehat{\lambda}}_a - \widehat{\theta}_b\overline{\widehat{\lambda}}_b), & \text{or} \\ (\overline{X}_a - \overline{X}_b)\widehat{\beta}_a + \overline{X}_b(\widehat{\beta}_a - \widehat{\beta}_b) + (\widehat{\theta}_a\overline{\widehat{\lambda}}_a - \widehat{\theta}_b\overline{\widehat{\lambda}}_b), & \end{cases}$$

where $\overline{Y}_g$, $\overline{X}_g$, and $\overline{\widehat{\lambda}}_g$ are, for each group g (a and b), the sample average of log-wages ($\sum_{N=1}^{n_g} Y_{gN}/n_g$), $1 \times K_Y$ vector of the average characteristics of individuals, and the mean selection bias correction term computed using probit estimates ($\sum_{N=1}^{n_g} \widehat{\lambda}_{gN}/n_g$), respectively. $\widehat{\beta}_g$ and $\widehat{\theta}_g$ are, for each group g (a and b), $K_Y \times 1$ vector and scalar of the second step OLS coefficients for X_{gN} and $\widehat{\lambda}_{gN}$, respectively.

The discrimination coefficients are also modified accordingly,

$$(12) \qquad \widehat{\mathrm{D}}_g = \exp\big(\overline{X}_g(\widehat{\beta}_a - \widehat{\beta}_b)\big) - 1$$

where $g = a$ or b, and $\widehat{\beta}$ is $K_Y \times 1$ vector of estimates of log-wage parameters from the second step OLS analysis.

We call the decomposition analysis which relies on Heckman's two-step method to answer the two questions, "how to obtain consistent estimates of log-wage parameters?" and "how to compute the selection bias?", a "selection bias correction (SBC) approach" to the decomposition analysis. It requires the computation of the selection bias correction term (λ_{gN}) to obtain consistent estimates of log-wage parameters and to measure the selection bias (Λ_{gN}).

[13]The modified decomposition equation is one of the many possible modifications summarized in Neuman and Oaxaca (1998).

2.3 *Generalized Selection Bias Approach: A Two Equation Model*

In this section, we propose a new approach, the "generalized selection bias (GSB) approach" to decomposition analysis based on joint estimation of log-wage equation (4) and selection equation (5) using the MLE method.[14]

The likelihood function of group g (a and b) for the joint estimation of log-wages and selection equations is[15]

$$(13) \quad L_g = \prod_{S_{gN}=1} \Phi\left(\frac{Z_{gN}\,\gamma_g + \mu_{v_{gN}|e_{gN}}}{\sigma_{v_g|e_g}}\right) \cdot \frac{1}{\sigma_{e_g}} \cdot \phi\left(\frac{e_{gN}}{\sigma_{e_g}}\right) \prod_{S_{gN}=0} \Phi\left(-\frac{Z_{gN}\,\gamma_g}{\sigma_{v_g}}\right),$$

where for each group g (a and b), $e_{gN} = Y_{gN} - X_{gN}\beta_g$, $\mu_{v_{gN}|e_{gN}} = e_{gN}\rho_{e_g v_g}\sigma_{v_g}/\sigma_{e_g}$, $\sigma_{v_g|e_g} = \sigma_{v_g}\sqrt{1-\rho_{e_g v_g}}$, $\sigma_{v_g} = 1$, and $N = 1, \ldots, N_g$.

By maximizing the likelihood function (13), we obtain consistent estimates of the log-wages and selection equation parameters (β_g, γ_g), and standard deviation of log-wages and correlation coefficient between e_g and v_g (σ_{e_g} and $\rho_{e_g v_g}$).[16] This

---

[14]The GSB approach is essentially independent of the estimation method as long as it can estimate consistent estimates of log-wage parameters by jointly estimating log-wages and selection equations to take correlation between them into account. In this paper, we focus on fully parametric classical MLE method as a method of joint estimation of log-wages and selection equations. This is mainly because of relative ease of estimation and popularity of the MLE method. We will briefly discuss the possibility of using semiparametric and Bayesian methods for the joint estimation and the modification of decomposition analysis at the end of this paper.

[15]Equation (13) is the functional expression of the following,

$$L_g = \prod_{S_{gN}=1} Pr(S_{gN}=1|e_{gN})Pr(e_{gN}) \prod_{S_{gN}=0} Pr(S_{gN}=0).$$

The derivation of the analytical formula of the selection bias correction term can be avoided in the MLE method because the likelihood for those of $S_{gN}=1$, $Pr(e_{gN}, S_{gN}=1)$, can be expressed as $Pr(S_{gN}=1|e_{gN})Pr(e_{gN})$, not $Pr(e_{gN}|S_{gN}=1)Pr(S_{gN}=1)$ in two equation model. $Pr(e_{gN}|S_{gN}=1)$ requires the analytical formula for the selection bias, which is $\rho_{e_g v_g}\sigma_{e_g}\lambda_{gN}$.

[16]The estimates obtained from joint estimation using the MLE method are not only consistent, but also have other desirable properties (they are asymptotically efficient and normally distributed).

gives the answer to the first question, "how can consistent estimates of log-wage parameters be obtained when there are selection issues?". This answer is common sense to economists.

What is not answered yet is the second question, "how can the selection bias in log-wage be measured?".[17] One may try to compute the selection bias according to its analytical formula, equation (8), using the MLE estimates (Dolton and Makepeace (1986, 1987)). We might call this a pseudo-SBC approach in the sense that it evaluates the analytical formula of the selection bias used in Heckman's two-step method.

There is another way, much easier and simpler than evaluating the analytical formula of the selection bias once the MLE estimates are available. The unconditional expectations of log-wages, equation (6), can be computed using the MLE estimates as follows,

$$\mathrm{E}(Y^*_{gN}|X_{gN}) = X_{gN}\widetilde{\beta}_g, \tag{14}$$

and the conditional expectations of log-wages, equation (7), may be written as

$$\mathrm{E}(Y^*_{gN}|X_{gN}, S_{gN}=1) = X_{gN}\widetilde{\beta}_g + \mathrm{E}(\widetilde{e}_{gN}|S_{gN}=1), \tag{15}$$

where $\widetilde{\beta}_g$ is $K_Y \times 1$ vector of the MLE estimates of log-wage parameters for group

[17] In contrast to the SBC approach which answers both questions at the same time in the second step, the GSB approach answers the first question first, and the second question later using the answer to the first question.

g (a and b) and $\widetilde{e}_{gN} = Y_{gN} - X_{gN}\widetilde{\beta}_g$, i.e., residuals computed using the MLE estimates of log-wage parameters (below called MLE residuals).

The selection bias, the difference between equations (14) and (15), is simply the conditional expectation of the MLE residuals, i.e., $\widetilde{\Lambda}_{gN} = \mathrm{E}(\widetilde{e}_{gN}|S_{gN} = 1)$.[18] We call $\widetilde{\Lambda}_{gN}$, the conditional expectation of the MLE residuals, a "generalized selection bias" (GSB).

The log-wages for selected sample, equation (9), may be written using GSB as

$$Y_{gN} = X_{gN}\widetilde{\beta}_g + \widetilde{\Lambda}_{gN} + \widetilde{\varepsilon}_{gN}, \tag{16}$$

where $\widetilde{\varepsilon}_{gN} = \widetilde{e}_{gN} - \widetilde{\Lambda}_{gN}$ and $\mathrm{E}(\widetilde{\varepsilon}_{gN}|X_{gN}, \widetilde{\Lambda}_{gN}, S_{gN} = 1) = 0$.

We can apply the Blinder-Oaxaca decomposition method to the log-wage differentials adjusted for the GSB, i.e., the log-wage differentials measured from the population regression equations (14) for two groups evaluated with the MLE estimates. The decomposition equation (2) is modified as follows using only the

[18]The selection bias can be evaluated according to the analytical formula (the pseudo-SBC approach), $\widetilde{\Lambda}_{gN} = \widetilde{\rho}_{e_g v_g}\widetilde{\sigma}_{e_g}\widetilde{\lambda}_{gN}$, where $\widetilde{\sigma}_{e_g}$ and $\widetilde{\rho}_{e_g v_g}$ are, for each group g (a and b), the MLE estimates of standard deviation of log-wages and correlation coefficient between e_g and v_g, respectively, and $\widetilde{\lambda}_{gN}$ is the selection bias correction term of individual $N = 1, \ldots, n_g$ in group g (a and b) evaluated using the MLE estimates of selection equation parameters.

The selection bias measured in the SBC approach is, in fact, the conditional expectation of the second step OLS residuals. That is, $\widehat{\Lambda}_{gN} = \mathrm{E}(\widehat{e}_{gN}|S_{gN} = 1) = \widehat{\theta}_g\widehat{\lambda}_{gN}$, where $\widehat{\theta}_g = \widehat{\rho_{e_g v_g}\sigma_{e_g}}$ and $\widehat{e}_{gN} = Y_{gN} - X_{gN}\widehat{\beta}_g$. Note that X_{gN} does not include $\widehat{\lambda}_{gN}$.

selected sample ($S_{gN} = 1$),

$$(17) \qquad \overline{Y}_a - \overline{Y}_b = \begin{cases} (\overline{X}_a - \overline{X}_b)\widetilde{\beta}_b + \overline{X}_a(\widetilde{\beta}_a - \widetilde{\beta}_b) + (\overline{\widetilde{\Lambda}}_a - \overline{\widetilde{\Lambda}}_b), & \text{or} \\ (\overline{X}_a - \overline{X}_b)\widetilde{\beta}_a + \overline{X}_b(\widetilde{\beta}_a - \widetilde{\beta}_b) + (\overline{\widetilde{\Lambda}}_a - \overline{\widetilde{\Lambda}}_b), \end{cases}$$

where $\overline{Y}_g$, $\overline{X}_g$, and $\overline{\widetilde{\Lambda}}_g$ are, for each group g (a and b), the sample average of log-wages ($\sum_{N=1}^{n_g} Y_{gN}/n_g$), $1 \times K_Y$ vector of the average characteristics of individuals, and the mean GSB ($\sum_{N=1}^{n_g} \widetilde{\Lambda}_{gN}/n_g$), respectively. $\widetilde{\beta}_g$ is $K_Y \times 1$ vector of the MLE estimates of log-wage parameters for group g (a and b).

In the SBC approach, the selection bias correction term (hence the selection bias) of each individual is computed for the decomposition analysis of wage differentials between group a and b. However, in the GSB approach, the computation of the GSB ($\widetilde{\Lambda}_{gN}$) itself for each individual is not required for the decomposition analysis of wage differentials between group a and b. The sample average of the GSB used in the decomposition equation (17) can be measured by the sample average of the MLE residuals since $\mathrm{E}(\widetilde{\varepsilon}_{gN}|X_{gN}, \widetilde{\Lambda}_{gN}, S_{gN} = 1) = 0$, that is, $\overline{\widetilde{\Lambda}}_g = \sum_{N=1}^{n_g} \widetilde{e}_{gN}/n_g$, where $\widetilde{e}_{gN} = Y_{gN} - X_{gN}\widetilde{\beta}_g$. Since the evaluation of the GSB at the individual level is not necessary when we compare wage differentials of two groups, we don't have to rely on the analytical formula for selection bias. The GSB approach eliminates the burden of deriving the analytical formula for the selection bias and computing it. It is the ability to skip the computation of the selection bias following the analytical formula that makes the GSB approach attractive.

The discrimination coefficients are also modified accordingly,

$$\widetilde{\mathrm{D}}_g = \exp\big(\overline{X}_g(\widetilde{\beta}_a - \widetilde{\beta}_b)\big) - 1 \tag{18}$$

where $g = a$ or b, and $\widetilde{\beta}$ is $K_Y \times 1$ vector of estimates of log-wage parameters from the MLE method.

In this section, we have shown the basic idea of the GSB approach using two equation model which does not rely on the selection bias correction term. Next section, we will discuss merits of the GSB approach and its multivariate extensions.

2.4 *Comparison of Two Approaches*

Table I summarizes how to compute the selection bias in two approaches, the SBC and GSB approaches. For exposition purposes, it also shows the pseudo-SBC approach and what the equivalent to the GSB would be if the second step OLS estimates from Heckman's two-step method are used.

To summarize, for the SBC approach:

(1) Estimate the parameters of the probability that $S^*_{gN} > 0$ using probit analysis.

(2) From the probit estimates of the sample selection parameters (γ_g, since $\sigma_{v_g} = 1$), calculate the selection bias correction term (λ_{gN}).

(3) Regress log-wages (Y_{gN}) on exogenous variables augmented with the estimated value of λ_{gN}.

(4) Estimate sample means of Y_{gN}, X_{gN}, and $\widehat{\lambda}_{gN}$.

(5) Compute the decomposition equation (11) and the discrimination coefficients (12) using the sample means of Y_{gN}, X_{gN}, and $\widehat{\lambda}_{gN}$, and the second step OLS estimates of log-wage parameters.

On the other hand, for the GSB approach:

(1) Write the likelihood function of the joint estimation (log-wages and selection equations) using the property of conditional and marginal distributions while avoiding the need to compute selection bias analytically.

(2) Maximize the likelihood, estimating β_g, γ_g, $\rho_{e_g v_g}$, and σ_{e_g}.

(3) Estimate the MLE residuals of log-wages, $\widetilde{e}_{gN} = Y_{gN} - X_{gN}\widetilde{\beta}_g$.

(4) Estimate sample means of Y_{gN}, X_{gN}, and $\widetilde{e}_{gN}$.

(5) Compute the decomposition equation (17) and the discrimination coefficients (18) using the sample means of Y_{gN}, X_{gN}, and $\widetilde{e}_{gN}$, and the MLE estimates of log-wage parameters.

Considering that the SBC and GSB approaches are based on Heckman's two-step and the MLE method, respectively, the comparison between two approaches is more or less equivalent to comparing the two estimation methods.

Heckman's two-step method has been a preferred method for empirical studies for last two decades. The popularity of Heckman's two-step method may be because it can be implemented using standard procedures (e.g., OLS and probit) provided by almost every statistical package. It also handles non-normal distribution of error in selection equation (Lee (1983)). However, it should be noted that

every model Heckman's two-step method is able to handle can be estimated using the MLE method. Furthermore, the GSB approach (the MLE method) may be preferred to the SBC approach (Heckman's two-step method) for following reasons.

First of all, the GSB approach will be the only feasible approach when Heckman's two-step method cannot be applied due to conceptual difficulties. One example is the truncation issue. The use of the probit analysis to obtain the coefficients of determinants in the first step is ruled out because we don't have observations for which $S_{gN} = 0$.[19] The GSB approach is the only available approach to decomposition analysis, though the pseudo-SBC approach is still available, when data has a truncation issue, since Heckman's two-step method cannot be used.

From theoretical viewpoint, we may argue that the GSB approach is preferred to the SBC approach because the GSB approach is more efficient since it is based on the full information MLE method. Even though the efficiency of the GSB approach (or the MLE method) can be easily accepted, many will be reluctant to switch to the GSB approach if the implementation of the MLE method is difficult. The practical aspects of the implementation of the GSB approach will be discussed for two possible cases where the analytical formula is relatively simple and very complicated, respectively.

As Heckman (1979, p. 155) notes, multivariate extensions of his two-step method are mathematically straightforward using the results of moment gen-

[19]See Bloom and Killingsworth (1985), Hausman and Wise (1977), and chapter 6 of Maddala (1983) for truncation issue.

erating function of truncated multivariate normal distribution.[20] The merit of Heckman's two-step method (ease of implementation), however, diminishes substantially if the selection issue becomes even slightly more complicated than the standard two equation model. For example, double selection requires computation of two λ's.[21] The computation of λ's requires a bivariate probit analysis in the first-step of Heckman's two-step method. One the other hand, for the GSB approach, the likelihood of joint estimation of log-wages and double selection equations may be expressed by the combination of conditional bivariate probit given the value of log-wages and marginal density of log-wages for individual whose log-wages are observed, and only bivariate probit for individual whose log-wages are not observed. Estimating log-wages and double selection equations jointly is often easier, or at least no more difficult than estimating only bivariate probit analysis of selection equations (see Co, Gang and Yun (forthcoming, 1999)). This might be because continuous log-wages provide variations which facilitate the estimation. The computational time taken for the joint estimation and the bivariate probit analysis for Heckman's two-step method is somewhat similar. Furthermore, since the the OLS analysis does not provide a correct variance of the estimates, we have to correct the variance of second-step OLS estimates. This requires non-trivial programming. Considering these factors, we argue that the GSB approach (the MLE method) will be preferred to the SBC approach

[20]See Tallis (1961) for the moment generating function of the truncated multivariate normal distribution. If multivariate normal distribution requires high dimensional numerical integration, simulation methods can be used. See Stern (1997).

