

Caudill, Sten B., Grootuis, Peter A., & Whitehead, John C. (2011). The development and estimation of a latent choice multinomial logit model with application to contingent valuation, *American Journal of Agricultural Economics*, 93(4):983-992. Publisher version available from <http://www.oxfordjournals.org/en/> (ISSN: 0002-9092) (DOI: 10.1093/ajae/aae030)

The Development and Estimation of a Latent Choice Multinomial Logit Model with Application to Contingent Valuation

Steven B. Caudill, Peter A. Grootuis, and John C. Whitehead

ABSTRACT

We offer a new approach to investigate hypothetical bias in contingent valuation using a latent choice multinomial logit model. To develop this model, we extend Dempster, Laird, and Rubin's 1977 work on the expectations maximization algorithm to the estimation of a multinomial logit model with missing information on category membership. Our model can be used to determine within-choice heterogeneity. Using data on the preservation of Saginaw wetlands in Michigan, we find evidence for two types of Yes responders in the data. We suggest that one set of Yes responders consists of yea-sayers who answer Yes to the hypothetical question but are less likely to pay the bid amount if it were real. We suggest that the second group of respondents does not suffer from hypothetical bias and are more likely to pay the bid amount if it were real. Even if the connection to hypothetical bias cannot be made, our method can be used in sensitivity analyses of willingness-to-pay estimates.

Key words: contingent valuation, hypothetical bias, individual heterogeneity, multinomial logit.

In order to investigate the possibility of hypothetical bias in the contingent valuation method, we develop a generalization of the multinomial logit (MNL) model we call the latent choice multinomial logit (LCMNL) model. This model allows for within-choice parameter heterogeneity. With cross-choice parameter constraints imposed, the LCMNL model can be interpreted as either a model for misclassification or a generalization of the pooling test of Cramer and Ridder (1991). The constrained version of the LCMNL model has been applied by Caudill (2006), Caudill and Groothuis (2005), and Caudill, Ayuso, and Guillen (2005).

The present application is the first use of the unconstrained version of the LCMNL model. Unlike the constrained LCMNL, the unconstrained version can locate parameter heterogeneity. An examination of a previous willingness-to-pay (WTP) application using the constrained LCMNL model can be used to highlight the present contribution to the literature. In an application of the constrained LCMNL, Caudill and Groothuis (2005) are able to determine whether the respondents who do not know if they are willing to pay are more like those respondents who are willing to pay, more like those who are not willing to pay, or a distinct group. In the present contingent valuation method using the unconstrained LCMNL model, we seek to determine whether the sample is characterized by two kinds of Yes responders—different from each other and different from No responders—each represented by different parameter vectors. Using contingent valuation data on the Saginaw wetlands in Michigan, we find two different types of Yes responders in the data. Information from a follow-up certainty question is available, and, using this information, we find that one group of Yes responders is very certain about WTP, and the other group of Yes responders is not very certain about WTP. This finding provides circumstantial evidence indicating that the second group of Yes responders is subject to hypothetical bias (Loomis 2011).

Hypothetical bias is the tendency for hypothetical WTP to overestimate real WTP (Blumenschein et al. 2008; Cummings et al. 1997; Cummings, Harrison, and Rutström 1995). Hypothetical bias is usually attributed to the presence of passive use values and to respondent unfamiliarity with paying for policies that provide passive use value. In fact, hypothetical bias has been found in studies examining willingness to pay for private goods, for which no passive use values should exist (List and Gallet 2001). Surprisingly, hypothetical bias is ignored in much of the literature on the contingent valuation method (Harrison 2006).

Hypothetical bias may arise because respondent answers to contingent valuation WTP questions have no real consequences other than a connection to the influence of government policy (Carson and Groves 2007). Respondents who state that they would pay for the policy change are not required to actually pay. Consequently, some respondents may state that they would pay for the policy when, in fact, they would not if placed in the real situation. There are at least two explanations for this behavior. The respondent may be trying to influence policy by signaling support (e.g., strategic bias), or the respondent may simply be responding affirmatively to please the interviewer (e.g., yea-saying). The presence of hypothetical bias leads to inflated WTP estimates. The implication is that these estimates from the contingent valuation method are upper bounds on benefits conferred in the context of benefit-cost analysis unless steps are taken to mitigate hypothetical bias (Whitehead and Blomquist 2006).

