UNIVERSITY OF ALBERTA

EXPLORATIONS OF THE USE OF RANDOM UTILITY MODELS IN NONMARKET BENEFIT ESTIMATION

by

PETER CHARLES BOXALL ©

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

IN

AGRICULTURAL ECONOMICS

DEPARTMENT OF RURAL ECONOMY
EDMONTON, ALBERTA

SPRING 1999

0-612-39507-3

ABSTRACT

The random utility model is frequently used to examine the economic value of environmental quality changes. This dissertation extends this model by developing and applying procedures that allow: i) examination of *ex ante* values of unknown attributes; ii) incorporation and understanding of sources of preference heterogeneity, in particular, heterogeneity over tastes; and iii) incorporation of endogenous feedback.

The first paper examines the hypothetical discovery of new attributes in recreation site choice models using joint revealed-stated preference data. The empirical application involved the discovery of aboriginal rock paintings along wilderness canoe routes in eastern Manitoba. A four year study of wilderness recreation trips included a stated preference experiment in which canoeists were asked if they would change site choices in response to the presence of two types of rock paintings: a "pristine" painting and another spoiled by human vandals. The resulting stated site preferences (with new attributes) were combined with the revealed site preferences (without the attributes) in the econometric analysis. The results suggest that preferences over the SP and RP models were not statistically different. Estimated welfare measures for the presence of "pristine" paintings range from $5.71 - $8.38 per trip, and are about 11-12 times greater than those for vandalized paintings.

The second paper develops a finite mixture approach to random utility models. In this analysis, attitudinal and other individual-specific information was used to develop probabilistic membership in latent classes. The probability of membership in these classes was estimated jointly with choice parameters using choice experiment data in a model that

examined wilderness park choice in the Precambrian Shield region. Four classes were found in the data, and these classes reflected different motivations for taking wilderness trips and different behaviour in response to environmental quality changes. Two approaches for assessing the welfare of environmental changes were developed that took into account membership in each class, and the distribution of welfare impacts among the classes.

The third paper examines congestion in wilderness recreation areas. In this analysis, a random utility model was developed in which utility for one individual is a function of the utility of other recreationists. This model formalizes the notion that an attractive quality change for one recreationist is also attractive to others, thereby increasing the chances of congestion. Thus, in this framework, congestion can be considered endogenous feedback. To develop this notion in a random utility model, an instrumental variables approach was utilized. An ordered logit model of anticipated congestion was developed in which anticipated congestion levels were estimated as a function of individual characteristics and some environmental quality variables. This model was applied to canoeing in a system of wilderness parks and predictions of anticipated congestion were developed. These predictions were used as instruments in park choice models and the results compared to other choice models in which reported congestion levels were used instead of the predictions. The comparisons revealed significant differences in welfare estimates from the two models.

ACKNOWLEDGMENTS

First and foremost I sincerely thank my wife Lois for supporting once again, the attainment of another graduate degree by her husband. She continues to encourage me to do my best and graciously tolerated the many absences, mood swings, and general whining that accompanies graduate training and research. My sons and I are truly lucky to have such a fine person providing the foundation for our family.

I also would like to thank, with great pleasure, the support, guidance, and friendship of my advisor, Vic Adamowicz. He has influenced many students at the university and I count myself fortunate to being one of that company.

This research could not have been completed without the assistance and friendship of David Watson and Jeffrey Englin. The three of us together constituted the research team that developed much of the information that was used in this thesis. In particular, "Backwoods Dave" was instrumental in developing and collecting wilderness recreation data and Jeff provided large doses of ideas, positive reinforcement, and had a significant impact on my econometrics training.

Many other colleagues also provided assistance to me. Bonnie McFarlane deserves special mention for helping me with Chapter 3 and for being a supportive associate during my graduate training. Drs. Joffre Swait, Mel McMillan, Wolfgang Haider, Grant Hauer, and Edward Morey provided me with many ideas and comments on the thesis research. I also thank Dan Mulrooney, Rod Drew, Jim Crone, Kelly Leavesley, John McQueen, Grant Williamson, Mandy Fisher, Randall Hoscheit and a host of summer students who assisted me in many ways.

Finally, this research would not have been possible without the support of my employer, the Canadian Forest Service. I thank Steve Price and Bill White who provided financial and moral support for my doctoral studies at the University of Alberta. The research in this thesis was funded by the Canada-Manitoba Partnership Agreement in Forestry, the Northern Ontario Development Agreement, and the Canadian Forest Service's core research budget.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES AND ILLUSTRATIONS

Figure 3. The effects of changing road access at Wabakimi Provincial Park on welfare and expected congestion of wilderness recreationists. Note that in this simulation the size of the park was held at 1550 km^2.

CHAPTER 1

An Introduction to the Thesis

Most economic models operate on the assumption that individuals make best use of opportunities and resources available and that they respond rationally and predictably to changes in the conditions they face. These changes can involve prices, wages, financial endowments, and modifications of the natural environment. Thus, economic theory uses behaviour as an underpinning to value goods and services. This behavioural footing sets economics apart from other social science disciplines such as sociology or psychology which rely largely on attitudes, in turn provided by what people say, not what they do. The behavioural basis for economic analysis is powerful in that behaviour can be observed; values are thus inferred by observing what people do in the context of what is available or attainable (or what they could have done).

Samuelson (1938) was the first to suggest that by merely observing individuals choosing among various bundles of goods, a theory of behaviour could be devised in which consumers maximize utility based on some principles of rationality. The concept involves consumers "revealing" their preferences for goods and services. This idea led researchers to develop a set of approaches to examine the demand for nonmarket goods (goods which do not have observed prices associated with them) where observed choices of these goods involve the indirect expenditure of valuable assets such as time and money. Thus the valuation of the nonmarket good or service has associated with it some market purchase. This linkage is called the assumption of weak complementarity, for it allows the isolation of the demand for the nonmarket good through complementary market purchases.

Methods which utilize this complementary linkage are called revealed preference (RP) methods and they are typically used to examine recreational values. The most popular of these is the travel cost model (TCM) where the nonmarket good (recreation) is assumed to be complementary to expenditures on travel (travel costs). The TCM uses visits to a set of recreation sites as choices and travel costs (both time and expenses) as the complementary market purchases.

Other methods, called stated preference (SP) methods, attempt to establish a hypothetical market and elicit from consumers their intentions to change behaviour in response to changes in prices associated with changes in environmental quality. The most popular of these methods is contingent valuation (CVM), but contingent behaviour and choice experiments are also in the domain of stated preference (SP) methods (Boxall et al. 1996).

While RP methods can be used to estimate the value of some nonmarket activities they have also been adapted to estimating the value of changes in the conditions surrounding the "consumption" of the nonmarket good or service. From the environmental management perspective this approach can be used to value changes in the quality of the environmental or resource conditions associated with management decisions. For example, three types of TCMs have been proposed to examine the value of quality changes associated with recreation use. These are the varying parameter model (Smith and Desvouges 1986), the hedonic TCM (Brown and Mendelsohn 1984), and the discrete choice TCM or random utility model (RUM) (Fletcher et al. 1990).

While each of these RP approaches has particular strengths and weaknesses, the use of RUMs is far more widespread. Extending this model is the subject of this thesis.

The RUM involves defining an individual's utility function on attributes that define choices (Manski 1977). Consumers are assumed to maximize utility by considering the bundles of attributes that characterize the choices available in some defined set of alternatives. The model examines the probability that a consumer chooses a bundle of attributes that reveals the highest level of utility to him or her. The random element of this model arises from the fact that the utility function is defined as a function of the observable attributes of an alternative plus a random error term that reflects unobserved attributes of the alternative or the individual. Thus, the RUM is particularly appealing because it is consistent with notions of utility as a function of environmental attributes and because it uses this notion to assess the ability of consumers to substitute between defined sets of choices. These features facilitate the determination of measures of economic welfare directly from the estimated model. However, as this study will show, RUMs also possess the ability to model complex behavioural processes.

The RUM, however, is also the theoretical underpinning of referendum CVM and choice experiments (Boxall et al. 1996; Hanemann 1984). With SP methods the choice alternatives include the actual conditions and also one or more hypothetical ones. The use of RUMs in these instances allows researchers to examine the effect on individual utility importance of changing the attributes of an experience or policy. In these cases scenarios, sites, or "states of the world" can be presented and individuals are allowed to reveal their preferences by choosing or ranking these. Note that since this theoretical framework can be used with both RP and SP information it is also possible to combine these data and estimate joint RP-SP RUMs (Adamowicz et al. 1994).

In this dissertation research three issues in the use of RUMs in environmental valuation are examined. The first, described in Chapter 2, is the effect of introducing a new attribute *ex post*, and predicting the effect of this attribute on the choice behaviour of an individual. This analysis incorporated both RP and SP information. The second issue involves the introduction of individual characteristics in the choice model in order to better understand and incorporate preference heterogeneity. In this case, economic information arises from an SP experiment and this is combined with attitudinal data. Finally, the third issue is the incorporation of endogenous perceptions of environmental attributes in RUMs. In this analysis, information involves RP data, but this is supplemented with perceptual information on some the attributes of the choice alternatives. In addressing all of these issues the empirical application of the procedures developed will be examined using wilderness recreation demand using Canadian data.

This thesis, however, is also about "pushing" the boundaries of economic analysis in the sense that economic theory and methods are combined with information, theory and methods from other social science disciplines. For example, the second chapter examines the effect of the discovery of aboriginal cultural artifacts and vandalism of these artifacts on recreation site choice. The third chapter borrows the theory of latent class analysis from sociology and takes attitudinal measurement methods from social psychology; these are introduced into the economic analysis of recreation demand. The fourth chapter utilizes sociological thinking about endogenous feedback in the examination of congestion in recreation site choice behaviour.

As noted, the following three chapters each deal with a different issue related to the use of RUMs for environmental valuation. The final chapter provides summary remarks and offers conclusions for future research related to RUMs.

LITERATURE CITED

Adamowicz, W.L., J. Louviere, and M. Williams. 1994. Combining stated and revealed preference methods for valuing environmental amenities. *Journal of Environmental Economics and Management* 26:271-296.

Boxall, P.C., W.L. Adamowicz, J. Swait, M. Williams and J.J. Louviere. 1996. A comparison of stated preference methods for valuing environmental goods. *Ecological Economics* 18:243-253.

