Year 2008 *Paper* 1075

Taste Indicators and Heterogeneous Revealed Preferences for Congestion in Recreation Demand

Felipe Vasquez Lavin
University of Conception, Chile

W. M. Hanemann
University of California, Berkeley

Taste Indicators and Heterogeneous Revealed Preferences for Congestion in Recreation Demand

Abstract

Researchers using revealed preference data have mostly relied on the Mixed Logit (ML) framework to model unobserved heterogeneity. In this paper, we suggest an extension of this model where we integrate direct measures of taste and revealed preferences, under a unified econometric setting, to describe heterogeneous preferences for congestion in recreation demand. ML is a random parameter discrete choice model, which decomposes the coefficients of the regression equation into a mean effect shared by all individuals in the sample, and a deviation with respect to this mean, specific to each individual. Within this structure, heterogeneity is summarized using a parametric density function for the coefficients of the model. From this distribution one can identify the portion of people who like or dislike an attribute of the good. On the other hand, taste indicators, represented in a like-dislike scale, constitute complementary information about the distribution of tastes in the population. We combine both sources of information to characterize preferences in our model. The traditional ML suggests almost 60% of people in the sample like crowded places while our integrated model implies almost 100% of the people dislike congestion. These results show the beneof using taste indicators to describe heterogeneous preferences for attributes describing alternatives of a choice set.

Taste Indicators and Heterogeneous Revealed Preferences for Congestion in Recreation Demand*

Felipe Vásquez Lavín
Department of Economics, University of Concepción, Chile.

Michael Hanemann
Department of Agricultural and Resource Economics,
University of California, Berkeley.

June 11, 2008

Abstract

Researchers using revealed preference data have mostly relied on the Mixed Logit (ML) framework to model unobserved heterogeneity. In this paper, we suggest an extension of this model where we integrate direct measures of taste and revealed preferences, under a unified econometric setting, to describe heterogeneous preferences for congestion in recreation demand. ML is a random parameter discrete choice model, which decomposes the coefficients of the regression equation into a mean effect shared by all individuals in the sample, and a deviation with respect to this mean, specific to each individual. Within this structure, heterogeneity is summarized using a parametric density function for the coefficients of the model. From this distribution one can identify the portion of people who like or dislike an attribute of the good. On the other hand, taste indicators, represented in a like-dislike scale, constitute complementary information about the distribution of tastes in the population. We combine both sources of information to characterize preferences in our model. The traditional ML suggests almost 60% of people in the sample like crowded places while our integrated model implies almost 100% of the people dislike congestion. These results show the benefits of using taste indicators to describe heterogeneous preferences for attributes describing alternatives of a choice set.

*The paper benefited from discussions with Kenneth Train, Jeff LaFrance, Ricardo Cavazos, and participants at the ARE summer 2007 workshop, and at the research seminar at the College of Natural Resources and the Environment, University of Massachusetts, Amherst.

1 Introduction

Preference heterogeneity plays an important role in the modern econometric analysis of discrete and discrete/continuous consumer choice behavior. Alternative econometric approaches – for example random coefficient logit versus classical logit – differ in how they characterize preference heterogeneity. From a policy perspective, too, preference heterogeneity can be important, whether in the policy context of designing different interventions that target different groups with different preferences and behavior patterns or in the political economy context of being able to identify winners and losers and deal with them appropriately. Similarly, for marketing purposes, the ability to chart preference heterogeneity can have important practical applications.

One way to identify this type of heterogeneity in the population is to use direct measures of tastes, sometimes called psychometric indicators or attitudinal measurements because they are intended to capture underlying psychological factors affecting decisions. The measurement of tastes may involve presenting people with a series of statements about their attitudes or perceptions of attributes of the commodity and ask them to identify how strongly they agree or disagree with each statement on an ordinal scale. The responses may be coded as a positive or negative value depending on whether the respondent agrees or disagrees with the statement. Such measures have been used broadly in the latent variables literature (Train et al., 1987a; Ben-Akiva et al., 1999a) where they were considered proxy variables for unobservable explanatory variables in a discrete choice model.

A more sophisticated approach is to estimate a random parameter discrete choice model using observed behavior (i.e., a revealed preference approach). A random parameter model is an extension of the traditional fixed parameter model in which coefficients associated with the explanatory variables are assumed to be identical for everybody in the sample. The sign of the coefficient on an attribute in a fixed parameter logit model indicates whether on average people like it or dislike it but, in this framework, it is impossible to know how large a proportion of the sample likes or dislikes it. In the random parameter framework, by contrast, coefficients of attributes are decomposed into a mean value, representing the average attitude over the whole sample, and a deviation from the mean which is specific to each person in the sample. Within this structure, heterogeneity is summarized using a parametric density function for the coefficients in the model; this distribution reflects how preferences are distributed across the sample, and from this, one can estimate the proportion of the sample who likes an attribute and the proportion who dislikes it (Train, 1998).

The most extensively used random parameter discrete choice model is the Mixed Logit model, which has been utilized to represent preference heterogeneity in many areas of applied econometrics including industrial organization (Berry, 1994; Berry et al., 1995, 1998; Nevo, 2000), marketing (Louviere et al., 2002; Wedel et al., 1999; Ben-Akiva et al., 1999a; 1999b) and nonmarket valuation (Train, 1998; Von Haefen et al., 2004).

Results in a *ML* model depend strongly on the assumptions made about the distribution of the unknown component of the coefficients. For some coefficients, there is a theoretical basis for assessing whether the estimated distribution of a coefficient on an attribute has a reasonable shape. However, on some occasions we do not know the theoretically correct sign or the appropriate shape of the distribution for the coefficient of interest. In other words, we do not know whether the outcome of the random parameter estimation reflects the true taste variation in the sample. Often, it is for those coefficients that we may be most likely to employ a random parameter specification; we want to estimate the entire distribution of the coefficient because we want to interpret this distribution as a reflection of taste variation in the population.

Here are examples of these two situations in the context of recreation demand models. First, consider a travel cost (price) coefficient, and assume the random component of this coefficient is normally distributed over the population. With this setup, it is often found that a portion of the coefficient's distribution lies in the positive part of the real line. This implies there is a nonzero probability of a positive price coefficient for some individuals, though this is not plausible theoretically. Given this prior belief about the sign of the coefficient, researchers solve this inconsistency constraining the support of the distribution so that the estimated model satisfies the theoretical restrictions, for example, using a lognormal distribution for the coefficient (Walker and Ben-Akiva, 2002; Train, 2003).

Second, consider a parameter associated with a variable measuring the crowding conditions in a recreational site. Again, using a normal distribution assumption for the unknown random component, we will probably obtain a portion of the estimated distribution in both the positive and negative parts of the real line. Evaluation of this result is more difficult because we do not have an unambiguous prior belief for the sign of the parameter. We could always develop arguments to support either sign.[1] On one hand, we could justify a negative sign for the coefficient by arguing people dislike crowds, which is typically called a congestion effect. On the other hand, a positive coefficient would suggest an agglomeration effect, i.e., people like to be in sites other people like (and are therefore crowded).

It could be argued that the same considerations involved in assessing a correct distribution for the price coefficient should also apply to the crowding coefficient, or whatever is the coefficient of interest. If so, we should proceed in the same way as with the distribution for the price coefficient, i.e., using a predetermined criterion or some external information (or possibly our own belief) regarding the correct location of the distribution, and impose the necessary restrictions to make it conform to this location. Unfortunately, unlike the price example, we do not have information coming from economic theory to specify how people should react to this attribute.

The common practice in random parameter models is to adjust the distribution of those coefficients for which we have prior beliefs, but to accept as correct the empirically estimated

[1] A similar situation can be found in the traditional fixed parameter estimation. Given the lack of theory or additional information about the relationship between an explanatory variable and the dependent variable, almost any sign of the coefficient and sometimes any magnitude of it can be explained by some creative relationship between the variables.

distributions of those coefficients for which we do not have such priors. In this way we treat different coefficients asymmetrically, and there is no reason to think that only some of the distributions are correct.

Forming the wrong conclusion about the congestion effect can have serious policy consequences. For example, if closing a site to sportfishing increases congestion at other sites, this sorting effect will be harmful for visitors to the other sites if they dislike congestion but beneficial if they like being with other people. The problem is we may not be able in practice, to identify the correct sign for the coefficient on crowding because, in the data, the crowding variable is highly correlated with other attributes of the site such as the abundance of fish or low travel cost.

The problem arises because, if we rely only on observed behavior, our inference suffers from the lack of an experimental design that would more completely reveal preferences. With observational data, we do not have control over the combinations of attributes among the choice alternatives and cannot provide a comprehensive set of alternatives involving all possible combinations of attribute levels. Certain combinations of attributes and alternatives will inevitably be missing from our data. Consequently, if we observe a person visiting a very crowded site, this does not necessarily mean he likes crowding; it might be that he tolerates crowded sites because other characteristics of the site make it attractive despite the crowding.

Asking people directly about their attitudes towards crowding and other attributes may help to overcome this informational limitation and may permit us to better characterize the distribution of tastes in the population. However, relying exclusively on psychometric questions to infer preferences does not take into account the fact that people do make decisions under constraints and are willing to make trade-offs among attributes. It seems desirable, therefore, to combine both sources of information to characterize preference heterogeneity.

In this article we integrate revealed preferences and taste indicators to account for heterogeneity of preferences among individuals in a Mixed Logit (*ML*) framework. Our model incorporates psychological aspects influencing the decision making process and, therefore, determine the observed outcomes of a discrete choice model (the model can be classified as an hybrid model following Ben-Akiva et al., 2002 and Walker and Ben-Akiva, 2002).

Ben-Akiva et al. and Walker and Ben-Akiva suggest the use of psychometric measurements or taste indicators to provide information on attitudes and perceptions about attributes of alternatives and use this information in the estimation process to enrich the model. This leaves open the questions of how these indicators have to be measured and to what extent they can provide additional information to identify the distributions of the parameters. These questions are the focus of the present research.

In our application, the ML model by itself indicates that around half of the people in the sample like crowding. However, the psychometric scale alone implies people strongly dislike crowding. Integrating the psychometric scale in the estimation of the choice model leads to quite a different conclusion compared to either on its on: nobody in the sample likes congestion.

In the next section we discuss briefly previous literature on congestion and compare them to ours. We then describe the theoretical background of random parameters in the context of ML. This section presents the econometric tools needed to estimate the model and to characterize the distribution of taste among individuals. It also describes several ways to incorporate additional information to characterize this distribution suggested in the literature. Section 4 describes the data, the estimated model and the main econometric results. Finally, we present a robustness analysis of these results, including different definitions of the beta coefficient, other distributions for the simulation process such as the lognormal and multivariate normal distributions. We also present an extension of the ML model known as the Generalized Mixed Logit model (Fiebig et al., 2007), which allows us to identify the portion of the variability in the coefficients of the model explained by the variation of the variance of the error term. The last part of this section performs a test on endogeneity of the crowding index.

2 Literature review

Previous literature on congestion has found both positive and negative signs for parameters associated with crowding (Cicchetti and Smith, 1973; Deyak and Smith, 1978; McConnell, 1977; Berrens et al., 1993; Boxall et al., 2003; Schuhmann and Schwabe, 2004).[2] As mentioned above, both results are justified as the existence of agglomeration or congestion effects. An alternative approach suggests that the effect at low levels of crowding is positive but after a certain threshold, crowding becomes detrimental to people's utility. Schuhmann and Schwabe (2004) estimate a model with linear and quadratic effects of crowding and find results consistent with this hypothesis; there is a positive linear effect and a negative quadratic effect.

Only a small number of these studies have estimated congestion effects using revealed preference methods (Deyak and Smith, 1978; Bell and Leeworthy, 1990) and most of them use stated preferences (McConnell, 1977; Boxall et al., 2003).