In the past, two approaches have been developed to mitigate the overstatement of hypothetical WTP: the *ex ante* approach and the *ex post* approach. The *ex ante* approach addresses hypothetical bias in the survey design stage. Respondents are variously (a) told that there are substitutes for the policy available (Loomis, Gonzalez-Caban, and Gregory 1994), (b) reminded that they are income constrained (Loomis et al. 1996), (c) asked to answer as if they were placed in an actual payment situation (Loomis et al. 1996), and/or (d) told that hypothetical bias is a significant problem and asked not to succumb to this type of respondent error (Cummings and Taylor 1999). The *ex post* approach addresses hypothetical bias with follow-up questions.

Respondents who indicate that they are willing to pay for the policy are asked to rate the certainty they have in their response. Respondent certainty is measured on a qualitative or quantitative scale where the low and high ends allow respondents to express their degree of certainty about their payment (Blumenschein et al. 2008; Champ et al. 1997). Based on the certainty of the respondent, hypothetical WTP responses can be recoded to avoid uncertain values.

Although both the *ex ante* and *ex post* approaches have been somewhat successful in the past, each has shortcomings. A problem with both approaches is that there is usually no way to detect whether they are effective, unless hypothetical WTP is compared with real WTP. Also, the *ex post* approach suffers from ad hoc cutoff assignments. Our approach, using the LCMNL model, complements the *ex ante* and *ex post* approaches by investigating respondent heterogeneity. We are able to statistically determine whether there is more than one type of Yes responder in the data. This finding of respondent heterogeneity, alone, allows for additional sensitivity analyses of WTP and can be used in conjunction with both the *ex post* and *ex ante* approaches to adjust WTP estimates and mitigate hypothetical bias.

A Latent Choice Multinomial Logit Model

In an important article, Dempster, Laird, and Rubin (1977) (henceforth DLR) show how the EM algorithm can be used to obtain maximum likelihood estimates from incomplete data. The first illustration discussed by DLR is a multinomial probability example using data from Rao (1955) on 197 animals divided into four categories. DLR consider the case where the original first category of animal is split into two new categories, but exactly which animals are assigned to which category is unknown. The resulting multinomial probability model is characterized by five categories with missing information on membership in the first and second categories. DLR show how the parameters of this model can be estimated by maximum likelihood using the EM algorithm. This article extends DLR's work on the use of the EM algorithm to the estimation of an MNL with within-choice parameter heterogeneity we call the LCMNL. A similar model has been estimated by Magder and Hughes (1997).

The usual MNL model is a special case of the LCMNL.¹ To present the LCMNL model and to facilitate comparisons with the MNL model, we adopt the language associated with the nested logit model. That is, we characterize the choice model in terms of branches and stems. Branches are observed alternatives and stems are unobserved alternatives nested inside branches. The researcher must determine *a priori* the number and location of the unobservable stems in the model. A branch may

contain any number of unobservable stems, including zero. Thus, the usual MNL model is an LCMNL having only observable branches, each with no stems.

As a point of departure, we begin with the usual MNL model, with m observable choices or branches. Probabilities in this model are given by

$$(1) \quad P_{ij} = \exp[X_i \beta_j] / D \text{ where}$$

$$D = \sum_{j=1}^m \exp[X_i \beta_j].$$

The data include a set of indicator or response variables denoting the choice made by each individual. Let $Y_{ij} = 1$ if individual i makes choice j and $Y_{ij} = 0$ otherwise. This leads to the usual log-likelihood function and the familiar first-order conditions

$$(2) \quad \log L = \sum_{i=1}^n \sum_{j=1}^m Y_{ij} \log P_{ij}$$

$$(3) \quad \frac{\partial \log L}{\partial \beta_j} = \sum_{i=1}^n (Y_{ij} - P_{ij}) X'_i, \text{ for } j = 1, \dots, m.$$

In the LCMNL model, there is the possibility that some of the observed branches contain unobserved stems. To develop this model, we denote the unknown parameters by β_{jk} , where j indicates "branch" and k indicates "stem." Probabilities in this model are given by

$$(4) \quad P_{ijk} = \exp[X_i \beta_{jk}] / D \text{ where}$$

$$D = \sum_{j=1}^m \sum_{k=1}^{s_j} \exp[X_i \beta_{jk}]$$

where m represents the number of branches in the model and s_j represents the number of stems associated with branch j . The construct, α_{ijk} , is the conditional probability of stem k given branch j calculated for each individual:

$$(5) \quad \alpha_{ijk} = \frac{P_{ijk}}{\sum_{k=1}^{s_j} P_{ijk}} = \frac{P_{ijk}}{P_{ij}} \text{ letting } \sum_{k=1}^{s_j} P_{ijk} = P_{ij}.$$

Note that if the values of α equal one for all individuals and all stems, the first-order conditions for the usual MNL model obtain. The sum of probabilities of all stems on a branch must equal the probability of the branch. If $s_j = 1$ for all branches (no branch has any stems), the usual MNL model obtains.