Brown, G.M. and R. Mendelsohn. 1980. The hedonic travel cost method. *Review of Economics and Statistics* 66:427-433.

Fletcher, J.J., W.L. Adamowicz, and T. Graham-Tomasi. 1990. The travel cost model of recreation demand: theoretical and empirical issues. *Leisure Sciences* 12:119-147.

Hanemann, W.M. 1984. Welfare evaluations in contingent valuation experiments with discrete responses. *American Journal of Agricultural Economics* 66:332-341.

Manski, C.F. 1977. The structure of random utility models. *Theory and Decision* 8:229-254.

Samuelson, P. A. 1938. A note on the pure theory of consumer's behaviour. *Economica* NS 5: 61–71 and 353–54.

Smith, V.K. and W. Desvouges. 1986. *Measuring Water Quality Benefits*. Kluwer-Nijhoff, Norwell, Mass.

CHAPTER 2

Valuing Undiscovered Attributes: A Combined Revealed-Stated Preference Analysis of North American Aboriginal Artifacts

INTRODUCTION

A challenge in the valuation of environmental amenities is the *ex ante* measurement of values associated with unknown goods and services. The behaviour inherent in any revealed preference information is, of course, associated with the *ex ante* situation. Newly discovered goods and services such as new species or cultural artifacts will result in new alternatives and/or new attributes of existing alternatives that may affect future behaviour. Since these goods or services are currently unknown or unavailable, there is no revealed preference information to use in their valuation. As a result, one is forced to rely upon stated preference information to examine the new attribute or alternative. Nevertheless, it is important to maintain consistency with revealed behaviour associated with the *ex ante* situation. The challenge is to acknowledge and exploit all of the information available when valuing newly discovered attributes.

This analysis examines the potential discovery of aboriginal rock paintings along wilderness canoe routes in the Precambrian Shield region in central Canada. The region contains about 400 paintings on rock faces along water courses (Rajnovich 1994), and some of these faces are located in popular canoeing areas (e.g. Quetico Provincial Park, Dewdney and Kidd (1962)). Many other areas in North America contain similar paintings or rock carvings (Grant 1983). Anthropological scholars call the paintings pictographs and the carvings petroglyphs because they represent picture writing, not necessarily works of art. These drawings were used to communicate among individuals or with the spirit world by the Aboriginal peoples.

Anthropologists believe some of the pictographs in the Shield region to be 2000 years old (Rajnovich 1994). Thus, there has been concern regarding their documentation and preservation for historical and cultural reasons. Of particular concern is the destruction of pictographs and petroglyphs by flooding caused by the construction of dams.[1] Discoveries of pictographs and other artifacts are still occurring as a result of proposed hydroelectric developments in parts of the Canadian Shield and new pictographs are catalogued periodically.

Pictographs in the Shield still have spiritual and cultural significance to Aboriginal peoples which is evidenced by the discovery of recent offerings of tobacco, clothing, and prayer sticks (pers. observ.; Dewdney and Kidd 1962). However, they are also sought by wilderness recreationists who consider them to be an important feature of a wilderness experience (Boxall unpublished). Pictographs are also promoted by ecotourism operators who use them to attract clients interested in cultural experiences. The discovery of pictographs may increase visitation levels to wilderness areas, and while generating recreation and tourism benefits, may also increase the chances of vandalism. Steinbring and Elias (1968), Dewdney and Kidd (1962) and others describe pictographs which have been shot at by hunters and where non-aboriginal individuals have spray

[1] For example, the Glen Canyon contained some spectacular petroglyphs which are now submerged beneath Lake Powell (Grant 1983).

painted their names or initials over them. Therefore, tensions exist among recreationists as to the value of pictographs to recreationists. There is also concern about the importance of them to aboriginal cultures, and about the risks of vandalism. Wilderness managers should be interested in forecasting demand to view pictographs, not only to estimate their benefits, but also to provide information on the probabilities of their defacement.

The paper proceeds by developing the random utility model RUM used in this analysis. This model directly incorporates revealed preference information about existing recreation site attributes and stated preference information about undiscovered site attributes. In this section the combined revealed–stated preference approach is developed. This is followed by a description of the data used in this analysis. The empirical section applies the combined revealed preference-stated preference model to the hypothetical discovery of aboriginal rock paintings along water courses in an actual wilderness area. The empirical application also examines the effect of vandalism on the benefits generated by the rock paintings.

THEORY

A variety of modelling frameworks has been proposed to analyse combined revealed and stated preference data (e.g. Cameron (1992); Englin and Cameron (1996); Adamowicz et al. 1994, 1997). The RUM framework of Adamowicz et al. (1994, 1997) is especially appealing in settings where most individuals make a single trip. In this setting the single trip nature of pure random utility models is less troubling than in other contexts (see Morey (1994)). The appeal of the RUM is its ability to handle substitution between site attributes and the direct measurement of economic welfare, while retaining the ability to test the consistency of the revealed and stated preference components of the model.

Consider a recreationist who makes a choice from a set of C possible sites. The probability (π) that site *j* will be visited is equal to the probability that the utility gained from visiting *j* is greater than or equal to the utilities of choosing any other site in C. In this framework it is assumed that the indirect utility function consists of the sum of two components: an observed component, V_j, and a random component e_j. The probability of selecting site j can be written as :

$$\pi(j) = \Pr\{ V_j + e_j \geq V_k + e_k ; \forall k \in C \}. \tag{1}$$

Empirical implementation of (1) requires the selection of a distribution to characterise the random component of the model (e_j). The conditional logit model can be used to estimate these probabilities if the random components of the indirect utility functions are assumed to be independently distributed with a Type-I Extreme Value distribution (Weibull).

This model is typically estimated with the observable component, V_j, expressed as a linear function of m site attributes and the cost of visiting a site. A new attribute introduced into this framework will take the form of an additional attribute in the indirect utility function:

$$V_j = \sum_{1}^{m} \beta_m X_{jm} + \alpha X_{jnew} + \gamma (Y_n - p_j), \tag{2}$$

where X_{jm} represents existing choice based attributes, $X_{j\,new}$ represents a new attribute, Y_n is the consumer's income, p_j is the cost of visiting site j, and β, α, and γ are unknown parameters. However, by definition in the case of *ex ante* measurement of the new attribute, the only revealed preference data available are based on behaviour that does not take into account the discovery of the new attribute. Consequently, the α parameter will be impossible to estimate. An estimable model requires situations where data exist on choices with and without the new attribute.

One alternative is to obtain a set of choice data from another location that include the new attribute and to transfer the values in a benefits transfer process. Alternatively, the *ex post* valuation of the attribute could be explored using revealed preference data after the discovery of the attribute. In neither case, however, can one tailor predictions of the effects of a discovery to a specific area or site. Since there is no market information about the effect of the new attribute on choice behaviour, the original revealed preference data must be augmented in this situation. One way to augment the revealed preference data is to add stated preference data. Stated preference data can be used to assess the change in intended behaviour that results from the introduction of new attribute. A potential solution to these problems is to apply a combined analysis where revealed and stated preference information for the same set of individuals is pooled.

An empirical issue is the appropriate combination of the revealed and stated preference data. If only revealed or stated preference data is used McFadden (1973) has shown that the choice probabilities take the form:

$$\exp \mu (V_j) / \sum_{k \in C} \exp \mu (V_k) \quad , \tag{3}$$

where μ is a scale parameter. Since this parameter is not identified in a single set of data it is typically normalised to 1. Once the variables in the deterministic component of the indirect utility function, V, are specified and a functional form is selected, the model becomes estimable using maximum likelihood methods.

When multiple data sets such as stated and revealed preference data are pooled, an important issue is the consistency of the data sets with each other. A useful measure of this consistency is a test of the equality of the scale parameters in the two data sets. An econometric test of the equality of the scale parameters can be constructed. This is done by normalising *one* scale parameter in (3) and letting the scale parameter from the other data set vary in the estimation process as shown by Swait and Louviere (1993), and Adamowicz et al. (1994; 1997). This method involves the notion that in any one data set μ is not identifiable, but that in any two (or more) datasets their ratio(s) can be identified (e.g. μ_1/μ_2). Thus, for a pooled set of revealed and single stated preference data this process involves the concatenation of the choice probabilities as follows:

$$\begin{aligned} RP: \quad \pi(j) &= \exp(\mu_{RP} V_j / \sum_{K \in C} \exp(\mu_{RP} V_k) \\ SP: \quad \pi(j) &= \exp(\mu_{SP} V_j / \sum_{K \in C} \exp(\mu_{SP} V_k) \quad , \end{aligned} \tag{4}$$

where μ_{RP} will be set to 1 and the μ_{SP} is a parameter to be estimated. Of course, this methodology can be used to extend the number of pooled data sets to any arbitrary size. If there were multiple new attributes one could extend the number of pooled data sets and concatenate choice probabilities.

In this paper the method is applied to three data sets. The three data sets correspond to a single set of revealed preference data and two stated preference data sets. The precise log likelihood function for this problem, given a sample of N individuals, is given by:

$$LL = \sum_{n=1}^{N(RP)} \sum_{k \in C} \ln \pi_n \{j \mid \beta\} + \sum_{n=1}^{N(SP_1)} \sum_{k \in C} \ln \pi_n \{j \mid \beta, \mu_{sp1}\} + \sum_{n=1}^{N(SP_2)} \sum_{k \in C} \ln \pi_n \{j \mid \beta, \mu_{sp2}\}.$$

The first part of the log likelihood function corresponds to the revealed preference data and the second two components correspond to the two stated preference data sets. While there really are three scale parameters, note that only two are estimated. These two estimates are then compared to the normalized scale parameter to determine whether they are statistically different from 1.0.

The Data

The study involves wilderness recreation in Nopiming Provincial Park, Manitoba (Figure 1). The park is a 1440 km^2 area located about 145 km east of Winnipeg and is situated in the Precambrian or Canadian Shield. The area contains numerous rock outcrops that can rise as much as 36 m above the surrounding countryside. These are a dominant feature and are sought after by recreationists. Pictographs are frequently found on rock outcrops along watercourses in this region, and while no paintings have been reported in the park, there are some in similar areas around Nopiming and in more remote areas in Ontario (Rajnovich, 1994). The park has several river systems that contain small rapids and waterfalls that are attractive to backcountry recreationists interested in canoeing and kayaking. Most of the park is forested. Jack pine is the most abundant tree species in the park, although considerable areas of black spruce, aspen and white spruce can be found.

The wilderness recreation in this park and the surrounding region has been carefully studied in recent years. This involved an economic assessment of the importance of fire, forest ecosystems and other features (Boxall et al. 1996; Englin et al. 1996). As a result there is a detailed inventory of features along canoe routes that was verified through intensive field work and GIS databases obtained from the provincial forest management agency. These inventories, and in particular the site visits, identified areas in Nopiming that could potentially have rock paintings.

A registration system was developed to provide an understanding of the frequency of visitation to the backcountry areas of the park. In 1995 the registrants were surveyed. The survey included a stated preference experiment in which backcountry visitors were asked to respond to the possible presence or discovery of rock paintings in the park. The survey sample was created using the names of the leaders of the recreation parties who registered for a backcountry trip in Nopiming Park in 1993 or 1994. The original sample of 661 registrants was reduced to 587 by eliminating incomplete addresses, and multiple trips by the same individual. There were, however, very few recreationists who took more than one trip in a year. The final sample included individuals from five Canadian provinces, and three American states.

The experiment involved presenting pictures of two pictographs to respondents. The first involved a "pristine" pictograph. This pictograph exists in a more remote wilderness area in Ontario northeast of Nopiming. The second involved a picture of a pictograph located in a remote area in northern Manitoba that had been defaced by vandals and appears to be weathered. These pictures and the stated preference questions used in the experiment are shown in Figure 2.

The survey design exploited the knowledge of historical trip behaviour. Each respondent was offered the chance to change his or her trip to another route to see a rock painting. Since the original trip was known, each respondent was offered the rock paintings at a site they had not visited during the study period (1991-1994). Thus, the experiment ensured that every respondent had an opportunity to change his or her original site choice to a different site. The pictographs were offered at two routes: the Seagrim Lake canoe route and the Manigotagan River route. These sites were chosen because they were the only routes in the park that had rock outcrops similar to those where pictographs are typically found.

The survey included a total of three mailouts. First, a questionnaire and cover letter was sent to the 587 individuals in early March 1995. Two weeks later a reminder post card was sent to any individual that had not responded. Finally, five weeks after the original mailout, a second questionnaire and cover letter was sent to nonrespondents. These procedures resulted in the return of 431 completed questionnaires which, adjusting for undeliverables (e.g. people moving etc.), represented a response rate of 81%.

The final data sets used for analysis consist of actual site choices for the respondents for 1993 and 1994 (revealed preference data) and their stated choices from the questionnaire (stated preference data). In this information, the choice set was limited to the eight major routes in the park. Any respondent whose actual trips were not to any of these eight routes was excluded from the analysis. This resulted in a final sample consisting of 386 respondents with complete trip data.

Boxall et al. (1996) identified 14 route attributes that explained the choice of a canoe route in the same park. These included travel costs, the incidence of historical fires, the amounts of various forest ecosystems, indicators of the level of physical effort required to complete the route, and some alternative specific constants. All of these variables and various combinations were initially examined as explanators of site choice, but only travel costs and three others (not including pictographs) could be identified in this study. The Boxall et al. study, however, used 20 routes in the analysis, of which eight were the same as those in the present study. Thus, the reduced number of routes suggests that only a subset of the full set of the independent variables used previously may be significant explanators of site choice.

Figure 3 illustrates the creation of the three data sets and the construction of the pictograph variables. In each set of data the eight routes form the choice set, and four independent variables form the route characteristics (X_{RP}). Since there are two types of pictographs, one dummy variable was formed for each type. The first involved the pristine pictograph (P in Fig. 3) which has a value of 1 at the route it was presented to the respondent in the questionnaire (Seagrim or Manigotagan) and a value of 0 otherwise. The defaced pictograph variable (D) was similarly constructed. Since there was no pictograph present in the RP choice data, both pictograph variables had values of 0 at each of the 8 routes.

The econometric analysis proceeded by constructing separate conditional logit models for each of the three datasets. Two additional models were estimated on the combination of the RP data with each of the SP datasets. In this estimation, the two sets of data were stacked as shown in Figure 3. Finally, a joint model estimated on the concatenation of all three datasets was performed. In each of the three joint models the likelihood function in equation (5) was adjusted to account for the number of datasets combined in the estimation.

RESULTS

The actual site choices of the respondents and their response to the SP experiment is shown in Table 1. Note that the Tulabi route was the most popular route actually chosen. About 42% of the respondents in the sample indicated they would change their actual route choice to another route to view a pristine painting. This change would occur regardless of the route where a painting was discovered (Seagrim and Manigotagan). However, only about 10% of the respondents stated that they would change their behaviour to view a defaced painting. The effect of the pictograph attributes on site choice is portrayed in Figure 4 where the cumulative increase in the number of trips to Seagrim and Manigotagan is shown relative to the availability of the two types of paintings.

The welfare measures associated with both the pristine and vandalised rock paintings are quantified in this analysis. In this analysis, the *a priori* hypothesis was that the pristine painting provides substantial positive benefits to the recreationists. This arises because the paintings enhance the attributes of some alternatives in the choice set. Thus, sites with paintings should exhibit an increased probability of visitation. It was further hypothesised that the vandalised paintings would not provide benefits as large as the pristine painting. However, defacement aside, the vandalised picture may still induce some change in trip behaviour by increasing the probability of visiting the sites with paintings.

Table 2 shows the parameters for six econometric models. The first three columns in the table report results for the individual RP, pristine pictograph SP, and defaced pictograph SP models. Column 4 shows the results of the model that combines RP data with the pristine SP data and column 5 shows the results of combining RP data with data SP data for the defaced pictograph. The last column shows the final model that includes the RP data together with both sets of the SP data.

In all of the models the parameters on travel costs to the route[2], hectares of recent burned areas, and hectares of black spruce old growth ecosystems are negative and significant. The parameters on hectares of white spruce growth and the single alternative

---

[2] Travel costs were estimated using the standard approach in the literature (e.g. Parsons and Needelman 1992). This involved two calculations: i) the out-of-pocket expenses estimated as $0.25 * the round trip distance in kilometers, and ii) the opportunity cost of time estimated as one quarter of a respondent's annual income divided by 2080 hours multiplied by the travel time. Travel time was estimated using an average speed of 90 km/h during the round trip distance.

specific constant for the Manigatogan canoe routes are positive and significant. These results are consistent with previous research on site choice behaviour in the park involving a larger set of canoe routes (Boxall et al. 1996) and with trip data from different years (1991 and 1992) and a smaller set of routes (Englin et al 1996) . In the RP model there is no parameter for pictographs because the paintings are not available. However, pictograph parameters are in the SP data. For the pristine pictograph model, the pictograph parameter is large and positive, while in the defaced pictograph model the parameter on the pictograph is smaller, but still positive. These findings are consistent with *a priori* expectations. In the three joint models the individual parameters are similar to those in the other models.

Tests of the equality of the restricted (joint) and unrestricted (single) models were conducted using likelihood ratio tests. These results are reported in Table 3. In each comparison the hypothesis of equality between models is not rejected at the 5% level of significance. In particular, the hypothesis of equality for the three-way joint model (RP+SPp+SPd) is not rejected. This suggests that the single RP and SP models share the same preference structures as the joint RP-SP models. Thus, unlike the finding of Adamowicz et al. (1994; 1997) it is **not** necessary to scale the SP data to the RP data. The ratios of scale parameters in these data are not significantly different than 1.0.

These specification tests support the use of the RP-SPp-SPd model to assess switching behaviour and the welfare effects of discovering pictographs at the two routes in the park. For this purpose, simulations were first conducted to assess the effect of the presence of the pristine pictographs on the probability of visiting each of the eight major routes in the park. Figure 5 shows the distribution of the mean probabilities for 3 scenarios. The scenario with no pictographs suggests that the Tulabi route is most favoured by the average individual in the sample. However, discovery of a pictograph at the Seagrim route increases the probability of taking a trip there by almost six times. A pictograph at the Manigotagan route increases the probability of taking a trip there about five fold. The presence of pictographs significantly changes the distribution of trip probabilities. A discovery at either of the two routes suggests considerable switching of site choices so that many more recreationists would visit those routes than would be the case otherwise.

Second, mean welfare measures associated with the two types of pictographs (calculated over the sample) were estimated using Hanemann's (1982) formula. The change in site choice behavior as a result of the presence of pristine pictographs increases the benefits of a canoe trip to the park by \$8.38 if the pictograph was located at Seagrim and \$5.71 if it was located on the Manigotagan. These increased benefits would fall to only \$0.73 and \$0.46 per trip if the painting was vandalised (Figure 6). Thus, at these routes, a pristine pictograph would provide about 11-12 times the benefits of a vandalised one. Simulating the presence of pictographs at the other six routes suggests that the magnitudes of the benefits are higher or lower, but the pattern of the difference between the pristine and defaced paintings is similar. The overall magnitudes of the benefits across the sites reflect the complementarity of the pictographs with other attractive or negative features of the routes used in the choice models.

DISCUSSION

One important challenge facing managers of public lands is the tension between use, overuse and risk. Clear cases in point are cultural resources such as the pictographs studied in this paper. In this analysis the value of pictographs to wilderness recreationists is examined. Pristine pictographs are quite valuable, in some cases as much as $8.00 per trip. This compares favorably with museum admission charges. A defaced pictograph, however, is worth about one twelfth of the pristine pictograph. This type of difference in values between vandalized and unvandalized attributes is frequently speculated to be the case in the literature (e.g. Harrison 1976), and this is empirically verified in this study. The contrast in values suggests that concerns over the effect of vandalism on tourism or recreational experiences are well founded. Of course, knowing the values of pictographs does not solve the management challenge associated with their presence. There remains the question of whether it is worthwhile forgoing the benefits associated with the pristine pictograph to reduce the risk of the pictograph being vandalized. Knowledge of the risk of vandalism is also needed to conduct a rational policy discussion of this issue. However, without estimates of values, no economic discussion of the merits of different policies can be conducted.

An additional problem faced by public land managers concerned with areas currently used by indigenous peoples is the protection of spiritual and cultural artifacts. In Canada information about the location of these features is collected at times from Aboriginal peoples, but this is frequently not released to the public for fear of depreciative behavior on the part of non-Aboriginal peoples. The approach used in this paper to examine switching behavior of recreationists in response to discoveries of pictographs would allow the forecasting of visitation levels. Managers could use these forecasts to design access plans or other policies to reduce the chances of discovery of places of significance to aboriginal peoples, and thus decrease the probabilities of vandalism or other depreciative behaviors.

Early work joining stated and revealed preference data in random utility models struggled to develop methods that tested the consistency of the behavior suggested by revealed and stated data. Quite often the two data sets were not consistent with one another unless one was scaled (e.g. Adamowicz et al. 1994, Ben-Akiva and Morikawa 1990). In this analysis of three data sets, the evidence supports the hypothesis that the stated and revealed data come from consistent behavioral models. This finding is likely to have resulted from several factors. One is the clarity of the good in question. Pictographs are well known to Canadians who live in the study region and these artifacts are well known to those who visit wilderness areas there. Secondly, the population of canoeists is sufficiently homogeneous to make simple specifications of the scale parameter possible. A more heterogeneous population may not provide the scaling results seen in this study. Finally, this study was undertaken as part of a larger effort focused on modeling wilderness site choice behavior in the Canadian Shield region. In this larger context the role of landscape features, ecosystem processes such as forest fires, and wilderness managerial features were understood. This knowledge helped to clarify the processes that are driving the choices of wilderness canoeists in the region. Furthermore, the detailed information base about the recreationists allowed the survey used in this study to be "custom designed" for each respondent. This design, in concert with the high level of knowledge of the factors affecting site choice behavior, may have contributed to the success of the modeling effort reported in this paper.

This paper raises a number of questions. For example, an important issue is the possibility of the model over-predicting switching behaviour. Ben-Akiva and Morikawa (1990) found this feature in their analysis of travel modes using a similar combined estimation approach. Furthermore, it is unclear whether the switching of routes would continue to occur in the long run. It is possible that dynamic forces may play a role in site choice behaviour with discoveries causing switching once or twice, but not permanently. An additional issue is the effect of discoveries on the levels of visitation. In the empirical example examined in this paper, very few individuals actually took more than one trip a year. Related research on these recreationists suggests that trip quantities do not vary with changes in route qualities (Englin et al. 1996). There is no reason to believe that the presence of pictographs would change this behavior. Nevertheless, this issue and the questions of dynamic behavior raised above represent worthy topics for future research in this area.

The objective of this paper was to improve the understanding of the demand by recreationists for Aboriginal artifacts in wilderness areas. The approach used to examine this issue was to analyze combined revealed and stated preference data. The results suggest that wilderness users would switch their site choices to areas with artifacts and that this would generate large recreation benefits. However, with increased levels of visitation comes an increased risk of vandalism which, in turn, affects the generation of benefits to recreationists and aboriginal peoples. The successful management of wilderness areas containing Aboriginal artifacts will require the development of extensive and reliable databases. The effectiveness of the stated preference design used in this study was only possible due to high levels of understanding of the users of the wilderness areas.

LITERATURE CITED

Adamowicz, W.L., J. Louviere, and M. Williams. 1994. Combining stated and revealed preference methods for valuing environmental amenities. *Journal of Environmental Economics and Management* 26:271-296.

Adamowicz, W.L., J. Swait, P.C. Boxall, J. Louviere, and M. Williams. 1997. Perceptions versus objective measures of environmental quality in combined revealed and stated preference models of environmental valuation. *Journal of Environmental Economics and Management* 32:65-84.

Ben-Akiva, M. and T. Morikawa. 1990. Estimation of switching models from revealed preferences and stated intentions. *Transportation Research A* 24A:485-495.

Boxall, P.C., D.O. Watson, and J. Englin. 1996. Backcountry recreationists' valuation of forest and park management features in wilderness parks of the western Canadian Shield. *Canadian Journal of Forest Research* 26:982-990.

Cameron, T.A. 1992. Combining contingent valuation and travel cost data for the valuation of nonmarket goods. *Land Economics* 68:302-317.

Dewdney, S. and K. Kidd. 1962. *Indian Rock Paintings of the Great Lakes.* University of Toronto Press, Toronto ON.

Englin, J., P.C. Boxall, K. Chakraborty, and D.O. Watson. 1996. Valuing the impacts of forest fires on backcountry forest recreation. *Forest Science* 42:450-455.

Englin, J. and T. A. Cameron. 1996. Augmenting travel cost models with contingent behavior data. *Environmental and Resource Economics* 7:133-147.

Grant, C. 1983. *The Rock Art of the North American Indians*. Cambridge University Press, Cambridge.

Hanemann, W. M., 1982. Applied welfare analysis with qualitative response models, Working Pap. No. 241. Univ. California, Berkeley, CA, 26 pp.

Harrison, A. 1976. Problems: vandalism and depreciative behavior. Chapter 24, in G.W. Sharpe (editor) *Interpreting the Environment*. J. Wiley and Sons Inc., New York, NY.

Maddala, F. S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge University Press, New York, NY.

McFadden, D. 1973. Conditional logit analysis of qualitative choice behavior. In: P. Zarembka (Editor), *Frontiers in Econometrics*. Academic Press, New York, NY. pp. 105-142.

Morey, E. R. 1994. What is consumer surplus per day of use, when is it a constant independent of the number of days of use, and what does it tell us about consumer surplus? *Journal of Environmental Economics and Management* 26:257- 270.

Parsons, G.R. and M.S. Needelman. 1992. Site aggregation in a random utility model of recreation. *Land Economics* 68:418-433.

Rajnovich, G. 1994. *Reading Rock Art: Interpreting the Indian Rock Paintings of the Canadian Shield.* Natural Heritage/Natural History Inc., Toronto, ON

Steinbring, J. and D. Elias. 1968. A key pictograph from the Bloodvein River, Manitoba. *American Antiquity* 33:499-501.

Swait, J. R. and J. 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.

TABLE 1
Actual and hypothetical site choices in response to aboriginal pictographs by wilderness recreationists at Nopiming Provincial Park, Manitoba.

Routes	Original number of trips	Number of surveys where pictographs were offered	Number of people switching for a pictograph	Number of people switching for a defaced pictograph
Tulabi	183			
Shoe	12			
Rabbit	58			
Seagrim	41	246	103	28
Gem	12			
Beresford	40			
Manigotagan 1	19	140	58	14
Manigotagan 2	21			
Total	386	386	161	42

TABLE 2

Parameters (standard errors) from Conditional Logit Models used to Examine Recreation Site Choice in Nopiming Provincial Park, Manitoba.

Variables	Single Models			Joint Models (Combined RP-SP)		
	RP	SPp[a]	SPd			
	No Pic	Pristine Pic	Defaced Pic	RP+SPp[b]	RP+SPd	RP+SPp+SPd
Travel Cost	-0.0704* (.0073)	-0.0628* (.0080)	-0.0579* (.0068)	-0.0703* (.0054)	-0.0648* (.0050)	-0.0656* (.0042)
Recent Burns	-0.2072* (.0208)	-0.1481* (.0027)	-0.1446* (.0215)	-0.1892* (.0165)	-0.1786* (.0150)	-0.1730* (.0131)
Black spruce old growth	-1.2931* (.1376)	-0.9224* (.1525)	-0.8421* (.1308)	-1.1245* (.1013)	-1.0697* (.0947)	-1.0204* (.0800)
White spruce old growth	5.6714* (.7268)	3.9250* (.8897)	3.5366* (.7243)	5.2194* (.5670)	4.6971* (.5148)	4.6046* (.4467)
Pristine pictograph		2.1379* (.1565)		2.3802* (.1312)		2.2985* (.1218)
Defaced pictograph			0.3279* (.1891)		0.5222* (.1783)	0.4673* (.1744)
Constant for Manigotagan Routes	2.6873* (.4719)	2.6465* (.4863)	2.0700* (.4220)	2.8845* (.3361)	2.4267* (.3159)	2.5807* (.2630)
Log L	-667.73	-641.28	-716.72	-1313.48	-1388.87	-2033.50

[a] SPp refers to the SP data with the pristine pictograph, SPd refers to the SP data with the defaced pictograph

[b] RP+SPp refers to the combined RP and SP data with the pristine pictograph, RP+SPd refers to the combined RP and SP data with the defaced pictograph etc.

TABLE 3

Tests of the hypothesis of parameter equality between the recreation site choice models.

Models	Log Likelihood	Likelihood Ratio Test	
		$\sum$ Log L	χ^2
No pictograph (RP)	-667.73		
Pristine Pictograph (SPp)[a]	-641.28		
Defaced Pictograph (SPd)	-716.72		
RP + SPp	-1313.48	-1309.01	8.94 [b]
RP + SPd	-1388.87	-1384.45	8.84 [b]
RP + SPp + SPd	-2033.45	-2025.73	15.44 [c]

[a] Definition of abbreviations are shown in Table 2.

[b] Critical χ^2 at P = 0.05, 7 df is 14.07

[c] Critical χ^2 at P = 0.05, 14 df is 23.68

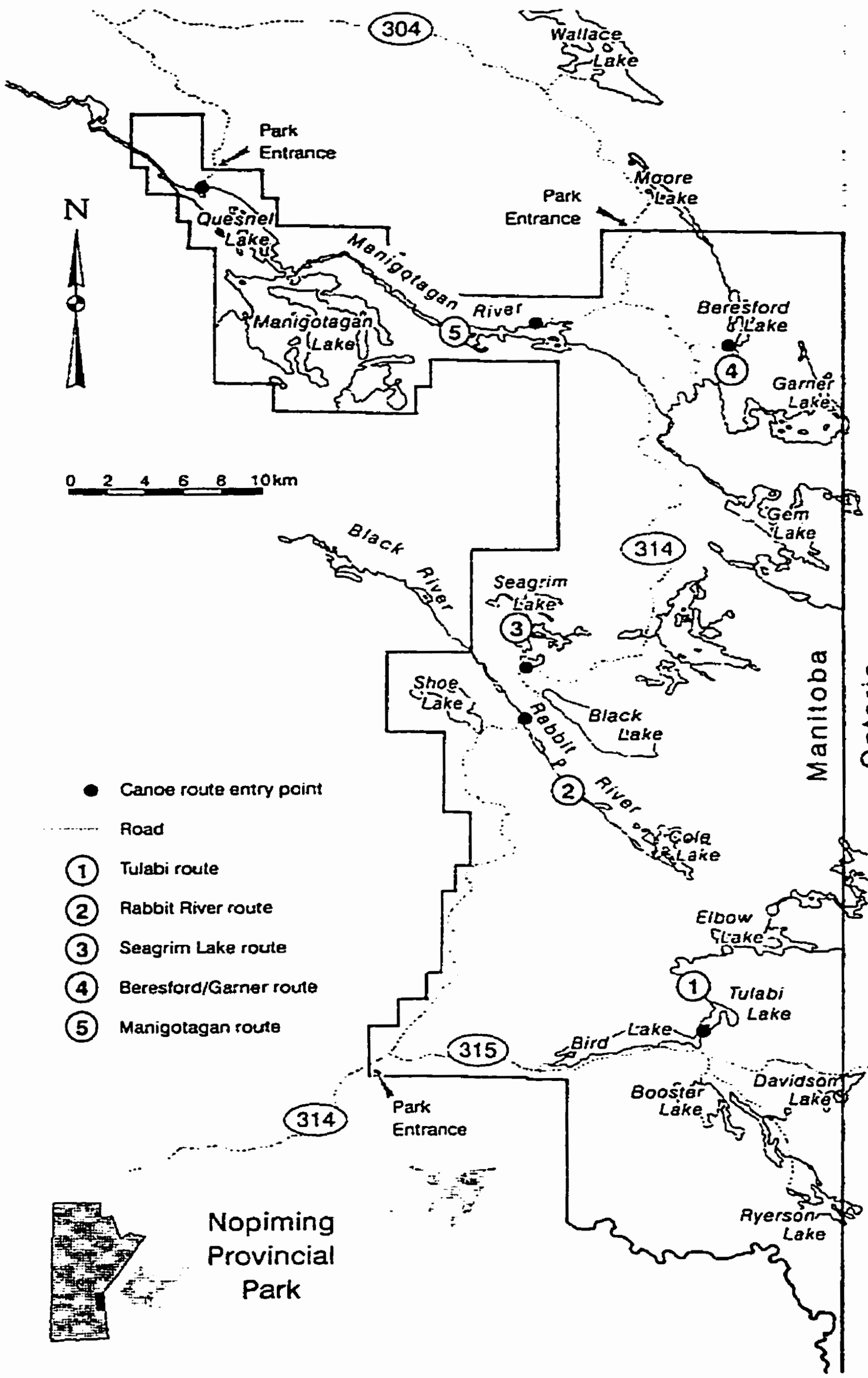


Figure 1. A map of Nopiming Provincial Park, Manitoba. Fieldwork suggested that rock faces along the canoe routes marked 3 and 5 would be suitable places for pictographs.

There is considerable evidence of ancient historical use of forests by Canada's aboriginal peoples. This can be seen in the form of pictographs or rock paintings. In Manitoba existing pictographs may be up to 2000 years old and are still highly visible. In these questions we wish to examine the influence of these paintings on your visits to backcountry areas. Please remember that these questions are hypothetical.

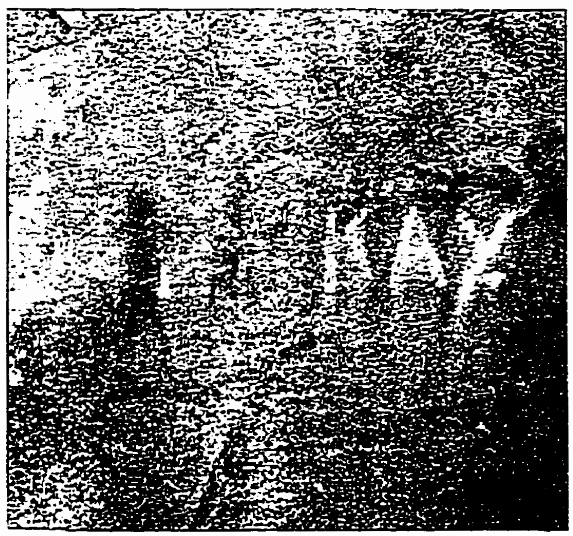
10 (a). Suppose a pictograph (rock painting) looking like **photograph E** was located on the Seagrim Lake route. Would you have visited this route **instead of the actual route you visited on your first trip to Nopiming in 1993 or 1994?**

YES _____ NO _____

10 (b). If your answer was yes above, would you still have changed your trip to the Seagrim Lake route if the rock painting looked like **photograph F** below.

YES _____ NO _____

Photograph E
Rock painting 1



Photograph F
Rock painting 2

Figure 2. The question and photographs used to generate the two stated preference data sets used in the analysis.

Choice Data	Attribute Data
$[RP]$	$[X_{RP}] \sim \begin{bmatrix} 0 \\ \cdot \\ \cdot \end{bmatrix} \sim \begin{bmatrix} 0 \\ \cdot \\ \cdot \end{bmatrix}$
$[SP_P]$	$[X_{RP}] \sim \begin{bmatrix} 0 \\ P \\ 0 \end{bmatrix} \sim \begin{bmatrix} 0 \\ \cdot \\ \cdot \end{bmatrix}$
$[SP_D]$	$[X_{RP}] \sim \begin{bmatrix} 0 \\ \cdot \\ \cdot \end{bmatrix} \sim \begin{bmatrix} 0 \\ D \\ 0 \end{bmatrix}$

Figure 3. An illustration of the stacking of data sets performed in the combined revealed and stated preference analysis of the pictographs.

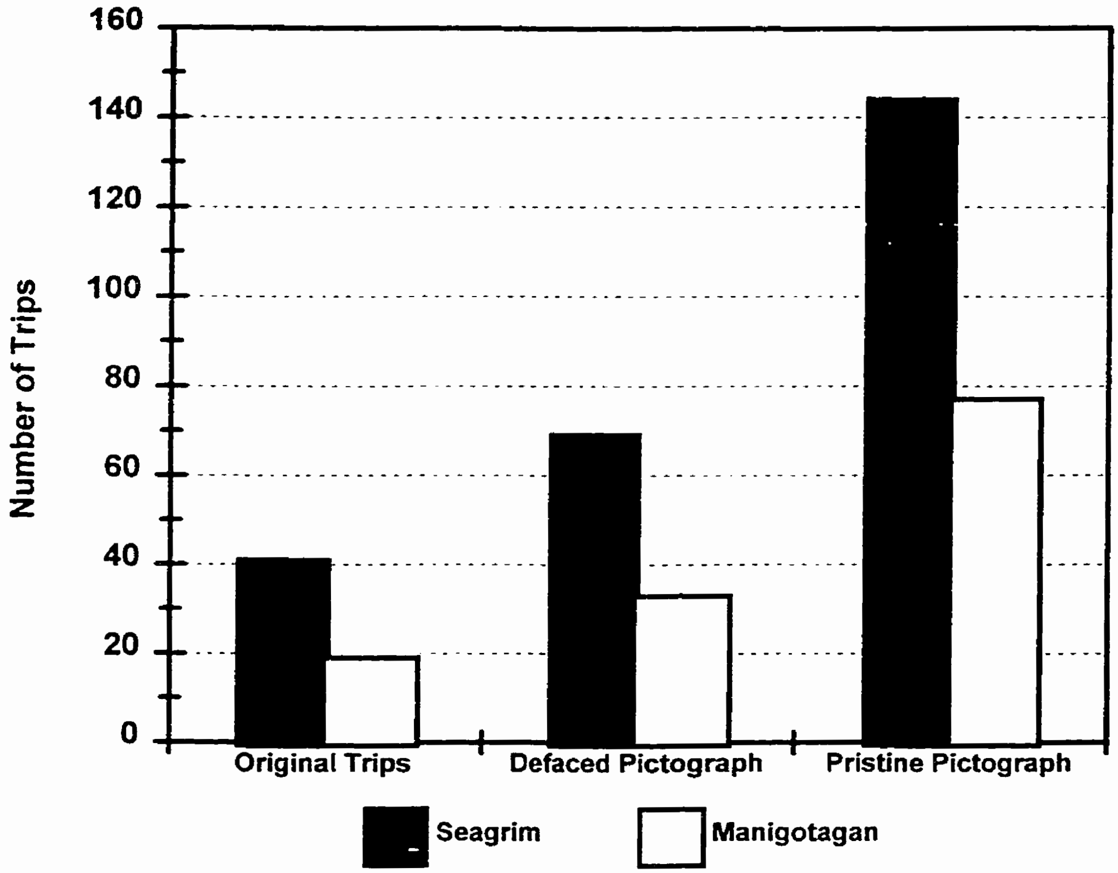


Figure 4. The frequency distribution of trips with and without pictographs at two canoe routes in Nopiming Provincial Park, Manitoba.

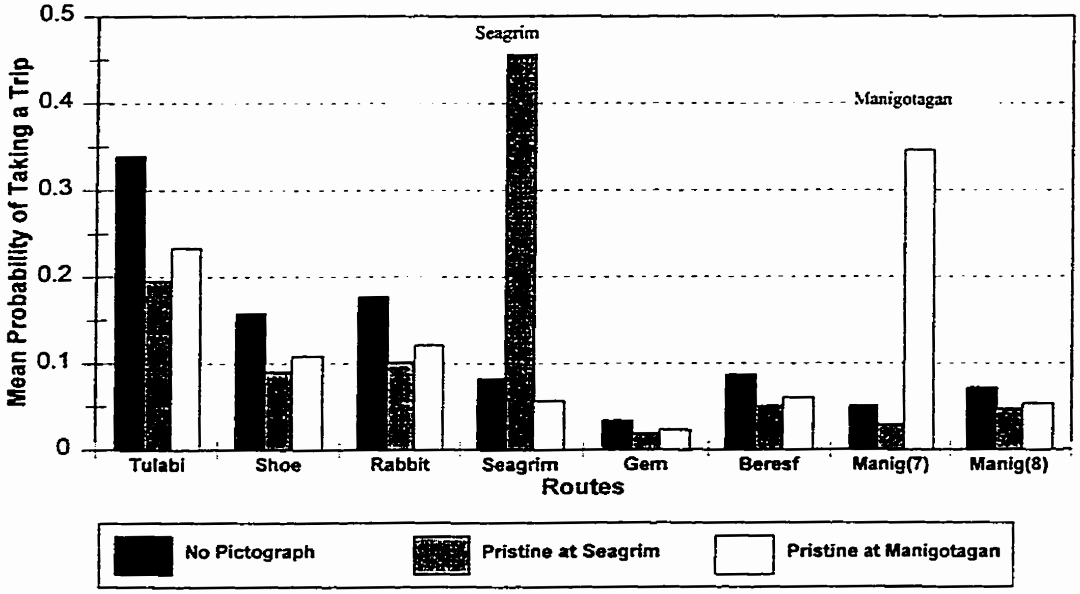
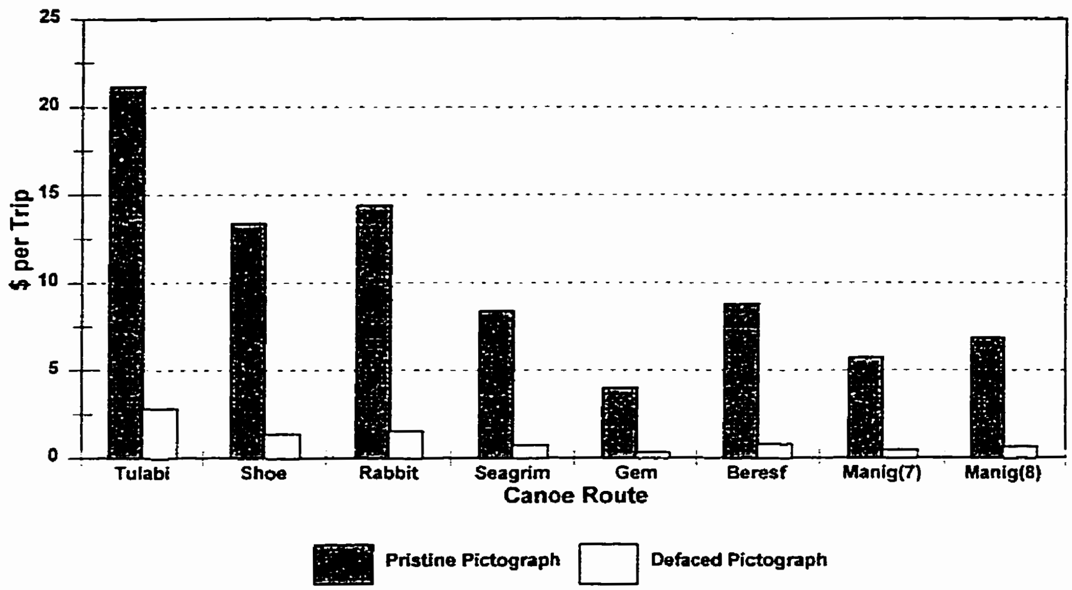


Figure 5. Predicted trip distributions with and without the pristine pictograph at both routes in Nopiming Provincial Park, Manitoba.



Fgure 6. Welfare measures associated with pristine and vandalised pictographs at two canoe routes in Nopiming Provincial Park, Manitoba.

CHAPTER 3

Understanding Heterogeneous Preferences in Random Utility Models: The Use of Latent Class Analysis

INTRODUCTION

Consumer preferences for goods and services are characterized by heterogeneity. Accounting for this heterogeneity in economic analysis will be useful in estimating unbiased models as well as for forecasting demand by including individual characteristics and providing a broader picture of the distribution of resource use decisions or policy impacts. However, many empirical economic analyses assume homogeneous preferences among consumers. Alternatively, previous analyses considered preference heterogeneity *a priori* by: 1) including demographic parameters in demand functions directly or through the utility function (e.g. Pollack and Wales 1992); or 2) by stratifying consumers into various segments and estimating demands separately on each stratum. For these analyses, economists traditionally focus on demographic variables.

There is empirical evidence that these methods identify sources of heterogeneity. For example, Famulari (1995) showed that stratifying households by demographic categories significantly improved tests of consistency with the axioms of revealed preference. Boxall et al. (1996) studied recreation demand in a traditional travel cost model framework[1] and found that stratification reduced the percentage and mean error of

[1] The traditional travel cost model considers price quantity data gathered from a sample of recreationists treating these as if this data were generated by a single consumer maximizing a constrained utility function.

violations of the choice axioms. These studies examined traditional demand analysis, rather than considering the individual choice behaviour typified by the random utility framework. Also there are few economic studies which examine individual-specific variables other than sociodemographic factors.

Heterogeneity is particularly difficult to examine in the random utility model because an individual's characteristics are invariant among a set of choices. In an econometric sense this means that the effect of individual characteristics are not identifiable in the probability of choosing commodities. In essence, the model parameters are the same for each sampled individual implying that different people have the same tastes over model components. These features have been examined by interacting individual-specific characteristics with various attributes of the choices (e.g. Adamowicz et al. 1997). Morey et al. (1993) take advantage of knowledge of income levels by explicitly incorporating them into the indirect utility function of their respondents. These methods are limited because they require *a priori* selection of key individual characteristics and attributes and only involve a limited selection of individual specific variables (e.g. income).

Another set of approaches called random parameter logit/probit models explicitly account in a sense for heterogeneity by allowing model parameters to vary randomly over individuals (e.g. Layton 1996; Train 1997, 1998). While these procedures explicitly incorporate and account for heterogeneity, they are not well-suited to explaining the sources of heterogeneity. In many cases these sources relate to the characteristics of individual consumers.

Two streams of research point to a role for individual-specific characteristics in explaining heterogeneity in choice. The first highlights the possible role of individual characteristics in affecting tastes. For example in Salomon and Ben-Akiva (1983) choice model development was preceded by multivariate cluster analysis of sociodemographic characteristics to determine relatively homogeneous segments of individuals. In this process, the series of choice models estimated separately for each cluster was statistically superior to a model which pooled the clusters.

A second avenue in explaining heterogeneity involves the scale factor described in the previous chapter in equation 3. Cameron and Englin (1997) explain heterogeneity by "parameterizing" this scale in binary logit models. In this case parameterizing heterogeneity in the choice model with demographic variables exhibited superior statistical properties over models which imposed homogeneity. However, it is not clear if the approach of Cameron and Englin in which individual characteristics enter as affecting scale is more appropriate than an alternative approach in which these features influence tastes (i.e. utility parameter differences). In the present paper it is assumed that heterogeneity affects tastes.

In any approach to incorporate heterogeneity into demand analysis there must be *a priori* knowledge of the elements of heterogeneity. Ideally, an effective procedure should utilize theory to provide a foundation for possible sources of heterogeneity. While these sources may include sociodemographics, theory may also point to other characteristics of individuals such as attitudes, perceptions, social influences, and past experiences.[2] Furthermore, while theory may provide an understanding of sources of heterogeneity, it would also be desirable to incorporate heterogeneity in the estimation of economic choice parameters. These features point to joint estimation of the explanators of heterogeneity and the explanators associated with attributes of choices.

A promising avenue for tackling these problems involves the use of latent variables. Latent variable approaches involve the logic that some unobserved or latent variable explains the behaviour of interest, but that one can only observe indicators (in essence other variables) which are functionally related to the latent variable. An analyst using this approach must assume that covariation among a set of observed variables is due to each variable's relationship to the latent variable. That is, the latent variable is actually the true source of the covariation.

The strategy of use of latent variables can be extended to consideration of latent classes. In this approach the latent construct represents a typology, classification, or series of segments which are constructed from a combination of observed constituent variables (McCutcheon 1987). Thus, latent class methods involve characterizing segments from a set of discrete observed measures such as attitudinal scales, or they can involve empirically testing whether a theoretically posed typology adequately fits a set of data (McCutheon 1987:8). This framework, when coupled with information on

---

[2] These individual features are commonly used in the marketing, transportation, and tourism literatures to define various market segments (Wind 1978). These types of features are usually referred to as psychographics.

preferences relating to consumer choice, offers an opportunity to both understand and incorporate preference heterogeneity in consumer demand analysis.

The Latent Segmentation Approach

McFadden (1986) recognized the prospect of using latent variables in understanding choice behaviour. He posed an integration of information from choice models with attitudinal, perceptual and socioeconomic factors using a latent variable system. While the observable outputs using this approach are predictions of choice or market behaviour, the underlying constructs of the choice decision process are more elaborate than traditional consumer demand theory. McFadden (1986) mentions that " the critical constructs in modeling the cognitive decision process are *perceptions* or beliefs regarding the products, generalized *attitudes* or values, *preferences* among products, decision protocols that map preferences into choices, and *behavioral intentions* for choice" (McFadden 1986:276). Thus, the problem for an analyst using this approach is to gather psychometric data to quantify the theoretical or latent constructs underpinning choice behaviour and then simulate this choice behaviour using attributes associated with the products of interest.

Swait (1994) utilized McFadden's idea to understand preferences for beauty aids. In this application latent segments were characterized by different degrees of sensitivity to product attributes. Swait utilized brand image ratings from a sample of consumers along eight psychometric dimensions as individual-specific information, and a set of repeated choices of preferred products from among five brands was taken as the choice information. Swait's (1994) model simultaneously conducted market segmentation and predicted choice of beauty product for the sample. This model, called a finite-mixture model in the statistical literature (Titterington et al. 1985), allows market segments to be related to characteristics of individual consumers such as psychographic or socioeconomic effects, but also elements of observed behaviour. This type of model may have considerable relevance to decision-makers in that it allows a degree of understanding of preference heterogeneity through incorporation of individual characteristics. It also accounts for preference heterogeneity to a degree by simultaneously estimating segment specific membership and choice parameters.

This paper applies this latent segmentation approach to a set of wilderness recreation park choice data. The foundation of this application is a model which incorporates motivations towards wilderness recreation and perceptions of environmental quality. The behavioral aspects of this study use information from a choice experiment involving wilderness park choice. In this experiment five environmental and managerial attributes were varied in the design. The analysis will assess simultaneously the influence of individual characteristics, motivational aspects, and the influence of choice-based attributes in the estimation of latent segments.

THE LATENT SEGMENTATION MODEL

In deriving the latent segmentation model, random utility theory is first employed to model choices among a set of substitutes or alternatives on a given choice occasion with each choice occasion assumed to be independent of the others. An individual (n) receives utility, U, from choosing an alternative (i) equal to $U_{ni}=U(X_{ni})$, where X_{ni} is a vector of the attributes of i. Utility is modelled as two components, where one portion is deterministic and depends on the attributes of the alternative, and the remainder is not. Thus, $U_{ni}=V_{ni}+\epsilon_{ni}$ where $V_{ni}= f(X_{ni})$ is the deterministic component and ϵ_{ni} a random component of the utility function.

In this model, individual n faces a choice of one alternative from a finite set C of sites. The probability (π) that alternative i will be visited is equal to the probability that the utility gained from its choice is greater than or equal to the utilities of choosing another alternative in C. Thus, the probability of choosing i is:

$$\pi_n(i) = \text{Prob}\ \{V_{ni} + \epsilon_{ni} \geq V_{nk} + \epsilon_{nk};\ i \neq k,\ \forall\ k \in C\}. \tag{1}$$

The conditional logit model, developed by McFadden (1974), can be utilized to estimate these probabilities if the random terms are assumed to be independently distributed Type-I extreme value variates. Substituting the attributes associated with each alternative into the deterministic portion of utility (V) and selecting a linear functional form allows the choice probabilities take the form:

$$\pi_n(i) = \frac{\exp \mu(\beta X_i)}{\sum_{k \in C} \exp \mu(\beta X_k)} \tag{2}$$

where μ is a scale parameter that is assumed to equal 1, and β is a vector of parameters. Note that in this model the vector β is not specific to an individual.

Now assume the existence of S segments in a population and that individual n belongs to segment s (s = 1,...,S). The utility function can now be expressed $V_{in|s} = \beta_s X_{in} + \epsilon_{in|s}$. In this expression the utility parameters are now segment specific and equation (2) becomes:

$$\pi_{n|s}(i) = \frac{\exp \mu_s(\beta_s X_i)}{\sum_{k \in C} \exp \mu_s(\beta_s X_k)} \tag{3}$$

where β_s and μ_s are segment-specific utility and scale parameters respectively.

Following Swait (1994), consider an unobservable or latent membership likelihood function M^* that can classify individuals into one of the S segments. The classification variables that influence segment membership are related to latent general attitudes and perceptions, as well as socioeconomic characteristics of the individuals. For a specific individual n, this function can be described by the following set of equations:

$$\begin{aligned} M^*_{ns} &= \Gamma_{ps} P^*_n + \Gamma_s S_n + \zeta_{ns} \\ P^*_n &= \beta_P P_n + \zeta_{nP} \end{aligned} \tag{4}$$

where M^*_{ns} is the membership likelihood function for n and segment s; P^*_n is a vector of latent psychographic constructs held by n; S_n is a vector of observed sociodemographic characteristics of individual n; P_n is a vector of observed indicators of the latent psychographic constructs held by n; Γ and β_p are parameter vectors to be estimated; and the ζ vectors represent error terms. Relating this function to the classical latent variables approach where observed variables are related to the latent variable, M^* can be expressed at the individual level as:

$$M^*_{ns} = \lambda_s Z_n + \zeta_{ns}, \; s=1,\ldots,S \tag{5}$$

where Z_n is a vector of both the psychographic constructs (P_n) and socioeconomic characteristics (S_n), and λ_s is a vector of parameters. This classification mechanism allows n to be placed in s if and only if :

$$M^*_{ns} = \max\{M^*_{nj}\}, \; j \neq s, \; s = 1,\ldots,S. \tag{6}$$

As Swait (1994) points out, these membership likelihood functions are random variates and one must specify the distribution of their error terms in order to use them in practice. Thus, following Swait (1994), Gupta and Chintagunta (1993), and Kamakura and Russell (1989), the error terms are assumed to be independently distributed across individuals and segments with Type I extreme value distribution and scale factor α. Incorporating these assumptions allows the probability of membership in segment s to be characterized by:

$$\pi_{ns} = \frac{\exp \alpha(\lambda_s Z_n)}{\sum_{s=1}^{S} \exp \alpha(\lambda_s Z_n)}, \tag{7}$$

This form is the multinomial logit model used by Schmidt and Strauss (1975) in which individual-specific characteristics rather than attributes of choices produce choice probabilities. Other functional forms could be chosen to represent the probability of segment membership. However, the following constraints must be met:

$$\sum_{s=1}^{S} \pi_{ns} = 1; \quad \text{and} \quad 0 \leq \pi_{ns} \leq 1. \tag{8}$$

To further develop the latent segment model define $\pi_{ns}(i)$ as the joint probability that individual n belongs to segment s and chooses alternative i. This can be expressed as the following product of the probabilities defined in equations (3) and (7): $\pi_{ns}(i) = \pi_{ns}$ $\pi_{n|s}(i)$. Thus, the probability that a randomly chosen individual n chooses i is given by:

$$\pi_n(i) = \sum_{s=1}^{S} \pi_{ns} \, \pi_{n|s}(i), \tag{9}$$

and substituting the equations for the choice (equation 3) and membership (equation 7) probabilities yields the expression:

$$\pi_n(i) = \sum_{s=1}^{S} \left[\frac{\exp \alpha(\lambda_s Z_n)}{\sum_{s=1}^{S} \exp \alpha(\lambda_s Z_n)}\right]\left[\frac{\exp \mu_s(\beta_s X_i)}{\sum_{k \in C} \exp \mu_s(\beta_s X_k)}\right]. \qquad \textbf{(10)}$$

This model allows the use of both choice attribute data and individual consumer characteristics to simultaneously explain choice behaviour. Note that the expression contains two types of logit formulations; one is a multinomial logit model which includes the segment membership parameters and the other is a conditional logit model which contains the segment specific utility parameters. Because these two formulations are mixed together this model is considered to be a mixed logit model in the literature (e.g. Titterington et al. 1985).

A number of features of this model are noteworthy. First, the observation that the ratio of probabilities of selecting any two alternatives (equation 10) would contain arguments that include the systematic utilities from other alternatives in the choice set is of note. This is the result of the probabilistic nature of membership in the elements of S. The implication of this result is that independence from irrelevant alternatives need not be assumed (Shonkwiler and Shaw 1997).

Second, there are two types of scale factors which cannot be estimated simultaneously. The α scale factor represents the scale across the segment membership function and as such is not identifiable. The μ_s terms denote the scale for the sth segment's utility function and in theory can be used to test hypotheses about scale and utility parameter equality across segments (Swait and Louviere 1993). These scale factors are only identifiable under conditions where the segment specific utility parameters are constrained to be equal (e.g. Adamowicz et al. 1997). However, this assumption of parameter equality across segments is contrary to the spirit of the latent segment model used here since a researcher would **not** want to impose utility parameter equality. Therefore, utilizing this model in empirical estimation requires that all of the scale factors in (10) are set equal to one.

Third, as Swait (1994) points out when $\lambda_s = 0$, $\beta_s = \beta$, and $\mu_s = \mu$ for each segment, equation (10) reduces to the conditional logit model in shown in (2). These conditions essentially impose homogeneity of preferences and are represented by the case in which there exists only one segment in which every individual in the data holds membership. Conversely, one could consider the case where each individual in a set of data can be considered a segment. Under this condition each respondent behaves as if their behaviour is consistent with a conditional logit model, but each individual has their own set of parameters. This situation can be represented by the random parameter logit/probit models (e.g. Layton 1996; Train 1997, 1998). Thus, the latent segmentation model represents a model located within a range of approaches. On one end of the range is the single segment case which assumes perfect homogeneity of preferences. On the other end is the case where each individual is considered a segment in which heterogeneity of preferences is, in a sense, completely accounted for. The potential advantage of the latent segment model in this series of approaches is its potential to explain and account for heterogeneity to some degree.

AN APPLICATION - WILDERNESS PARK CHOICE IN CENTRAL CANADA

A Framework for Wilderness Recreation Decisions

In understanding the selection of wilderness areas for recreation trips, a framework of choice and segmentation was developed based on the path diagrams in McFadden (1986) and Swait (1994). The framework, shown in Figure 1, incorporates latent constructs in boxes shaded with grey while the white boxes represent observable variables. This model utilizes psychographic features that relate to motivations for taking a wilderness recreation trip. The observable motivational indicators are related to latent motivations, and these, in concert with an individual's sociodemographic characteristics, influence the likelihood of membership in one or more latent classes or segments. When observable motivational indicators are available, this part of the framework can be represented by equation (7).

The other components in Figure 1 are related to the attributes of the available wilderness choices and consist in part of actual or objective characteristics of the places one could choose to go. However, some of these characteristics may be influenced by past visits, contact with media, levels of wilderness experience etc., and these elements may result in the formation of perceptions of wilderness features. Perceptions of attribute qualities have been revealed as an important influence in choice behaviour by Adamowicz et al. (1997). Both objective and subjective components of wilderness choice attributes, along with sociodemographic characteristics, may influence wilderness recreation preferences. This part of the framework is represented by equation (3).

Putting the psychographic and sociodemographic characteristics together with the objective and subjective wilderness attributes enables the implementation of the latent segmentation model. Thus, the decision protocol is represented through equation (10). The result of the model is the probability of choosing a wilderness area from available wilderness choices. A final set of influences on this choice, however, result from exogenous features such as the closure of wilderness areas due to forest fires or other stochastic events.

Empirical Application

This framework for understanding wilderness park choice was applied to recreationists who use a set of five wilderness parks in eastern Manitoba (Nopiming and Atikaki Provincial Parks), western Ontario (Woodland Caribou, Quetico and Wabakimi Provincial Parks) and northern Minnesota (Boundary Waters Canoe Area (BWCA)). Recreational use of these parks has been considered a demand system in previous research (Boxall et al. 1999; Englin et al. 1998) indicating that a sample of visitors to these areas would consider them as elements of a recreation choice set. The parks represent a range of development, entry restrictions, congestion levels, and management intervention. They cater to a relatively heterogeneous market and have a number of management issues which require knowledge about the characteristics of people who use them and the "products" or features desired for recreation trips. Thus, the application of the latent segment model to visitors in these areas would be of considerable value to park managers.

During 1995 a sample of 1000 visitors to Nopiming and Atikaki Provincial Parks in Manitoba, and Woodland Caribou, Quetico, and Wabakimi Provincial Parks in Ontario were drawn from park registrations or on-site registrations administered by the Canadian Forest Service.[3] About 71% of individuals in this sample were from Quetico, about 18% were from Woodland Caribou, 10% were from both Manitoba parks, and about 1% were from Wabakimi. This distribution was selected because it approximately represented the levels of visitation across the five parks (see Boxall et al. 1999).

A questionnaire was developed that gathered information about opinions of wilderness management, levels of past visitation to all of the parks, descriptions of a typical wilderness trip, and sociodemographic characteristics. Three additional pieces of information were collected that were used in the latent segment model. The first involved a series of 20 statements which represented reasons why the individual visited backcountry or wilderness areas. Respondents were asked to rate the level of importance of each statement on a 5 point Likert scale ranging from "Not at all important" to Very important." The statements used for this purpose were derived from research by Crandall (1980) and Beard and Ragheb (1983) on leisure motivations. The scores of the respondents were used to derive a scale to measure motivations for visiting wilderness areas.

---

[3] Registrations from the Boundary Waters Canoe Area were not available and thus visitors to this park could not be included in the sample.

The second was the application of a choice experiment which required respondents to consider choosing among five wilderness areas for a trip next season, or the option of not taking a trip. The choice experiment employed the actual park names as choice options (hence a "branded" choice) where the two Manitoba parks were combined into one and the Boundary Waters Canoe Area was available as one of the choices. Five attributes each consisting of four levels were developed based on three years of previous research on wilderness recreation in the area, and discussions with park managers, recreationists, and academics. These attributes were: (1) the fee per day per person; (2) the chances of entry into the park as a result of entry or quota restrictions; (3) the type of campsite available; (4) the level of development related to human habitation and access; and (5) the total number of encounters with other wilderness recreation groups per day. These attributes and their levels are described in Table 1.

A choice scenario (Figure 2) consisted of six alternatives (five parks and the stay at home option).[4] Statistical design methods were used to structure the presentation of the levels of the five attributes in the scenario. In this presentation the levels of attributes of one alternative (the BWCA) were held constant while those of the other four parks were varied in the design. The attributes and levels for these four parks were constructed from a $4^5 \times 4^5 \times 4^5 \times 4^5 \times 2$ orthogonal main-effects design, yielding 64 possible combinations of the levels (or choice sets). This number was considered too large a task for a respondent to complete so the 64 combinations were blocked into 8 versions of the

[4] Note that this design incorporates two choice questions. Only the first choice was analyzed in this study.

questionnaire with eight choice scenarios presented in each version.

The third body of information from the questionnaire was the answers to a series of questions aimed at gathering respondents' current perceptions of the levels of the attributes in each park. These levels and attributes were the same as those described in Table 1.

The questionnaire was mailed to the sample of 1000 recreationists. After adjusting for non-deliverables, the response after one post card reminder and a second follow-up questionnaire was 80%. Further adjustment of the respondents for item non-response resulted in a final sample of 620 individuals who provided complete data for the measurement of motivations, sociodemographic characteristics, and information on 4892 choices.

Econometric Model

The first step in developing the latent segmentation model involved the analysis of the motivational indicators. This entailed a factor analysis of the 20 statements on reasons for taking a trip. The factor analysis provided estimates of the latent motivational constructs which enter the membership likelihood function. Because these statements were developed *a priori* to assess motivations, this involved a confirmatory approach. The scores from the 20 statements were factor analyzed using principal component analysis with varimax rotation. Components were extracted until eigenvalues were less than or equal to 1.0.

The factor analysis identified four components of motivations for taking a wilderness trip which accounted for virtually all of the variation. These motivational components were labeled based on magnitudes of the loadings of individual statements shown in Table 2. The first component was called "challenge and freedom" because statements relating to this factor loaded highly in this factor (shaded grey in Table 2). The second factor was labeled "nature appreciation". The third factor involved statements relating to family and friends and was labeled "social relationships". The fourth factor was called "escape from routine".

Scores for the four factors were then calculated for each individual in the sample yielding four variables to be included in the Z_n vector in equation (10). An additional variable added to this vector was a dummy variable which equaled one if a respondent's trip length typically was 3 or less days. This variable was selected to capture sociodemographic effects that may influence trip characteristics and that may not be related to the factor scores. Other sociodemographic features of respondents could have been chosen for inclusion in this vector, but the complexity of the model and the degree of estimation required limited the set of variables for inclusion. However, to explore the role of sociodemographic features in segment membership, a posterior analysis of the characteristics of latent segments was performed. This will be described below. Thus, in summary five variables and an intercept were included in the Z_n vector.

The X_i vector consisted of the attributes associated with the parks presented in the choice task. These variables entered the latent segment model through their impact on the utility function. Recall that there are five attributes, each with four levels. The attributes were effects coded as described by Louviere (1988) and Adamowicz et al. (1994).

Estimation of the λ_s and β_s parameter vectors was performed via maximum likelihood in GAUSS using the BFGS algorithm. For the λ_s vector, the parameters for one of the segments must be normalized to zero to permit identification of segment membership parameters for the other segments. The log likelihood function was:

$$\ln L(\lambda, \beta \mid S) = \sum_{n=1}^{N} \sum_{\forall m} \delta_{ni} \ln\left[\sum_{s=1}^{S} \frac{\exp \alpha(\lambda_s Z_n)}{\sum_{s=1}^{S} \exp \alpha(\lambda_s Z_n)} \times \frac{\exp \mu_s(\beta_s X_i)}{\sum_{i=1}^{6} \exp \mu_s(\beta_s X_i)}\right] \quad \textbf{(11)}$$

where N refers to the 620 individuals who provided complete information, m represents the 4892 choice sets for which choice data were provided, i represents the alternatives from the choice experiment, and δ_{ni} equals 1 if individual n chose i and 0 otherwise. The other symbols are described above. In this procedure independence was assumed across the set of choices from each respondent, and the scale parameters (α and μ_s) were set equal to 1.

In estimating latent segment models the number of segments, S, cannot be defined. Thus, S must be imposed by the investigator and statistical criterion must be used to select the "optimal" number of segments in a set of estimations where the number of segments imposed varies in each estimation. At issue in this process is that while one expects improvement in the log likelihood values as additional segments are added to the model, the model fits must be "penalized" for the increase in the number of parameters that are added due to additional segments. Thus, following Kamakura and Russell (1989), Gupta and Chintagupta (1994) and Swait (1994) three criteria were used to assist in determining the size of S. These were: the minimum Akaike Information Criterion (AIC), the maximum Bayesian Information Criterion (BIC) (Allenby 1990), and the maximum of a modification of McFadden's ρ^2 called the Akaike Likelihood Ratio index (or ρ bar^2)(see Ben-Akiva and Swait 1986). Their calculation is shown in the first row of Table 3. As Swait (1994) describes, these criteria should be used as a guide to determine the size of S; conventional rules for this purpose do not exist and judgement and simplicity play a role in the final selection of the size of S.

RESULTS AND DISCUSSION

Choosing the Number of Segments

In estimating the models, 1, 2, 3, 4, 5 and 6-segment solutions were attempted. Table 3 summarizes the aggregate statistics for these models as well as a single segment model. The log likelihood values at convergence (column 3) reveal improvement in the model fit as segments are added to the procedure, particularly with the 2, 3, and 4 segment models. This is evident in the ρ^2 values which increase from the base of 0.197 to 0.244 with the 6 segment model. This information supports the hypothesis of the existence of latent segments, but does not suggest how many segments are in the data.

The other statistics in Table 3 must be inspected to answer this question.

Inspection of columns 5 to 7 in Table 3 support four segments as the optimal solution in the data. First, while the AIC values grow smaller as the number of segments increases, the change in AIC is markedly smaller for the 4- to 5-segment and 5- to 6-segment solutions than the 2- to 3- and 3- to 4-segment solutions. Second, the ρbar^2 statistics exhibit a similar pattern in that improvement in the values is reduced beyond the 4-segment model. Finally, the maximum BIC statistic is clearly associated with the 4 segment model. It is noteworthy that the BIC values fall when additional segments are added to the 4-segment solution.

Characterizing the Segments

The segment membership (λ_s) parameters for the 4-segment solution are displayed in Table 4. Note that the parameters for the first segment are equal to 0 which results from their normalization during estimation. Thus, the other three segments must be described relative to this first segment. Segment 2 was labeled "weekend challengers" because the dummy variable on short or weekend-long trips was relatively large and positive, and the parameter on motivations relating to challenge and freedom was the same. For segment 3 the short trip dummy was close to zero, but the variable on motivations relating to nature appreciation was positive and was the largest over all 4 of the segments. For this reason, this segment was labeled "nature nuts". Segment 4 was classified as "wilderness trippers" because the short trip dummy variable was large and negative. Finally segment 1 was labeled "escapists" due to the fact that the motivational factor on escape from routine was negative for the other 3 segments. Despite the labels, however, the diversity of influences on segment membership is striking. For example, motivations relating to social relationships are positive for one segment, but negative for two others. Only for escape from routine and nature appreciation are the directions of the effects similar across segments.

The utility function parameters (β_s) for the 4-segment solution are displayed in Table 5. Also shown are the parameters for a 1 segment solution for comparison. The parameters on entry fees are negative for each segment which is consistent with economic theory. Parameters for the chances of entry are variable across the segments, suggesting that this effect is characterized by heterogeneity. The 4 segment model implies that weekend challengers and wilderness trippers would seek parks with high chances of entry, while escapists and nature nuts prefer areas with low chances of entry. Individuals in these latter segments might choose places with lower chances of entry because these areas may offer the experiences they are seeking due to the restrictions on the number of visitors. These effects can be compared to the single segment model in which suggests all individuals would prefer areas with high chances of entry.

The utility parameters associated with the campsite type, levels of development and encounter variables were plotted in Figure 3 to show the differences among the segments. These plots identify the sensitivity of segments to changes in the levels of these three sets of variables. Campsite type seems not to be a large influence on park choice for any of the four segments. Yet it is noteworthy that in the single segment model, two of the campsite parameters are not statistically significant while these parameters are significant in each of the four segments. However, development and encounter levels appear to have an important effect on choice behaviour. Recreationists in segment 3 (nature nuts) would be strongly negatively affected by higher levels of development. