Using Attitudes to Characterize Heterogeneous Preferences

Dr. Rebecca Moore
Warnell School of Forestry and Natural Resources
University of Georgia

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Orlando, FL, July 27-29, 2008.

Conventional stated preference studies use standard logit models to estimate preference parameters. While easy to implement, there are two disadvantages to this approach. First, the standard logit model assumes homogenous preferences (or at most *explained* and *deterministic* heterogeneity). Second, typically only a small portion of the available information is used in the econometric analysis. The purpose of this paper is to explore the potential for using the attitudinal data typically available to the researcher to *explain* preference heterogeneity[1].

Using data from a contingent valuation (CV) study of Green Bay water clarity improvements, I estimate three models of heterogeneous preferences. Two of these models use latent class (LC) methods and the other takes a random parameter logit (RPL) approach. All three models consider the attitudinal data when estimating preference parameters. The first LC approach follows that of Morey, Thacher and Breffle (2006), where the attitude and choice data are both driven by underlying preferences, which are identified by exogenous class membership. The second LC approach is based on that of Boxall and Adamowicz (2002), in which the attitudinal data condition preference class membership, which then drives the choice decision. The RPL model I estimate is structurally similar to the mixture model of Morey and Rossmann (2003), though that paper does not consider attitude data. In this RPL model, preference parameters are considered random variables with means conditional on attitudinal data. These three papers are discussed in more detail in Section II. The models used in these papers were developed for repeated stated choice data rather than the dichotomous choice CV data analyzed in this paper. Sections III and IV adapt the models to the current context.

Using data from a mail survey, I estimate the preference parameters and expected WTP associated with improving water clarity in Green Bay, Wisconsin allowing for heterogeneous

[1] Throughout the paper I will refer to "attitude questions" and "attitudinal data". I use these terms generically to refer to many of the standard auxiliary questions, including questions about attitudes, beliefs, interests, prior experience, etc.

preferences. All three models suggest that preferences for the water quality improvement project are heterogeneous, but the three models are structurally different and make different assumptions regarding the correct use of attitude data. As a result, the estimated taste parameters differ between models (though not always significantly). However, there is little difference in the expected WTP implied by each of the three models. From a policy perspective, this is evidence that we do not need to be overly concerned with how the attitude data are included, only that they are included at all.

II. Heterogeneous Preferences in Non-Market Valuation

There are a variety of ways to allow preferences to vary across the population. The most common way is to directly include individual characteristics in the WTP function, which can significantly increase its predictive power (Haab and McConnell 2003; Morey et al 2002; Poe and Bishop 1999). A similar method is to use observable demographic data to divide the sample into different groups, and estimate unique preference parameters for each group. This approach is easy and common, but very restrictive in assuming heterogeneity is deterministic; everyone with the same characteristics is required to have the same preferences (Morey and Rossmann 2003).

Latent class (LC) models are another approach to modeling heterogeneous preference. LC models assume that the population consists of a finite number of classes or groups of people. Preferences are allowed to vary between groups, but all individuals within a class are assumed to have identical preferences. The key feature of these models is that class membership is unobserved by the analyst. LC methods are common in other fields (for example, Swait 1994; Tittering et al. 1985; McLachlan and Peel 2000), and are increasingly being used by environmental economists (Swait 1994; Boxall and Adamowicz 2002; Morey, Thacher, and

Breffle 2006; Breffle, Morey and Thacher 2005; Provencher and Bishop 2004; Provencher, Barenklau, and Bishop 2002; Scarpa and Thiene 2005).

Random Parameters Logit (RPL) models are a natural alternative to LC models. RPL models introduce preference heterogeneity by "individualizing" preferences; each individual is understood to have a possibly unique set of preference parameters. The analyst cannot observe individual preferences, but assumes some distribution of preferences across the population, treating each individual's preferences as a random draw from this distribution. Train (1998, 2003) provides an explanation of the method and there are several examples of RPL in the existing applied economics literature (*e.g.,* Train 1998; Chen and Coslett 1998; Morey and Rossmann 2003). One traditional drawback of the RPL approach is that the source of the heterogeneity remains unknown. However, as Morey and Rossmann (2003) demonstrate, it is possible to condition the distribution of the preference parameters on individual characteristics. Based on their application, the authors find that the mixture model dominates the traditional RPL model, but more importantly, it identifies a particular group of individuals who are made worse off by increased preservation of the statues. Morey and Rossmann do not consider attitudinal data in their mixture model. The RPL model I estimate in this paper uses the Morey and Rossmann mixture model, but conditions the parameter distributions on attitudes as well as demographics.

2.1 *Preference heterogeneity and attitudes*

The results of a study by Bishop and Heberlein (1979) provide some of the earliest evidence that attitudes might influence contingent behavior. That study used a split sample design, in which one group received actual cash offers to sell their goose hunting license, while the other group received hypothetical, but similar offers. They found that the commitment to hunting, measured with an attitude scale, was the strongest influence in the hypothetical decision process. Though slow to catch on, including attitudes as explanatory variables in stated preference studies has become more frequent, but heterogeneity in most of these studies is constrained to be deterministic (Barro et al 1996; Luzer and Cosse 1998).

McFaddan (1986) suggested that attitudes and beliefs could be used to understand and estimate individuals' preferences for different market goods. This approach has been adopted outside of economics (e.g., Swait 1994; McCutcheon 1996; Yamaguchi 2000), but despite the potential benefit this perspective could bring to the non-market valuation literature, there are very few environmental economic studies in which attitudinal data are used to explain unobserved preference heterogeneity. From these limited studies, there are two distinct routes for incorporating attitudinal data. Boxall and Adamowicz (2002) use LC models to understand the recreation site decisions of visitors to five wilderness parks in Canada. The survey used in that study included both attitude questions and a stated choice experiment in which each respondent faced eight sets of six possible combinations of park attributes. In their model, the latent preference classes are identified by unique taste parameters related to the attributes of the choice decision, as in most LC models. However, Boxall and Adamowicz condition the probability that an individual belongs to a particular class on the individual's demographic characteristics (as is common) *and* his responses to the attitude questions. In summary, attitudes define classes, and classes define choice preferences.

Breffle, Morey, and Thacher (2005) take an alternative approach in their application of a LC model to a study of angler preferences in Green Bay, WI. The survey for this study included 15 attitude questions and eight stated preference paired-comparison questions regarding preferred fishing conditions. In this paper, the attitudinal and choice data are assumed to come from the same underlying exogenous preferences. Class membership is not dependent on attitudes, rather, attitudes are dependent on class. Morey, Thacher, and Breffle (2006) estimates a LC model in which the choice data are ignored and only attitudes parameters are estimated. Breffle, Morey and Thacher (2005) extend the model to include joint estimation of the choice and attitude parameters.

III. The Latent Class (LC) Models

In this paper I compare two attitude-based LC models. The first approach, which I will call the MTB model, is based on that of Morey, Thacher, and Breffle (2006) and Breffle, Morey, and Thacher (2005). It assumes preference class membership is exogenous and jointly determines both the respondent's stated choice and attitude responses. The second approach, the BA model, is based on that of Boxall and Adamowicz (2002) and assumes that the respondent's attitude responses are a signal of their underlying, latent preferences and so can be used to determine class membership. Respondent's class then determines the individual's choice specific taste parameters. Both the MTB and BA models were originally designed for studies in which individual's make repeated choices. In this section I develop the models in the context of a CV decision in which each respondent faces a single dichotomous choice regarding their willingness to pay for a specified good. For consistency, I follow the notation of MTB as closely as possible throughout all three models.

3.1 *The MTB model*

Consider an individual facing a dichotomous choice CV question with an offer amount of $\$T$, whose willingness to pay (WTP) for the good is defined by

$$WTP_i = \beta_i w_i + \varepsilon_i \tag{2.1}$$

where w_i is a vector of attributes of the good being valued, β_i is a conformable vector of individual i's taste parameters and ε_i is a value known to the respondent but unobserved by the analyst. Note that in equation (2.1), demographic and other attitude variables do not impact WTP directly. Instead, the model will allow these factors to affect WTP indirectly through the preference parameters, β.

Assume the population consists of C different classes, or groups of individuals, and each class is identified by a unique set of preference parameters β^c. If ε_i is treated as an iid Gumbel-distributed random variable, the probability of observing individual i's response to the CV question, given i's membership in class c is given by

$$\Pr(r_i \mid c) = \left[\frac{1}{1+e^{-\beta^{c(i)}w_i}}\right]^{r_i} \left[\frac{e^{-\beta^{c(i)}w_i}}{1+e^{-\beta^{c(i)}w_i}}\right]^{1-r_i} \tag{2.2}$$

where r_i equals 1 if the individual answered "YES" and 0 if the individual answered "NO", $c(i)$ identifies the class to which the individual belongs, and $\beta^{c(i)}$ is the parameter vector associated with this class.

Suppose that in addition to the choice data, r, and the attribute values, w, the analyst also observes demographic and attitudinal data for each individual. Let z_i be a vector of socio-demographic characteristics of individual i. Traditionally, the attitudinal data come from a series of questions with Likert-scale answers. If there are Q attitude questions, with S possible answers per question, we can denote the respondent's answers as the vector x_i, which has (Q x S) elements, so that x_{iqs} equals 1 if respondent i chose answer s in response to question q, and equals 0 otherwise. The MTB model assumes that the attitude and choice data are jointly determined by the individual's underlying preferences, so that the attitude responses of members of the same preference class should have a higher correlation than those of individuals from different preference classes. However, once you account for class membership, attitude responses are independent, so then the probability of observing individual i's attitude response pattern, given their membership in class c is given by

$$\Pr(x_i \mid c) = \prod_{q=1}^{Q} \prod_{s=1}^{S} \left(\pi_{qs|c}\right)^{x_{iqs}} \tag{2.3}$$

where $\pi_{qs|c}$ is the probability that an individual in class c chooses answer s to attitude question q. Given Equations (2.2) and (2.3), the joint probability of the observed responses on the attitudinal scales and the CV question for individual i is

$$\begin{aligned} L_i = \Pr(x_i, r_i : z_i) &= \sum_{c=1}^{C} \Pr(c : z_i) \Pr(x_i \mid c) \Pr(r_i \mid c) \\ &= \sum_{c=1}^{C} \Pr(c : z_i) \prod_{q=1}^{Q} \prod_{s=1}^{S} \left(\pi_{qs|c}\right)^{x_{iqs}} \left[\frac{1}{1+e^{-\beta^{c(i)} w_i}}\right]^{r_i} \left[\frac{e^{-\beta^{c(i)} w_i}}{1+e^{-\beta^{c(i)} w_i}}\right]^{1-r_i} \end{aligned} \tag{2.4}$$

where $\Pr(c : z_i)$ is the probability that individual i is a member of class c, given his demographic characteristics, z_i, $\Pr(x_i \mid c)$ is given in (2.3), and $P(r_i \mid c)$ is given in (2.2). The analyst's goal is to find the values of $\pi_{qs|c}$, β^c, and $\Pr(c : z_i)$ that maximize the log likelihood function

$$\ln L = \sum_{i=1}^{N} \ln(L_i) \tag{2.5}$$

Without observing the probability of being in a class, we cannot estimate the parameters directly. Instead I rely on the Sequential Estimation approach detailed in Breffle, Morey, and Thacher (2005). This approach is a variant of the expectation-maximization (E-M) algorithm (Dempster et al., 1977) designed for maximum likelihood estimation under incomplete information. In the E-M algorithm, the analyst first calculates the expected value of the unobserved information, assumes these are the true values, and then maximizes the likelihood function based on these expected values. Using the resulting parameter estimates, the expected values of the unobserved values are recalculated and the process continues until the likelihood function does not change. Breffle, Morey, and Thacher detail two alternative methods for estimating the maximum likelihood parameters, which they label Sequential estimation and FIML. Based on the application in that paper, the authors report similar qualitative results of the two solution methods. However, the sequential estimates are much easier to obtain. As the purpose of my paper is to compare the qualitative implications of model selection, I estimate the model with the sequential method and briefly outline the process below.

Sequential estimation occurs in two stages. In stage one, only the attitudinal data are considered. The likelihood function for the attitudinal data is

$$L_a(\Pr(c:z), \pi_{qs|c}) = \prod_{i=1}^{N}\left[\sum_{c=1}^{C}\Pr(c:z)\prod_{q=1}^{Q}\prod_{s=1}^{s}\left(\pi_{qs|c}\right)^{x_{iqs}}\right] \tag{2.6}$$

The values of $\pi_{qs|c}$ that maximize equation (2.6) are

$$\pi_{qs|c} = \frac{\sum_{i=1}^{N}\Pr(c:z_i \mid x_i)x_{iqs}}{\sum_{i=1}^{N}\Pr(c:z_i \mid x_i)}. \tag{2.7}$$

where $\Pr(c:z_i \mid x_i)$ is the conditional probability of being in class c given one's responses to the attitude questions, x. This is an estimate of the number of times individuals in class c answered s to question q. If the demographic variables take a finite number of discrete values, as they do in the application discussed below, the class membership probabilities that maximize equation (2.6) are

$$\Pr(c:z_i) = \frac{1}{N_{z_i}}\sum_{i=1}^{N}\Pr(c:z_i \mid x_i) \tag{2.8}$$

where N_{z_i} is the number of respondents with the same demographic characteristics as respondent i.

Again, there is incomplete information in this problem, specifically, $\Pr(c : z_i \mid x_i)$ is an element of both (2.7) and (2.8) but is unobserved by the analyst, and the E-M algorithm must be used as follows. First, make an initial guess of the (N x 1) matrix $\Pr(c : z_i \mid x_i)$. Using these values, and equations (2.7) and (2.8), calculate $\pi_{qs|c}$ and $\Pr(c : z_i)$. Use equation (2.6) to calculate the likelihood value given these values. Second, using Bayes theorem, updated estimates of $\Pr(c : z_i \mid x_i)$ can be calculated by

$$\Pr(c : z_i \mid x_i) = \frac{\Pr(c : z_i) \prod_{q=1}^{Q} \prod_{s=1}^{S} \left(\pi_{qs|c}\right)^{x_{1qs}}}{\sum_{c=1}^{C} \Pr(c : z_i) \prod_{q=1}^{Q} \prod_{s=1}^{S} \left(\pi_{qs|c}\right)^{x_{1qs}}} \tag{2.9}$$

The process is repeated until the change in the likelihood value is less than some specified threshold. This stage of the Sequential Estimation provides an estimate of $\Pr(c \mid x_i)$ conditional only on the attitudinal data. The second stage of the estimation takes these values as given and estimates β^c, the taste parameters associated with the CV decision. The likelihood function of this stage is

$$L_{CV}(\beta^{c(i)}) = \prod_{i=1}^{N} \left[\sum_{c=1}^{C} \Pr(c \mid x_i) \left[\frac{1}{1+e^{-\beta^{c(i)} w_i}} \right]^{r_i} \left[\frac{e^{-\beta^{c(i)} w_i}}{1+e^{-\beta^{c(i)} w_i}} \right]^{1-r_i} \right] \tag{2.10}$$

Given Pr(c|x) from stage 1, estimates of β can be found directly.