[21]See Fishe, Trost and Lurie (1981), Ham (1982), and Tunali (1986) for double selection.

(Heckman's two-step method) for double selection issues.

The main difference between the two approaches is that the GSB approach does not rely on the analytical formula of the selection bias.[22] Hence the GSB approach alleviates the burden of computing the λ_{gN}'s, especially in a complicated selection model. In practice, correcting the variance of the second-step OLS estimates will be also very cumbersome as the selection issues become complicated. This implies that the GSB approach will be very useful if the computation of the selection bias correction term is difficult analytically or computationally or both.

An example might be found in some issues of the piecewise budget line constraint model which studies the effects of tax and subsidy, fixed costs associated with labor market participation, etc. Though the main interest of the piecewise budget line constraint model is estimating labor supply, Heckman and MaCurdy (1981) pointed out this model could be applied to estimate (level) wage in addi-

[22]One might argue that the analytical formula of the selection bias is necessary for the decomposition analysis if a researcher believes that $\widehat{\theta}_a\overline{\widehat{\lambda}}_a - \widehat{\theta}_b\overline{\widehat{\lambda}}_b$ in the equation (11) or its equivalent in the GSB approach should be decomposed further to $\widehat{\theta}_a(\overline{\widehat{\lambda}}_a - \overline{\widehat{\lambda}}_b) + (\widehat{\theta}_a - \widehat{\theta}_b)\overline{\widehat{\lambda}}_b$. This is still in debate (Neuman and Oaxaca (1998)).

We may justify not pursuing further decomposition from two perspectives. The selection bias represents the effects of unobserved characteristics of an individual on the log-wages. The unobserved characteristics will be the combination of many factors, not just one characteristic of the person. If the λ_{gN} is considered as another exogenous variable, then it is not clear whether λ_{gN} and its coefficient can be treated like any other observed exogenous variables and their coefficients because each observed exogenous variable represents only one aspect of the individual's characteristics.

Another argument is that, in decomposition or discrimination analysis, we are interested in the difference in the unconditional expectations of log-wages, not the difference in the conditional expectations. Note that the selection bias correction term is not included in the population regression equation (6). The selection bias correction term affects only the conditional expectations of log-wages. It will be easy to understand our point if we interpret the unconditional and conditional expectations of log-wages as the expectations of offered and observed log-wages. The decomposition or discrimination analysis studies the difference in offered log-wages per se, not the difference in observed log-wages.

tion to hours by endogenizing the wage.[23] Heckman and MaCurdy (1981) try to extend Heckman's two-step method to analyze complex selection models which are usually analyzed using the MLE method. However, it seems that the extension of Heckman's two-step method to piecewise budget line constraint model has not gained popularity. It is the MLE method that is used in the piecewise budget line constraint models.

Previous studies on wage differentials are restricted to the simple selection issue, usually with only one binary selection. The lack of diversity of the selection issues studied in previous papers might be because only the SBC approach has been available. The GSB approach is theoretically superior and computationally practical. It will extend the scope of studies by providing a general framework for decomposition analysis. The implementation of the GSB approach requires only the conditional and marginal density functions of multivariate normal distribution in most of cases which can be easily evaluated thanks to the recent developments in computing technology. The necessity of using the MLE method for the joint estimation is not an extra burden, since there are so many studies in which the MLE method is preferred.

3. RACIAL WAGE DISCRIMINATION AMONG WOMEN

We apply the GSB approach to racial wage differentials between white and

[23]See Hausman (1985) and Moffitt (1986, 1990) for the general discussion of the piecewise budget line constraint model. See Yun (1998) for endogenizing wages in this context.

other races using the female sample from March 1995 Current Population Survey. The source of selection is participation for the first illustration, and tobit type hours of work for the second illustration. In the second illustration, we show decompositions when data is either censored or truncated.

3.1 *Data*

The female sample used in the empirical study is drawn from the March 1995 Current Population Survey. 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.[24] 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 II describes the variables used for our study.

Table III shows means and standard deviations of variables used in the decomposition analysis. The characteristics of working women are different from 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 have a higher non-labor

[24]The information on last year's earnings are used to compute non-labor income. For details of computation of the non-labor income, see Table II.

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 non-working and working samples. Though white women have 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 among non-working 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. We study what portion of racial wage differentials can be explained by difference in the characteristics and by difference in the coefficients using the GSB approach to decomposition analysis proposed in section 2. The MLE method, the basis of our decomposition analysis in these illustrations, is implemented using both Gauss CML (constrained maximum likelihood) program and the SAS Non Linear Programming procedure (SAS Institute, 1997).[25]

3.2 *Illustration 1: Selection Bias Due to Participation Choice*

The selection bias due to the participation decision is a well-studied subject

[25]For illustration 1, Limdep is also used to double-check. Limdep reports both estimates of Heckman's two-step and the MLE methods.

in the area of both labor supply and wage determination.[26] Most papers on wage differentials, especially those on the gender gap, address sample selection bias arising from the participation decision using Heckman's two-step method.[27]

We illustrate the decomposition of racial wage differentials from both SBC and GSB approaches.[28] This illustration is a direct application of two equation model explained in the section 2.

Women are partitioned into two groups according to their race, whites ($g = w$) and other races ($g = o$). Hence group a and b in section 2 are whites (w) and other races (o), respectively. For the selection equation, we have a participation equation. Equations (4) and (5) are respectively latent log-wages and participation equations. Women will participate in the labor market if S^*_{gN} in equation (5) has a positive value.

Tables IV and V show the estimates of log-wages and participation equations, respectively. First of all, the correlation coefficient between the errors of log-wage and participation equations is significantly different from zero for white women,

[26]Participation is usually defined to include employment or unemployment. 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.

[27]There are numerous studies which use Heckman's two-step method to correct selection bias caused by the participation decision; for example, Reimers (1983), Hoffman and Link (1984), Dolton and Makepeace (1986, 1987), Blau and Beller (1988), Vella (1988), Dolton, Makepeace, and van Der Klaauw (1989), Wright and Ermisch (1991), Choudhury (1993), Wellington (1993), Baker, Benjamin, Cegep, and Grant (1995), Joshi and Paci (1998), and Schaffner (1998).

[28]Dolton and Makepeace (1986, 1987) estimate the log-wages and participation equations using both Heckman's two-step and the MLE methods. To the best of our knowledge, Dolton and Makepeace (1986, 1987) are the only papers which use the MLE method in the context of decomposition analysis. However, they use the MLE estimates to compute the so-called λ and its coefficient, i.e., they adopt the pseudo-SBC approach.

and not significantly different from zero for other races.[29] As shown in Table IV, the estimates of log-wage parameters from Heckman's two-step and the MLE methods are not much different from the OLS estimates except for the constant term. The MLE estimates are, however, closer to the OLS estimates than are Heckman's two-step estimates. Most of the estimates of the log-wage parameters for white women are significant, but estimates for other race women are not very significant. Nonetheless, the estimates for both groups of women have the expected signs. The determinants of participation, shown in Table V, also 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.

We compute the selection bias from both SBC and GSB approaches. Nowadays, it is a simple exercise to compute the λ_{gN} term and its coefficient in Heckman's two-step method. From the SBC approach, the sample average of the selection bias ($\overline{\widehat{\Lambda}}_g = \widehat{\rho_{e_g v_g} \sigma_{e_g}} \overline{\widehat{\lambda}}_g$) is 0.03832 and 0.07575 for white and other race women, respectively. From the GSB approach, the sample average of the GSB ($\overline{\widetilde{\Lambda}}_g = \sum_{N=1}^{n_g} \widetilde{e}_{gN}/n_g$) is 0.02360 and 0.02374 for white and other race women, respectively.

[29] If the covariance ($\sigma_{e_g v_g}$) or correlation coefficient ($\rho_{e_g v_g}$) is zero, then there is no selection bias. In that case, the OLS estimates from the second step of Heckman's two-step method are equal to those of simple OLS without selection bias correction term (λ_{gN}), reported at the first column of Table IV, because the estimate for the λ_{gN} is zero ($\rho_{e_g v_g} \sigma_{e_g} = 0$). Also, in that case, the estimates from the MLE method are identical to the simple OLS estimates for log-wage parameters, reported at the first column of Table IV, and probit estimates for the participation equation (i.e., the first step probit estimates in Heckman's two-step method), reported at the first column of Table V.

For illustration purposes, we compute the selection bias following the analytical formula ($\rho_{e_g v_g}\sigma_{e_g}\lambda_{gN}$) using the MLE estimates (the pseudo-SBC approach). First, the sample average of the selection bias following the analytical formula using the MLE estimates ($\widetilde{\rho}_{e_g v_g}\widetilde{\sigma}_{e_g}\overline{\widetilde{\lambda}}_{gN}$) is 0.02363 and 0.02375 for white and other race women, respectively. The discrepancy between the sample average of the GSB and the selection bias derived using the analytical formula is 0.00003 and 0.00001 for white and other race women, respectively.[30] Second, on the way to computing selection bias from the analytical formula using the MLE estimates, we compute $\rho_{e_g v_g}\sigma_{e_g}$, the coefficient for λ_{gN} in Heckman's two-step method. For white and other race women, the values are respectively 0.084 and 0.062, much smaller than those from Heckman's two-step method (0.136 and 0.199 for white and other race women, respectively). The sample average of the selection bias correction term derived from the Heckman's two-step method ($\overline{\widehat{\lambda}}_g$) is 0.28096 and 0.38152 for white and other race women, respectively. The sample average of the selection bias correction term using MLE estimates ($\overline{\widetilde{\lambda}}_g$) are 0.28106 and 0.38154 for white and other race women, respectively. Since $\overline{\widehat{\lambda}}_g$ and $\overline{\widetilde{\lambda}}_g$ are not different substantially, we may conclude that the difference between the sample average of selection bias derived from the SBC approach ($\overline{\widehat{\Lambda}}_g$) and the GSB approach ($\overline{\widetilde{\Lambda}}_g$) comes from the different estimates for the correlation and the standard deviation of the log-wages.

[30]The discrepancy might be caused either because the MLE method fails to obtain the "true" optimization, or because there is a precision problem in computing the λ_{gN}, since computing λ_{gN} in extreme area where probability is very close to zero or 1 might be problematic. See McCullough and Vinod (1999).

Table VI shows the decomposition results. As it is clear from equations (11) and (17), we do not decompose the difference in the sample mean of the selection bias ($\widehat{\theta}_w\overline{\widehat{\lambda}}_w - \widehat{\theta}_o\overline{\widehat{\lambda}}_o$ or $\overline{\widehat{\Lambda}}_w - \overline{\widehat{\Lambda}}_o$) further into difference in the mean of λ_g and the difference in the $\theta = \rho_{e_g v_g}\sigma_{e_g}$. Some earlier papers using the SBC approach include $\widehat{\theta}_g$ and $\overline{\widehat{\lambda}}_g$ as simply another coefficient on another variable (albeit one pertaining to unobserved characteristics).[31] We treat the difference in the GSB (or the selection bias in Heckman's two-step method) as a separate component in decomposition analysis.[32] We have already discussed the reasons for not pursuing further decomposition from theoretical viewpoint in section 2.4. The other main, more practical, reason is that we want to have a consistent decomposition equation regardless of the selection issue. We can find the analytical form of the selection bias easily in this illustration ($\Lambda_{gN} = \rho_{e_g v_g}\sigma_{e_g}\lambda_{gN}$), but we may have difficulty in deriving selection bias analytically in many cases. This means we cannot decompose the selection term into coefficient and unobserved characteristics. We believe the consistency gained from our treatment outweighs the gain from this refinement.[33]

As shown in Table VI, decomposition results from the conventional analysis using simple OLS estimates and the GSB approach show very similar patterns.

[31]Dolton and Makepeace (1986), and Joshi and Paci (1998), among others.

[32]Reimers (1983), Wright and Ermisch (1991), and Ermisch and Wright (1992) follow this direction.

[33]Since we devise the GSB method to avoid the difficulty of finding the analytical formula for selection bias, we do not recommend dividing the selection bias into coefficients and unobserved characteristics even if there is only a single selection. But in the case where researchers believe that the refinement is worth doing, they can refine the decomposition equation using the analytical formula of the selection bias evaluated at the MLE estimates (the pseudo-SBC approach), as Dolton and Makepeace (1986, 1987) did.

Difference in characteristics explains half of log-wage differentials (about 57% (46%) when the coefficients of log-wage parameters of other race (white) women are used), and difference in coefficients explains other half of the log-wage differentials (about 43% (54%) when the average characteristics of white (other race) women are used as weights). This is because the difference in the GSB is virtually non-existent.[34] The discrimination coefficients (in percentage) from the conventional analysis and the GSB approach (D_g and $\widetilde{D}_g$ in equations (3) and (18), respectively) are about 3.4% and 4.3% when characteristics of white and other race women are respectively used as weights. In contrast, decomposition results from the SBC approach shows larger discrimination than do those from both the conventional analysis and the GSB approach. The discrimination coefficients (in percentage) from the SBC approach ($\widehat{D}_g$ in equation (12)) are 7.0% and 8.1% when characteristics of white and other race women are respectively used as weights.[35] This is because differences in the selection bias between white and other race women are large and negative, which means log-wages of other race women are more likely increased due to their unobserved characteristics.[36]

In this section, we discussed a very simple selection model which can be easily analyzed using both SBC and GSB approaches. In next illustration, we will show the application of the GSB approach when the choice is not a participation but

[34]The sample average of the GSB is 0.02360 and 0.02374 for white and other race women, respectively.

[35]If we include the difference in the coefficients of the λ_{gN} as part of discrimination, the discrimination coefficients (in percentage) become 5.1% and 5.6%, still higher than those from both the conventional analysis and the GSB approach.

[36]The sample average of the selection bias is 0.03832 and 0.07575 for white and other race women, respectively.

tobit type hours.

3.3 *Illustration 2: Selection Bias Due to Hours Choice*

In this section, the choice set is a continuous variable (hours of work), unlike in the previous section where choice set was discrete (i.e., participation or not). The purpose of this illustration is not to claim that selection should be modeled as continuous hours for the study of wage differentials, but to show that the GSB approach can be easily applied as long as the joint estimation of log-wages and selection equations is available.[37] We will analyze two specifications: joint estimation of log-wages and hours with censoring (using whole sample) and truncation (using only working sample).[38]

We assume that each person has latent log-wages (Y^*_{gN}) and her own optimal hours of work (H^*_{gN}) specified separately for each race group as follows,

$$Y^*_{gN} = X_{gN}\beta_g + e_{gN}, \tag{19}$$

$$\begin{aligned} H^*_{gN} &= \alpha_g Y^*_{gN} + Z_{gN}\gamma_g + v_{gN} \\ &= \underbrace{\alpha_g X_{gN}\beta_g + Z_{gN}\gamma_g}_{\widehat{H}_{gN}} + \underbrace{\alpha_g e_{gN} + v_{gN}}_{\eta_{gN}} \end{aligned} \tag{20}$$

[37]However, we might claim the efficiency of using continuous hours as a choice variable since hours will give us more information than just the binary variable participation or not.

[38]See Heckman (1974), Wales and Woodland (1980), and Mroz (1987) for details. Mroz (1987) refers to censoring and truncation models as tobit and conditional tobit, respectively. There is another similar specification, the so-called generalized tobit which consists of both participation and hour choices and log-wage equation. See Mroz (1987) for detail. Gang and Yun (1999) estimate the generalized tobit using the MLE method.

where X_{gN} and Z_{gN} are respectively $1 \times K_Y$ and $1 \times K_H$ vectors of socio-economic characteristics of individual N in group g (w and o); coefficients β_g and γ_g are $K_Y \times 1$ and $K_H \times 1$ vectors of parameters; coefficient α_g is a scalar of parameter; $\mathrm{E}(e_{gN}) = 0$, $\mathrm{E}(v_{gN}) = 0$, $\mathrm{E}(e_{gN}^2) = \sigma_{e_g}^2$, $\mathrm{E}(v_{gN}^2) = \sigma_{v_g}^2$, $\mathrm{E}(e_{gN}\, v_{gN'}) = \sigma_{e_g v_g}$ if $N = N'$ and zero if $N \neq N'$; $\widehat{H}_{gN}$ and η_{gN} are deterministic and stochastic components of the optimal hours, respectively; $N = 1, \ldots, N_g$.

We observe the log-wages and hours only when H^* is positive for n_g individuals, and not working otherwise for $(N_g - n_g)$ individuals, that is, $Y_{gN} = Y_{gN}^*$ if $H_{gN}^* > 0$ and missing otherwise, and $H_{gN} = \max(H_{gN}^*, 0)$, where Y_{gN} and H_{gN} are respectively observed log-wages and hours.[39] This model assumes continuous labor supply and ignores factors such as fixed time or money costs associated with working which cause discontinuity of labor supply.[40]

For joint estimation of hours and log-wages, the MLE method has been frequently used in previous studies. We estimate hours and log-wages jointly for data on the whole sample (censoring specification) by maximizing following likelihood

[39]If we ignore the choice of hours, i.e., if we consider only whether H^* is positive or not, then this model reduces to a simple participation choice model discussed in sections 2 and 3.2.

[40]See Killingsworth (1983) pp. 23-28 and Hausman (1980).

function,[41]

$$L_g = \prod_{H_{gN}>0} \frac{\phi\left(\frac{v_{gN}-\mu_{v_{gN}|e_{gN}}}{\sigma_{v_g|e_g}}\right)}{\sigma_{v_g|e_g}} \cdot \frac{\phi\left(\frac{e_{gN}}{\sigma_{e_g}}\right)}{\sigma_{e_g}} \prod_{H_{gN}=0} \Phi\left(-\frac{\widehat{H}_{gN}}{\sigma_{\eta_g}}\right), \tag{21}$$

and for working women only (truncation specification) by maximizing following likelihood,[42]

$$L_g = \prod_{H_{gN}>0} \frac{\phi\left(\frac{v_{gN}-\mu_{v_{gN}|e_{gN}}}{\sigma_{v_g|e_g}}\right)}{\sigma_{v_g|e_g}} \cdot \frac{\phi\left(\frac{e_{gN}}{\sigma_{e_g}}\right)}{\sigma_{e_g}} \Bigg/ \Phi\left(\frac{\widehat{H}_{gN}}{\sigma_{\eta_g}}\right), \tag{22}$$

where $e_{gN} = Y_{gN} - X_{gN}\beta_g$ and $v_{gN} = H_{gN} - \widehat{H}_{gN} - \alpha_g e_{gN}$; the conditional mean and standard deviation are respectively $\mu_{v_{gN}|e_{gN}} = e_{gN}\rho_{e_g v_g}\sigma_{v_g}/\sigma_{e_g}$ and $\sigma_{v_g|e_g} = \sigma_{v_g}\sqrt{1-\rho^2_{e_g v_g}}$; $\rho_{e_g v_g} = \sigma_{e_g v_g}/(\sigma_{e_g}\sigma_{v_g})$ and $\sigma_{\eta_g} = \sqrt{\alpha_g^2\sigma_{e_g}^2 + 2\alpha_g\rho_{e_g v_g}\sigma_{e_g}\sigma_{v_g} + \sigma_{v_g}^2}$.

Tables VII and VIII show the estimates of log-wages and hour parameters.[43]

---

[41]Equation (21) is functional expression of the following,

$$L_g = \prod_{H_{gN}>0} Pr(v_{gN}|e_{gN}) \cdot Pr(e_{gN}) \prod_{H_{gN}=0} Pr(\eta_{gN} \leq -\widehat{H}_{gN}) \ .$$

[42]Equation (22) is functional expression of the following,

$$L_g = \prod_{H_{gN}>0} Pr(v_{gN}|e_{gN}) \cdot Pr(e_{gN}) \Big/ Pr(\eta_{gN} \geq -\widehat{H}_{gN}) \ .$$

[43]We do not include education in the hours equation because many estimates have the wrong signs when education is included. This might be a multicollinearity problem between education and predicted log-wages ($X_{gN}\beta_g$) when education variable is included in the hours equation. The correlation coefficient is about 0.953 (white women) and 0.926 (other race women) for the censoring specification, and about 0.947 (white women) and 0.900 (other race women) for the truncation specification when the predicted log-wages are computed using the MLE estimates. The omission of education variable is not rare in simultaneous equations. Hausman and Wise (1977, p. 931) assume that "these attributes (education, I.Q., and occupation training) of individuals, given their wage rates, do not affect their choices between labor and leisure." Blundell, Duncan and Meghir (1992) cite identification as a reason why education and other variables are excluded from the hour equation. The estimation results when education is included in the

Even though the correlation coefficient ($\rho_{e_g v_g}$) is small, that does not mean that the selection issue does not exist. Since the hours equation has an error term ($\eta_{gN} = \alpha_g e_{gN} + v_{gN}$) which includes the stochastic component of log-wages, there is always correlation between log-wages and hours, i.e., $\rho_{\eta_g e_g} \neq 0$, even if $\rho_{e_g v_g} = 0$ as long as $\alpha_g \neq 0$ (see Moffitt (1984) for this point).

The estimates of the log-wage parameters reported in Table VII show the similarities between the two specifications, and even with those estimated in previous section (Table IV). However, in censoring model for other race women, the MSA and the regional variables (West, South, and Midwest) have different estimates from other specifications. Estimates of the constant term of both white and other race women in censoring specification are quite different from the simple OLS reported in Table IV and other specifications (the truncation model reported Table VII, and the participation choice model reported in Table IV).

In contrast, the hour equation reported in Table VIII shows quite a difference between two specifications: censoring (with whole sample) and truncation (with working sample). Mroz (1987) also reports differences in estimates between censoring (tobit) and truncation (conditional tobit) specifications. For example, in the censoring specification, the coefficients for log-wages and non-labor income are about two to three times larger than those in truncation specification (Mroz (1987), Tables IX and X), while our results are about three to four times larger in censoring specification. Mroz (1987, p. 790) concludes "the hours of work de-

hours equation are available upon request.

cisions made when the woman is in the labor force appear quite distinct from her labor force participation decision."[44] The signs of significant estimates are reasonable except that of the marriage variable in the censoring model; it becomes positive and very large.[45]

Table IX shows the labor supply elasticities evaluated at average white woman's hours and wages. The elasticities have the expected signs, and their magnitudes lie within the ranges reported in previous papers.[46] The uncompensated wage elasticity is measured using $\alpha_g / \overline{H}_w$ where $\overline{H}_w$ is the average hours of white working women. The uncompensated wage elasticity of white women is larger than that of other race women when we measure the elasticity using OLS and the truncation specifications. However, the reverse is true when the censoring specification is used. The income elasticity is measured using $\delta_g \overline{W}_w$, where δ_g is the coefficient of labor supply parameter for non-labor income, and $\overline{W}_w$ is the average (level) wages of white working women. The income elasticity also shows the same pattern; the income elasticity of white women is larger in absolute terms than that of other race women when OLS and truncation specifications are used; the opposite is true when the censoring specification is used.

Table X shows the decomposition results. As expected, the decomposition

[44]However, Wales and Woodland (1980) report similar estimates between two specifications from an experiment with generated data.

[45]We estimate the censoring specification, without the marriage variable in the hours equation. The results of the log-wage equation are similar to those with the marriage variable in the hours equation. The estimates for remaining variables in the hours equation are similar to those with the marriage variable in the hours equation. These results are also available upon request.

[46]Previous studies of the elasticity of female labor supply are summarized in Killingsworth and Heckman (1986) and Killingsworth (1983).

from the truncation specification is very similar to that from OLS. The sample average of the GSB is negligible in both white and other race women in the truncation specification. The discrimination coefficients (in percentage) are also very similar to those of OLS; 3.4% (4.4%) when the characteristics of white (other race) women are used as weights. However, the sample average of the GSB is big when the censoring specification is used. The fact that other race women have larger GSB than do white women (0.069 for whites, 0.147 for other races) in the censoring specification results in larger discrimination. The discrimination coefficients (in percentage) are 11.22% (evaluated using whites' characteristics) and 11.93% (evaluated using other races' characteristics).

In this section, we have illustrated the GSB approach to decomposition analysis using tobit type specifications. These specifications are studied quite often in labor supply literature, but not in the context of wage differentials and discrimination. The illustration in this section shows the potential of the GSB approach: liberation from the limited selection models currently available for wage differential and discrimination analysis.

4. CONCLUSION

The major contribution of this paper is finding a new and flexible way to measure the effects of selection on the log-wages. In this context, we offer a general approach to performing decomposition analysis when selection effects are present. We call the difference between unconditional and conditional expectations of the

log-wages a generalized selection bias (GSB) when the two expectations are measured using the estimates from the joint estimation of the whole model (log-wages and selection equations) using the MLE method. The unconditional and conditional expectations are, respectively, the deterministic component of log-wages, and the deterministic component plus the conditional expectations of the stochastic component of log-wages, where the deterministic component is computed using the estimates from the joint estimation. That is, the GSB is the expectation of the residuals estimated from the joint estimation. It is appropriate to apply the Blinder-Oaxaca decomposition method to the log-wage differentials adjusted for the GSB. The GSB approach to decomposition analysis is not only easy to implement and flexible enough to apply to virtually any kind of selection issue, but also efficient because it uses full information. We have illustrated GSB approach by applying it to the racial wage differentials among women using data from the Current Population Survey.

In our illustrations, the fully parametric classical MLE method has been used to estimate log-wages and selection equations jointly. However, the GSB approach is not restricted to the fully parametric classical MLE method. It is not crucial which estimation technique is used for the joint estimation of log-wages and selection equations as long as the joint estimation provides the consistent estimates of log-wages. We will now briefly discuss two other methods which have their own merits.

The argument for computing the GSB in this paper ignores the possibility

that we may misspecify the error distributions of log-wages or selection equation or both. In the situation where we assume a normal distribution, but the true distribution is, say, log-normal, our estimates will be biased, and so will the GSB. Also, we cannot obtain the correct discrimination coefficients. We can avoid the problem of misspecification of the error distribution by adopting a distribution free estimation method, a semiparametric method.[47] Since semiparametric methods usually estimate parameters up to a scale factor, it is difficult to estimate the constant term. Hence the difference in the constants, which is thought to be a part of discrimination, cannot be calculated. The only solution is including the difference in the constant coefficients into the GSB. The decomposition formula will be the same, only the concept of the GSB is extended to include the difference in the constant terms.

Bayesian methods can also be used for our joint estimation. Since Bayesian estimation gives us the (posterior) distribution of coefficients, mean values of coefficients could be used for computing the GSB and the decomposition analysis (17), presuming that mean values are consistent estimates of population parameters. Unlike semiparametric method, this approach will not change the interpretation of the GSB or discrimination. Interesting development in recent Bayesian estimation is the Bayesian sampling method. It estimates the (posterior) distribution of coefficients of highly complicated models by utilizing Markov Chain Monte Carlo (MCMC) simulation methods. We can estimate the distribution of each compo-

[47]See Powell (1994) for a survey of semiparametric methods.

nent of the decomposition analysis (17), by evaluating them from the sampled estimates in each sampling round.[48]

The GSB approach can be implemented in conjunction with any kind estimation method; MLE, semiparametric, and Bayesian estimation methods. With the progress of computing technology, the GSB approach is able to handle virtually any selection issue. The GSB approach is a basic tool for wage differentials and discrimination analysis. It is conceptually simple, and practically versatile.

REFERENCES

Baker, Michael, Dwayne Benjamin, Andrée Desaulniers Cegep, and Mary Grant (1995): "The Distribution of the Male/Female log-wages Differential, 1970–1990," *Canadian Journal of Economics*, 28, 479-501.

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

Blau, Francine D. and Andrea H. Beller (1988): "Trends in log-wages Differentials by Gender, 1971-1981," *Industrial and Labor Relations Review*, 41, 513-529.

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

Bloom, David E. and Mark R. Killingsworth (1982): " Pay Discrimination Research and Litigation: The Use of Regression," *Industrial Relations*, 21, 318-339.

——— (1985): "Correction for Truncation Bias Caused by A Latent Truncation Variable," *Journal of Econometrics*, 27, 131-135.

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

[48]See Tanner (1996) and Chib and Greenberg (1996).

Blundell, Richard, Alan Duncan and Costas Meghir (1992): "Taxation in Empirical Labour Supply Models: Lone Mothers in the UK," *Economic Journal*, 102, 265-278.

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.

Chib, Siddhartha and Edward Greenberg (1996): "Markov Chain Monte Carlo Simulation Methods in Econometrics," *Econometric Theory*, 12, 409-431.

Choudhury, Sharmila (1993): "Reassessing the Male-Female Wage Differential: A Fixed Effects Approach," *Southern Economic Journal*, 60, 327-340.

Co, Catherine Y., Ira N. Gang, and Myeong-Su Yun (forthcoming): "Returns to Returning," *Journal of Population Economics*.

——— (1999): "Switching Model with Self-Selection: Self-Employment in Hungary," manuscript, Rutgers University.

Dolton, Peter J. and Gerald H. Makepeace (1986): "Sample Selection and Male-Female log-wages Differentials in the Graduate Labour Market," *Oxford Economic Papers*, 38, 317-341.

——— (1987): "Marital Status, Child Rearing and log-wages Differentials in the Graduate Labour Market," *Economic Journal*, 97, 897-922.

Dolton, Peter J., Gerald H. Makepeace, and W. van Der Klaauw (1989): "Occupational Choice and log-wages Determination: The Role of Sample Selection and Non-Pecuniary Factors," *Oxford Economic Papers*, 41, 573-594.

Ermisch, John F. and Robert E. Wright (1992): "Differential Returns to Human Capital in Full-time and Part-time Employment," in *Issues in Contemporary Economics Vol. 4 Women's Work in the World Economy* ed. by Nancy Folbre, Barbara Bergmann, Bina Agarwal and Maria Floro. New York: New York University Press, 195-212.

Fishe, Raymond P.H., Robert P. Trost, and Philip M. Lurie (1981): "Labor Force log-wages and College Choice of Young Women: An Examination of Selectivity Bias and Comparative Advantage," *Economics of Education Review*, 1, 169-191.

Gang, Ira N. and Myeong-Su Yun (1999), "The Gender Gap during Rapid Transition," manuscript, Rutgers University.

Green, William H. (1995): *LIMDEP: Version 7.0 User's Manual*, Plainview, NY: Econometric Software, Inc..

Ham, John C. (1982): "Estimation of Labour Supply Model with Censoring Due to Unemployment and Underemployment," *Review of Economic Studies*, 49, 335-354.

Härdle, Wolfgang and Oliver Linton (1994): "Applied Nonparametric Methods," in *Handbook of Econometrics, Vol. IV*, ed. by Robert F. Engel and Daniel L. McFadden. Amsterdam: Elsevier Science B.V., 2295-2339.

Hausman, Jerry (1980): "The Effect of Wages, Taxes, and Fixed Costs on Women's Labor Force Participation," *Journal of Public Economics*, 14, 161-194.

——— (1985) "The Econometrics of Nonlinear Budget Sets," *Econometrica*, 53, 1255-1282.

Hausman, Jerry A. and David A. Wise (1977): "Social Experimentation, Truncated Distributions, and Efficient Estimation," *Econometrica*, 45, 919-938.

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

——— (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 153-161.

Heckman, James J. and Thomas E. MaCurdy (1981): "New Methods for Estimating Labor Supply Functions: A Survey," *Research in Labor Economics*, 4, 65-102.

Hoffman, Saul D. and Charles R. Link (1984): "Selectivity Bias in Male Wage Equations: Black-White Comparisons," *Review of Economics and Statistics*, 66, 320-324.

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

Killingsworth, Mark R. (1983): *Labor Supply*, Cambridge: Cambridge University Press.

Killingsworth, Mark R. and James J. Heckman (1986): "Female Labor Supply: A Survey," in *Handbook of Labor Economics, Vol. I*, ed. by Orley Ashenfelter and Richard Layard. Amsterdam: Elsevier Science B.V., 103-204.

Lee, Lung-Fei (1983): "Generalized Econometric Models with Selectivity," *Econometrica*, 51, 507-512.

Maddala, G. S. (1983): *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press.

McCullough, B. D. and H. D. Vinod (1999): "The Numerical Reliability of Econometric Software," 37, 633-665.

Moffitt, Robert (1984): "The Estimation of a Joint Wage-Hours Labor Supply Model," *Journal of Labor Economics*, 2, 550-566.

——— (1986): "The Econometrics of Piecewise-Linear Budget Constraints," *Journal of Business and Economic Statistics*, 4, 317-328.

——— (1990): "The Econometrics of Kinked Budget Constrains," *Journal of Economic Perspectives*, 4, 119-139.

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.

Neuman, Shoshana and Ronald L. Oaxaca (1998): "Estimating Labor Market Discrimination with Selectivity Corrected Wage Equations: Methodological Considerations and An Illustration from Israel," paper presented at CEPR European Summer Symposium in Labour Economics and Migration, July 12, 1998.

Neumark, David (1988): "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination," *Journal of Human Resource*, 23, 279-295.

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

Oaxaca, Ronald L. and Michael R. Ransom (1988): "Searching for the Effect of Unionism on the Wages of Union and Nonunion Workers," *Journal of Labor Research*, 9, 139-148.

——— (1994): "On discrimination and the Decomposition of Wage Differentials," *Journal of Econometrics*, 61, 5-21.

Powell, James L. (1994): "Estimation of Semiparametric Models," in *Handbook of Econometrics, Vol. IV*, ed. by Robert F. Engel and Daniel L. McFadden. Amsterdam: Elsevier Science B.V., 2443-2521.

Reimers, Cordelia W. (1983): "Labor Market Discrimination Against Hispanic and Black," *Review of Economics and Statistics*, 65, 570-579.

Rosen, Sherwin (1986): "The Theory of Equalizing Differences," in *Handbook of Labor Economics, Vol. I*, ed. by Orley Ashenfelter and Richard Layard. Amsterdam: Elsevier Science B.V., 641-692.

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

Schaffner, Julie Anderson (1998): "Generating Conditional Expectations from Models with Selectivity Bias: Comment," *Economics Letters*, 58, 255-261.

Stern, Steven (1997): "Simulation-Based Estimation," *Journal of Economic Literature*, 35, 2006-2033.

Tallis, G. M. (1961): "The Moment Generating Function of the Truncated Multinormal Distribution," *Journal of Royal Statistical Society*, 23, series B, 223-229.

Tanner, Martin A. (1996): *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*, New York: Springer-Verlag.

Tunali, Insan (1986): "A General Structure of Models of Double-Selection and an Application to a Joint Migration/log-wages Process with Remigration," *Research in Labor Economics*, 8(B), 235-283.

Vella, Frank (1988): "Generating Conditional Expectations from Models with Selectivity Bias," *Economics Letters*, 28, 97-103.

Wales, T. J. and A. D. Woodland (1980): "Sample Selectivity and the Estimation of Labor Supply Functions," *International Economic Review*, 21, 437-468.

Wellington, Alison J. (1993): "Changes in the Male/Female Wage Gap, 1976-85," *Journal of Human Resources*, 28, 383-411.

Wright, Robert E. and John F. Ermisch (1991): "Gender Discrimination in the British Labour Market: A Reassessment," *Economic Journal* 101, 508-521.

Yun, Myeong-Su (1998): "Part-Time Work: Wage Differentials and Female Labor Supply," Working Paper, No. 98-35, Department of Economics, Rutgers University.

TABLE I

TWO APPROACHES TO THE DECOMPOSITION ANALYSIS

Selection Bias Correction Approach		
Error (e_{gN})	Two-Step	(MLE)
Selection Bias Correction Term	$\widehat{\lambda}_{gN}$	$\widetilde{\lambda}_{gN}$
(a) Conditional Expectations	$\widehat{\theta}_g\widehat{\lambda}_{gN}$	$\widetilde{\rho}_{e_g v_g}\widetilde{\sigma}_{e_g}\widetilde{\lambda}_{gN}$
(b) Unconditional Expectations	0	0
Selection Bias: (a) - (b)	$\widehat{\theta}_g\widehat{\lambda}_{gN}$	$\widetilde{\rho}_{e_g v_g}\widetilde{\sigma}_{e_g}\widetilde{\lambda}_{gN}$
Log-wages (Y^*_{gN})	Two-Step	(MLE)
Selection Bias Correction Term	$\widehat{\lambda}_{gN}$	$\widetilde{\lambda}_{gN}$
(a) Conditional Expectations	$X_{gN}\widehat{\beta}_g + \widehat{\theta}_g\widehat{\lambda}_{gN}$	$X_{gN}\widetilde{\beta}_g + \widetilde{\rho}_{e_g v_g}\widetilde{\sigma}_{e_g}\widetilde{\lambda}_{gN}$
(b) Unconditional Expectations	$X_{gN}\widehat{\beta}_g$	$X_{gN}\widetilde{\beta}_g$
Selection Bias: (a) - (b)	$\widehat{\theta}_g\widehat{\lambda}_{gN}$	$\widetilde{\rho}_{e_g v_g}\widetilde{\sigma}_{e_g}\widetilde{\lambda}_{gN}$
Generalized Selection Bias Approach		
Error (e_{gN})	(Two-Step)	MLE
(a) Conditional Expectations	$\mathrm{E}(\widehat{e}_{gN}\|S_{gN}=1)$ $=\widehat{\Lambda}_{gN}$	$\mathrm{E}(\widetilde{e}_{gN}\|S_{gN}=1)$ $=\widetilde{\Lambda}_{gN}$
(b) Unconditional Expectations	0	0
Selection Bias: (a) - (b)	$\widehat{\Lambda}_{gN}$	$\widetilde{\Lambda}_{gN}$
Log-wages (Y^*_{gN})	(Two-Step)	MLE
(a) Conditional Expectations	$X_{gN}\widehat{\beta}_g + \widehat{\Lambda}_{gN}$	$X_{gN}\widetilde{\beta}_g + \widetilde{\Lambda}_{gN}$
(b) Unconditional Expectations	$X_{gN}\widehat{\beta}_g$	$X_{gN}\widetilde{\beta}_g$
Selection Bias: (a) - (b)	$\widehat{\Lambda}_{gN}$	$\widetilde{\Lambda}_{gN}$

[a] $\widehat{e}_{gN} = Y^*_{gN} - X_{gN}\widehat{\beta}_g$, and $\widetilde{e}_{gN} = Y^*_{gN} - X_{gN}\widetilde{\beta}_g$. See text for the notations.

TABLE II

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 III
MEAN CHARACHERISTICS 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)
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 IV

LOG-WAGES: ILLUSTRATION 1

	Whites					
	OLS		Two-Step		MLE	
	Est.	(s.e.)	Est.	(s.e.)	Est.	(s.e.)
Constant	−0.214	(0.161)	−0.357*	(0.178)	−0.303	(0.165)
Age	0.059**	(0.008)	0.060**	(0.008)	0.060**	(0.008)
Age^2/100	−0.067**	(0.009)	−0.069**	(0.010)	−0.068**	(0.010)
Education	0.098**	(0.003)	0.103**	(0.004)	0.101**	(0.003)
MSA	0.089**	(0.019)	0.090**	(0.017)	0.089**	(0.019)
West	−0.023	(0.025)	−0.024	(0.025)	−0.024	(0.025)
South	−0.104**	(0.022)	−0.104**	(0.023)	−0.104**	(0.022)
MidWest	−0.105**	(0.022)	−0.106**	(0.022)	−0.105**	(0.022)
λ			0.136*	(0.057)		
σ_e					0.457**	(0.006)
ρ_{ev}					0.184*	(0.072)
Adjusted R^2	0.254		0.255			
	Others					
	OLS		Two-Step		MLE	
	Est.	(s.e.)	Est.	(s.e.)	Est.	(s.e.)
Constant	−0.102	(0.373)	−0.391	(0.509)	−0.192	(0.390)
Age	0.039*	(0.018)	0.042*	(0.021)	0.040*	(0.018)
Age^2/100	−0.036	(0.022)	−0.037	(0.025)	−0.036	(0.022)
Education	0.103**	(0.008)	0.113**	(0.013)	0.106**	(0.009)
MSA	0.105*	(0.051)	0.106*	(0.045)	0.105*	(0.051)
West	0.053	(0.063)	0.047	(0.071)	0.051	(0.063)
South	−0.101	(0.053)	−0.104	(0.056)	−0.102	(0.053)
MidWest	−0.032	(0.066)	−0.040	(0.069)	−0.035	(0.065)
λ			0.199	(0.217)		
σ_e					0.504**	(0.014)
ρ_{ev}					0.124	(0.162)
Adjusted R^2	0.229		0.230			

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

TABLE V

PARTICIPATION: ILLUSTRATION 1

	Whites			
	Two-Step		MLE	
	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			
	Two-Step		MLE	
	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 VI

DECOMPOSITION ANALYSIS: ILLUSTRATION 1

(a) Observed Log-Wage Differentials	
$E(Y_w)$ - $E(Y_o)$	0.077 (100.00%[a])
OLS	
Component	Logarithm (%)
(b) Difference in Characteristics	
(b.a) $(\overline{X}_w - \overline{X}_o)\widehat{\beta}_o$	0.043 (56.24%)
(b.b) $(\overline{X}_w - \overline{X}_o)\widehat{\beta}_w$	0.035 (45.22%)
(c) Difference in Coefficients	
(c.a) $\overline{X}_w(\widehat{\beta}_w - \widehat{\beta}_o)$	0.034 (43.75%)
(c.b) $\overline{X}_o(\widehat{\beta}_w - \widehat{\beta}_o)$	0.042 (54.78%)
Two-Step	
Component	Logarithm (%)
(b) Difference in Characteristics	
(b.a) $(\overline{X}_w - \overline{X}_o)\widehat{\beta}_o$	0.047 (60.94%)
(b.b) $(\overline{X}_w - \overline{X}_o)\widehat{\beta}_w$	0.036 (47.31%)
(c) Difference in Coefficients	
(c.a) $\overline{X}_w(\widehat{\beta}_w - \widehat{\beta}_o)$	0.067 (87.69%)
(c.b) $\overline{X}_o(\widehat{\beta}_w - \widehat{\beta}_o)$	0.078 (101.32%)
(d) Difference in Selection Bias	
$\widehat{\rho_{e_w v_w}\sigma_{e_w}}\overline{\lambda}_w - \widehat{\rho_{e_o v_o}\sigma_{e_o}}\overline{\lambda}_o$	-0.037 (-48.63%)
MLE	
Component	Logarithm (%)
(b) Difference in Characteristics	
(b.a) $(\overline{X}_w - \overline{X}_o)\widetilde{\beta}_o$	0.044 (57.75%)
(b.b) $(\overline{X}_w - \overline{X}_o)\widetilde{\beta}_w$	0.036 (46.54%)
(c) Difference in Coefficients	
(c.a) $\overline{X}_w(\widetilde{\beta}_w - \widetilde{\beta}_o)$	0.033 (42.43%)
(c.b) $\overline{X}_o(\widetilde{\beta}_w - \widetilde{\beta}_o)$	0.041 (53.64%)
(d) Difference in the GSB	
$(\overline{\widetilde{\Lambda}}_w - \overline{\widetilde{\Lambda}}_o)$	-0.0001 (-0.18%)

[a] Percentage of observed differentials contributed by component is in parentheses.

TABLE VII

LOG-WAGES: ILLUSTRATION 2

	Whites			
	Censoring		Truncation	
	Est.	(s.e.)	Est.	(s.e.)
Constant	-0.525^{**}	(0.169)	-0.224	(0.161)
Age	0.063^{**}	(0.008)	0.059^{**}	(0.008)
$\text{Age}^2/100$	-0.071^{**}	(0.010)	-0.067^{**}	(0.009)
Education	0.109^{**}	(0.003)	0.098^{**}	(0.003)
MSA	0.069^{**}	(0.019)	0.089^{**}	(0.018)
West	-0.019	(0.026)	-0.014	(0.025)
South	-0.093^{**}	(0.023)	-0.093^{**}	(0.023)
MidWest	-0.071^{**}	(0.024)	-0.099^{**}	(0.023)
σ_e	0.479^{**}	(0.008)	0.454^{**}	(0.006)
ρ_{ev}	-0.034	(0.042)	-0.065	(0.039)
	Others			
	Censoring		Truncation	
	Est.	(s.e.)	Est.	(s.e.)
Constant	-0.713	(0.427)	-0.101	(0.373)
Age	0.042^{*}	(0.020)	0.039^{*}	(0.018)
$\text{Age}^2/100$	-0.035	(0.024)	-0.036	(0.022)
Education	0.129^{**}	(0.010)	0.103^{**}	(0.008)
MSA	0.049	(0.058)	0.103^{*}	(0.051)
West	0.103	(0.069)	0.048	(0.063)
South	-0.043	(0.059)	-0.103	(0.053)
MidWest	-0.025	(0.072)	-0.034	(0.065)
σ_e	0.571^{**}	(0.030)	0.502^{**}	(0.014)
ρ_{ev}	0.033	(0.105)	-0.043	(0.086)

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

TABLE VIII

HOURS EQUATION: ILLUSTRATION 2

	Whites					
	OLS		Censoring		Truncation	
	Est.	(s.e.)	Est.	(s.e.)	Est.	(s.e.)
Constant	34.295**	(3.079)	6.153	(5.534)	33.315**	(3.119)
Age	−0.105	(0.158)	−0.037	(0.285)	−0.167	(0.162)
$Age^2/100$	0.058	(0.191)	−0.101	(0.346)	0.132	(0.195)
Marriage	−0.474	(0.474)	1.713*	(0.870)	−0.418	(0.475)
Children < 6	−2.644**	(0.359)	−8.494**	(0.618)	−2.664**	(0.360)
Children 6–18	−1.649**	(0.253)	−2.064**	(0.462)	−1.650**	(0.253)
Family Size	−0.054	(0.190)	−0.357	(0.350)	−0.028	(0.193)
Non-Labor Inc.	−0.039**	(0.008)	−0.157**	(0.015)	−0.042**	(0.009)
Log-Wages	4.025**	(0.298)	15.121**	(1.049)	4.962**	(0.639)
σ_v			16.857**	(0.293)	8.440**	(0.108)
Adjusted R^2	0.126					
	Others					
	OLS		Censoring		Truncation	
	Est.	(s.e.)	Est.	(s.e.)	Est.	(s.e.)
Constant	28.885**	(5.539)	2.508	(13.112)	28.321**	(5.676)
Age	0.136	(0.279)	−0.532	(0.660)	0.109	(0.281)
$Age^2/100$	−0.189	(0.338)	0.601	(0.800)	−0.159	(0.340)
Marriage	−0.121	(0.913)	5.246*	(2.080)	−0.087	(0.865)
Children < 6	0.235	(0.593)	−4.348**	(1.268)	0.241	(0.600)
Children 6–18	−0.408	(0.400)	−0.654	(0.933)	−0.404	(0.418)
Family Size	−0.005	(0.246)	−0.122	(0.587)	0.003	(0.282)
Non-Labor Inc.	−0.012	(0.022)	−0.213**	(0.051)	−0.015	(0.023)
Log-Wages	3.209**	(0.514)	18.100**	(2.267)	3.708**	(1.126)
σ_v			18.541**	(0.952)	7.236**	(0.199)
Adjusted R^2	0.054					

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

TABLE IX

ELASTICITY: ILLUSTRATION 2

		Whites	
	OLS	Censoring	Truncation
Compensated Wage Elasticity	0.564	2.238	0.620
Uncompensated Wage Elasticity	0.107	0.402	0.132
Total Income Elasticity	−0.457	−1.836	−0.488
		Others	
	OLS	Censoring	Truncation
Compensated Wage Elasticity	0.230	2.981	0.271
Uncompensated Wage Elasticity	0.085	0.482	0.099
Total Income Elasticity	−0.145	−2.499	−0.172

[a] Compensated wage elasticity is $\alpha_g/\overline{H}_w - \delta_g\overline{W}_w$, where $g = w$ and o, that is uncompensated wage elasticity minus total income elasticity.

[b] α and δ are respectively the coefficients for log-wages and non-labor income in hours equation.

[c] Elasticity is evaluated at the average hours and wages of white working women.

TABLE X

DECOMPOSITION ANALYSIS: ILLUSTRATION 2

(a) Observed Log-Wage Differentials	
$E(Y_w)$ - $E(Y_o)$	0.077 (100.00%[a])
Censoring	
Component	Logarithm (%)
(b) Difference in Characteristics	
(b.a) $(\overline{X}_w - \overline{X}_o)\widetilde{\beta}_o$	0.049 (63.11%)
(b.b) $(\overline{X}_w - \overline{X}_o)\widetilde{\beta}_w$	0.042 (54.90%)
(c) Difference in Coefficients	
(c.a) $\overline{X}_w(\widetilde{\beta}_w - \widetilde{\beta}_o)$	0.106 (138.21%)
(c.b) $\overline{X}_o(\widetilde{\beta}_w - \widetilde{\beta}_o)$	0.113 (146.42%)
(d) Difference in the GSB	
$(\overline{\widetilde{\Lambda}}_w - \overline{\widetilde{\Lambda}}_o)$	-0.078 (-101.32%)
Truncation	
Component	Logarithm (%)
(b) Difference in Characteristics	
(b.a) $(\overline{X}_w - \overline{X}_o)\widetilde{\beta}_o$	0.044 (56.75%)
(b.b) $(\overline{X}_w - \overline{X}_o)\widetilde{\beta}_w$	0.034 (43.71%)
(c) Difference in Coefficients	
(c.a) $\overline{X}_w(\widetilde{\beta}_w - \widetilde{\beta}_o)$	0.033 (43.24%)
(c.b) $\overline{X}_o(\widetilde{\beta}_w - \widetilde{\beta}_o)$	0.043 (56.28%)
(d) Difference in the GSB	
$(\overline{\widetilde{\Lambda}}_w - \overline{\widetilde{\Lambda}}_o)$	0.00001 (-0.01%)

[a] Percentage of observed differentials contributed by component is in parentheses.