The differences in the sign of the parameter in these applications and the prevalence of stated preference studies can be explained by two factors. First, there is not a clear consensus about how to measure congestion in recreation demand models (Shelby, 1980; Jakus and Shaw, 1997; Boxall et al., 2003). Some authors use objective measures of congestion such as intensity of use, defined as aggregate demand per unit of area, parking spots available, waiting times for using resources like boat ramps or climbing walls, number of encounters with other users, etc., while other researchers focus on subjective measures or perceived measures of congestion. Second, Boxall et al. (2003) argue the problem with revealed preference estimations of congestion effects is that aspects such as expectations, experience

[2] Earlier literature discussing congestion from a theoretical perspective are Anderson and Bonsor (1974), Freeman and Havemann (1977), Cesario (1980), Smith (1981), and McConnell (1988).

and mitigation behavior are confounded with the measure of congestion and it is difficult to disentangle the effect of congestion from other aspects of the decision process due to its endogeneity.

None of revealed preference studies mentioned above use a discrete choice model similar to the one presented in this paper. Recently, Timmins and Murdock (2007) developed a method to estimate a revealed preference discrete choice model taking into account the endogeneity of congestion. Endogeneity in their model is the result of people sorting across sites, in other words, congestion is affected by similar unobservable attributes determining site selection. They apply this model to estimate welfare measures for fishing activities while O'Hara (2006), using a similar approach, estimates the effect of congestion in the demand for rock climbing. In order to capture part of the unobservable site attributes affecting selection Murdock (2006) suggests the estimation of a full set of alternative specific constants. The inclusion of this full set of alternative constants implies parameters varying only across alternatives are not identified. Their methodology relies on an instrumental variable regression approach that requires a large number of choice alternatives.

Finally, Freeman and Havemann (1977) and McConnell (1988) have shown theoretically that heterogeneity in willingness to pay and in aversion to crowding affect the optimal price required to achieve a socially optimum level of congestion. Furthermore, in their results, heterogeneity also determines the composition of users of the sites, since users with lower price sensitivity and higher aversion to crowding switch their demand to sites with higher entrance fees.

Previous literature on congestion has either ignored heterogeneity in preference for crowding or dealt with it through the estimation of different models for different groups of individuals, or different coefficients for different activities or sites. All of these approaches to congestion belong to the fixed parameter framework. Even though the sorting model of Timmins and Murdock (2007) can be extended to a random parameter model, they do not provide any estimation of heterogeneity for the congestion parameter and Murdock's (2006) empirical results on heterogeneity are unusual and contradict expected theoretical results.

Our paper provides a simpler way to deal with heterogeneity of preferences based on actual measures of crowding together with indicators of attitudes toward crowding. The model is easier to estimate and produces qualitatively similar results to those in the literature mentioned above. Additionally, our methodology can be applied with any number of alternatives as long as one has an explicit measure of congestion. Our congestion measure can be classified as an actual (no subjective) measure because it has been measured independently of the anglers by somebody outside of the sample. Therefore, following Jakus and Shaw (1997) it can be considered exogenous to the individual.[3]

Once the model has been estimated, any welfare calculation should take into account the sorting over sites produced by changes in the number of available sites or in the quality of them. We provide estimates of welfare measures under several assumptions about the sorting effects.

3 Model Specification

The traditional random parameter model (Train, 1998; 2003) starts with the definition of a utility function for individual n in period t given that he chose alternative j, denoted by

$$U_{njt} = \beta_n' x_{njt} + \varepsilon_{njt}, \tag{1}$$

where x_{njt} is a vector of observed attributes of the alternatives, β_n is a vector of unobserved coefficients varying randomly over individuals and ε_{njt} is an identically and independently distributed extreme value error term, independent of β_n. The vector of coefficients is defined as $\beta_n = b + \eta_n$, where b is the average effect and η_n represents an individual's deviation with respect to the average, in other words, β_n is characterized by a probability distribution,

[3] We will test this hypothesis in our estimation process.

$\beta_n \sim g(\beta\,|\theta)$, where θ is a set of parameters defining the distribution $g(.)$. Replacing this definition in (1) the utility function is

$$U_{njt} = (b + \eta_n)'\, x_{njt} + \epsilon_{njt} = b' x_{njt} + \eta_n' x_{njt} + \varepsilon_{njt}.$$

Given the assumption for the error term ε_{njt}, and conditioning on the random component η_n, the probability that individual n chooses alternative j in period t is

$$L_{njt} = \frac{e^{\beta_n' x_{njt}}}{\sum_i e^{\beta_n' x_{nit}}}, \tag{2}$$

which is the formulation of a conditional logit model.

Typically, the researcher observes a panel of N individuals making decisions over T periods. Grouping all decision occasions for the same individual, the joint probability that individual n chooses a sequence of alternatives $y_{n\mathbf{j}} = \langle y_{nj_1 1}, ..., y_{nj_T T}\rangle$ is given by

$$\mathbf{L}_{n\mathbf{j}}(\beta) = L_{ny_{nj1}1} * ... * L_{ny_{njT}T} = \prod_{t=1}^{T} L_{ny_{njt}t} = \prod_{t=1}^{T}\prod_{j}^{J}\left(\frac{e^{\beta_n' x_{njt}}}{\sum_i e^{\beta_n' x_{nit}}}\right)^{y_{njt}}, \tag{3}$$

with y_{njt} equal to 1 if person n chose alternative j in period t and 0 otherwise. Since we do not observe η_n, the unconditional probability is obtained integrating out this random component, that is,

$$P_{n\mathbf{j}} = \int_\beta \mathbf{L}_{n\mathbf{j}}(\beta) g\left(\beta|\,\theta\right) d\beta = \int_\beta \prod_{t=1}^{T}\prod_{j}^{J}\left(\frac{e^{\beta_n' x_{njt}}}{\sum_i e^{\beta_n' x_{nit}}}\right)^{y_{njt}} g\left(\beta|\,\theta\right) d\beta,$$

and the log of the likelihood function is

$$\ell = \ln L(\beta) = \sum_{n}^{N} \ln\left(\int_\beta \prod_{t}^{T}\prod_{j}^{J} \left(L_{njt}\right)^{y_{njt}} g\left(\beta|\,\theta\right) d\beta\right).$$

The integral in the likelihood function is approximated using simulation techniques, i.e., we take R random draws from the distribution $g\left(\beta|\,\theta\right)$, and the integral inside the likelihood function, which represents the probability $P_{n\mathbf{j}}$, is replaced by

$$\check{P}_{n\mathbf{j}} = \frac{1}{R}\sum_{r=1}^{R}\left(\prod_{t=1}^{T}\prod_{j}^{J}\left(\frac{e^{\beta^{r\prime}x_{njt}}}{\sum_i e^{\beta^{r\prime}x_{nit}}}\right)^{y_{njt}}\right)$$

where β^r denotes the r-th random draw. The estimated parameters $\hat{\beta}$ obtained by maximizing this likelihood function are called the simulated maximum likelihood estimators.

The random coefficient β_n can be defined as a function of observed variables with the purpose to correctly capture observed sources of variation in individuals' preferences. For instance, Harris and Keane (1999), Berry (1994) and Berry et al. (1995) incorporate individuals' information in the definition of β_n, so that the randomness of this coefficient depends on some observable information and not just on the assumptions about the random component η_n. In the ML model including individual characteristics will turn into interactions between characteristics and attributes of the good.

Ben-Akiva et al. (1999a; 1999b) and Walker and Ben-Akiva (2002) suggest including taste indicators in the model. The simplest way to incorporate these indicators is to treat them as additional explanatory variables and plug them directly into the definition of the random parameter β_n as in Harris and Keane (1999).[4] In summary, the first attempt to improve the description of heterogeneity is to include both taste indicators and demographic information in the definition of β_n, that is

$$\beta_n = b + \omega D_n + \rho H_n + \eta_n, \tag{4}$$

where D_n are demographic variables and H_n are variables reflecting individual's opinions

[4] See Walker and Ben-Akiva (2002) for a full literature review on different ways to extend the traditional random utility model.

about the attributes of the goods (taste indicators). The utility function becomes

$$U_{njt} = (b + \omega D_n + \rho H_n + \eta_n)' x_{njt} + \varepsilon_{njt} \tag{5}$$

$$= \omega D_n x_{njt} + \rho H_n x_{njt} + (b + \eta_n) x_{njt} + \varepsilon_{njt}. \tag{6}$$

This model includes new sources of heterogeneity through interactions between individual's information and attributes, $D_n x_{njt}$ and between indicators and attributes, $H_n x_{njt}$.

Ben-Akiva et al. (1999a; 1999b) and Walker and Ben-Akiva (2002) suggest a different way to exploit the information provided by taste indicators. Their proposed model deals with latent explanatory variables; thus, it applies in cases where the researcher does not observe all the relevant explanatory variables, and the taste indicators provide a measure of how people feel about the unobserved variables. In our case we do observe the relevant explanatory variables, but we want to represent how each individual feels about this attribute; in other words, we want to describe the distribution of the coefficients associated with these variables. We adjust the Ben-Akiva and Walker model to fit our objective.

The structural model propose by Ben-Akiva and Walker includes the utility equation given in (1) and the equation that characterize the random parameters given in (4) but without taste indicators, i.e.,

$$\beta_n^* = h(D_n, \varpi) + \eta = b + \omega D_n + \eta_n, \tag{7}$$

D_n is the matrix of explanatory variables including any relevant information except taste indicators. Ben-Akiva et al. define measurement equations that allow one to identify the structural model. In the choice case we have already defined these equations, which are the indicators y_{nit} explained above. This measurement provides the choice probability equation given by either (2) and (3) depending on whether we have a cross section or a panel data.

Finally, we have to define the taste indicator measurement equations

$$I_S = w\left(W, \beta_n^*; \alpha\right) + \upsilon \quad \upsilon \sim D\left(0, \Sigma_\upsilon\right),$$

these indicator equations are the responses to S psychometric questions that capture with some level of certainty the individuals' true tastes, which we want to identify. The definition includes any functional form for $w(.)$ and also any set of explanatory variables, W, beside the definition of β_n^*. Putting all these equations together we obtain a model similar to the *ML* described above, but with the inclusion of an additional element capturing the information provide by the taste indicator equations. The probability that an individual n chooses a sequence of alternatives, $y_{n\mathbf{j}}$, is

$$P_{n\mathbf{j}}\left(y_{n\mathbf{j}} \left| x_{n\mathbf{j}}, D_n \right.; \varpi, \eta, \varepsilon\right) = \int_{\beta^*} \mathbf{L}_{n\mathbf{j}}\left(y \left| x_{n\mathbf{j}}, \beta^* \right.; \varpi, \eta, \varepsilon\right) g\left(\beta^* \left| D_n \right.; \varpi, \eta\right) d\beta^*.$$

Compared to the previous choice probability formula we only introduced new notation to show this probability depends on the coefficients of the function defining β^*, i.e., coefficients in ϖ, on the error components of this equation η, and the error definition of ε.

Now introducing the indicator component into the model and assuming independence among error terms in the three equations, the joint distribution of $y_{n\mathbf{j}}$ and I is

$$f\left(y_{n\mathbf{j}}, I \left| x_{n\mathbf{j}}, D_n, W \right.; \varpi, \alpha, \eta, \upsilon, \varepsilon\right) =$$

$$\int_{\beta^*} \mathbf{L}_{n\mathbf{j}}\left(y \left| x_{n\mathbf{j}}, \beta^* \right.; \varpi, \eta, \varepsilon\right) f\left(I \left| W, \beta^* \right.; \varpi, \eta\right) g\left(\beta^* \left| D_n \right.; \varpi, \eta\right) d\beta^*,$$

where $f\left(I \left| W, \beta^* \right.; \varpi, \eta\right)$ is the density function associated with the indicator equations. Again this integral can be calculated by simulation. Assuming an extreme value distribution for ε, and a linear form for the indicators, $I_s = \alpha_s' \beta_n^* + \upsilon_s$, with υ_s normally distributed, we obtain

$$
\begin{aligned}
f\left(y_{n\mathbf{j}}, I\left|x_{n\mathbf{j}}, D_n, W\right. ; \varpi, \alpha, \eta, \upsilon, \varepsilon\right) &= \int\limits_{\beta^*} \prod_{t=1}^{T} \prod_{j}^{J} \left(\frac{e^{\beta_n^{*\prime} x_{njt}}}{\sum_i e^{\beta_n^{*\prime} x_{nit}}} \right)^{y_{njt}} * \\
& \prod_{r=1}^{r} \frac{1}{\sigma_{\upsilon_s}} \phi \left[\frac{I_s - \alpha_s' \beta_n^*}{\sigma_{\upsilon_s}} \right] g\left(\beta^* \left|D_n\right. ; \gamma, \eta\right) d\beta^*,
\end{aligned}
$$

which is calculated using simulations as

$$
\begin{aligned}
f\left(y_{n\mathbf{j}}, I\left|x_{n\mathbf{j}}, D_n, W\right. ; \varpi, \alpha, \eta, \upsilon, \varepsilon\right) &= \frac{1}{R} \sum_{r} \left(\prod_{t=1}^{T} \prod_{j}^{J} \left(\frac{e^{\beta^{*r\prime} x_{njt}}}{\sum_i e^{\beta^{*r\prime} x_{nit}}} \right)^{y_{njt}} \right) * \\
& \prod_{s=1}^{S} \frac{1}{\sigma_{\upsilon_s}} \phi \left[\frac{I_s - \alpha_s' \beta^{*r}}{\sigma_{\upsilon_s}} \right].
\end{aligned}
$$

Clearly, there are many ways to capture heterogeneity. The simplest way is to include individual characteristics interacting with attributes of the good that have fixed coefficients. The coefficients are fixed, but since the characteristic varies among individuals, the attribute's effect is specific to each individual. This is analogous to incorporating individual characteristics D_n and/or taste indicators H_n in the definition of the random parameters as shown in equation (5) since these variables will enter in the model as interactions with attributes of the alternatives with a fixed coefficient. Second, together with the interactions we could include a random component in the model that captures unobserved individual effects. Finally, we could use indicators, assuming they partly describe the true latent parameter β^*, and incorporate their distributions directly in the definition of the likelihood function. In this way the taste indicators are not a direct part of the random parameter definition but instead they contribute to calibrating the estimated distribution of the random parameter in order to make this distribution consistent with how people perceived these attributes.

Any of these modeling strategies provides a taste distribution for the whole sample; however, it might be useful to describe a distribution of taste for a specific consumer or group of consumers. We could follow Revelt and Train (1999) and Train (2003, ch. 11) to characterize a conditional distribution of taste for a subgroup of individuals. Their procedure narrows the area where a particular consumer lies in the original distribution of tastes using information of his/her previous choices. Therefore, we can predict the mean value of beta for each individual in the sample, and from this deduce whether he likes or dislikes the particular attribute.

Revel and Train (1998) distinguish between the distribution of tastes in the population and the distribution of tastes in a subpopulation of people who made a particular sequence of choices. We denote the former distribution by $g(\beta\,|\theta\,)$, and the latter distribution by $h(\beta\,|i,x,\theta\,)$, following their notation. From our estimates we can only say that an individual's parameter lies somewhere in the support of the distribution $g(\beta\,|\theta\,)$ without specifying whether he is in the positive or negative part of the distribution. The distribution $h(\beta\,|i,x,\theta\,)$ will provide better information about the relative position of an individual or group of individuals sharing the same chosen alternatives over time. Naturally, the quality of the information derived from the data depend on the number of observed choices available for each individual, with a greater precision for cases with several observed choices per individual.

The distribution of the subpopulation is obtained using the Bayes' Rule. The joint density of y_n and β can be written as

$$f\left(y_n,\beta\,|x_n,\theta\,\right)=P\left(y_n\,|x_n,\beta\,\right)g(\beta\,|\theta\,)=h\left(\beta\,|y_n,x_n,\theta\,\right)P\left(y_n\,|x_n,\theta\,\right),$$

then

$$h\left(\beta\,|y_n,x_n,\theta\,\right)=\frac{P\left(y_n\,|x_n,\beta\,\right)g(\beta\,|\theta\,)}{P\left(y_n\,|x_n,\theta\,\right)}.$$

From this distribution Train (2003) shows how to derive statistics conditional on y_n. For example, the mean is given by

$$\bar{\beta}_n=\int\beta h\left(\beta\,|y_n,x_n,\theta\,\right)d\beta=\int\beta\frac{P\left(y_n\,|x_n,\beta\,\right)g(\beta\,|\theta\,)}{P\left(y_n\,|x_n,\theta\,\right)}d\beta=\frac{\int\beta P\left(y_n\,|x_n,\beta\,\right)g(\beta\,|\theta\,)d\beta}{\int P\left(y_n\,|x_n,\beta\,\right)g(\beta\,|\theta\,)d\beta},\quad(8)$$

which can also be estimated using simulation, i.e. we get a random draw from $g(\beta\,|\theta\,)$ and estimate $\check{\beta} = \sum w^r \beta^r$ where $w^r = P\left(y_n\,|x_n, \beta^r\,\right) / \sum_r P\left(y_n\,|x_n, \beta^r\,\right)$. For this simulation we need to consider $g(\beta\,|\theta\,)$ depends on θ. The distribution for these estimators is normal with mean $\hat{\theta}$ and variance $\hat{W}$. These estimators are the outcomes from the maximization of the likelihood function. We could replace these point estimators in $\bar{\beta}_n(.)$ and take random draws from $g(\beta\left|\hat{\theta}\right.)$ to calculate the value of $\bar{\beta}_n$, or we could take a random draw from $N(\hat{\theta}, \hat{W})$, and calculate $\bar{\beta}^r$ based on this random value (θ^r), i.e., take the random draws of $g(\beta\,|\theta\,)$ after replacing θ^r on it and repeat this procedure R times.

4 Application

In the standard revealed preference approach with panel data, we observe N individuals making a recreational trip to one of a number of different destination zones in a period of time t, during T periods. Our data consists of 434 Alaskan anglers who made fishing trips to catch king salmon during the summer of 1986 (18 weeks). For each individual n, we have information on his travel cost (TC_{nj}) to each of the 21 sites where king salmon is available. We also have information on the quality of fishing (Q_{jt}) at each site on each decision occasion. This is based on detailed sportfishing advisories published each week by the Alaska Department of Fish & Game (ADFG) describing fishing conditions at sites around the state. They define the fishing in qualitative terms using adjectives and descriptors, rather than predicting a specific catch rate per hour of effort. Accordingly, we view this as an ordinal measure of fishing quality. Based on the advisories, our variable is coded as an eight-level indicator of fishing quality, starting from 1 ("no fish are available") to 8 ("excellent fishing"). The fish advisories are specific to different types of fish; we use here the advisory for king salmon. Additionally, we have a three-level indicator of crowding conditions (CR_{jt}) at site j in period t, with 0 for not crowded, 1 for somewhat crowded and 2 for very crowded. This is based on information provided by ADFG. There is also a dummy variable indicating whether people have a cabin in site j (CA_{nj}) and a measure of the amount of king salmon harvested in site j in the previous year (HAR_j). Finally, there are several individual information variables that will be described below. We use weekly data because fishing opportunities vary significantly from week to week as different runs of fish return to their spawning sites. The weekly variation in fishing is reflected in both the weekly fishing quality index for each site and also the specific set of sites at which king salmon is available week by week.

For an indication of the main characteristics of the data consider figure 1 which shows average values of quality, crowding and travel cost for each site in the sample. Sites are sorted according to quality in an increasing order, and to account for the differences in magnitudes of the three variables we plot the relative difference of each characteristic with respect to its mean (that is, $[x_i - \bar{x}] / \bar{x}$).

There are several patterns we can highlight from this graph. First, sites with lower travel cost and lower level of quality have consistently lower level of crowding (for example, sites 3, 4, and 5). Second, high quality and low cost always have high level of crowding (sites 10, 11, 13, 14, 16 and 21, among others). In other words, crowding moves in an opposite direction to price and in the same direction as quality. Third, the positive effect of quality is offset by cost at sites with a very high travel cost, especially at sites with the highest levels of quality. For instance, for sites 17 and 20 the travel cost is so high that it totally dominates the effect of quality. Lastly, there are a few exceptions to these patterns; site 2 has a low level of quality, a high price, and a high level of crowding.

In summary, the level of crowding is highly correlated with the other two explanatory variables and there is substitution among these characteristics. In other words, people tolerate higher levels of congestion because the quality of the fishing activities is good in that area. This is especially important with king salmon because many people want to catch a trophy size fish and, to do this, they have to go to sites with good fishing quality and put up with the congestion there.

The particular definition of the utility function in equation (1) is

$$U_{njt} = \beta_n^1 \ln(TC_{njt}) + \beta_n^2 * Q_{jt} + \beta_n^3 * CR_{jt} + \beta^4 * CA_{nj} + \beta^5 \ln(HAR_j) + \varepsilon_{njt}.$$

For identification purposes we assume only the first three coefficients are random. Even though the model can be estimated with all the coefficients being random, there have been concerns in the random parameter literature about the identification of models in that case (Walker, 2002; Ben-Akiva et al., 2001).

Table 1 presents the results of a fixed coefficient, ML, and the integrated ML model excluding any additional explanatory variable in the definition of beta.[5] While this table contains the main result of the paper, the tables following it show several variants of the simple models in order to check the robustness of our main result. The numbers on the first row of the table describe the different models, follows by a row with the name of the model and a summary of the coefficient and/or the components of the model. For instance, the first column denoted by (1), shows the fixed parameter model where we estimate only the coefficient b, while column (2) presents results for the ML in which we have estimated the coefficient b and used simulation to integrate out the component η. The third column (3) adds the letter I which means we have integrated the taste indicators.

In all of these models we assume the coefficients have independent normal distributions. We use 100 Halton sequences as the sampling procedure for the R draws needed in the calculation of the integral discussed above.[6] In these regressions, all coefficients are statistically significant except for the crowding coefficient in the fixed parameter model. The fixed coefficient model and the integrated models yield a negative sign for the crowding coefficient while the ML model suggests a completely different conclusion about people's preferences

[5] It useful to have in mind that in the random parameter estimation there are two parameters to estimate; the mean value and the standard deviation. Additionally, as in any other estimation we obtain the standard error for those parameters. These standard errors are used for the calculation of the t statistics.

[6] Actually, a Halton sequence is a systematic sampling procedure instead of a random procedure, which increases the coverage and reduces the variance of the simulator. See Train (1999; 2003) for a discussion on Halton sequences and other topics on simulation.

for crowding levels.

In comparison with the fixed parameter model, the random parameter models improve our information about the crowding parameter because they show a significant standard deviation for the coefficient, which implies a significant level of unobserved taste heterogeneity in the population. In the *ML* model, all the distributions have part of their mass on both sides of the real line. For the price coefficient, we observed a probability of 10% that an individual has a nonnegative value for the parameter, which is inconsistent with economic theory. Similarly from the distribution for the quality indicator we conclude that around 9% of the sample prefers sites with lower quality. These outcomes arise because we have not constrained the distributions to be only in the positive or negative part of the real line. We could do this using a log normal distribution for those coefficients or we could just discard the portion of the distribution that is outside of the expected range.[7]

As noted above, the crowding coefficient is not significant in the fixed parameter regression; however, it is significant in the *ML* estimation with a standard deviation that is also significant. Using results from the *ML* model, figure 2 shows the distribution $g(\beta\,|\theta)$ under the name *ML* distribution. These results imply that around 60% of the population likes crowding. If this were the case, it would explain why this coefficient is not significant in the fixed parameter estimation; negative values for some individuals cancel out with positive values of other individuals.

In the integrated models we use several taste indicators obtained in a Likert scale or like-dislike ranking. In one of these questions individuals were asked about how desirable was to fish in a site with few other fishermen around, the results are presented in table 2. For most of the people, crowded places are undesirable (Notice that -2 means people like uncrowded places). In figure 2, we also show the distribution of this ranking using the mean and the variance on table 2 and assuming a normal distribution for it (under the title psychometric

[7] Same results are obtained with normal draws, so they are not presented in the analysis.

information).

There is an inconsistency between what people mentioned in the survey about their preferences for congestion and what we infer from the estimation of the *ML* model. There are several ways to deal with this inconsistency. First, we could ignore the information from table 2 arguing people do not answer this type of questions carefully, and what really matters is what they do and not what they say. This is an acceptable strategy only if we have a clear notion about how the lack of an experimental design affects the possibility of recovering the true taste parameters. As mentioned above, in revealed preference studies with observational data, individuals do not face all possible combinations of alternatives and attributes. Therefore it may not be possible to recover the unconstrained taste distribution. The integrated *ML* model takes this into account by incorporating the taste indicators in the estimation of the model.

We use six indicators. All of them have a five-level response ranging from very undesirable to very desirable, or from definitively no to definitively yes. Two of them are directly related to crowding. For instance, the variable $NOCROWD$ represents the answer about how desirable is "A site with few other fisherman around." The variable $CHCROWD$ shows how well the statement "We usually go out of our way to avoid sites crowded with other fishermen" fits with an individual. Two indicators are indirectly related to crowding since they describe attributes that we expect to be present together with lower levels of crowding. These variables are $BEAUTY$ and $WILD$, describing a site with exceptional beauty and a wilderness area, respectively. Finally, there are two indicators which are related to the size of the fish caught, $Trophy$ describes the answer to the statement "a good chance to catch trophy-sized fish" and $Limit$ describes the statement "a good chance to catch your limit."

These indicators are incorporated in the equations $I_s = \alpha_s \beta^*_{cr} + \nu_s = \alpha_s (b + \eta) + \nu_s$ in the particular case presented in table 1. For identification purposes the standard deviation of one of these equations has been fixed to one. The main qualitative result in this model formulation is 99.9% of the population dislikes crowding.

Figure 2, also includes the plot of the distribution of the integrated model. These new estimates correctly indicate people do not like crowded places. Additionally, actual crowding does not have a large negative impact, compared to what respondents stated in the psychometric questions. Furthermore the variance is smaller than in the *ML* estimations suggesting not a lot of variation in the way people feel about crowding.

We also estimate the same models using a broader definition of the parameter for crowding, $\beta^*_{cr} = f(D_h; \omega) + \eta$, where we add several individuals' characteristics as explanatory variables: size of the household, age, gender, education, number of years people have been fishing in Alaska, a dummy variable taking value 1 if people know many fishing places, number of years as a resident in Alaska and income. The utility function has the same variables defined before. Results are presented in table 3. The mean value of the crowding coefficient depends on the demographic information when we use interactions between demographic variables and crowding levels in the estimation. We estimate this mean as $E(\tilde{\beta}^*_{cr}) = \hat{\omega}'\bar{X}$ and its standard deviation by $SD(\tilde{\beta}^*_{cr}) = \sqrt{var(\eta)} = \hat{\sigma}_\eta$. Using this procedure we find a mean value of -0.017 in the fixed coefficient estimation, of 0.338 for the *ML* model and -0.35 for the integrated model, all of them are very close to the value in the models without interactions. The standard errors for these coefficients are estimated using a delta method and are 0.0444, 0.1013 and 0.029, respectively. In this estimation, all of the coefficients of the choice model and the indicator equations are significant, while only the size of the household, age, age squared, and the number of years fishing in Alaska are significant in the equation for β^*_{cr}.

The estimated models have significant implications for the distribution of preferences for crowding and for the estimation of welfare measures. In the fixed parameter case without demographics, the coefficient is the same for all the individuals in the sample. When demographics are considered, each individual has his own crowding coefficient but it reflects only observed heterogeneity. For the random parameters cases, we calculate the mean value of the distribution for each individual using Revelt and Train's approach. In other words we calculate the mean $\bar{\beta}$ of the distribution $h(\beta\,|i, x, \theta)$. Figure 3 presents three histograms of the predicted individuals' coefficients calculated using the regressions with demographic information. The fixed parameter model and the *ML* model predict some portion of the sample with a positive coefficient for congestion. The fixed parameter model predicts around 53% of the population will have a positive value for this coefficient while the *ML* model predicts about 64% of the population with a positive value. In contrast, the integrated model predicts only a negative coefficient.

For any given individual there are two possible types of outcomes. First, it could be the case the three models predict the same sign for the mean of the coefficient (which must be negative since the integrated model only predicts a negative sign) but they differ only in the magnitude of the coefficients. Second, the models predict different signs for the mean of the coefficient. However, a comparison of values could be misleading because the coefficients in different conditional logit models could have different scale parameters.[8] Additionally it could also be argued that the relevant comparison is whether or not a specific individual has the same predicted position in the distribution.

In order to perform a meaningful comparison, we identify the number of individuals that have the same position in the distribution. A simple approach is to identify whether or not the individual has exactly the same position in the distribution. With this very strict criterion we found only 1 person lies in the same position in the distributions of the fixed model and the integrated model, none of them lies in the same position between the integrated model and the *ML* model, and only two people lie in the same position between the fixed model and the *ML* model. Using a less strict criterion we check for the number of people lying in the same quintile. In this case 22% of the population belongs to the same quintile between

[8] We discuss this specific topic in detail in the next section using a generalization of the *ML* model.

the fixed and the integrated model, 24% between the *ML* and the integrated model and 35% between the *ML* and the fixed model.

The effect of these differences in welfare measures depends on the possible sorting effect after a change in the number of sites or the quality of them. There is a direct effect and an indirect effect of closing a site. First, closing a site that is highly crowded will generate a higher economic loss for models that assume congestion is a desirable attribute of the site. Second, in the presence of sorting effects the closure of a site will increase the level of congestion in other sites and this will decrease the total loss if congestion is desirable.

To address this issue we perform the following exercise. First we estimate a welfare measure associated with closing a site that is generally highly crowded, namely the lower Kenai River from Cook Inlet to Soldotna Bridge (a prime fishing site for king salmon), and we assume there is no effect on crowding at any other sites. Second we assume people move to the next closest site, which is the rest of Kenai River, and this raises the level of congestion at that site by 1 unit for each person who had visited the lower Kenai River. Third, we assume the sorting affects all sites, increasing the crowding level in one unit for those sites that are not already at the highest level of congestion.[9]

Table 4 shows the welfare measures for each of these scenarios and for the three estimations that include demographic variables. In scenario 1, Both *ML* and the fixed coefficient estimation overstate the welfare measure in comparison with the integrated model. Basically, these two models imply the site is more desirable from an individual's perspective because it is crowded. Closing it implies a higher loss than if the site were not crowded. The second scenario includes a sorting effect but only to the closest site, and it also allows the crowding level to go beyond its original scale. Now the *ML* and fixed coefficient estimation underestimate the welfare loss since an increase in the crowding level of the other place is a benefit for some individuals, which goes in the opposite direction of the initial effect. In

[9] For example, if a site has a level 2 for congestion then it will remain at the same level, but if the site is at level 0 then it will increase to 1.

our third scenario, we take into account that the original crowding level cannot be higher than 2, therefore we increase the crowding level only when it is less than 2, but in this case we assume all the sites are affected by the sorting effect. Since all sites are affected by the policy, the integrated model provides the highest cost of the event, since the direct and indirect effect go in the same direction. The difference between the fixed coefficient model and the *ML* model are explained by the relative number of people being affected either in a positive or negative way by the change. The fixed model predicts a lower number of people with a positive coefficient; therefore the indirect effect can also be detrimental to people's welfare.

4.1 Robustness Analysis

In this section we consider some additional factors that bear on the robustness of the results presented above. We show the following results. First, the traditional approach which incorporates the taste indicators as explanatory variables in the definition of beta together with the demographic variables does not change the qualitative results of the fixed parameter or *ML* model. Second, we provide estimations with other distributions for the coefficients including lognormal and multivariate normal distributions, where we capture dependency patterns in the *ML* model. We discuss the effect of these changes on the estimated distribution of the coefficient associated with crowding. Third, we continue our analysis with a discussion of the possible effect of scale heterogeneity on the estimation of the *ML* and integrated model. In this subsection we address the possibility our results were driven by heterogeneity in the variance of the error terms of the choice model. Finally, we test endogeneity of our measure of crowding, discuss the inclusion of a full set of site specific constants to account for unobservable attributes of the alternatives and compare our results to current solution suggested in the literature for this problem.

4.1.1 Taste indicators in the definition of beta

A simple way to use the information in the taste indicators would be to incorporate them as explanatory variables in the definition of beta together with the demographic variables, i.e. defining beta as $\beta_n = b + \gamma D_n + \rho H_n + \eta_n$. This is similar to the approach followed by Harris and Keane (1999). They have attitudinal data for both unobserved and observed variables and they use the indicators in the definition of the coefficients for the observed attributes of the good and the latent variables as well. In our setting we are dealing only with observed attributes, therefore, we include the indicators only in the definition of the beta associated with the observed measure of congestion. These results are presented in table 5. The inclusion of the taste indicators in the definition of beta does not change the sign of the crowding coefficient. The mean values, evaluated at the mean of the explanatory variables, are 0.16 and 0.77 for the fixed and *ML* model, respectively.

4.1.2 Lognormal and dependent distributions

Following the prior beliefs about congestion, prices and quality, we could specify the model with distributions that allow only positive or negative values for the coefficients. Lognormal distributions are the way to do this. Additionally we could use a simulation process that takes into account the possible correlation among the distributions of the parameters. We performed these estimations using the following alternatives: independent normal distributions for price and quality and lognormal for crowding, multivariate normal distributions for the random coefficients, independent log normal distributions for the three random parameters, dependent normal distributions for price and quality and lognormal distribution for crowding, among other combinations.

The estimates using a lognormal distribution only for the crowding coefficient and independent normal distribution for price and quality is presented in table 6. We also compare the predicted lognormal distribution of the crowding coefficient and the distribution derived from the integrated model in figure 4. The lognormal distribution is almost a spike at zero showing the use of this distribution does not provide any useful information about taste heterogeneity for congestion. The mean of the distribution is -0.398, which is very close to the mean in the integrated model and its variance is 322.05, with a standard deviation of 17.946 which is significantly larger than the variance of all other models. The mode is -4.340×10^{-6}, virtually zero and the median is -8.8265×10^{-3}. Therefore, most of the people in the sample do not care about crowding conditions. Estimation of a model with a lognormal distribution for all the random parameters did not converge, which is common in applications of *ML* models.

On the other hand, the inclusion of dependency among distributions of the random parameters does not solve the sign problem in the crowding coefficient nor does it change the results in the integrated model. We estimate a *ML* (see table 7) with dependent distributions and the results are similar to those found with the independent distributions. The coefficients s12, s13 and s23 are the corresponding covariance between the distributions. We tried several other combinations of normal and lognormal distributions with dependent error structures but none of them changed the sign of the crowding coefficient.

4.1.3 Generalized Mixed Logit Model

Heterogeneity can also be associated with the scale component of the distribution function. As Ben-Akiva and Lerman (1985) and Louviere (2001) show, when the ε_{ijt}'s are independently and identically extreme value type I random variables then the variance of the random component is inversely related with the scale parameter. This scale parameter is fixed to 1 in traditional logit models since it cannot be identified separately from the parameters of the vector β. The cumulative distribution of ε_{ijt} is $F(\varepsilon_{ijt}) = \exp\left[-\exp\left(-\lambda\varepsilon_{ijt}\right)\right]$ for $\lambda > 0$, with λ the scale parameter of the distribution, and $Var(\varepsilon_{ijt}) = \sigma_{\varepsilon}^2 = \pi^6/6\lambda^2$ (see Louviere et al., 2000, page 142). Making the scale parameter explicit in the model, the probability of individual n choosing alternative j in period t is

$$L_{njt} = \frac{e^{\lambda \beta'_n x_{njt}}}{\sum_i e^{\lambda \beta'_n x_{nit}}}.$$

One way to identify the scale parameter is to define λ as a function of observable characteristics of the alternatives, the context of the choice experiment or individuals' characteristics (Dellaert et al., 1999). Another way is to combine different sources of information such as different data sets or stated and revealed preferences.

As part of the specification of the model, the scale parameter could also capture part of the heterogeneity, i.e. it could be an individual specific parameter, that is $\lambda = \lambda_n$ (Swait and Louviere, 1993; Louviere, 2001; Louviere et al., 2002). Recently, Fiebig et al. (2007) have suggested a generalized mixed logit (GML) model which captures both sources of heterogeneity and includes several possible combinations of scale and mean heterogeneity. In this case β_n is defined as

$$\beta_n = \lambda_n \beta + \gamma \eta_n + (1 - \gamma) \lambda_n \eta_n, \ \gamma \in (0, 1).$$

The new definition of the parameter beta encompasses several particular cases of heterogeneity. If $\lambda_n = 1$ for all n then we have the traditional ML model with $\beta_n = \beta + \eta_n$. For all other values of λ_n, if $\gamma = 0$ then $\beta_n = \lambda_n (\beta + \eta_n)$, in this case both the mean effect and the individual's deviation with respect to the mean are affected by the scale heterogeneity. If $\gamma = 1$,then $\beta_n = \lambda_n \beta + \eta_n$, and only the mean effect is affected by the scale. Finally if $\gamma \in (0, 1)$, the portion $(1 - \gamma)$ of the deviation with respect to the mean is affected by the scale heterogeneity, while the portion γ is not. With this definition of beta the probability of individual n choosing alternative j in period t is

$$L_{njt} = \frac{e^{[\lambda_n \beta + \gamma \eta_n + (1-\gamma) \lambda_n \eta_n]' x_{njt}}}{\sum_i e^{[\lambda_n \beta + \gamma \eta_n + (1-\gamma) \lambda_n \eta_n]' x_{nit}}}.$$

Estimation of this model requires the definition of a distribution for λ_n, in this case a lognormal distribution seems appropriate since the scale parameter is positive. Additionally it requires imposing some restrictions on the distribution for identification purposes. We follow Fiebig et al. and define $\lambda_n = \exp(\bar{\lambda} + \tau\eta_\lambda)$, with $\eta_\lambda \sim N(0,1)$ and $\bar{\lambda} = -\tau^2/2$ in order to impose that $E(\lambda_n) = 1$.

The inclusion of λ_n in the model only allows one to decompose heterogeneity between mean and scale heterogeneity but it does not help one to identify the correct sign for the coefficients since λ_n can only be positive. In other words, the inclusion of λ_n affects the absolute values of the betas but not their sign. In the integrated model we use the following definition

$$\beta_n^* = h(D_n, \varpi) + \eta = \lambda_n\beta + \gamma\eta_n + (1-\gamma)\,\lambda_n\eta_n + \omega D_n. \tag{9}$$

Table 8 shows the results of this generalized model. Below the names of the models we include the coefficient λ to show we are estimating the GML and the integrated GML, which contains the psychometric information. Estimation of the last model requires an additional identification restriction that consists in fixing the gamma coefficient; this value was fixed to the value obtained in the simple GML. In the *GML* model we impose the condition that $\gamma \in (0,1)$ using $\gamma = \exp(\gamma^*)/(1+\exp(\gamma^*))$, and the variance of γ is calculated using the delta method, i.e., $var(\gamma) = [\gamma\,(1-\gamma)]^2\,var(\gamma^*)$. The *GML* model contributes to enhance our understanding of heterogeneity since both τ and γ are significant, suggesting there is scale heterogeneity as well as a mean heterogeneity and that the scale coefficient affects both the beta coefficient and the random component of β_n. Notice that even though the coefficients of the *GML* models are significant they do not significantly change neither the crowding coefficient nor its standard deviation.

4.1.4 Endogeneity of Crowding Variables

Perhaps the main problem we have to address is the possible endogeneity of our crowding index. While there are several sources of endogeneity in a discrete choice model (Louviere et al., 2005), in this section we focus on the case of endogenous explanatory variables, i.e. situations in which the explanatory variable is correlated with the unobserved component of the utility function. A plausible explanation of the positive sign of congestion in previous papers is the endogeneity of this variable, which could be correlated with omitted relevant characteristics of the sites. According to Jakus and Shaw (1997) the degree of endogeneity is related to how congestion is measured in empirical work. Jakus and Shaw classified measures of congestion in four categories; actual, expected, anticipated and perceived congestion. For them, actual measures of congestion, measured independently of the consumer by an outside member of the sample, could be considered exogenous. Our measure of congestion falls in this category, therefore it could be considered exogenous. Nevertheless, any measure of congestion will at least partially reflect people's decisions about which site to visit. In other words, even if a measure of congestion is obtained from outside the sample and is independent of consumer choices in the sample, it could still be correlated with some components of the unobserved variables of the location that make it endogenous. In our previous analysis we assume congestion is exogenous, in this section we test this hypothesis in order to give some additional support to our prior belief.

There are two ways to deal with endogeneity in discrete choice model. First, the Berry, Levinsohn and Pakes (1995) approach that uses aggregate market information and a linearization of the discrete choice model suitable for traditional instrumental variable techniques. This approach is used by Timmins and Murdock (2007). The second approach was proposed by Villas-Boas and Winer (1999). They estimate a regression equation of the endogenous variable against some instruments together with the choice probability equation.

Timmins and Murdock use the expected share of anglers who choose each site as a measure of congestion in the recreational sites of the sample. The distinctive characteristic of this application is the large number of sites from which people can choose. The endogeneity of the congestion variable together with the high correlation of congestion with other explanatory variables makes it quite challenging to correctly identify the sign of congestion. They use a two step estimation procedure following a similar method used in the industrial organization literature (Nevo, 2000). In the first step one estimate mean utilities together with interactions between site characteristics and individual attributes; in the second step, the mean utilities are used to estimate a linear instrumental variable (IV) regression. However, the large proportion of site shares that are zero requires the use of a quantile IV regression. A negative sign for the congestion parameter is obtained only with the use of this quantile estimation.

The fact that the second stage is a linear IV regression model whose properties depend on the size of the choice set and the large proportion of zero market shares makes this procedure unsuitable for discrete choices where there is a limited number of choice alternatives (See O'Hara 2006 for an example of this problem). If there is only a small number of alternatives in the choice set, the second stage estimation may not be an appropriate solution.

Therefore, we test endogeneity following Villas-Boas and Winer's (1999) suggestion for models with micro data, which uses instrumental variables to test the endogeneity of the variable. We use two instruments in our estimation, one of them is motivated by Villas-Boas and Winer and the other follows Timmins and Murdock. These instruments are the lag of the crowding levels and a function of the exogenous attributes of other alternatives, in this case we use the sum of the quality levels excluding the own level of quality.

Villas-Boas and Winer do not assume a random parameter model, instead they assume the utility model has an additional error component capturing unobserved sites characteristics. In this sense their model is similar to the models suggested by Murdock and Timmins (2007) and Murdock (2006).

Extending these models to a random parameter approach only implies an additional layer of integration, the utility function is

$$U_{ijt} = X'_{ijt}\beta_n + \xi_{jt} + \varepsilon_{ijt},$$

where there is an additional component, ξ_{jt}, which they assume is distributed normal with mean zero and variance σ_{ξ}^2. This is exactly the same error structure used by Timmins and Murdock which is employed in a second step linear IV regression to estimate the coefficient of variables varying only across alternatives. In contrast Villas-Boas and Winer use a more structural approach defining a second equation for the potentially endogenous variable, in this case crowding rate

$$CR_{jt} = CR(W_{ij}; \alpha) + \nu_{jt},$$

with $\nu_{jt} \sim N\left(0, \sigma_{\nu}^2\right)$, and W_{ij} are instrumental variables. The correlation between this error term and the error term ξ_{jt} captures the possible endogeneity of the variable, that is $E\left(\nu_{jt}\xi_{jt}\right) = \rho\sigma_{\nu}\sigma_{\xi}$, additionally it is assumed that $E\left(\nu_{j't'}\nu_{jt}\right) = 0$ and $E\left(\nu_{j't'}\xi_{jt}\right) = 0$. Under these assumptions the conditional probability of choosing a sequence of alternatives is given by

$$P\left(y_{n\mathbf{j}}\left|\xi_{jt}, \nu_{jt}, \eta_n\right.\right) = \prod_t^T \prod_j^J \left(\frac{e^{\theta\left(X'_{ij_{it}t}\beta_n + \xi_{jt}\right)}}{\sum_j e^{\theta\left(X'_{ijt}\beta_n + \xi_{jt}\right)}}\right)^{y_{njt}}$$

We still have a random parameter in the definition $\beta_n = b + \eta_n$, the joint distribution of $y_{n\mathbf{j}}$ and CR_{jt} would be

$$P\left(y_{n\mathbf{j}}, CR_{jt}\right) = \int_{\xi_{jt}} \int_{\beta} \prod_t^T \prod_j^J \left(\frac{e^{\theta\left(X'_{ij_{it}t}\beta_n + \xi_{jt}\right)}}{\sum_j e^{\theta\left(X'_{ijt}\beta_n + \xi_{jt}\right)}}\right)^{y_{njt}} f\left(\nu_{jt}\right) f\left(\xi_{jt}\left|\nu_{jt}\right.\right) g(\beta\left|\theta\right.) d\beta d\nu_{jt} d\xi_{jt}$$

This integral is calculated by simulation bearing in mind the error components ξ_{jt} and $f\left(\nu_{jt}\right)$ are the same for each individual but vary over periods of time. The joint distribution of $f\left(\xi_{jt}, \nu_{jt}\right)$ is

$$f\left(\xi_{jt}, \nu_{jt}\right) = \left(\frac{1}{\sigma_{\nu_{jt}}\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\nu_{jt} - \bar{\nu}}{\sigma_{\nu_{jt}}}\right)^2\right\}\right) \cdot \left(\frac{1}{\sigma_{\xi_{jt}}\sqrt{(1-\rho^2)}\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2_{\xi_{jt}}(1-\rho^2)}\left(\xi_{jt} - \bar{\xi} - \rho\frac{\sigma_{\xi_{jt}}}{\sigma_{\nu_{jt}}}(\nu_{jt} - \bar{\nu})\right)^2\right\}\right)$$

and since $\bar{\nu} = 0$

$$f\left(\nu_{jt}\right) = \left(\frac{1}{\sigma_{\nu_{jt}}\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\nu_{jt}}{\sigma_{\nu_{jt}}}\right)^2\right\}\right) = \frac{1}{\sigma_{\nu_{jt}}}\phi\left(\frac{CR_{jt} - \alpha_{0j} - \alpha_{1j}W_{jt}}{\sigma_{\nu_{jt}}}\right)$$

and given $\bar{\xi} = 0$ we can deduce that

$$\left(\xi_{jt} \,|\nu_{jt}\right) \sim N\left(\rho\frac{\sigma_{\xi_{jt}}}{\sigma_{\nu_{jt}}}\nu_{jt}, \sigma^2_{\xi_{jt}}\left(1-\rho^2\right)\right)$$

Table 9 presents the results of this model where the instruments are denoted by *lagcrowding* and *quality(-1)* which represent the lag of crowding and the sum of the quality index excluding the own quality. The variable sigma 1 and sigma 2 represent the variance $\sigma_{\xi_{jt}}$ and $\sigma_{\nu_{jt}}$, respectively. From the fact that the rho coefficient is not significant we can reject the hypothesis of endogeneity of the crowding variable.

Another way to capture unobserved attributes of the site is to estimate a full set of site specific constants. We estimate the *ML* and the integrated model with this set of site specific constant, i.e., with a set of dummy variables for each alternative as in Nevo (2000) and Murdock (2006). Since in the data all the variables except the previous harvest (HAR) vary at least according to two out of the three elements – time, people and sites – we are able to identify these dummy variables only excluding the harvest variable. The coefficient for this variable can be recovered afterwards by regressing the specific constant against the log of harvest, as was suggested by Murdock. Including the site specific constant controls for any unobserved site effect and reduces the probability of having correlation between the observed variable and unobserved attributes. Estimation with dummy variables for each site did not affect the conclusions of the model.

5 Conclusions

We show a simple way to use taste indicators in the traditional *ML* model, which contributes to describe heterogeneity of preferences over attributes of the goods. This methodology is especially useful for those attributes that could affect individuals' utility in a positive and negative direction. In the particular case of congestion, the traditional *ML* model reports results inconsistent with the way people feel about this site's attribute. The methodology presented in this paper brings together two different pieces of information; on one hand the observed behavior and, on the other, the expressed attitude toward the attributes of interest. Our results suggest researchers could obtain a better representation of preferences if they combine these two sources of information.

References

[1] Anderson , F. And N. Bonsor (1974). "Allocation, Congestion and the Valuation of Recreational Resources" *Land Economics* 50:51-57.

[2] Bell, F. and V. Leeworthy (1990). "Recreational Demand by Tourist for Saltwater Beach Days" *Journal of Environmental Economics and Management* 18:189-205.

[3] Ben-Akiva, M. and S. Lerman (1985). Discrete Choice Analysis: Theory and Application to Travel Demand, Cambridge Mass: MIT Press.

[4] Ben-Akiva, M., D. McFadden, T. Gärling, D. Gopinath, J. Walker, D. Bolduc, A. Börsch-Supan, P. Delquié, O. Larichev, T. Morikawa, A. Polydoropoulou and V. Rao (1999a). "Extended Framework for Modeling Choice Behavior" *Marketing Letters* 10(3):187-203.

[5] Ben-Akiva, M., J. Walker, A. Bernardino, D. Gopinath, T. Morikawa and A. Polydoropoulou (1999b). "Integration of Choice and Latent Variable Models" Working Paper, MIT.

[6] Ben-Akiva, M., D. Bolduc and J. Walker (2001). "Specification, Identification, & Estimation of Logit Kernel (or Continuous Mixed Logit) Model" Working Paper, MIT.

[7] Ben-Akiva, M., D. McFadden, K. Train, J. Walker, C. Bhat, M. Bierlaire, D. Bolduc, A. Boersch-Supan, D. Brownstone, D. Bunch, A. Daly, A. de Palma, D. Gopinath, A. Karlstrom and M. Munizaga (2002). "Hybrid Choice Models: Progress and Challenges" *Marketing Letters* 13(3):163-175.

[8] Berrens, R., O. Bergland and R. Adams (1993). "Valuation Issues in an Urban Recreational Fishery: Spring Chinook Salmon in Portland, Oregon" *Journal of Leisure Research* 25(1):70-83.

[9] Berry, S. (1994). "Estimating Discrete-Choice Models of Product Differentiation" *Rand Journal of Economics* 25(2):242-262.

[10] Berry, S., J. Levinsohn and A. Pakes (1995). "Automobile Prices in Market Equilibrium" *Econometrica* 63(4):841-890.

[11] Berry, S., J. Levinsohn and A. Pakes (1998). "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market" NBER Working Papers Series No. 6481 Cambridge, MA.

[12] Boxall, P. , K. Rollins, K. and J. Englind (2003). "Heterogeneous Preferences for Congestion during a Wilderness Experience" *Resource and Energy Economics* 25:177-195.

[13] Cesario, F. (1980). "Congestion and the Valuation of Recreation Benefits" *Land Economics* 56(3):331-338.

[14] Cicchetti, C. and V. Smith (1973). "Congestion, Quality Deteriorations, and Optimal Use: Wilderness Recreation in the Spanish Peaks Primitiva Area" *Social Science Research* 2:15-30.

[15] Dellaert, B., J. Brazzel and J. Louviere (1999). "The Effect of Attribute Variation on Consumer Choice Consistency" *Marketing Letters* 10(2):139-147.

[16] Deyak, T. and V. Smith (1978). "Congestion and Participation in Outdoor Recreation: A Household Production Function Approach" *Journal of Environmental Economics and Management* 5:63-80.

[17] Fiebig, D., M. Keane, J. Louviere and N. Wasi (2007). "The Generalized Multinomial Logit Model" Working paper University of Technology, Sydney.

[18] Freeman, M. and R. Havemann (1977). "Congestion, Quality Deterioration, and Heterogeneous Preferences" *Journal of Public Economics* 8:225-232.

[19] Harris, K. and M. Keane (1999). "A model of Health Plan Choice: Inferring Preferences and Perceptions from a Combination of Revealed Preference and Attitudinal Data" *Journal of Econometrics* 89:131-157.

[20] Jakus, P. and D. Shaw (1997). "Congestion at Recreation Areas: Empirical Evidence on Perceptions, Mitigating Behavior and Management Preferences" *Journal of Environmental Management* 50:286-401.

[21] Louviere, J. (2001). "What if Consumer Experiments Impact Variances as well as Means? Response Variability as a Behavioral Phenomenon" *Journal of Consumer Research* 28:506-511.

[22] Louviere, J., D. Hensher and J. Swait (2000). *State Choice Models; Analysis and Application* Cambridge University Press.

[23] Louviere, J., R. Carson, A. Ainslie, J. Deshazo, T. Cameron, R. Kohn and T. Marley (2002). "Dissecting the Random Component of Utility" *Marketing Letters* 13:3,177-193.

[24] Louviere, J., K. Train, M. Ben-Akiva, C. Bhat, D. Brownstone, T. Cameron, R. Carson, J, Deshazo, D. Fiebig, W. Green, D. Hensher and D. Waldman (2005). "Recent Progress on Endogeneity in Choice Modeling" *Marketing Letters* 16(3/4): 255-265.

[25] McConnell, K. (1977). "Congestion and Willingness to Pay: A Study of Beach Use" *Land Economics* 53(2):185-195.

[26] McConnell, K.(1988). "Heterogeneous Preferences for Congestion" *Journal of Environmental Economics and Management* 15:251-258.

[27] Murdock, Jennifer (2006). "Handling Unobserved Site Characteristics in Random Utility Models of Recreation Demand" *Journal of Environmental Economics and Management* 51:1-25

[28] Nevo, A. (2000). "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand" *Journal of Economics and Management Strategy* 9:513-48.

[29] O'Hara, M. (2006). "Nobody Goes There Anymore....It's Too Crowded" Working paper, Binghamton University, Department of Economics.

[30] Revel, D. and K. Train (1998). "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level" *The Review of Economics and Statistics* 80(4):647-657.

[31] Schuhmann, P. and K. Schwabe (2004). "An Analysis of Congestion Measures and Heterogeneous Angler Preferences in a Random Utility Model of Recreational Fishing" *Environmental and Resource Economics* 27:429-450.

[32] Shelby, B. (1980). "Crowding Models for Backcountry Recreation" *Land Economics* 56(1):43-55.

[33] Smith, K. (1981). "Congestion, Travel Cost Recreational Demand Models, and Benefit Evaluation" *Journal of Environmental Economics and Management* 8:92-96.

[34] Swait, J. and J. Louviere (1993). "The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models" *Journal of Marketing Research* 30: 305-314.

[35] Timmins, C. and J. Murdock (2007). "A Revealed Preference Approach to the Measurement of Congestion in Travel Cost Models" *Journal of Environmental Economics and Management* 53(2):230-249

[36] Train, K., D. McFadden and A. Goett (1987a). "Consumer Attitudes and Voluntary Rate Schedules for Public Utilities" *The Review of Economics and Statistics 69(3):383-391.*

[37] Train, K., D. McFadden and M. Ben-Akiva (1987b). "The Demand for Local Telephone Service: A Fully Discrete Model of Residential Calling Patterns and Service Choices" *Rand Journal of Economics 18(1):109-123.*

[38] Train, K. (1998). "Recreation Demand Models with Taste Differences Over People" *Land Economics* 74(2):230-39.

[39] Train K. (2003). *Discrete Choice Model with Simulation* Cambridge University Press.

[40] Vásquez-Lavín, F. and M. Hanemann (2007). "Utility Consistent Frameworks to Model Corner Solutions with Large Demand Systems" working paper, ARE University of California at Berkeley.

[41] Villas-Boas J. and R. Winer (1999). "Endogeneity in Brand Choice Models" Management Science 45(10):1324-1338.

[42] Von Haefen, R., D. Phaneuf and G. Parson (2004). "Estimation and Welfare Analysis with Large Demand System" *Journal of Business & Economic Statistics 22(2):194-205.*

[43] Walker, J. (2002). "The Mixed Logit (or Logit Kernel) Model: Dispelling Misconceptions of Identification" *Transportation Research Record* 1805:86-98.

[44] Walker, J. and M. Ben-Akiva (2002). "Generalized Random Utility Model" *Mathematical Social Sciences* 43:303-343.

[45] Wedel, M., W. Kamakura, N. Arora, A.Bemmaor, J. Chiang, T. Elrod, R. Johnson, P. Lenk, S. Neslin and C. Poulsen (1999). "Discrete and Continuous Representations of Unobserved Heterogeneity in Choice Modelling" *Marketing Letters* 10:3,219:232.

Table 1. Discrete Choice Models

	(1) Fixed	(2) Mixed Logit		(3) Integrated ML	
	b	b, η		b, η, l	
explanatory variable	**coeffic.**	**coeffic.**	**s.d.**	**coeffic.**	**s.d.**
Log(TC)	**-0.91**	***-1.13***	***0.87***	***-1.18***	***0.76***
t-value	-27.86	-16.29	12.06	-20.51	14.50
Quality	**0.24**	***0.26***	***0.19***	***0.31***	***0.25***
t-value	12.76	10.27	4.51	11.50	7.79
Crowding	-0.03	***0.36***	***1.31***	***-0.35***	***0.12***
t-value	-0.70	3.64	11.58	-12.22	10.12
Cabin	**2.13**	***2.61***	-	***2.41***	-
t-value	15.58	14.25	-	22.89	-
Log(harvest)	**0.38**	***0.43***	-	***0.39***	-
t-value	12.61	11.76	-	24.82	-
Indicators					
Nocrowd	-	-	-	*1.00*	***1.03***
t-value	-	-	-	-	18.88
Chose No crowded	-	-	-	***3.88***	***0.60***
t-value	-	-	-	11.36	27.78
Wild	-	-	-	***2.65***	***0.78***
t-value	-	-	-	10.39	26.28
Beauty	-	-	-	***3.02***	**10.90**
t-value	-	-	-	*0.69*	26.38
Trophy	-	-	-	***2.10***	***0.88***
t-value	-	-	-	9.79	26.22
Limit	-	-	-	***3.40***	***0.75***
t-value	-	-	-	11.18	33.16

Bold numbers are significant at a 5 % level.

