Center for Research on Economic and Social Theory

and

Department of Economics

Working Paper Series

The Computation of Opportunity Costs in Polytomous Choice Models with Selectivity

Lung-fei Lee

January, 1993
Number 93-04

DEPARTMENT OF ECONOMICS
University of Michigan
Ann Arbor, Michigan 48109-1220

January, 1993

The Computation of Opportunity Costs in Polytomous Choice Models with Selectivity

by

Lung-fei Lee
Department of Economics, University of Michigan, Ann Arbor, MI 48109-1220

ABSTRACT

This article provides general formulas for the computation of opportunity costs (or forgone earnings) of unchosen alternatives in sample selection models with binary or polytomous choices. With observed choice probabilities and outcomes of chosen alternatives, the opportunity costs of unchosen alternatives can be evaluated. The derived formulas for the general model involve only single integrals. For the probability indexed sample selection models with conditional logit choice probabilities, which include the polytomous choice model of Lee [1983], the formulas have simple closed form expressions. These formulas are useful for empirical studies.

J.E.L. Classification: C30, C34, C35

Key Words and Phrases:
Discrete choice model, sample selection, polytomous choices, selection bias, conditional logit model, conditional expectation, opportunity cost, forgone earning.

Address:
Lung-fei Lee, Department of Economics, Lorch Hall, The University of Michigan, Ann Arbor, Michigan 48109-1220.

1. Introduction

In Lee [1983], sample selection models with polytomous choices have been formulated. The approach is based on some order statistics on the decision criteria and the use of some distributional transformations. The model has been used in several studies in the empirical economics literature (see, for examples, Trost and Lee [1984], Falaris [1987], Haveman et al [1988], Hensher and Milthorpe [1987], Wong [1986], and Wong and Liu [1988], among others). The model specification has originally emphasized on the correction of sample selection biases in continuous outcome equations. For some model analysis, one might be interested in the computation of opportunity costs for alternatives that are not chosen by an individual. The computation of such costs is not obvious and has neither been considered in Lee [1983] nor in the existed literature. This issue is also of interest for many other parametric or semiparametric models where the choice probabilities and the observed outcome equations have been specified or estimated without the specification of a joint distribution for all the distrubances in the system. In addition to the model in Lee [1983], the model in Olsen [1980] based on least squares correction or semiparametric models in Robinson [1988], Powell [1987], and Ichimura and Lee [1991] are examples.

As an illustration, let us briefly review the polytomous choice model with selectivity in Lee [1983]. The formulas derived in subsequent sections, however, will have broader applicability. A general sample selection model with m alternatives and m continuous outcome equations:

$$y_j = x_j\beta_j + \sigma_j u_j, \tag{1.1}$$

and

$$y_j^* = z_j\gamma_j + \eta_j, \quad j = 1, \cdots, m, \tag{1.2}$$

where u_j and η_j, for all j, are independent of x, where x consists of all distinct exogenous variables in the vectors $x_1, \cdots, x_m$ and $z_1, \cdots, z_m$, and have zero means and unit variances. The dependent variable or outcome y_j is observed if and only if the alterative j is chosen. The alternative j is chosen if and only if

$$y_j^* > \max\{y_k^* : k = 1, \cdots, m;\ k \neq j\}. \tag{1.3}$$

Equations (1.2) and (1.3) imply a polytomous choice model. Popular polytomous choice models are the conditional logit model of McFadden [1973] and the nested logit model (McFadden [1978]). Given the choice probabilities and the marginal distributions of u_j, $j = 1, \cdots, m$, Lee [1983] has suggested approaches for the correction of sample selection biases in the estimation of the outcome equations (1.1) based on order statistics and some distributional transformations. Define the random variable ϵ_j:

$$\epsilon_j = \max\{y_k^* : k = 1, \cdots, m;\ k \neq j\} - \eta_j. \tag{1.4}$$

The inequality (1.3) can then be rewritten as

$$z_j\gamma_j > \epsilon_j, \tag{1.5}$$

which is similar to a binary choice criterion for the alternative j. Let $z\gamma = (z_1\gamma_1, \cdots, z_m\gamma_m)'$ and

$$D_j = \begin{pmatrix} -I_{j-1} & \begin{matrix} 1 \\ \vdots \end{matrix} & 0 \\ 0 & \begin{matrix} \vdots \\ 1 \end{matrix} & -I_{m-j} \end{pmatrix}$$

is an $(m-1) \times m$ matrix, where I_s denotes the identity matrix of dimension s. Then

$$D_j z\gamma = (z_j\gamma_j - z_1\gamma_1, \cdots, z_j\gamma_j - z_{j-1}\gamma_{j-1}, z_j\gamma_j - z_{j+1}\gamma_{j+1}, \cdots, z_j\gamma_j - z_m\gamma_m). \tag{1.6}$$

Let $G_j(D_j z\gamma) = P(z_j\gamma_j - z_k\gamma_k > \eta_k - \eta_j, k \neq j, k = 1, \ldots, m|x)$ be the specified choice probability for the alternative j given x. The distribution of ϵ_j will be implicitly implied by G_j. The distribution of ϵ_j at c conditional on x is

$$F_j(c|x) = Prob[\epsilon_j < c|x] \\ = G_j(c - z_1\gamma_1, \cdots, c - z_{j-1}\gamma_{j-1}, c - z_{j+1}\gamma_{j+1}, \cdots, c - z_m\gamma_m). \tag{1.7}$$

The model specification in Lee [1983] differs from the conventional parametric specification as follows. In a conventional specification, a joint distribution of ϵ_j and u_j or, more explicitly, the joint distribution of $(\eta_1, \cdots, \eta_m)$ and u_j is specified. The marginal distibutions of u_j and ϵ_j are implied by the joint distibution. In the Lee's specification, marginal distributions of u_j and ϵ_j are first specified, and the model is then completed with a specification of a joint distribution for u_j and ϵ_j with specified marginal distributions. The advantage of the latter specification is that familiar distribution for the choice probabilities and distributions for the latent outcome equations can be imposed. When u_j is a standard normal variable, Lee [1983] suggested the use of the normality transformation. The ϵ_j can be transformed into a standard normal variate ϵ_j^* by $\epsilon_j^* = \Phi^{-1}(F_j(\epsilon_j|x))$, where Φ is the standard normal distribution function. The u_j and ϵ_j^* are then asssumed to be bivariate normally distributed with zero means, unit variances and correlation coefficient ρ_j. For this case, it follows that

$$E(y_j|x, \epsilon_j^*) = x_j\beta_j + \sigma_j\rho_j\epsilon_j^*, \tag{1.8}$$

and

$$E(y_j|x, j \text{ is chosen }) = x_j\beta_j - \sigma_j\rho_j\frac{\phi(\Phi^{-1}(G_j(D_j z\gamma)))}{G_j(D_j z\gamma)}, \tag{1.9}$$

where ϕ is the standard normal density function. If u_j were not normally distributed, other transformation rather than the normal transformation might be desirable for computational simplicity. If the marginal distribution of u_j were unknown, flexible functional specifications such as the bivariate Edgeworth expansion might also be used (see, Lee [1982] for details). Equations such as (1.9) provide the regression equations for the estimation of the outcome equations given the selected samples. Some simple two-stage consistent estimation methods or the maximum likelihood method can be applied to estimate the model. The above models can be regarded as special cases of a choice probability indexed sample selection model in that for each alternative, the sample selection bias correction term is a function of the choice probability of that alternative. Sample selection models with a binary decision rule belong to this category. As an example, the sample selection model with polytomous choices in Dubin and McFadden [1984] does not belong to this category. In a recent Monte Carlo study by Schmertmann [1991], there are evidences that the sample selection model of Dubin and McFadden [1984] is more likely subject to the multicollinearity issue and, on the other hand, the model in (1.8) is more likely subject to misspecification bias. The Lee [1983] article has emphasized on the specification of selection bias corrected outcome equations. After the estimation of unknown parameters in the model, one might be interested in the computation of opportunity costs (or forgone earnings) for unchosen alternatives given the individual's choice. Specifically, one might be interested in the computation of the conditional expectations $E(y_k|x, j \text{ is chosen })$ with $k \neq j$. The computation of such terms is not obvious from the model specification above.

The issues of computing the opportunity costs for the unchosen alternatives are general. In this article, we consider a general sample selection model where the disturbances in the latent equations are independent of all the exogenous variables in the model. We assume that the choice probabilities and the observed outcome equations for the chosen alternatives are estimable from observed data. We shown that these quantities provide sufficient information for the derivation of opportunity costs (or forgone expected outcomes) for the unchosen alternatives. These formulas may be useful for parametric and/or semiparametric models.

This article is organized as follows. In Section 2, we show that opportunity costs are identifiable from the bias corrected outcome equations and choice probabilities. General formulas for opportunity costs can be derived. Section 3 shows that for the probability indexed sample selection model with conditional logit choice probabilities, opportunity cost formulas have simple closd form expressions. Section 4 summaries our conclusions.

2. Selection Biases and Opportunity Costs in General Sample Selection Models

The correction of selection biases for the observed outcome equations in (1.2) for the sample selection model of (1.1) and (1.2) depends on the following conditional expectations:

$$E(u_j|x, \eta_k - \eta_j \leq z_j\gamma_j - z_k\gamma_k, \ k = 1, \cdots, m), \qquad j = 1, \cdots, m. \tag{2.1}$$

Given (2.1) and choice probabilities, the problem is to derive the conditional expectations $E(u_l|x, \eta_k - \eta_j \leq z_j\gamma_j - z_k\gamma_k, \ k = 1, \cdots, m)$ for $l \neq j$. Denote

$$v_j = (\eta_1 - \eta_j, \cdots, \eta_{j-1} - \eta_j, \eta_{j+1} - \eta_j, \cdots, \eta_m - \eta_j)', \qquad j = 1, \cdots, m. \tag{2.2}$$

The function G_j, where $G_j(c) = P(v_j < c)$ for any $c \in R^{m-1}$, corresponds to the distribution of v_j. For sample selection models where the u and v's are independent of all the exogenous variables in x, specification of (2.1) is, in general, equivalent to the specification of $E(u_j|v_j \leq c)$ for any $c \in R^{m-1}$.

It is interesting to note that $E(u_j|v_j = c)$ can be derived from $E(u_j|v_j \leq c)$ and the choice probability function $G_j(c)$. Let $f_j(u, v)$ denote the joint density of (u_j, v_j) and $g_j(v)$ denote the density function of v_j. It follows that

$$\begin{aligned} E(u_j|v_j \leq c) &= \frac{1}{G_j(c)}\int_{-\infty}^{c_{m-1}} \cdots \int_{-\infty}^{c_1}\int_{-\infty}^{\infty} u f_j(u, v) du dv \\ &= \frac{1}{G_j(c)}\int_{-\infty}^{c_{m-1}} \cdots \int_{-\infty}^{c_1} E(u_j|v_j) g_j(v_j) dv_j, \end{aligned} \tag{2.3}$$

where $c = (c_1, \cdots, c_{m-1})'$. By the fundamental theorem of Calculus, (2.3) implies that

$$E(u_j|v_j = c) = \frac{1}{g_j(c)}\frac{\partial^{m-1}[E(u_j|v_j \leq c)G_j(c)]}{\partial c_1 \cdots \partial c_{m-1}}. \tag{2.4}$$

With the knowledge of $E(u_j|v_j)$ and G_l, $E(u_j|v_l \leq r)$ for $r \in R^{m-1}$ can then be derived as follows. From the discrete choice component of the model, since only the differences $\eta_j - \eta_k$, where $j, k = 1, \cdots, m$, are relevant in the determination of choice probabilities, $\bar{\eta} = (\eta_1 - \eta_m, \cdots, \eta_{m-1} - \eta_m)$ can be taken as the basic vector of disturbances. (2.2) defines a one-to-one mapping from $\bar{\eta}$ in R^{m-1} to v_j in R^{m-1} for each j. Thus there exists a one-to-one transformation from v_l to v_j for each pair of l and j:

$$v_j = A_{lj}v_l, \tag{2.5}$$

where A_{lj} is a $(m-1) \times (m-1)$ matrix constructed from the expansion of a $(m-2) \times (m-2)$ identity matrix by inserting a $(m-1)$-dimensional row vector $(0, \cdots, 0, -1, 0, \cdots, 0)$, where -1 is in the $(j-1)$th position, as the lth row of A_{lj} and by inserting a $(m-1)$-dimensional column vector $(-1, -1, \cdots, -1)'$ as the $(j-1)$th column of A_{lj} if $l < j$; otherwise, by inserting $(0, \cdots, 0, -1, 0, \cdots, 0)$, where -1 is in the jth position, as the $(l-1)$th row of A_{lj}, and by inserting a $(m-1)$-dimensional column vector $(-1, -1, \cdots, -1)'$ as the jth column of A_{lj} if $j < l$. Explicitly, for $l < j$,

$$A_{lj} = \begin{pmatrix} I_{l-1} & 0 & \begin{matrix}-1\\ \vdots\end{matrix} & 0 \\ 0 & 0 & -1 & 0 \\ 0 & I_{j-1-l} & \vdots & 0 \\ 0 & 0 & \begin{matrix}\vdots\\ -1\end{matrix} & I_{m-j} \end{pmatrix} \tag{2.6}$$

and, for $j < l$,

$$A_{lj} = \begin{pmatrix} I_{j-1} & \begin{matrix}-1\\ \vdots\end{matrix} & 0 & 0 \\ 0 & \vdots & I_{l-1-j} & 0 \\ 0 & -1 & 0 & 0 \\ 0 & \begin{matrix}\vdots\\ -1\end{matrix} & 0 & I_{m-l} \end{pmatrix}. \tag{2.7}$$

For any vector $r \in R^{m-1}$, $E(u_j|v_l \leq r)$ can be computed from the following formula:

$$
\begin{aligned}
E(u_j|v_l \leq r) &= \frac{1}{G_l(r)} \int_{-\infty}^{r_{m-1}} \cdots \int_{-\infty}^{r_1} E(u_j|v_l = c) g_l(c) dc \\
&= \frac{1}{G_l(r)} \int_{-\infty}^{r_{m-1}} \cdots \int_{-\infty}^{r_1} E(u_j|v_j = A_{lj} c) g_l(c) dc \\
&= \frac{1}{G_l(r)} \int_{-\infty}^{r_{m-1}} \cdots \int_{-\infty}^{r_1} \nabla^{m-1} [E(u_j|v_j \leq A_{lj} c) G_j(A_{lj} c)] \frac{g_l(c)}{g_j(A_{lj} c)} dc,
\end{aligned} \tag{2.8}
$$

where ∇^{m-1} denotes the $(m-1)$th order partial differentiation operator with respect to the m-1 different arguments involved in relevant functions [see (2.4)]. This formula can be simplified. From (2.5), $g_l(v_l) = g_j(A_{lj} v_l)$ because the Jacobian of the transformation A_{lj} is unity. Hence

$$
E(u_j|v_l \leq r) = \frac{1}{G_l(r)} \int_{-\infty}^{r_{m-1}} \cdots \int_{-\infty}^{r_1} \nabla^{m-1} [E(u_j|v_j \leq A_{lj} a) G_j(A_{lj} a)] da. \tag{2.9}
$$

Consider the case $j < l$. Define the transformation $c = A_{lj} a$. Explicitly, from (2.5) with (2.7),

$$
A_{lj} a = (a_1, \cdots, a_{j-1}, a_{j+1}, \cdots, a_{l-1}, 0, a_l, \cdots, a_{m-1})' - a_j (1, \cdots, 1)',
$$

where $a = (a_1, \cdots, a_{m-1})'$. Since the Jacobian of A_{lj} is unity, by the transformation of variables, (2.9) implies

$$
\begin{aligned}
& E(u_j|v_l \leq r) \\
&= \frac{1}{G_l(r)} \int_{-r_j}^{\infty} \int_{-\infty}^{r_{m-1}+c_{l-1}} \cdots \int_{-\infty}^{r_l+c_{l-1}} \int_{-\infty}^{c_{l-1}} \int_{-\infty}^{r_{l-1}+c_{l-1}} \cdots \int_{-\infty}^{r_{j+1}+c_{l-1}} \int_{-\infty}^{r_{j-1}+c_{l-1}} \cdots \int_{-\infty}^{r_1+c_{l-1}} \\
& \qquad \nabla^{m-1} [E(u_j|v_j \leq c) G_j(c)] dc_1 \cdots dc_{l-2} dc_l \cdots dc_{m-1} dc_{l-1} \\
&= \frac{1}{G_l(r)} \int_{-r_j}^{\infty} \nabla_{l-1} [E(u_j|v_j \leq (r_1, \cdots, r_{j-1}, r_{j+1}, \cdots, r_{l-1}, 0, r_l, \cdots, r_{m-1})' + c_{l-1}(1, \cdots, 1)') \\
& \qquad \cdot G_j((r_1, \cdots, r_{j-1}, r_{j+1}, \cdots, r_{l-1}, 0, r_l, \cdots, r_{m-1}) + c_{l-1}(1, \cdots, 1))] dc_{l-1} \\
&= \frac{1}{G_l(r)} \int_{-\infty}^{r_j} \nabla_{l-1} [E(u_j|v_j \leq (r_1, \cdots, r_{j-1}, r_{j+1}, \cdots, r_{l-1}, 0, r_l, \cdots, r_{m-1})' - t(1, \cdots, 1)') \\
& \qquad \cdot G_j((r_1, \cdots, r_{j-1}, r_{j+1}, \cdots, r_{l-1}, 0, r_l, \cdots, r_{m-1}) - t(1, \cdots, 1))] dt \\
&= \frac{1}{G_l(r)} \int_{-\infty}^{r_j} \nabla_{l-1} [E(u_j|v_j \leq A_{lj}(r_1, \cdots, r_{j-1}, t, r_{j+1}, \cdots, r_{m-1})') \\
& \qquad \cdot G_j(A_{lj}(r_1, \cdots, r_{j-1}, t, r_{j+1}, \cdots, r_{m-1})')] dt,
\end{aligned} \tag{2.10}
$$

where ∇_s denotes the partial differentiation operator with respect to the sth argument,i.e., $\nabla_s [E(u_j|v_j \leq c) G_j(c)] = \frac{\partial}{\partial c_s} [E(u_j|v_j \leq c) G_j(c)]$ with $c = (c_1, \cdots, c_{m-1})'$. Similarly, if $l < j$,

$$
A_{lj} a = (a_1, \cdots, a_{l-1}, 0, a_l, \cdots, a_{j-2}, a_j, \cdots, a_{m-1})' - a_{j-1}(1, \cdots, 1)',
$$

and

$$
\begin{aligned}
E(u_j|v_l \leq r) = & \frac{1}{G_l(r)} \int_{-\infty}^{r_{j-1}} \nabla_l [E(u_j|v_j \leq A_{lj}(r_1, \cdots, r_{j-2}, t, r_j, \cdots, r_{m-1})') \\
& \cdot G_j(A_{lj}(r_1, \cdots, r_{j-2}, t, r_j, \cdots, r_{m-1})')] dt.
\end{aligned} \tag{2.11}
$$

Thus given G_l, G_j, and $E(u_j|v_j \leq c)$, $E(u_j|v_l \leq r)$ can be evaluated. The expected opportunity cost for the outcome y_j given the alternative l being chosen is

$$
E(y_j|x, l \text{ is chosen }) = x_j \beta_j + \sigma_j E(u_j|x, v_l \leq D_l z \gamma). \tag{2.12}
$$

The formulas (2.10) and (2.11) can be used to evaluate $E(u_j|x, v_l \leq D_l z \gamma)$. Whether they have closed form expressions or not will depend on more specific structures of the selection model. In the subsequent section, we will show that, for the probability indexed sample selection model with the conditional logit choice probabilities, closed form expressions of (2.10) and (2.11) can be derived. However, closed form expressions do not seem to exist for general polytomous choice models. For the latter cases, (2.10) and (2.11) can be numerically computed by standard subroutines for a single integral.

Finally, we note that for the model with a binary decision rule the formulas (2.10) and (2.11) can always be simplified. For $m = 2$, $A_{lj} = -1$ for all l and j with $l \neq j$, and

$$
\begin{aligned}
E(u_j|v_l \leq r) &= \frac{1}{G_l(r)} \int_{-\infty}^{r} \nabla [E(u_j|v_j \leq -t) G_j(-t)] dt \\
&= \frac{1}{G_l(r)} \{E(u_j) - E(u_j|v_j \leq -r) G_j(-r)\} \\
&= -E(u_j|v_j \leq -r) \frac{G_j(-r)}{G_l(r)},
\end{aligned} \tag{2.13}
$$

by using the property $E(u_j) = 0$. The derivation of the opportunity costs in the model with a binary decision rule is conceptually straightforward because there is only a single disturbance in the decision component of the model.

3. Probability Indexed Sample Selection Models with Conditional Logit Choice Probabilities

Probability indexed sample selection models are referred here to the models of (1.1) and (1.2) with

$$E(u_j|v_j \le c) = \psi_j(G_j(c))/G_j(c), \tag{3.1}$$

where ψ_j is a function of the probability $G_j(c)$, which will in general depend on the alternative j. We note that as c tends to infinity, there will be no selection bias in (1.1). As $E(u_j) = 0$, it following that $\psi_j(1) = 0$. The sample selection models in Lee [1983] are in this category [see (1.9)]. With this type of models, closed form expressions for (2.10) and (2.11) can be derived when choice probabilities are the conditional logit probabilities of McFadden [1973].

The conditional logit model corresponds to

$$G_j(c) = \frac{1}{1+\sum_{k=1}^{m-1} e^{-c_k}}, \qquad j = 1,\cdots,m. \tag{3.2}$$

With (3.1) and (3.2),

$$E(u_j|v_j \le c)G_j(c) = \psi_j\left(\frac{1}{1+\sum_{k=1}^{m-1} e^{-c_k}}\right), \qquad j = 1,\cdots,m, \tag{3.3}$$

and

$$\begin{aligned}&\frac{\partial[E(u_j|v_j \le c)G_j(c)]}{\partial c_i}\\ &= \nabla\psi_j\left(\frac{1}{1+\sum_{k=1}^{m-1} e^{-c_k}}\right)\cdot\frac{e^{-c_i}}{\left(1+\sum_{k=1}^{m-1} e^{-c_k}\right)^2}, \qquad j=1,\cdots,m;\ i=1,\cdots,m-1,\end{aligned} \tag{3.4}$$

where $\nabla\psi_j(t) = \partial\psi_j(t)/\partial t$. For $j < l$, (3.3) implies that

$$\begin{aligned}&E(u_j|v_j \le A_{lj}(r_1,\cdots,r_{j-1},t,r_{j+1},\cdots,r_{m-1})')G_j(A_{lj}(r_1,\cdots,r_{j-1},t,r_{j+1},\cdots,r_{m-1})')\\ &= \psi_j\left(\frac{1}{1+\sum_{k\ne j}^{m-1} e^{-r_k+t}+e^t}\right).\end{aligned} \tag{3.5}$$

The $(l-1)$ component of $A_{lj}(r_1,\cdots,r_{j-1},t,r_{j+1},\cdots,r_{m-1})'$ is simply $-t$ from (2.7). Hence it follows from (3.4)

$$\begin{aligned}&\nabla_{l-1}[E(u_j|v_j \le A_{lj}(r_1,\cdots,r_{j-1},t,r_{j+1},\cdots,r_{m-1})')G_j(A_{lj}(r_1,\cdots,r_{j-1},t,r_{j+1},\cdots,r_{m-1})')]\\ &= \nabla\psi_j\left(\frac{1}{1+\sum_{k\ne j}^{m-1} e^{-r_k+t}+e^t}\right)\cdot\frac{e^t}{(1+\sum_{k\ne j}^{m-1} e^{-r_k+t}+e^t)^2}.\end{aligned} \tag{3.6}$$

Similarly, for $l < j$,

$$\begin{aligned}&\nabla_l[E(u_j|v_j \le A_{lj}(r_1,\cdots,r_{j-2},t,r_j,\cdots,r_{m-1})')G_j(A_{lj}(r_1,\cdots,r_{j-2},t,r_j,\cdots,r_{m-1})')]\\ &= \nabla\psi_j\left(\frac{1}{1+\sum_{k\ne j-1}^{m-1} e^{-r_k+t}+e^t}\right)\cdot\frac{e^t}{(1+\sum_{k\ne j-1}^{m-1} e^{-r_k+t}+e^t)^2}.\end{aligned} \tag{3.7}$$

Therefore, (2.10) and (2.11) imply that, for all l and j with $l \ne j$,

$$E(u_j|v_l \le r) = \frac{1}{G_l(r)}\int_{-\infty}^{r_i}\nabla\psi_j\left(\frac{1}{1+\sum_{k\ne i}^{m-1} e^{-r_k+t}+e^t}\right)\cdot\frac{e^t}{(1+\sum_{k\ne i}^{m-1} e^{-r_k+t}+e^t)^2}dt, \tag{3.8}$$

where $i = j$ for $j < l$, but $i = j-1$ for $l < j$. Since

$$\frac{d}{dt}\psi_j\left(\frac{1}{1+\sum_{k\ne i}^{m-1} e^{-r_k+t}+e^t}\right) = -\nabla\psi_j\left(\frac{1}{1+\sum_{k\ne i}^{m-1} e^{-r_k+t}+e^t}\right)\cdot\frac{(\sum_{k\ne i}^{m-1} e^{-r_k}+1)e^t}{(1+\sum_{k\ne i}^{m-1} e^{-r_k+t}+e^t)^2},$$

it follows from (3.8) that

$$\begin{aligned}E(u_j|v_l \le r) &= -\frac{1}{1+\sum_{k\ne i}^{m-1} e^{-r_k}}\cdot\frac{1}{G_l(r)}\int_{-\infty}^{r_i} d\psi_j\left(\frac{1}{1+\sum_{k\ne i}^{m-1} e^{-r_k+t}+e^t}\right)\\ &= -\frac{1}{1+\sum_{k\ne i}^{m-1} e^{-r_k}}\cdot\frac{1}{G_l(r)}\cdot\psi_j\left(\frac{1}{1+\sum_{k\ne i}^{m-1} e^{-r_k+r_i}+e^{r_i}}\right),\end{aligned} \tag{3.9}$$

for all l, j with $l \ne j$.

For the sample observation with characteristics x, $r = D_l z\gamma$, which implies that $r_j = z_l\gamma_l - z_j\gamma_j$ for $j < l$, but $r_{j-1} = z_l\gamma_l - z_j\gamma_j$ for $l < j$ [see (1.6)]. Hence

$$\begin{aligned}&E(u_j|x, l \text{ is chosen })\\ &= E(u_j|x, z_l\gamma_l+\eta_l \ge z_k\gamma_k+\eta_k,\ \ k=1,\cdots,m)\\ &= -\frac{1}{1+\sum_{k\ne j,l}^{m} e^{-(z_l\gamma_l-z_k\gamma_k)}}\cdot(1+\sum_{k\ne l}^{m} e^{-(z_l\gamma_l-z_k\gamma_k)})\cdot\psi_j\left(\frac{1}{1+\sum_{k\ne j,l}^{m} e^{z_k\gamma_k-z_j\gamma_j}+e^{z_l\gamma_l-z_j\gamma_j}}\right)\\ &= -\frac{\sum_{k=1}^{m} e^{z_k\gamma_k}}{\sum_{k\ne j}^{m} e^{z_k\gamma_k}}\cdot\psi_j\left(\frac{e^{z_j\gamma_j}}{\sum_{k=1}^{m} e^{z_k\gamma_k}}\right)\\ &= -\frac{1}{\left(1-\frac{e^{z_j\gamma_j}}{\sum_{k=1}^{m} e^{z_k\gamma_k}}\right)}\cdot\psi_j\left(\frac{e^{z_j\gamma_j}}{\sum_{k=1}^{m} e^{z_k\gamma_k}}\right), \qquad j=1,\cdots,m;\ l=1,\cdots,m;\ j\ne l.\end{aligned} \tag{3.10}$$

It is interesting to note from (3.10) that the conditional expectation $E(u_j|x, l$ is chosen), where $l \ne j$, depends only on the probability that the alterative j is not chosen and is invariant with l. This invariance property can be regarded as a theoretical structure or restriction on probability indexed sample selection models with conditional logit choice probabilities. Another interesting implication of (3.10) is that when there are positive selection biases, i.e., $E(u_j|x, j$ is chosen $) > 0$, for all j, the expected opportunity costs for the unchosen alternatives will always be less than population averages, i.e., $E(u_l|x, j$ is chosen $) < 0$ for all l, $l \ne j$. This property generalizes the one of the classical sample selection model with a binary choice decision rule and normal disturbances (Heckman [1979]).

4. Conclusion

This article provides a general method for the computation of opportunity costs for sample selection models with polytomous choices. It has been shown that the opportunity costs for the unchosen alternatives can be identified and computed from choice probabilities and observed outcomes of the chosen alternatives. The derived formulas involve single integrals for the general sample selection model. These formulas may be useful for empirical studies.

We have also shown that, for probability indexed sample selection models with conditional logit choice probabilities, which include the models in Lee [1983], the formulas have simple closed form expressions. Such formulas illustrate some of the theoretical structure or restrictions imposed on these models.

ACKNOWLEDGEMENTS

The author appreciates having financial support under NSF grant no. SES-9296071 for his research. The issues in this article are motivated by Professor Michael Finch of the School of Public Health at the University of Minnesota.

Reference

1. Dubin, J. and D. McFadden (1984), "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption", *Econometrica* 52: 345-362.
2. Falaris, E.M., (1987), "A Nested Logit Migration Model with Selectivity", *International Economic Review* 28: 429-443.
3. Haveman. R., B. Wolfe, and J. Warlick (1988), "Labor Market Behavior of Older Men: Estimates from a Trichotomous Logit Model", *Journal of Public Economics* 36: 153-175.
4. Heckman, J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica* 47: 153-161.
5. Hensher, D.A. and F.W. Milthorpe (1987), "Selectivity Correction in Discrete-Continuous Choice Analysis", *Regional Science and Urban Economics* 17: 123-150.
6. Ichimura, H. and L.F. Lee (1991), "Semiparametric least squares estimation of multiple index models: single equation estimation", Ch. 1 in *Nonparametric and Semiparametric Methods in Econometrics and Statistics: Proceedings of the Fifth International Symposium in Economic Theory and Econometrics*, ed. by W.A. Barnett, J.L. Powell and G.E. Tauchen, Cambridge U. Press, New York, N.Y.
7. Lee, L.-F. (1982), "Some Approaches to the Correction of Selectivity Bias", *Review of Economic Studies* 49: 355-372.
8. Lee, L.-F. (1983), "Generalized Econometric Models with Selectivity", *Econometrica* 51: 507-512.
9. McFadden, D. (1973), "Conditional Logit Analysis of Qualitative Choice Behavior", in *Frontiers in Econometrics*, ed. by P. Zarembka. New York: Academic Press.
10. McFadden, D. (1978), "Modelling the Choice of Residential Location", in *Spatial Interaction Theory and Residential Location*, ed. by A. Karlquist et al. Amsterdam: North-Holland.
11. Olsen, R.J. (1980), "A least squares correction for selectivity bias", *Econometrica*, 48: 1815-1820.
12. Powell, J.L. (1987), "Semiparametric estimation of bivariate latent variable models", discussion paper no. 8704, Social Systems Research Institute, U. of Wisconsin, Madison.
13. Robinson, P.M. (1988), "Root-n-consistent semiparametric regression", *Econometrica* 56: 931-954.
14. Schertmann C.P. (1991), "Selectivity Bias Correction Methods in Polychotomous Sample Selection Models", manuscript, Department of Economics, Florida State University.
15. Trost, R.P. and L.-F. Lee (1984), "Technical Training and Earnings: A Polytomous Choice Model with Selectivity", *Review of Economics and Statistics* 66: 151-156.
16. Wong, Y.C. (1986), "Entrepreneurship, Marriage, and Earnings", *Review of Economics and Statistics* 68: 693-699.
17. Wong, Y.C. and P.W. Liu (1988), "The Distribution of Benefits Among Public Housing Tenants in Hong Kong and Related Policy Issues", *Journal of Urban Economics* 23: 1-20.