3.2 *The BA model*

The fundamental difference between the MTB model and the BA model lies in how the attitudinal data are used. In the BA model, using the current notation, class membership is dependent upon both the attitude and sociodemographic data, so that

$$\Pr(c : z_i, x_i) = \frac{e^{\gamma_z^{c(i)} z_i + \gamma_x^{c(i)} x_i}}{\sum_{d=1}^{C} e^{\gamma_z^{d(i)} z_i + \gamma_x^{d(i)} x_i}} \tag{2.11}$$

where z_i is a vector of sociodemographic characteristics for respondent i, x_i is a matrix of attitudinal responses for respondent i, and $\gamma_z^{c(i)}$ and $\gamma_x^{c(i)}$ are conformable vectors of parameters to be estimated. Based on this definition for $\Pr(c : z_i, x_i)$, the probability of the observed responses on the CV question for individual i is then

$$\begin{aligned} L_i = \Pr(r_i : z_i, x_i) &= \sum_{c=1}^{C} \Pr(c : z_i, x_i) \Pr(r_i \mid c) \\ &= \left[\sum_{c=1}^{C} \frac{e^{\gamma_z^{c(i)} z_i + \gamma_x^{c(i)} x_i}}{\sum_{d=1}^{C} e^{\gamma_z^{d(i)} z_i + \gamma_x^{d(i)} x_i}} \left[\frac{1}{1 + e^{-\beta^{c(i)} w_i}} \right]^{r_i} \left[\frac{e^{-\beta^{c(i)} w_i}}{1 + e^{-\beta^{c(i)} w_i}} \right]^{1-r_i} \right] \end{aligned} \tag{2.12}$$

The analyst then finds the values of $\gamma_z^{c(i)}$, $\gamma_x^{c(i)}$, and β^c to maximize equation (2.12). If the attitude and demographic variables take on a finite number of discrete values, as they do in the application discussed below, the estimation does not require use of the E-M algorithm.

3.3 *MTB vs. BA*

A careful inspection of equations (2.4) and (2.12) will show that the BA and MTB approaches will generate different preferences classes and therefore different welfare results. Conceptually, the model developed in BA is one in which an individual's preferences for the good being valued depend on both demographic characteristics and attitudes about the good expressed in the attitude data. In other words, "deep" latent preferences are the source of attitudes, which are themselves determinants, along with sociodemographic variables, of the activity-specific preferences governing the CV decision. In comparison, the MTB model applies to the case in which “deep” preferences are the source of both activity-specific attitudes and activity-specific preferences. It is not immediately evident which of the two approaches is correct. It is often assumed that preferences are invariant (Tversky and Kahneman 1986), which leans towards the MTB model. On the other hand, the BA model is supported by a study by Pouta (2004) which suggests that in some cases, the process of answering attitude questions can impact preference parameters. In that stated preference study, half of the sample answered attitude questions in addition to the choice questions, the other half only answered the choice questions. The presence of the attitude questions increased the probability of choosing the “environmentally friendly” option and reduced the respondents’ sensitivity to the bid amount. This could mean that the process of answering the attitude questions directly impacts the preference parameters driving the individual CV decision.

In a CV study, we are interested in the respondent’s behavior regarding the valuation decision. Preferences, as defined by most economists, are what drive the individual’s choice between two bundles. Preferences are a construct; they can not be seen or observed but neoclassical economic theory is based on their existence. In a similar manner, social psychologists have long been interested in how attitudes are formed and the link between attitudes and behaviors. Attitudes are often broken into three categories- cognitive attitudes, affective attitudes, and behavioral intentions. Wilson (2000) relates these categories to the difference between “feeling, thinking, and acting”. Cognitive attitudes are commonly conceptualized as information or beliefs, and reflect thoughts or knowledge about an environmental good. Affective attitudes reflect feelings or emotions related to the good. These two components of attitudes are thought to influence the CV decision, which is a behavioral intention. Returning to the economic notion of preferences, the difference between the BA and

MTB model lies in how attitudes are related to preferences. Zajonc and Markus (1982) argue that affective and cognitive attitudes act independently and the relative importance of affective or cognitive attitudes on the formation of preferences is context dependent.

The attitude literature also contains competing theories on the link between attitudes and behaviors (Kim and Hunter 1993; Kraus 1995). The theories of reasoned action (Ajzen and Fishbein 1980) and planned behavior (Azjen 1987) posit that in order to predict a specific behavior from attitudes, a behavioral intention *specific to the same behavior* must be used (Eagly and Chaiken 1993). Under this theory, if CV questions are to be used to predict actual behavior, only attitudes related to paying are relevant, because this is the behavior of interest (Jorgensen et al 1999). An alternative viewpoint is that the nature of the attitude moderates the ability of the attitude to predict behavior. Following this theory, attitude strength is often cited as an important moderating variable; the stronger the attitude, the better it predicts behavior (Raden 1985). Returning to the role of attitudes in CV studies, under this theory any questions that measure some aspect of attitude strength are relevant. Possibilities included measures of the extremity of the affective attitudes (Ableson 1995), the importance of the environmental good (Krosnick et al. 1993), the certainty of one's attitudes (Raden 1985), the consistency between affective and cognitive attitudes towards the good (Millar and Tesser 1989), and many others.

Can the attitude research of social psychologists be used to identify either the BA or MTB model as the correct model? Not really. The attitude literature does not deem either of the conceptual models incorrect; they are both plausible. However, the value of exploring this line of research is that the choice of model should be informed by the nature of the attitude questions. There is enormous variation in the types of questions frequently found in CV studies that could be considered as "attitude questions" or could be answered on a Likert-style "attitude scale". In just the application discussed below the “attitude questions” could be sorted into many categories. For example, some questions deal specifically with the good of interest (“I am very concerned about the effects of runoff on Green Bay.”), while others reflect more general attitudes (“I support environmental protection.”). Some questions reflect affective attitudes (“The possibility that fishing will decline worries me.”), while others show cognitive attitudes (“It is inevitable that water quality will get worse.”). In some cases, the object of the attitude is the good, while in other cases it is the payment or action (“I object to new taxes”). Likert-scales are also often used to measure previous behavior. One question in the Green Bay water clarity study asked “How often do you fish from a boat on the Bay of Green Bay?”. Answers were given on a 5-point scale where a 1 indicated “Never” and a 5 indicated “Very Often”. In this case, a higher frequency of fishing could be considered an aspect of attitude strength.

Given this great variation in the types of questions whose answers could reasonably be deemed “attitudinal data”, it is highly possible that both the BA and MTB models are correct at least some of the time. Of greater relevance to this paper, is whether or not the choice of models will have a significant impact on the qualitative results of the analysis. The results of the application below suggest that this is not the case.

IV. The Random Parameter Logit (RPL) Model

The RPL model estimated in this paper is based on the mixture model found in Morey and Rossmann (2003). In that paper, the mixture model was developed to combine the features of RPL with a model of *explained* heterogeneity in which the parameters of the base model are deterministic functions of the observable characteristics of the individual. Though the Morey and Rossman paper does not consider responses to attitude questions, this extension is straightforward. In this section I present the model adapted for CV decision using the notation of the previous section to the extent possible.

Consider again an individual facing a dichotomous choice CV question, with offer amount $*T*, whose WTP is defined as

$$WTP_i = \beta_i w_i + \varepsilon_i \,. \tag{3.1}$$

This is identical to the WTP function in equation (2.1). Rather than assume homogenous preferences or classes of preferences, some or all of the elements of β are treated as normally distributed random variables. In our case, the mean and variance parameters of these random variables are functions of the observed demographic characteristics of the individual and the individual's responses to the attitude questions. For example, suppose there are three elements in the attribute vector, *w*, and two of the associated parameters are considered to vary across the population. The individual's WTP function in equation (3.1), can be rewritten as

$$\begin{aligned} WTP_i &= \beta_{1i} w_{1i} + \beta_{2i} w_{2i} + \beta_3 w_3 + \varepsilon_i \\ &= \left(\eta_{1i} + \gamma_2 z_i + \gamma_3 x_i\right) w_{1i} + \left(\eta_{2i} + \gamma_5 z_i + \gamma_6 x_i\right) w_{2i} + \beta_3 w_3 + \varepsilon_3 \end{aligned} \tag{3.2}$$

where *z* is a vector of demographic characteristics, *x* is a matrix of attitude responses, and $\eta_i = (\eta_{1i}, \eta_{2i})$ is an individual specific parameters unobserved by the analyst. The individual error term η is treated as a bivariate normal random variable with mean $\left(\gamma_1, \gamma_4\right)$ and variance $\Sigma = (\sigma_1^2, \sigma_4^2, \sigma_{1,4})$. The two unobserved error terms, η and ε are assumed independent.

With the standard assumptions on ε, the probability of observing individual *i*'s response to the CV question is

$$\Pr(r_i \mid \beta_i) = \left[\frac{1}{1 + e^{-\beta_i(\eta_i, \gamma | z_i, x_i) w_i}}\right]^{r_i} \left[\frac{e^{-\beta_i(\eta_i, \gamma | z_i, x_i) w_i}}{1 + e^{-\beta_i(\eta_i, \gamma | z_i, x_i) w_i}}\right]^{1 - r_i} \tag{3.3}$$

Note that this is similar to equation (2.2), however the structure of the WTP function is quite different. In the RPL model, the taste parameters, β_i are individual specific rather than class specific, and are partially determined by attitude and demographic parameters, γ. The individual likelihood function is

$$L_i\left(\beta_i \mid z_i, x_i\right) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \Pr(r_i \mid \beta_i)\phi\left(\beta_i \mid \gamma, \Sigma\right) d\beta \tag{3.4}$$

where the first term of the integrand is the probability of observing the CV response conditional on a given value of parameter vectors, as given in equation (3.3), and ϕ is the normal probability distribution function. The second term is the probability of individual *i* holding those particular parameters, given his demographics and attitude responses. The integration is then taken over all possible values of the taste parameter vector β_i.

It is well know that maximum likelihood estimation of RPL models requires simulation, as equation (3.4) has no analytical solution. Instead, the likelihood value is simulated by drawing *D* random draws from the multivariate normal distribution, calculating the likelihood value for each draw, and taking the average likelihood value. Train (2003) provides instructions and the GAUSS code for RPL estimation developed by Train, Revelt, and Ruud, can be found on Train's website[2]. The simulated maximum likelihood estimation for the RPL model estimated in this paper will produce consistent and efficient estimates of the distributional parameters associated with the taste parameters, β, including γ, and Σ.

4.1 *LC vs RPL*

The individual likelihood functions for the MTB LC model, BA LC model, and the RPL model can be seen in equations (2.4), (2.12), and (3.4), respectively. There are clear structural differences between the LC models and the RPL model; however, all three models take into

[2] http://elsa.berkeley.edu/Software/abstracts/train0196.html

account the CV choice data, *r*, the demographic data, *z*, and the attitude data, *x*. Conceptually, the inclusion of *x* in the RPL model as specified most closely matches the BA approach, in which deep underlying preferences control responses to attitude questions, and these responses then partially determine the specific taste parameters associated with the good being valued. As discussed in Provencher and Moore (2006), neither the LC nor RPL approach is statistically dominant, as they are not nested, and depending on the conditions of the model they are equally difficult to execute. Provencher and Bishop (2004) provide the only out-of-sample comparison of the forecasting abilities of the LC and RPL models. They find that if the total numbers of estimated parameters is held constant, neither approach forecasts significantly better than the other.

Like the choice between the BA and MTB models, the choice of the general modeling strategy must depend on the context of the application. With RPL models the estimated correlation between parameters (the off-diagonal elements of the matrix Σ in the model described above) are often constrained to zero to reduce the total number of parameters estimated in the model. In an application where the parameters are highly collinear, this would be a false assumption and lead to incorrect parameter estimates. In such an application, it would be preferable to treat these parameters as perfectly collinear, as is essentially the case of LC models. In the Boxall and Adamowicz paper, an RPL model was estimated in addition to the LC model, for comparison. However, in that paper, the distribution of the random parameters was not conditioned on the individuals attitude responses or demographics, so it is not surprising that the authors concluded that the RPL model also “identified heterogeneity, but captured it in a different way” than the LC model (Boxall and Adamowicz, p. 441). To the best of my knowledge, the application presented here represents the first attempt to compare RPL and LC models in which the preference heterogeneity of both models is at least partially explained by attitude responses.

V. Application to Green Bay Water Clarity

The models described above were used to estimate the preferences of northeastern Wisconsin property owners for pollution reduction in the bay of Green Bay. A mail survey was sent to a sample of 500 bayfront and 500 inland property owners in the four county region bordering Green Bay. The sample was stratified by county. The survey presented a plan to improve water clarity in Green Bay by four feet in all places and included maps depicting current water clarity and the water clarity resulting from the improvement. These maps were produced with remotely-sensed clarity data and allowed the respondents to see the effect of the improvement near their own property. The survey also included a dichotomous choice CV question asking respondents if they would vote for a referendum to increase property taxes in order to realize the water clarity improvement. Six bid amounts from $100 to $1000 annually were used. Further details of this data can be found in Moore, Provencher, and Bishop (2007). The remainder of this section focuses on the data used to estimate the models identified above.

In this application, the choice data, r, are the individual responses to the CV question. The attributes of the good varied across respondents so that the attribute vector, w, contains four variables: a constant, the inverse distance to the bay for all inland properties, ${}^{1}\!/_{d_i}$, the inverse of the initial water clarity level near the individual's property ${}^{1}\!/_{q_{oi}}$, and the offer amount, T. Respondent incomes were sorted into three income quantiles. Two dummy variables are used to indicate an individual's income in the middle (I_2) or highest (I_3) quantiles. The demographic characteristics also include a dummy variable indicating whether the property is bayfront ($w = 1$) or inland ($w = 0$).

The attitude data for this application come from the responses to 17 questions. Six of these had 4 possible answers, one question had 5 possible answers, and 10 had 6 possible answers. These questions, listed in Appendix A, include many types of attitude questions. With this information, the three models are specified as follows.

The MTB individual likelihood function is

$$L_{iMTB\_GB}=\sum_{c=1}^{C}\left[\Pr(c:w,I_2,I_3)\prod_{q=1}^{6}\prod_{s=1}^{4}\left(\pi_{qs|c}\right)^{x_{iqs}}\prod_{s=1}^{5}\left(\pi_{qs|c}\right)^{x_{iqs}}\prod_{q=1}^{10}\prod_{s=1}^{6}\left(\pi_{qs|c}\right)^{x_{iqs}}\left(\frac{1}{1+e^{-v_i}}\right)^{r_i}\left(\frac{e^{-v_i}}{1+e^{-v_i}}\right)^{1-r_i}\right] \tag{4.1}$$

$$where\ v_i=\beta_1^{c(i)}+\beta_2^{c(i)}\left(1/d_i\right)+\beta_3^{c(i)}\left(1/q_{0i}\right)+\beta_4^{c(i)}T$$

For the BA model, the individual likelihood function is

$$L_{iBA\_GB}=\sum_{c=1}^{C}\left[\frac{e^{m_i^{c(i)}}}{\sum_{a=1}^{C}e^{m_i^{a(i)}}}\left(\frac{1}{1+e^{-v_i}}\right)^{r_i}\left(\frac{e^{-v_i}}{1+e^{-v_i}}\right)^{1-r_i}\right]$$

$$where\quad m_i^{c(i)}=\gamma_1^{c(i)}+\gamma_2^{c(i)}w_i+\gamma_3^{c(i)}I_{2i}+\gamma_4^{c(i)}I_{3i}+\gamma_5^{c(i)}H_{2i}\left(x_i\right)+\gamma_6^{c(i)}H_{3i}\left(x_i\right) \tag{4.2}$$

$$v_i=\beta_1^{c(i)}+\beta_2^{c(i)}\left(1/d_i\right)+\beta_3^{c(i)}\left(1/q_{0i}\right)+\beta_4^{c(i)}T$$

where $H_2(x_i)$ and $H_3(x_i)$ are dummy variables categorizing the attitude responses of individual *i*, which are detailed below.

For the RPL model,

$$L_{iRPL\_GB}=\frac{1}{D}\sum_{d=1}^{D}\left[\left(\frac{1}{1+e^{-\left(\beta_1^d+\beta_2^d\left(1/d_i\right)+\beta_3^d\left(1/q_{0i}\right)+\beta_4^dT\right)}}\right)^{r_i}\left(\frac{e^{-\left(\beta_1^d+\beta_2^d\left(1/d_i\right)+\beta_3^d\left(1/q_{0i}\right)+\beta_4^dT\right)}}{1+e^{-\left(\beta_1^d+\beta_2^d\left(1/d_i\right)+\beta_3^d\left(1/q_{0i}\right)+\beta_4^dT\right)}}\right)^{1-r_i}\right]$$

$$\begin{aligned}\beta_1^d&=\eta_{1i}^d+\gamma_2^dw_i+\gamma_3^dI_{2i}+\gamma_4^dI_{3i}+\gamma_5^dH_{2i}\left(x_i\right)+\gamma_6^dH_{3i}\left(x_i\right)\\ \beta_2^d&=\eta_{2i}^d+\qquad+\gamma_8^dI_{2i}+\gamma_9^dI_{3i}+\gamma_{10}^dH_{2i}\left(x_i\right)+\gamma_{11}^dH_{3i}\left(x_i\right)\\ \beta_3^d&=\eta_{3i}^d+\gamma_{13}^dw_i+\gamma_{14}^dI_{2i}+\gamma_{15}^dI_{3i}+\gamma_{16}^dH_{2i}\left(x_i\right)+\gamma_{17}^dH_{3i}\left(x_i\right)\\ &\left(\eta_1,\eta_2,\eta_3\right)\sim N\left(\gamma_1,\gamma_7,\gamma_{12};\sigma_1^2,\sigma_7^2,\sigma_{12}^2,\sigma_{1,7},\sigma_{1,12},\sigma_{7,12}\right)\end{aligned} \tag{4.3}$$

where D is the number of draws of β. As is typically done, the marginal utility of money, β_4 is not allowed to vary across individuals, but the other taste parameters β_1, β_2, and β_3 are treated as random variables. In equation (4.3), β_1^d, β_2^d, and β_3^d are the dth random draw from a multivariate normal distribution with mean $\gamma = (\gamma_1, \gamma_7, \gamma_{12})$ and variance $\Sigma = (\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_{1,2}, \sigma_{1,3}, \sigma_{2,3})$.[3]

5.1 Defining Attitude Groups (H) in the BA and RPL models

The attitude data for individual *i*, consists of $\sum_{q=1}^{Q} S_q$ binary variables, where S_q is the number of possible responses to question q. This means there are 89 attitude variables. Due to computational limitations, it is not possible to condition class membership (in the case of the BA model), or the distribution of taste parameters (in the RPL model) on all of these variables individually. In their original paper, Boxall and Adamowicz use factor analysis to reduce the number of variables. In their study, scores from 20 attitude questions were analyzed using principal component analysis and four underlying components were identified. Based on their attitude responses, individual scores for each component were calculated and used to condition class membership. In their study, the 20 attitude questions were all of a similar type, asking about motivations driving the decision to visit a wilderness area. This is not the case in the application analyzed here. In addition, factor analysis can only be used with interval or scale data, not with categorical data. Some of the attitude questions in the Green Bay analysis allow for an “I don’t know” response, and of course, any of the questions could be unanswered, which I consider a valid response category. For these reasons, factor analysis could not be used.

[3] Note that the demographic variable, *w*, is not included in the definition of the taste parameter on the inverse distance variable. The distance variable has already been multiplied by a water dummy, in that inverse distance is zero for bayfront properties. Including the additional term creates a perfectly collinear data matrix and so the model cannot be estimated.

Instead, a cluster analysis was used to reveal natural groupings within the data. This process sorts the individual respondents into different groups based on the similarity of their responses to the attitude questions. In this respect it is similar to the attitude only LC model, estimated equation (2.6), but instead of z, class membership is dependent on x. But again, the size of the x matrix prohibits the estimation of that type of LC model. Also, unlike the LC model, the cluster analysis treats group assignment as deterministic (Everitt 1993). The SPSS two-step cluster analysis procedure is capable of processing very large data sets with both categorical and continuous data. In addition, the optimal number of clusters can be endogenously determined based on the Baysian Information Criterion (BIC), which is described below. Using this procedure, a cluster analysis resulted in the creation of three attitude groups in the Green Bay study. The variables $H_h(x_i)$ in equations (4.2) to (4.3) above equal 1 if individual i is a member of attitude group h, and zero otherwise, for h = (2, 3).[4] Individuals within each group have similar values of x_i. Although this maintains the underlying premise of the BA model in that attitude responses influence class membership, they do so through the assignment to one of the H attitude groups. This is a potentially significant departure from the original specification used by Boxall and Adamowicz.

5.2 *The Choice of C in the LC Models*

There are no statistical tests capable of identifying the correct number of preference groups in a LC model, but in most studies a variety of statistical indicators are calculated and compared across models with different values of C (Wedel and Kamakura 2000). Commonly used are the Bayesian information criterion (BIC), Akaike information criterion (AIC), the consistent AIC (CAIC) and the corrected AIC (AIC_C) (Boxall and Adamowicz 2002). These

[4] The results of the clustering process can be sensitive to the order of individuals in the data file. The data were randomly sorted three times and a cluster analysis performed each time. Based on the similarity of the results, the final attitude groups appear robust.

criteria are only guidelines in the selection of *C* (Swait 1994). The AIC tends to be biased towards high values of *C*, but with small sample sizes, the BIC is biased towards low values of *C* (McLachlan and Peel 2000).

Table 1 shows selection statistics for both LC models for various values of *C*. The MTB Green Bay (MTB GB) model was estimated with 1, 2, 3, 4, and 5 preference classes, and all criteria favor the 4 class model. The BA Green Bay (BA GB) model was estimated with 1, 2, 3, and 4 preference classes, but with this model, all criteria favor only 2 preference classes. This suggests a clear difference between the two LC approaches.

5.3 *Estimation Results and Welfare Estimates*

All three models were estimated using GAUSS and MAXLIK. For the RPL model, the GAUSS code used in the Morey and Rossmann paper[5], was adapted for this purpose. The preferred MTB model has 4 preference classes. The first four columns of Table 2 highlight some of the difference between the classes in terms of their attitudinal data. In general, members of class 1 are occasional boaters or anglers who generally want to spend more on a variety of public goods, but think water clarity in Green Bay is okay and are very pessimistic about whether or not the improvement plan will make a difference. Class 2 contains most of the individuals who did not answer many of the attitude questions. Members of class 3 are most likely to enjoy "soft" recreational uses of the Bay, like non-motorized boating and picnicking. They are also specifically concerned about water quality in Green Bay, and place this above other public goods in importance. Members of class 4 make use the bay and shoreline the most, but tend to think water quality is okay and do not prioritize environmental goods above other public goods.

---

[5] This code is available at Morey's website, http://www.colorado.edu/economics/morey/dataset.html

In comparison, the preferred BA model contains only two preference classes. The characteristics of these classes are reported in the last two columns of Table 2. With only two classes, and with class membership conditioned on attitude responses, the comparisons are much more definitive. Members of class 1 use the bay and shore line for recreation more often, think water quality is worse and are more concerned about it, and are less pessimistic that the improvement plan will be make a difference. Members of class 2 are more concerned about roads and economic growth than members of class 1.

While the BA model creates classes in which the attitudes are clearly distinct across groups, the MTB model is better able to assign members to each class. Table 3 reports the estimated probability of class membership conditioned on the appropriate response variables for each model. Every individual has a nonnegative conditional probability of being in each class but estimation results are considered more reliable if individuals are predicted to have a high probability of belonging to one particular class rather than a relatively equal probability of belonging to several classes. The final column of Table 3 lists the 25th percentile of the conditional class membership probabilities for each model. These numbers imply that with the MTB model, 75% of all members of classes 1 and 2 have over a 97% probability of belonging to their respective class. These two classes include over half of the respondents. The conditional membership probabilities for the BA model are slightly lower, though still very high. With respect to this study, the BA model appears to estimate fewer preference classes, but cannot assign class membership as decisively as the MTB model.

The estimated parameters for each model are shown in Tables 4, 5, and 6. Interpretation of the MTB parameters in Table 4 is fairly straightforward. In Table 5, the first four rows indicate the taste parameters for the two preference groups. The next six rows show the estimated parameters that translate demographic and attitude characteristics into class membership. In the estimation, these parameters must be set equal to 0 for one of the classes in order to identify the parameters of the other classes. As such, the estimates of γ shown in Table 5 are standardized estimates. Also shown in the last two rows of Tables 4 and 5 are point estimates of the expected WTP conditional on class, E[WTP|c], and the unconditional expected WTP, E[WTP]. These are calculated as

$$\begin{aligned} E[WTP \mid c] &= \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{-\beta_4^c}\left(\beta_1^c + \beta_2^c\left(\tfrac{1}{d_i}\right) + \beta_3^c\left(\tfrac{1}{q_{0i}}\right)\right)\right) \\ E[WTP] &= E\left[E[WTP \mid c]\right] = \sum_{c=1}^{C}\left(E[WTP \mid c]\Pr(c)\right) \end{aligned} \tag{4.4}$$

where E[WTP|c] is equal to the sample mean of individual WTP given everyone is in class *c*. E[WTP] is then the mean of these conditional values, weighted by the probability of being in each class. The specific form of the individual WTP function (the term being summed in the top equation) comes from equation (2.1).

Table 6 presents the results of the RPL model, in which the taste parameters are treated as random variables, except for the marginal utility of income, β_4 which is held constant. The estimated γ variables represent the impact of the demographic and attitude variables on the estimated mean of the taste parameters, as shown in equation (4.5). The estimates of σ are estimates of the variance-covariance matrix of the taste parameters. For example, the point estimate of the population variance of β_1 is $\sigma_1^2 = 0.031$, with a standard error of 0.229. This reflects the distribution of β_1 across the population. It is <u>not</u> the variance of the point estimate of β_1. Similarly, the estimated covariance between β_1 and β_2 is $\sigma_{1,2} = -0.0392$, with a standard error of 3.572. The unconditional E[WTP] is given in the last row of Table 6. Note that this value depends only on the estimated means of the random parameters and not on their estimated variance, so that

$$
\begin{aligned}
&E[WTP] = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\beta_4}\left(\beta_{1i} + \beta_{2i}\left({}^{1}\!/\!_{d_i}\right) + \beta_{3i}\left({}^{1}\!/\!_{q_{0i}}\right)\right) \\
&\textit{where} \qquad \beta_{1i} = \gamma_1 + \gamma_2 w_i + \gamma_3 I_{2i} + \gamma_4 I_{3i} + \gamma_5 H_{2i}(x_i) + \gamma_6 H_{3i}(x_i) \\
&\qquad\qquad\quad \beta_{2i} = \gamma_7 + \qquad\quad + \gamma_8 I_{2i} + \gamma_9 I_{3i} + \gamma_{10} H_{2i}(x_i) + \gamma_{11} H_{3i}(x_i) \\
&\qquad\qquad\quad \beta_{3i} = \gamma_{12} + \gamma_{13} w_i + \gamma_{14} I_{2i} + \gamma_{15} I_{3i} + \gamma_{16} H_{2i}(x_i) + \gamma_{17} H_{3i}(x_i)
\end{aligned}
\tag{4.6}
$$

The differences in Tables 4, 5, and 6 highlight the impact of model selection on the quantitative results of the estimation. The MTB model identifies 4 preference classes while the BA model only identifies one. All three models produce different estimates of the taste parameters, β, though the estimates are similar across models. Many of the estimated parameters in the RPL model are not significant, while the majority of the LC model parameters are. The conditional E[WTP|c] is shown to vary wildly for different classes, ranging from -$299.76 to $855.13 in the MTB model and from -$196.17 to $739.35 in the BA model. Based on the results of the RPL model, a bayfront property owner with an income in the highest quantile could have a WTP as high as $712.12. At the same time, an inland property owner, whose property is 10 km from the bay and whose income is in the lowest quantile, could have a WTP as low as -$529.20. While the numbers don't match up exactly, this range of individual WTP is consistent with those of the LC models. In fact, the unconditional E[WTP], which is an estimate of the WTP of an individual drawn at random from the entire population, is almost identical in the two LC models, with a value of $299.11 in the MTB model, $291.49 in the BA Model, and only slightly lower in the RPL model ($266.00). In terms of the policy implications of the analysis, the models appear to produce very similar results.

VI. Conclusions

The purpose of this paper was to investigate the impact of different models of attitude based preference heterogeneity on the qualitative conclusions of the study. Three current models were chosen from the literature and adapted for use with dichotomous choice CV data. Based on the application presented in this paper, it does not appear as if the choice of modeling approach will significantly impact point estimates of the expected WTP for an environmental good. The E[WTP] estimates from the Green Bay data are almost identical for the three models. Using these results as evidence, one could reasonably argue that significant energy should not be spent trying to identify the correct model, so long as all of the available information is used in the estimation.

In placing the three models side by side, this paper the paper allows for easy comparison of the conceptual models underlying the high powered econometrics. There is conflicting opinion, amongst economists and social psychologists, regarding the correct use of attitudinal data. The literature on attitude theory developed by social psychologists and only touched upon here identifies many types of attitudes and posits several theories on the causal linkages between the various attitude types and observed behavior. For an economist wanting to incorporate the "attitude" questions from their stated preference survey into the estimation, it is tempting to lump anything other than choice data and demographics together as "attitudes". Admittedly, the application in this paper includes a wider range of attitude question types than the two previous studies referenced. But support for the inclusion of each of these questions can be found in the attitude literature. To make the situation more confusing, the distinction between different types of data can be quite fuzzy. For example, is boat ownership a demographic variable, as it was considered by Morey, Thacher, and Breffle, in their study of fishing preferences? Or is it a reflection of the attitude that boating is important? In the end, these questions should be carefully considered and the context of the attitudinal data should be a driving force in model selection.

VII. References

Ableson, R. (1995), "Attitude Extremity." In *Attitude Strength: Antecedents and Consequences*, edited by R. Petty and J. Krosnick. Lawrence Erlbaum Associates: New Jersey.

Ajzen, I. and M. Fishbein (1980), *Understanding Attitudes and Predicting Social Behavior*. Prentic-Hall: Englewood Cliffs, NJ.

Ajzen, I. (1987), "Attitudes, Traits, and Actions: Dispositional Prediction of Behavior in Personality and Social Psychology." *Experimental Social Psychology* **20:** 1-63.

Barro, S., M. Manfredo, T. Brown, and G. Peterson (1996), "Examination of the Predictive Validity of CVM Using an Attitude-Behavior Framework." *Society and Natural Resources* **9(2):** 111-24.

Bishop, R. and T. Heberlein (1979), "Measuring Values of Extramarket Goods: Are Indirect Measures Biased?" *American Journal fo Agricultural Economics* **61(5):** 926-930.

Boxall, P.C. and W.L. Adamowicz (2002), "Understanding Heterogeneous Preferences in Random Utility Models: A Latent Class Approach." *Env. Res.Econ.* **23(4):** 421-446.

Breffle, W., E. Morey, and J. Thacher (2005), "Combining Attitudinal and Choice Data to Improve Estimates of Preferences and Preference Heterogeneity: A FIML, Discrete-Choice, Latent-Class Model." Unpublished Manuscript.

Chen, H. and S. Cosslett (1998), "Environmental Quality Preference and Benefit Estimation in Multinomial Probit Models: A Simulation Approach." *Amer. J. Agric. Econ.* **80:** 512-20.

Dempster, A., N. Laird, and D. Rubin (1977), "Maximum Likelihood from Incomplete Observations." *Journal of the Royal Statistical Society: Series B*, **39**: 1-38.

Eagly, A. and S. Chaiken, (eds.) (1993), *The Psychology of Attitudes*. Harcourt Brace Jovanovich College Publishers: New York, NY.

Everett, B. (1993), *Cluster Analysis*. Edward Arnold: London.

Haab, T. and K. McConnell (2003), *Valuing Environmental and Natural Resources: The Econometrics of Non-Market Valuation*. Edward Elgar: Northhampton, MA.

Jorgensen, B., G. Syme, B. Bishop, and B. Nancarrow (1999), "Protest Responses in Contingent Valuation." *Environmental and Resource Economics* **14:** 131-50.

Kim, M. and J. Hunter (1993), "Attitude-Behavior Relations: A Meta-Analysis of Attitudinal Relevance and Topic." *Journal of Communication* **43(1):** 101-42.

Krosnick, J., D. Boninger, Y. Chuang, M. Berent, and C. Carnot (1993), "Attitude Strength: One Construct or Many Related Constructs?" *J. Personality and Social Psy.* **65(6):** 1132-51.

Kraus, S. (1995) "Attitudes and the Prediction of Behavior: A Meta-Analysis of the Empirical Literature." *Personality and Social Psychology Bulletin* **21(1):** 58-76.

Luzar, J. and K. Cosse (1998), "Willingness to Pay or Intention to Pay: The Attitude-Behavior Relationship in Contingent Valuation." *Journal of Social Issues* **27(3):** 427-455.

McCutcheon, A. (1996), "Multiple Group Association Models with Latent Variables: An Analysis of Secular Trends in Abortion Attitudes, 197201988." *Sociological Methodology* **26:** 79-111.

McFaddan, D. (1986), "The Choice Theory Approach to Market Research", *Marketing Science* **5:** 275-297.

McLachlan, G. and D. Peel (2000), *Finite Mixture Models*. John Wiley & Sons: New York, NY.

Millar, M. and A. Tesser (1989), "The Effects of Affective-Cognitive Consistency and Thought on the Attitude-Behavior Relation." *J. of Experimental Social Psychology* **25:** 189-202.

Morey, E., K. Rossmann, L. Chestnut and S. Ragland (2002), "Modeling and Estimating WTP for Reducing Acid Deposition Injuries to Cultural Resources: Using Choice Experiments in a Group Setting to Estimate Passive-Use Values." In *Valuing Cultural Heritage*, edited by S. Navrud and R. Ready, Edward Elgar Publishing.

Morey, E. and K. G. Rossmann (2003), "Using Stated-Preference Questions to Investigate Variations in Willingness to Pay for Preserving Marble Monuments: Classic Heterogeneity, Random Parameters, and Mixture Models." *Journal of Cultural Economics* **27(3-4):** 215-229.

Morey, E., J. Thatcher, and W. Breffle (2006), "Using Angler Characteristics and Attitudinal Data to Identify Environmental Preference Classes: A Latent-Class Model." *Env. Res. Econ.* **34:** 91-115.

Poe, G. and R. Bishop (1999), "Valuing the Incremental Benefits of Groundwater Protection when Exposure Levels are Known." *Env. Res. Econ.* **42(3):** 425-436.

Pouta, E. (2004), "Attitude and Belief Questions as a Source of Context Effect in a Contingent Valuation Survey." *Journal of Economic Psychology* **25:** 229-242.

Provencher, B., K. Barenklau, and R. Bishop (2002), "A Finite Mixture Logit Model of Recreational Angling with Serially Correlated Random Utility." *American Journal of Agricultural Economics* **84(4):** 1066-75.

Provencher, B. and R. Bishop (2004), "Does Accounting for Preference Heterogeneity Improve the Forecasting of a Random Utility Model? A Case Study." *Journal of Environmental Economics and Management* **48(1):** 793-810.

Provencher, B. and R. Moore (2006), “A Discussion of ‘Using Angler Characteristics and Attitudinal Data to Identify Environmental Preference Classes: A Latent-Class Model’.” *Environmental and Resource Economics* **34:** 117-124.

Raden, D. (1985), “Strength-Related Attitude Dimensions.” *Social Psychology Q.* **48(4):** 312-30.

Scarpa, R. and M. Thiene (2005), “Destination Choice Models for Rock Climbing in the Northeastern Alps: A Latent Class Approach Based on Intensity of Preferences.” *Land Economics* **81(3):** 426-444.

Swait, J. (1994), “A Structural Equation Model of Latent Segmentation and Product Choice for Cross-Sectional Revealed Preference Choice Data.” *J. of Retailing and Consumer Services* **1:** 77-89.

Tittering, D., A. Smith, and U. Makov (1985), *Statistical Analysis of Finite Mixture Distributions*. John Wiley & Sons: New York, NY.

Train, K. (1998) *“Recreation Demand Models with Taste Differences over People” Land Economics* **74:** 230-39.

Train, K. (2003) *Discrete Choice Methods with Simulations*. Cambridge University Press: Cambridge, U.K.

Tversky, A. and D. Kahneman (1986), “Rational Choice and the Framing of Decisions.” *Journal of Business* **59:** 265-278.

Wedel, M. and W. Kamakura (2000), *Market Segmentation: Conceptual and methodological Foundataions* (2nd edition). Kluwer Academic Publishers.

Wilson, M.A. (2000), “Rethinking Scope Sensitivity and Contingent Valuation Surveys: Strong Environmental Attitudes and Contingent Economic Values.” Ph.D. Dissertation, University of Wisconsin – Madison.

Yamaguchi, K. (2000), "Multinomial Logit Latent-Class Regression Models: An Analysis of the Predictors of Gender-Role Attitudes among Japanese Women." *American Journal of Sociology* **105(6):** 1702-1740.

Zajonc, R. and H. Markus (1982), "Affective and Cognitive Factors in Preferences." *Journal of Consumer Research* **9:** 123-31.

Table 1. Selection Criteria for the Number of Classes

Model	C	#P	LL	AIC	CAIC	AICc	BIC
MTB Green Bay (N = 567)	1	4	-354.088	716.176	733.545	716.283	366.769
	2	8	-276.942	569.883	604.620	570.206	302.303
	3	12	-266.444	556.889	608.994	557.547	304.487
	4	**16**	**-251.058**	**534.115**	**603.589**	**535.230**	**301.781**
	5	20	-250.878	541.756	628.5985	543.4514	314.2816
BA Green Bay (N = 567)	1	4	-354.585	717.170	734.538	717.277	367.265
	2	**13**	**-287.822**	**601.644**	**658.092**	**602.405**	**329.034**
	3	22	-285.184	614.369	709.896	616.402	354.928
	4	31	-283.335	628.671	763.277	632.626	381.611

The preferred models are highlighted in bold.

Table 2. Attitude responses by class for the BA and MTB models.

	MTB				BA	
	Class 1	**Class 2**	**Class 3**	**Class 4**	**Class 1**	**Class 2**
% Responding “Do More”						
Make Government More Efficient	65	49	75	83	76	69
Improve Education	37	30	51	68	62	40
Improve Roads and Highways	22	23	15	35	24	26
Encourage Economic Growth and Jobs	64	40	50	60	51	64
Clean Up Pollution	43	57	92	90	93	54
% Responding "Never”						
Shoreline Fishing	48	32	38	14	27	36
Fishing from a Boat	40	23	35	8	20	31
Motor Boating	45	30	25	9	15	36
Sail Boating	93	53	54	72	62	83
Canoeing/Kayaking	89	53	43	64	52	79
Picnicking	40	19	18	19	18	32
Walking/hiking/jogging	31	23	11	12	9	29
Biking	55	38	30	33	31	48
% Responding “Agree” or “Strongly Agree”						
I do not think the government would really spend the money on Green Bay.	59	11	33	56	44	52
I am pessimistic about whether such programs will actually work	64	30	18	48	33	56
% Who are “Very Concerned” about the effects of runoff on Green Bay	23	57	94	67	86	32

Table 3. Conditional probability of class membership.

	Class	Mean (%)	25^{th} percentile (%)
MTB model Pr(c:z|x,r)	1	99.45	99.99
	2	95.08	97.88
	3	93.91	92.79
	4	91.13	87.04
BA model Pr(c:z,x|r)	1	91.79	87.47
	2	93.59	92.46

Table 4. Parameter estimates and associated standard errors: MTB model

	Class 1	Class 2	Class 3	Class 4
Constant	-0.337 (0.787)	2.001 (0.487)	0.005 (0.937)	-1.048 (0.657)
$1/d_i$	4.212 (2.938)	0.930 (1.943)	10.256 (8.161)	-4.901 (10.232)
$1/q_{0i}$	0.861 (9.603)	19.014 (7.309)	2.628 (8.465)	-0.398 (10.832)
T_i	-1.721 (1.213)	-3.566 (0.606)	-3.949 (1.250)	-3.886 (1.802)
E{WTP|c}	-120.98	855.13	87.90	-299.76
Unconditional E{WTP} = $299.11				

Table 5. Parameter Estimates and associated standard errors: BA model.

		Class 1	Class 2
Taste Parameters, β	Constant	1.607 (0.412)	-1.025 (0.449)
	$1/d_i$	2.221 (2.750)	4.542 (3.709)
	$1/q_{0i}$	14.045 (5.780)	4.641 (5.839)
	T_i	-3.261 (0.524)	-3.492 (1.050)
Class Parameters, γ	Constant	-	0.243 (0.586)
	w	-	-1.969 (0.920)
	I_2	-	-0.988 (0.742)
	I_3	-	-3.075 (1.545)
	H_2	-	6.953 (2.127)
	H_3	-	2.909 (1.138)
	E{WTP|c}	739.35	-196.17
Unconditional E{WTP} = $291.49			

Table 6. Parameter estimates and associated standard Errors: RPL Model.

		Variable	Estimate	St. Error
β_1	γ_1	Constant	0.751	0.448
	γ_2	water	0.429	0.426
	γ_3	I_2	-0.017	0.406
	γ_4	I_3	0.701	0.500
	γ_5	H_2	-2.285	0.490
	γ_6	H_3	-1.297	0.758
β_2	γ_7	$1/d$	2.857	2.943
	γ_8	$I_2(1/d)$	2.649	4.711
	γ_9	$I_3(1/d)$	-7.455	5.149
	γ_{10}	$H_2(1/d)$	7.074	5.510
	γ_{11}	$H_3(1/d)$	-4.075	9.081
β_3	γ_{12}	$(1/q_0)$	4.841	6.170
	γ_{13}	$water(1/q_0)$	0.842	6.301
	γ_{14}	$I_2(1/q_0)$	4.440	6.684
	γ_{15}	$I_3(1/q_0)$	4.448	8.146
	γ_{16}	$H_2(1/q_0)$	-4.986	7.539
	γ_{17}	$H_3(1/q_0)$	2.874	12.384
β_4		*Offer*	-2.903	0.365
Σ	σ_1^2	$Var(\beta_1)$	0.031	0.229
	σ_2^2	$Var(\beta_2)$	0.939	1.694
	σ_3^2	$Var(\beta_3)$	-2.123	2.587
	$\sigma_{1,2}$	$Cov(\beta_1, \beta_2)$	-0.392	3.572
	$\sigma_{1,3}$	$Cov(\beta_1, \beta_3)$	-3.389	1.721
	$\sigma_{2,3}$	$Cov(\beta_2, \beta_3)$	-0.457	1.794
Unconditional E{WTP} = $266.00				

Appendix. Attitude Questions for the Green Bay study.

1. For each issue, compared to what is being done now in the Green Bay area, do you think we should be doing less (1), about the same (2), or more? (1 = “Do Less”, 2 = “Do about the Same”, 3 = “Do More”)

   Make government more efficient
   Improve education
   Improve roads and highways
   Encourage economic growth and jobs
   Clean up pollution

2. In recent years, how often, if at all, have you or others in your household participated in the following activities on the Bay of Green Bay? (1 = “Never”, 5 = “Very Often”)

   Fishing from a boat
   Motor Boating (not fishing)
   Sail boating
   Canoeing/Kayaking

3. In recent years, how often, if at all, have you or others in your household walked, hiked, or jogged along the shoreline of Green Bay? (1 = “Never”, 5 = “Very Often”)

4. Overall, compared to other places you are familiar with, how would you rate the water clarity in the Bay of GB?

   _________Excellent
   _________Good
   _________Fair
   _________Poor
   _________I don’t know

5. How do you personally feel about the effects of runoff?

   _________ I am not at all concerned about the effects of runoff on Green Bay.
   _________ I am somewhat concerned about these effects.
   _________ I am very concerned about these effects.
   _________ I don’t know.

6. Thinking about water clarity in Green Bay, which of the options below do you think we as a state should choose.

   _________ Do less and spend less on runoff control; this would lead to further reductions in water clarity in future years.
   _________ Do and spend about the same as we are now, which would lead to continuation of conditions approximately like those depicted in the map Water Clarity in Green Bay-Now.
   _________ Do more and spend more on runoff control to get to conditions like those in the map Water Clarity-With More Runoff Control.

7. Please let us know how strongly you agree or disagree with each of the following statements (1 = “Strongly Agree”, 2 = “Agree”, 3 = “Neutral”, 4 = “Disagree”, 5 = “Strongly Disagree”)

   I would rather see the money used for other environmental purposes.
   I object to new taxes for any purpose including this one.
   I do not think the government would really spend the money on Green Bay.
   I am pessimistic about whether such programs will actually work.