Simulation assuming Independent multivariate normal distributions

Table 2. A site with few other fishermen around

	Value	frequency	%
very desirable	-2	215	49.2
desirable	-1	180	41.19
no opinion	0	34	7.78
undesirable	1	6	1.37
very undesirable	2	2	0.46
mean	-1.373	variance	0.526871

Table 3. Discrete Choice Models: Demographics

	(1) Fixed b, D_h	(2) Mixed Logit b, D_h, η		(3) ML Integrated b, D_h, η, I	
explanatory variable	**coeffic.**	**coeffic.**	**s.d.**	**coeffic.**	**s.d.**
Log(TC)	***-0.92***	***-1.12***	***0.87***	***-1.18***	***0.77***
t-value	-27.95	-16.88	13.23	-20.23	14.15
Quality	***0.24***	***0.26***	***0.19***	***0.31***	***0.26***
t-value	12.68	10.71	5.16	11.33	7.68
Crowding	***-1.33***	*0.10*	***1.32***	***-0.69***	***0.11***
t-value	-2.26	0.08	12.22	-5.82	9.84
Cabin	***2.15***	***2.62***	-	***2.42***	-
t-value	15.60	20.02	-	22.68	-
Log(harvest)	***0.39***	***0.44***	-	***0.39***	-
t-value	12.86	20.28	-	22.92	-
Parameters in definition of beta					
Size of household	*0.39*	*0.56*	-	***0.09***	-
t-value	1.17	0.74	-	1.70	-
Age	***43.52***	*14.02*	-	***9.95***	-
t-value	1.64	0.22	-	1.99	-
Age*age	***-0.65***	*-0.46*	-	***-0.09***	-
t-value	-2.16	-0.63	-	-1.64	-
Gender	*0.11*	*-0.15*	-	*0.03*	-
t-value	0.90	-0.53	-	1.24	-
Education	***0.15***	*0.33*	-	*0.00*	-
t-value	1.71	1.33	-	-0.26	-
Years fishing in Alaska	*0.59*	*0.14*	-	***0.18***	-
t-value	1.49	0.17	-	2.58	-
Know many fishing sites	*0.44*	*0.20*	-	*-0.01*	-
t-value	1.27	0.25	-	-0.12	-
Years as a resident	*-0.08*	*-0.18*	-	*0.01*	-
t-value	-1.22	-0.88	-	0.77	-
Income	*-0.01*	*0.25*	-	*0.01*	-
t-value	-0.06	0.91	-	0.22	-
Indicators					
Nocrowd	-	-	-	*1.00*	***1.03***
t-value	-	-	-	-	18.63
Chose No crowded	-	-	-	***3.84***	***0.59***
t-value	-	-	-	11.35	27.98
Wild	-	-	-	***2.63***	***0.77***
t-value	-	-	-	10.41	25.73
Beauty	-	-	-	***2.98***	***0.69***
t-value	-	-	-	10.94	26.27
Trophy	-	-	-	***2.07***	***0.88***
t-value	-	-	-	9.81	26.05
Limit	-	-	-	***3.35***	***0.76***
t-value	-	-	-	11.20	32.68

Bold numbers are significant at a 5 % level.

Table 4 . Welfare Losses after Closing Kenai River

	Scenario 1	Scenario 2	Scenario 3
Fixed	-64810.75	-86744.04	-91758.92
ML	-62592.22	-67228.04	-63267.80
Integrated	-59918.88	-90087.15	-112015.77

Scenario 1 : no sorting effects

Scenario 2: People move to the closest place

Scenario 3: People move to all other places

Table 5. Discrete Choice Models: demographics and indicators

	(1) Fixed	(2) Mixed Logit	
	b, D_h, H_h	b, D_h, H_h, η	
explanatory variable	**coeffic.**	**coeffic.**	**s.d.**
Log(TC)	**-0.91**	**-1.13**	**0.86**
t-value	-54.03	-16.77	12.98
Quality	**0.24**	**0.26**	**0.20**
t-value	13.32	10.40	5.51
Crowding	-0.43	**0.39**	**1.29**
t-value	-0.99	0.27	11.65
Cabin	**2.15**	**2.61**	-
t-value	23.05	19.99	-
Log(harvest)	**0.40**	**0.44**	-
t-value	23.27	20.08	-
Parameters in definition of beta			
Size of household	0.20	0.30	-
t-value	0.70	0.38	-
Age	13.99	13.99	-
t-value	0.74	0.22	-
Age*age	**-0.39**	-0.55	-
t-value	-1.84	-0.75	-
Gender	0.04	-0.22	-
t-value	0.41	-0.75	-
Education	**0.19**	0.38	-
t-value	2.37	1.49	-
Years fishing in Alaska	**0.68**	0.17	-
t-value	2.27	0.20	-
Know many fishing sites	**0.59**	0.25	-
t-value	2.24	0.31	-
Years as a resident	-0.07	-0.15	-
t-value	-1.04	-0.75	-
Income	0.00	0.25	-
t-value	0.05	0.87	-
Nocrowd	**0.25**	**0.30**	-
t-value	6.54	2.52	-
Chose No crowded	0.01	0.04	-
t-value	0.12	0.26	-
Wild	0.07	0.08	-
t-value	1.55	0.56	-
Beauty	-0.03	-0.14	-
t-value	-0.72	-1.06	-
Trophy	-0.02	-0.11	-
t-value	-0.68	-0.97	-
Limit	0.02	0.03	-
t-value	0.55	0.26	-

Bold numbers are significant at a 5 % level.

Table 6. ML LogNormal

	(1) Mixed Logit	
explanatory variable	**coeffic.**	**s.d.**
Log(TC)	**-1.14**	**0.80**
t-value	-19.77	13.07
Quality	**0.28**	**0.20**
t-value	11.61	5.57
Crowding	**-4.73**	**2.76**
t-value	-6.14	7.32
Cabin	**2.43**	-
t-value	15.00	-
Log(harvest)	**0.40**	-
t-value	11.73	-

Bold numbers are significant at a 5 % level.

Table 7. ML with Dependent normals

	(1) Mixed Logit	
explanatory variable	**coeffic.**	**s.d.**
Log(TC)	**-1.08**	**0.92**
t-value	-12.93	12.31
Quality	**0.26**	**0.20**
t-value	10.19	4.41
Crowding	**0.35**	**1.17**
t-value	3.53	7.44
Cabin	**2.59**	-
t-value	14.06	-
Log(harvest)	**0.44**	-
t-value	11.68	-
s12	0.01	-
t-value	0.44	-
s13	**0.46**	-
t-value	3.50	-
s23	0.25	-
t-value	0.89	-

Bold numbers are significant at a 5 % level.

Table 8. Discrete Choice Models: GML

	(1) GML λ, b, η		(2) Integrated GML λ, b, η, l	
explanatory variable	**coeffic.**	**s.d.**	**coeffic.**	**s.d.**
Log(TC)	**-1.33**	**0.94**	**-1.20**	**0.75**
t-value	-11.45	10.76	-21.61	14.25
Quality	**0.29**	**0.25**	**0.31**	**0.26**
t-value	9.11	4.69	11.17	8.02
Crowding	**0.40**	**1.46**	**-0.36**	**0.11**
t-value	4.14	10.48	-12.13	9.41
Cabin	**3.75**	-	**2.45**	-
t-value	9.71	-	20.29	-
Log(harvest)	**0.67**	-	**0.41**	-
t-value	10.66	-	23.61	-
Tau (τ)	**0.82**	-	**-0.14**	-
t-value	8.93	-	-3.65	-
Gamma (γ)	0.48	-	0.48	-
t-value	7.36	-	-	-
Indicators				
Nocrowd	-	-	1.00	**1.03**
t-value	-	-	-	18.79
Chose No crowded	-	-	**3.81**	**0.60**
t-value	-	-	11.35	28.21
Wild	-	-	**2.61**	**0.77**
t-value	-	-	10.40	26.34
Beauty	-	-	**2.98**	**0.68**
t-value	-	-	10.95	26.32
Trophy	-	-	**2.07**	**0.88**
t-value	-	-	9.78	26.19
Limit	-	-	**3.35**	**0.75**
t-value	-	-	11.22	32.82

Bold numbers are significant at a 5 % level.

Table 9. Discrete Choice Models: Testing Endogeneity

explanatory variable	**coeffic.**	**s.d.**
Log(TC)	**-1.17**	**0.94**
t-value	-16.93	12.46
Quality	**0.26**	**0.24**
t-value	9.32	6.21
Crowding	**0.51**	**1.46**
t-value	4.34	11.53
Cabin	**2.76**	
t-value	16.79	
Log(harvest)	**0.53**	
t-value	15.02	
alfa0	**0.21**	
t-value	18.87	
Lag of Crowding	**8.25**	
t-value	77.10	
Quality (-1)	**-0.11**	
t-value	-3.28	
Sigma 1	**0.50**	
t-value	5.34	
Sigma 2	**0.50**	
t-value	137.18	
rho	-0.10	
t-value	-0.97	

Bold numbers are significant at a 5 % level.

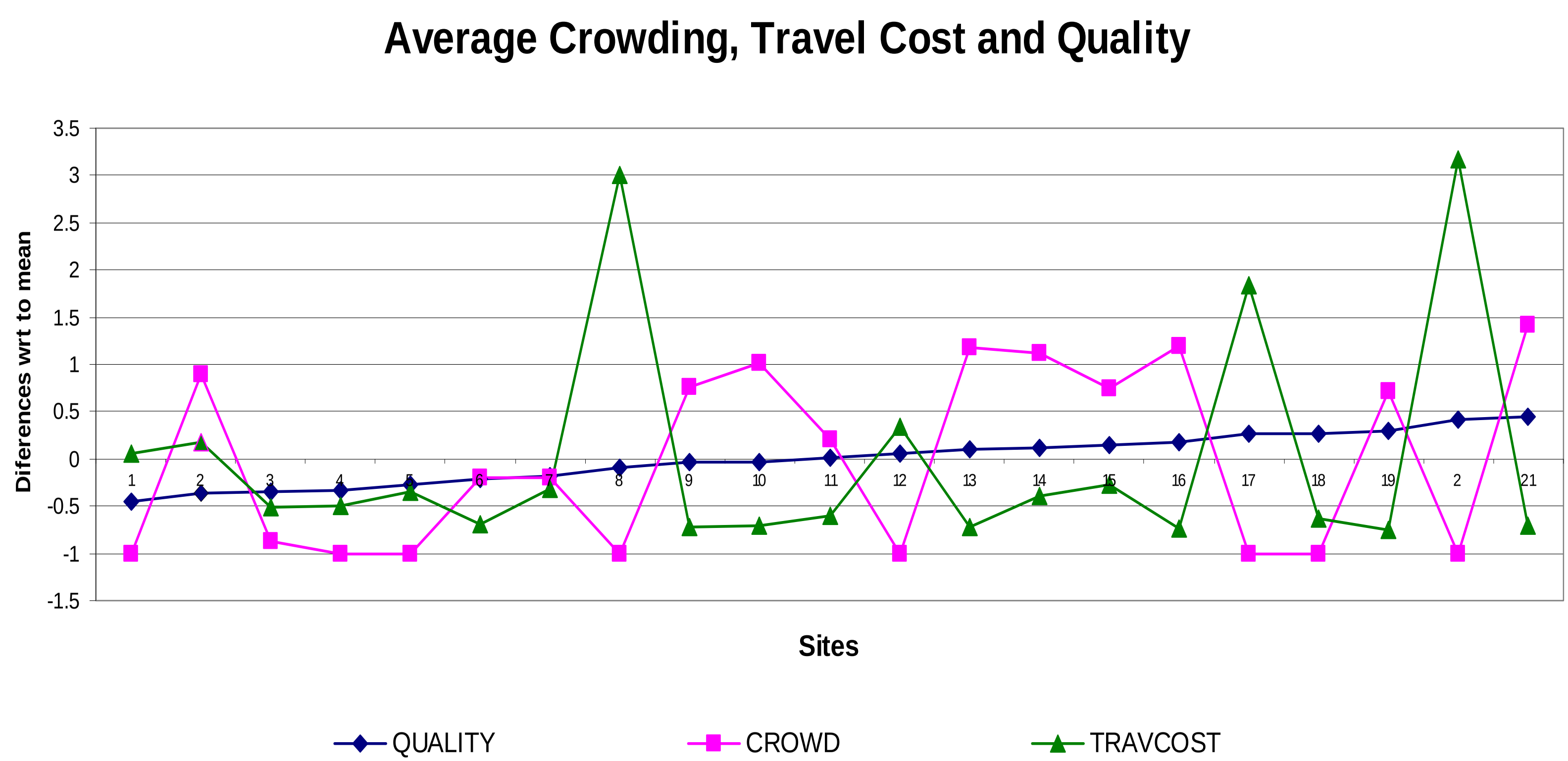
Average Crowding, Travel Cost and Quality
Diferences wrt to mean
3.5
3
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
2
21
Sites
QUALITY
CROWD
TRAVCOST

Figure 1

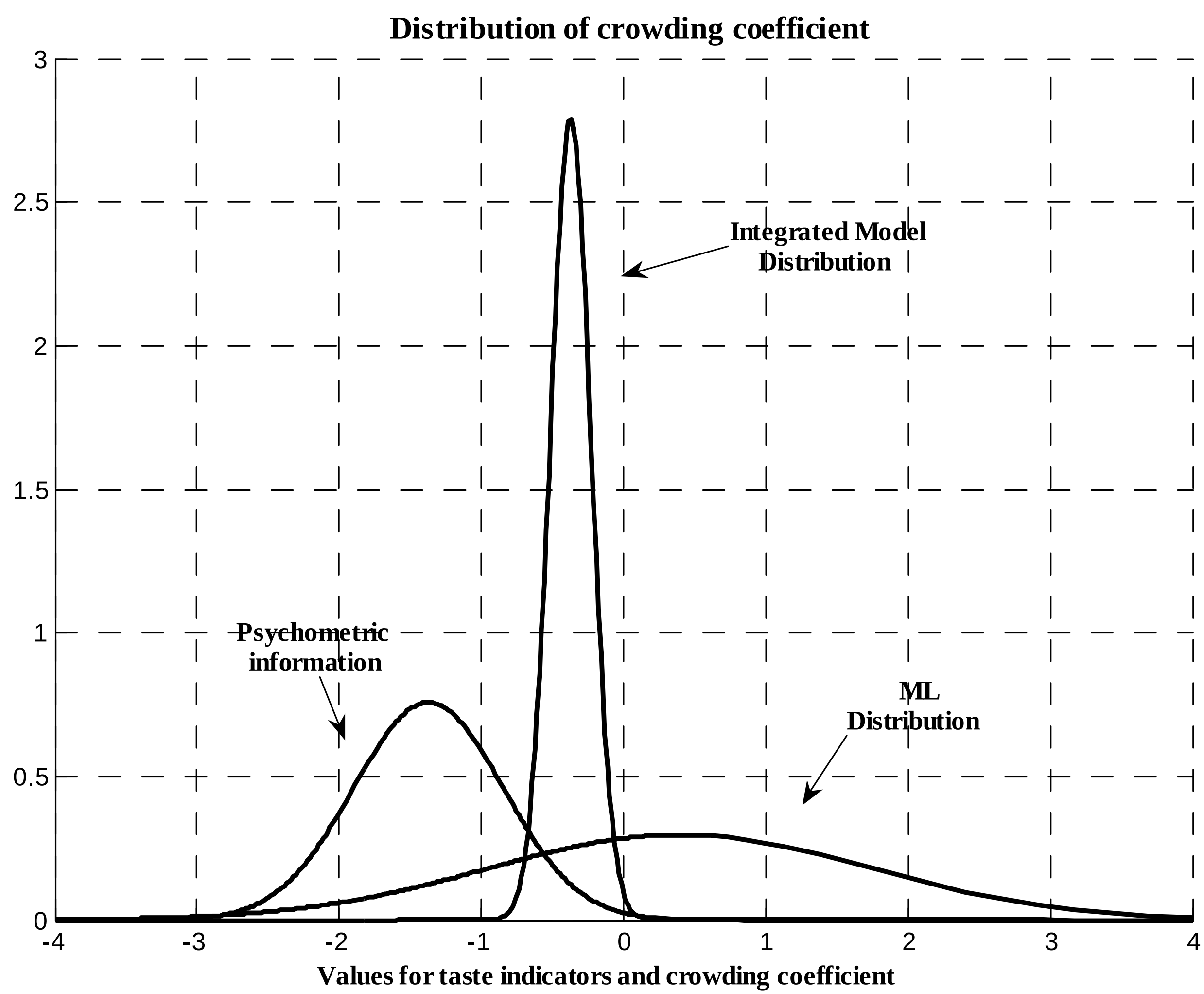
Distribution of crowding coefficient
Integrated Model Distribution
Psychometric information
ML Distribution
Values for taste indicators and crowding coefficient
3
2.5
2
1.5
1
0.5
0
-4
-3
-2
-1
0
1
2
3
4

Figure 2

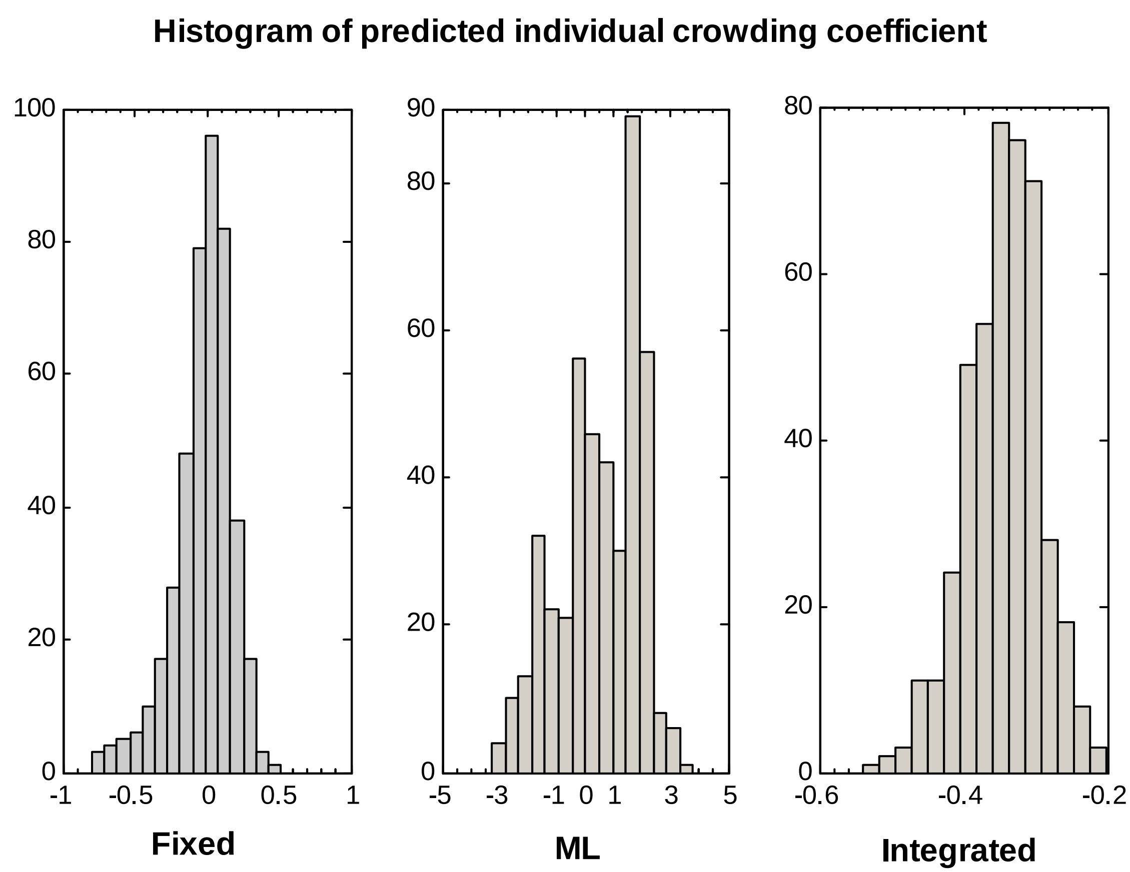
Histogram of predicted individual crowding coefficient
100
80
60
40
20
0
-1
-0.5
0
0.5
1
Fixed
90
80
60
40
20
0
-5
-3
-1
0
1
3
5
ML
80
60
40
20
0
-0.6
-0.4
-0.2
Integrated

Figure 3

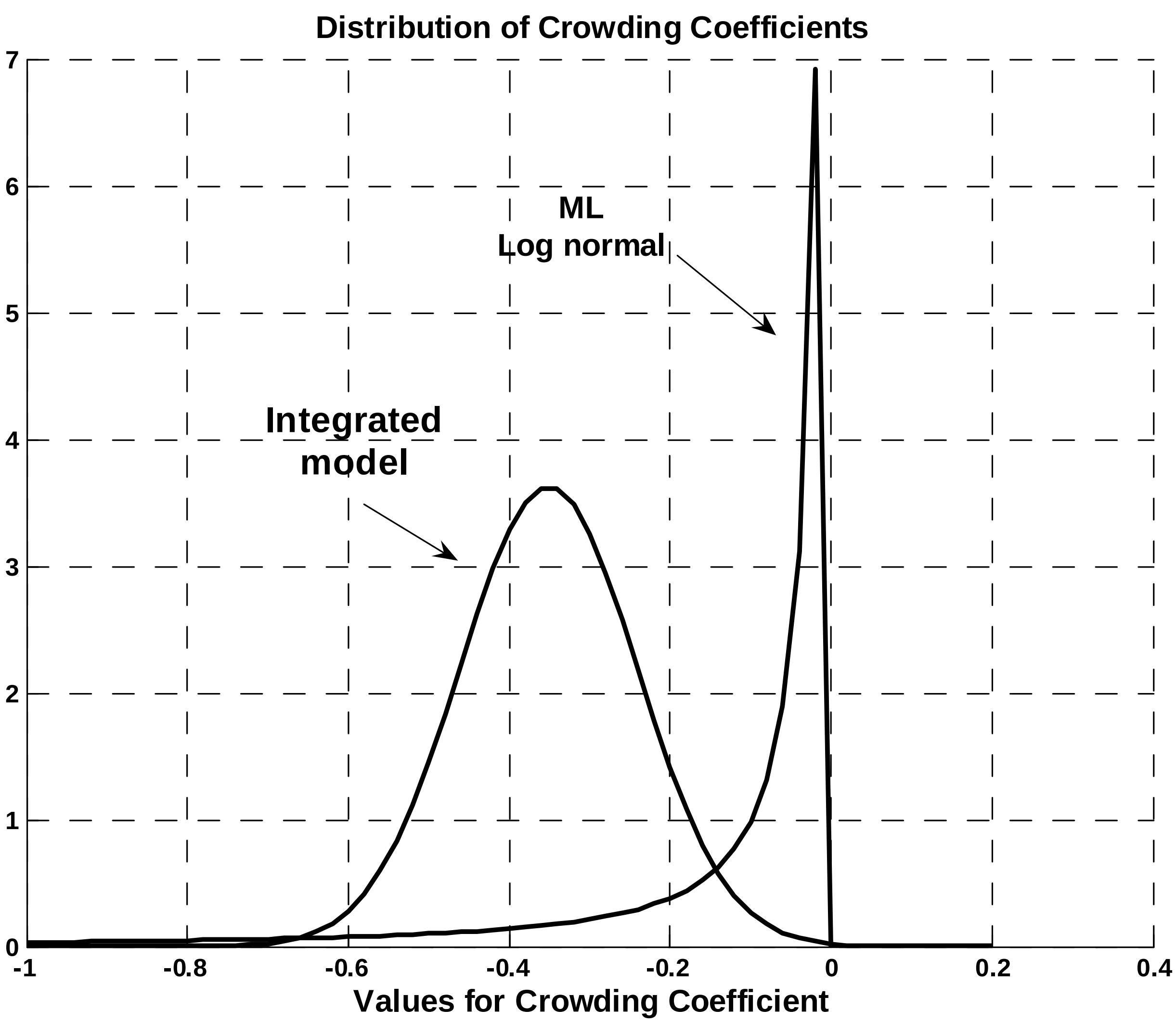
Distribution of Crowding Coefficients
ML
Log normal
Integrated
model
Values for Crowding Coefficient
7
6
5
4
3
2
1
0
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4

Figure 4