The log-likelihood function for the LCMNL model is based on probabilities like those in equation (4) and is given by

$$(6) \quad \log L = \sum_{i=1}^n \sum_{j=1}^m Y_{ij} \log \left[\sum_{k=1}^{s_j} P_{ijk} \right].$$

The indicator variable, Y_{ij} , indicates only which branch is chosen, as stem membership is unknown. That is why the logarithm of the sum of the probabilities on a branch appears in the likelihood function. Once again, it is clear that if $s_j=1$ for all branches, the log-likelihood function for the usual MNL model given in equation (2) obtains.

Maximization of the likelihood function for the LCMNL requires the first derivatives with respect to the unknown parameters in the model. The first derivatives can be shown to equal

$$(7) \quad \frac{\partial \log L}{\partial \beta_{jk}} = \sum_{i=1}^n \alpha_{ijk} (Y_{ij} - P_{ij}) X'_i, \text{ for} \\ j = 1, \dots, m \quad k = 1, \dots, s_j,$$

where the construct, α_{ijk} , is defined as in equation (5).

A Specific Case

The model is applied to contingent valuation WTP data. Individuals respond to WTP questions by indicating *Yes* or *No*. These are the two observable responses or branches. In the model presented below, we allow for two types of *Yes* responders. These are the two stems associated with the *Yes* branch. This allows within-response heterogeneity. Using follow up certainty questions, we suggest that one group of *Yes* responders—those with less conviction about their responses—might be subject to hypothetical bias.

To begin, we detail the estimation of an LCMNL model with two branches, where one branch contains two stems. First we examine an MNL model in which individuals make one of three choices. Probabilities in this model are given by

$$(8) \quad P_{ij} = \frac{\exp(X_i \beta_j)}{\sum_{j=1}^3 \exp(X_i \beta_j)} \quad j = 1, \dots, 3$$

where β_j is the parameter to be estimated and X_i is a vector of exogenous variables. The usual normalization applies so that $\beta_3 = 0$. The alternative selected in this model is represented by the usual set of dummy variables: Y_{i1} , Y_{i2} , and Y_{i3} , each taking the value zero or one, indicating that an alternative was or was not selected. The log-likelihood function in this MNL model is given by

$$(9) \quad \log L = \sum_{i=1}^n (Y_{i1} \log(P_{i1}) + Y_{i2} \log(P_{i2}) \\ + Y_{i3} \log(P_{i3}))$$

where n is the sample size. Maximum likelihood estimation of the parameters in this model is routine. The first-order conditions for maximization are

$$(10) \quad \frac{\partial \log L}{\partial \beta_j} = \sum_{i=1}^n (Y_{ij} - P_{ij})X'_i = 0 \quad j=1,2.$$

To develop the corresponding LCMNL model, let us assume that there are two observable choices. One of the observable choices is populated by individuals characterized by different parameter vectors; that is, within choice parameter heterogeneity is present. In other words, the model has two branches and one branch has two stems. We denote the probabilities associated with the first branch by P_1 and probabilities on the stems associated with the second branch by P_{21} and P_{22} . The probability definitions are essentially the same as with the MNL model with three observed choices but the likelihood function has changed. This change occurs because only the branch choice, Y_2 , is observed. The resulting incomplete-data/observed log-likelihood function is a special case² of that given in equation (6):

$$(11) \quad \log L = \sum_{i=1}^n [Y_{i1} \log(P_{i1}) + Y_{i2} \log(P_{i21} + P_{i22})].$$

The first-order condition with respect to the parameters associated with the first branch is the same as in the MNL model:

$$(12) \quad \frac{\partial \log L}{\partial \beta_1} = \sum_{i=1}^n (Y_{i1} - P_{i1})X'_i = 0.$$

The first-order condition with respect to the parameters associated with the first stem on the second branch is given by

$$(13) \quad \frac{\partial \log L}{\partial \beta_{21}} = \sum_{i=1}^n \alpha_{i21} (Y_{ij} - P_{ij})X'_i = 0,$$

where $\alpha_{i21} = \frac{P_{i21}}{P_{i21} + P_{i22}}, i = 1, \dots, n.$

The first-order condition (if it were needed) with respect to the parameters associated with the second stem on the second branch is similarly defined.

Estimation of the model by maximum likelihood is relatively straightforward if the EM algorithm is used. In the expectations step of the EM algorithm, the latent or missing variables are replaced by their conditional expectations, given the data and initial parameter estimates. The likelihood function is then maximized to obtain new parameter values. These values can be used to obtain new conditional expectations, and the process is repeated.

To use the EM algorithm to estimate the simple LCMNL model, we denote the set of unobservable indicator variables associated with the two stems on the second branch by Y_{i21} and Y_{i22} . These dummy variables take the value 1 if the observation is

associated with that stem and zero otherwise. Using these unobservable variables, the complete data log-likelihood in our simple problem can be written

$$(14) \quad \log L = \sum_{i=1}^n (Y_{i1} \log(P_{i1}) + Y_{i2}^* \log(P_{i21}) + Y_{i2}^* \log(P_{i22})).$$

The likelihood function above characterizes the estimation problem as a missing data problem. If the values of unobserved Y^* were known, estimation would be as simple as estimating an MNL model. In the expectations, or “E,” step of the EM algorithm, the values of unobserved Y^* are replaced with their conditional expectations, given the data and values of the unknown parameters. The conditional expectation is the probability of a stem, given the branch. These conditional expectations or probabilities are well known in the logit model and are given by

$$(15) \quad E(Y_{i2}^* | Y_{i2} = 1) = \frac{\exp(X_i \beta_{21})}{\exp(X_i \beta_{21}) + \exp(X_i \beta_{22})}$$

$$E(Y_{i2}^* | Y_{i2} = 0) = \frac{\exp(X_i \beta_{22})}{\exp(X_i \beta_{21}) + \exp(X_i \beta_{22})}.$$

With the conditional expectations inserted into the log-likelihood function, the maximization, or “M,” step of the EM algorithm maximizes the log-likelihood function. New parameter values are obtained, and the “E” step and the “M” step are repeated. This process continues until the likelihood function is maximized. Once the maximum has been found, standard errors are calculated using one iteration of the algorithm of Berndt et al. (1974).

One can see that the EM algorithm is embedded in the first-order conditions given in equations (7) and (13) above. The values of α are, for each individual, the conditional expectation of each stem, given the branch. Thus, the α -values are the conditional expectations given in equation (15). The EM algorithm uses initial parameter estimates to calculate the α -values in equations (7) or (13) and then maximizes the likelihood function. New parameter values are generated, new α -values, and the process is repeated. The program to estimate the LCMNL model, written in the IML matrix language in SAS, is available in a supplementary appendix online.

In order to compare the binomial logit estimation results with the LCMNL estimation results, several information criteria are used. In a recent article, Sarstedt and Schwaiger (2008) conducted a Monte Carlo experiment of the ability of several information criteria to detect latent class models. Each of these measures is based on the value of the likelihood function. Following Sarstedt and Schwaiger, we calculate the Akaike information criterion (AIC), the consistent Akaike information criterion (CAIC), the Bayesian Information Criterion (BIC), and the sample-size adjusted AIC (ABIC) for both the binomial logit and LCMNL models.

Estimation Results

We apply our method to contingent valuation data on the preservation of wetlands in the Saginaw Bay watershed in Michigan. The data have been previously examined by Whitehead et al. (2009). In this article, the contingent valuation question involves a management plan for purchasing wetlands for preservation in the Saginaw Bay watershed. The purchase plan was offered at six different price levels from \$25 to \$200, and respondents were asked about their WTP. The data set contains 281 responses. Fifty-five percent of respondents indicated a willingness to pay the specified amount (vs. 45% who did not). A description of the variables used, along with summary statistics, is given in the first and second columns of table 1. *Lnbid* is the natural log of the eight different price levels, and *acres* is the number of acres of wetlands that are to be purchased to preserve. *Travcost* is the cost to the respondent of travel to the Saginaw Bay area. *Subcost* is the cost to travel to a substitute site for recreation, either Traverse City on Lake Michigan or Alpena on Lake Huron, whichever is closer for the respondent. *Income* is the income of the respondent, while *member* is a dummy variable equal to one if the respondent is a member of an environmental group, and *likely* is a dummy

Table 1. Summary Statistics, Binomial Logit, and LCMNL Estimation Results

Variable Names and Definitions	Mean (SD)	Logit Estimates	LCMNL Estimates	LCMNL Estimates
		No Response	No Response	YES Response
<i>Intercept</i>	—	-2.292 (2.18)	-1.490 (0.20)	-0.332 (0.04)
<i>Lnbid</i> (natural log of bid amount, from \$5 to \$150)	4.372 (0.69)	1.058 (4.94)	1.631 (1.81)	0.659 (0.64)
<i>Acres/1000</i> (no. of acres purchased for preservation)	2.546 (1.42)	0.040 (0.41)	1.931 (1.24)	2.102 (1.41)
<i>Travcost/10</i> (travel cost to the Saginaw bay area)	5.031 (2.93)	0.085 (1.18)	0.656 (0.82)	0.592 (0.74)
<i>Subcost/100</i> (travel cost to a substitute recreational site)	1.467 (0.47)	-0.717 (1.35)	-8.608 (0.96)	-8.451 (0.95)
<i>Income/10</i> (household income)	5.252 (2.84)	-0.138 (2.18)	0.604 (0.95)	0.828 (1.30)
<i>Member</i> (1 if a member of an environmental or conservation organization)	0.406 (0.49)	-0.670 (2.33)	0.772 (0.55)	1.712 (1.13)
<i>Likely</i> (1 if thought the preservation plan was likely to succeed)	0.477 (0.50)	-1.217 (4.32)	-0.872 (0.87)	0.400 (0.34)
-logL	—	193.475	152.156	—

Note: LCMNL = latent choice multinomial logit model; SD = standard deviation. Numbers in parentheses are absolute values of *t*-ratios.

variable equal to one if the respondent thought enough people would donate to preserve the Saginaw wetlands.

We begin with the estimation of a simple logit model of the WTP decision. In this analysis, the dependent variable is coded 1=No, so a positive sign on a coefficient indicates that increases in the variable increase the probability of a No response. The estimation results are given in the third column of table 1. Several of the coefficients are statistically significant. In addition to the intercept, the coefficients of *Lnbid*, *income*, *member*, and *likely* are statistically significant at the $\alpha = .10$ level or better and have the

expected sign. These results are estimated for the sake of comparison, so our discussion is brief as we turn our attention to the LCMNL estimation results.

The LCMNL model estimated has two branches, with one branch having two stems. In our first effort to estimate the model, we allow for two types of *No* response. Our algorithm fails to converge to a unique maximum.³ We next estimate the model allowing for the possibility of two types of *Yes* responders. The LCMNL model does indeed find two types of *Yes* responders. We designate these two types of responders *yes* and **YES** (shortly we hope to associate *yes* with a weak *Yes* response and **YES** with a strong *Yes* response).

The LCMNL estimation results are given in the fourth and fifth columns of table 1. The fourth column of table 1 presents the estimated coefficients for the “No” responders. Only the coefficient of *Inbid* is significantly different from zero at the $\alpha = .10$ level or better. The results from estimating the parameters associated with the **YES** responders are given in the fifth column of table 1. In the **YES** equation, none of the estimated coefficients is significantly different from zero at the $\alpha = .10$ level or better. Despite the lack of statistical significance of the individual coefficients, all of the information criteria discussed previously favor the LCMNL model over the binomial logit model. We next examine the marginal effects from the binomial logit and LCMNL models.

The marginal effects associated with the models estimated are given in table 2. For the simple logit model, the marginal effects exhibit the same pattern of statistical significance as the estimated coefficients. These marginal effects are given in the second column of table 2. Marginal effects associated with *Inbid*, *income*, *member*, and *likely* are statistically significant at the $\alpha = .10$ level or better and have the expected sign. The marginal effects associated with the LCMNL model are given in the third, fourth, and fifth columns of

Table 2. Marginal Effects (MEs) for the Saginaw Data

Variable Names and Definitions	MEs Binary Logit <i>No</i> Response	MEs LCMNL <i>No</i> Response	MEs LCMNL YES Response	MEs LCMNL <i>yes</i> Response
<i>Intercept</i>	-0.565 (2.17)	-0.307 (0.67)	0.205 (0.32)	0.102 (0.12)
<i>Inbid</i> (natural log of bid amount, from \$5 to \$150)	0.261 (4.93)	0.280 (2.71)	-0.158 (0.73)	-0.123 (0.42)
<i>Acres/1000</i> (no. of acres purchased for preservation)	0.001 (0.41)	0.084 (0.35)	0.111 (0.75)	-0.195 (0.50)
<i>Travcost/10</i> (travel cost to the Saginaw bay area)	0.021 (1.18)	0.051 (0.70)	0.010 (0.18)	-0.062 (0.50)
<i>Subcost/100</i> (travel cost to a substitute recreational site)	-0.177 (1.35)	-0.547 (0.54)	-0.288 (0.43)	0.835 (0.50)
<i>Income/10</i> (household income)	-0.034 (2.18)	-0.006 (0.06)	0.073 (1.89)	-0.067 (0.52)
<i>Member</i> (1 if a member of an environmental or conservation organization)	-0.165 (2.36)	-0.130 (0.63)	0.241 (2.19)	-0.111 (0.48)
<i>Likely</i> (1 if thought the preservation plan was likely to succeed)	-0.292 (4.57)	-0.291 (3.83)	0.254 (1.35)	0.037 (0.22)
Σ <i>predicted probabilities</i>	—	0.548	0.342	0.110

Note: Numbers in parentheses are absolute values of *t*-ratios.

table 2. The third column presents the marginal effects associated with the *No* responders. Only the marginal effects associated with *Inbid* and *likely* are statistically significant at the $\alpha = .10$ level or better. The marginal effects associated with the **YES** responders are given in the fourth column of table 2. Marginal effects associated with *income* and *member* are statistically significant at the $\alpha = .10$ level or better and have the expected signs. The marginal effects associated with the *yes* responders are given in the fifth column of table 2. None of the individual marginal effects is statistically significant. Even so, the information criteria discussed previously indicate that the LCMNL model is preferred, so we next investigate the nature of the *Yes* response heterogeneity using a follow-up certainty question.

Is Hypothetical Bias Present?

In the Saginaw Wetland contingent valuation surveys, two groups of *Yes* respondents are found using the LCMNL model, and for this data set a follow-up certainty question is available. In this section, we investigate whether responses on the certainty question are correlated with either type of *Yes* response discovered by the LCMNL procedure. We obtain predicted probabilities from the LCMNL model for each response: *No*, *yes*, and **YES**. Using the follow-up certainty question available, we construct a dummy variable, *Certain*, based on the answer to the following question, "On a scale of 1 to 10, how sure are you that you would make the onetime donation?" Individuals who answered *Yes* to the contingent valuation question and scored 7 or higher on a ten-point scale on the follow-up question are assigned *Certain*=1, otherwise *Certain*=0. An auxiliary regression model is estimated in which *Certain* is the dependent variable, and the lone independent variable is the conditional probability of giving a **YES** response, given *Yes* obtained from the LCMNL estimation results. This independent variable is denoted P_{YES} . We estimate the following regression model

$$(16) \quad \textit{Certain} = 0.523 + 0.265P_{YES}(4.81)(1.98)$$

The coefficient of P_{YES} is positive and statistically significant at the $\alpha = .05$ level (t -ratio= 1.98). This result indicates that those in the **YES** group are more certain of their WTP (and that those in the *yes* group are less certain). This evidence is consistent with the *yes* group being associated with hypothetical bias, as identified by the *ex post* approach.

In order to shed more light on differences between the **YES** and *yes* regimes, we examine

Table 3. Weighted Means (SDs) of Independent Variables for the Saginaw Data

Variable	YES	<i>yes</i>
<i>Lnbid</i>	4.177 (0.41)	3.950 (0.21)
<i>Acres/1000</i>	2.787 (0.83)	1.342 (0.15)
<i>Travcost/10</i>	5.065 (1.72)	5.825 (1.00)
<i>Subcost/100</i>	1.534 (0.27)	1.664 (0.13)
<i>Income/10</i>	6.191 (1.56)	5.116 (0.91)
<i>Member</i>	0.559 (0.29)	0.290 (0.15)
<i>Likely</i>	0.630 (0.28)	0.629 (0.16)

the weighted means of the independent variables for those individuals predicted to belong to those groups. The means for these individuals are given in table 3. In comparison to the yes responders, those individuals associated with the **YES** response are offered more land for preservation, have higher incomes, and are more likely to belong to an environmental organization. These differences seem to be consistent with a group more likely to actually pay the bid amount and not subject to hypothetical bias (the **YES** group).

An examination of the WTP estimates from the models estimated on the Saginaw data set is useful. From the binomial logit model, the estimated mean WTP is \$63. Using the LCMNL results, if the sample is assumed to be composed entirely of **YES** responders, the estimated mean WTP is \$43. If the same is assumed for the yes responders, the estimate is \$3. A weighted average of the two yields an estimate of \$33, which is considerably lower than that given by the binomial logit model. Even if the WTP estimates for the two Yes response groups are summed, the result, \$46, is still much lower than the \$63 estimate obtained from the simple logit model. Accounting for response heterogeneity in the Yes response leads to a lower mean WTP estimate in the Saginaw case.

Conclusions

This article uses the expectations maximization algorithm of Dempster, Laird, and Rubin (1977) to estimate a multinomial logit model with missing information, which we call the latent choice multinomial logit model. The LCMNL model allows for within-choice heterogeneity and in that respect is similar to latent class and mixed logit. The LCMNL model is applied to contingent valuation data based on responses to questions about the preservation of wetlands in the Saginaw Bay area. In the application of the LCMNL model to the WTP data, we find two types of

Yes responders. One type is much wealthier and more environmentally conscious than the other. Using a follow-up certainty question, we show that one of the Yes groups identified by the LCMNL model has much more conviction about their WTP than the other Yes category. This evidence supports the existence of hypothetical bias in one of the Yes response groups.

Although the LCMNL model cannot be used to test for hypothetical bias, it is a valuable tool for investigating WTP. In the data set we examine, the Yes response heterogeneity is correlated with response certainty, which has been shown to be related to hypothetical bias. We have statistical evidence supporting the existence of within-response heterogeneity. Locating heterogeneity, whatever the underlying source, is important in and of itself because adjustments can be made to WTP estimates obtained from dichotomous choice models (logit or probit). As a method to detect response heterogeneity, the LCMNL model can be used along with the *ex ante* and *ex post* approaches for hypothetical bias reduction. Response heterogeneity can exist even among those responders "certain" of their WTP and in the complete absence of hypothetical bias. However, the greatest value of the LCMNL approach may lie in applications to older contingent valuation data sets for which *ex ante* and/or *ex post* adjustments are not available.

For these data sets, the LCMNL model offers the opportunity to locate response heterogeneity (perhaps due to hypothetical bias, perhaps not) and adjust WTP estimates.

Notes

1 Like the latent class logit model (see Greene and Hensher 2003; Oh, Choi, and Kim 2003; Shen 2009; van Rosmalen, van Herk, and Groenen 2010) and the mixed logit model, the LCMNL allows for heterogeneity. The latent class logit model is composed of MNLs that differ across individuals according to class membership, with class membership being unknown. The LCMNL has no unknown classes but does permit within-choice heterogeneity. That is, individuals making the same choices can be characterized by different sets of regression parameters.

2 In the absence of any restrictions on the model parameters, there is an identification problem. This occurs because the value of the likelihood function in equation (11) is unaffected by any reordering of probabilities associated with stems on the second branch. That is, the value of the likelihood function is unchanged if the second branch is characterized as $(P_{i21} + P_{i22})$ or $(P_{i22} + P_{i21})$. This phenomenon is well known in the literature on latent class models and is called “label switching” (see Aitken and Rubin 1985). In order to prove consistency, the model parameters must be in the same order each time, which is not guaranteed with the LCMNL model. However, this problem can be easily solved after estimation by rearranging the parameter vector. We do so in our model by imposing the constraint

$$\sum_i \alpha_{j1} < \sum_i \alpha_{j2} < \dots < \sum_i \alpha_{jJ}$$

which has the effect of listing parameters associated with the largest stem on a branch first, those associated with the next largest second, and so on.

3 Our model is related to a finite mixture model. A common problem with the estimation of mixture models is that multiple local maxima of the likelihood function exist (see Titterton, Smith, and Makov 1985, p. 95). The consequence is that different starting values for the parameters lead to different parameter estimates. This problem occurs when we examine the *No* responses.

References

- Aitken, M., and D. B. Rubin. 1985. Estimation and Hypothesis Testing in Finite Mixture Models. *Journal of the Royal Statistical Society, Series B* 47(1): 67–75.
- Berndt, E. R., B. H. Hall, R. E. Hall, and J. A. Hausman. 1974. Estimation and Inference in Nonlinear Structural Models. *Annals of Economic and Social Measurement* 3(4): 653–665.
- Blumenschein, K., G. C. Blomquist, M. Johannesson, N. Horn, and P. Freeman. 2008. Eliciting Willingness to Pay without Bias: Evidence from a Field Experiment. *Economic Journal* 118(525): 114–137.
- Carson, R., and T. Groves. 2007. Incentive and Informational Properties of Preference Questions. *Environmental & Resource Economics* 37(1): 181–210.
- Caudill, S. B. 2006. A Logit Model with Missing Information Illustrated by Testing for Hidden Unemployment in Transition Economies. *Oxford Bulletin of Economics and Statistics* 68(5): 665–677.
- Caudill, S. B., M. Ayuso, and M. Guillen. 2005. Fraud Detection Using a Multinomial Logit Model with Missing Information. *Journal of Risk and Insurance* 72(4): 539–550.
- Caudill, S. B., and P. A. Groothuis. 2005. Modeling Hidden Alternatives in Random Utility Models: An Application to Don't Know Responses in Contingent Valuation. *Land Economics* 81(3): 445–454.
- Champ, P. A., R. C. Bishop, T. C. Brown, and D. W. McCollum. 1997. Using Donation Mechanisms to Value Nonuse Benefits from Public Goods. *Journal of Environmental Economics and Management* 33(2): 151–162.
- Cramer, J. S., and G. Ridder. 1991. Pooling States in the Multinomial Logit Model. *Journal of Econometrics* 47(2–3): 267–272.
- Cummings, R. G., S. Elliot, G. W. Harrison, and J. Murphy. 1997. Are Hypothetical Referenda Incentive Compatible? *Journal of Political Economy* 105(3): 609–621.
- Cummings, R. G., G. W. Harrison, and E. E. Rutström. 1995. Homegrown Values and Hypothetical Surveys: Is the Dichotomous Choice Approach Incentive- Compatible? *American Economic Review* 85(1): 260–266.
- Cummings, R. G., and L. O. Taylor. 1999. Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method. *American Economic Review* 89(3): 649–665.

Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. Maximum Likelihood Estimation from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society, Series B* 39(1): 1–38.

Greene, W. H., and D. A. Hensher. 2003. A Latent Class Model for Discrete Choice Analysis: Contrasts with Mixed Logit. *Transportation Research: Part B: Methodological* 37(8): 681–698.

Harrison, G.W. 2006. Experimental Evidence on Alternative Environmental Valuation Methods. *Environmental and Resource Economics* 34(1): 125–162.

Hausman, J. A., J. Abrevaya, and F. M. Scott-Morton. 1998. Misclassification of a Dependent Variable in a Discrete response Setting. *Journal of Econometrics* 87(2): 239–269.

List, J. A., and C. A. Gallet. 2001. What Experimental Protocol Influence Disparities between Actual and Hypothetical Stated Values: Evidence from a Meta-Analysis? *Environmental and Resource Economics* 20: 241–254.

Loomis, J. 2011. What's to Know About Hypothetical Bias in Stated Preference Valuation Studies? *Journal of Economic Surveys* 25: 363–370.

Loomis, J. B., T. Brown, B. Lucero, and G. Peterson. 1996. Improving Validity Experiments of Contingent Valuation Methods: Results of Efforts to Reduce the Disparity of Hypothetical and Actual Willingness to Pay. *Land Economics* 72(4): 450–461.

Loomis, J.B., A. Gonzalez-Caban, and R. Gregory. 1994. Do Reminders of Substitutes and Budget Constraints Influence Contingent Valuation Estimates? *Land Economics* 70(4): 499–506.

Magder, L. S., and J. P. Hughes. 1997. Logistic Regression When the Outcome is Measured with Uncertainty. *American Journal of Epidemiology* 146(2): 195–203.

Oh, M., J. Choi, and D. Kim. 2003. Bayesian Inference and Model Selection in Latent Class Logit Models with Parameter Constraints. *Journal of Applied Statistics* 30: 191–204.

Rao, C.R. 1955. Estimation and Tests of Significance in Factor Analysis. *Psychometrika* 20(2): 93–111.

Sarstedt, M., and M. Schwaiger. 2008. Model Selection in Mixture Regression Analysis—A Monte Carlo Simulation Study. In *Data Analysis, Machine Learning and Applications: Proceedings of the 31st Annual Conference of the Gesellschaft für Klassifikation*, ed. C. Preisach, H. Burkhardt, L. Schmidt-Thieme, R. Decker, 61–68. Berlin and Heidelberg: Springer.

Shen, J. 2009. Latent Class Model or Mixed Logit Model? A Comparison by Transport Mode Choice Data. *Applied Economics* 41(22): 2915–2924.

Titterton, D., A. Smith, and V. Makov. 1985. *Statistical Analysis of Finite Mixture Distributions*. New York: John Wiley & Sons.

van Rosmalen, J., H. van Herk, and P. Groenen. 2010. Identifying Response Styles: A Latent-Class Bilinear Multinomial Logit Model. *Journal of Marketing Research* 47(1): 157–172.

Whitehead, J. C., and G. C. Blomquist. 2006. Contingent Valuation and Benefit-Cost Analysis. In *Handbook on Contingent Valuation*, ed. A. Alberini and J. R. Kahn, 66–91. Cheltenham, UK: Edward Elgar Publishing, pp.

Whitehead, J.C., P.A. Groothuis, R. Southwick, and P. Foster-Turley. 2009. Measuring the Economic Benefits of Saginaw Bay Coastal Marsh with Revealed and Stated Preference Methods. *Journal of Great Lakes Research* 35(3): 430–437.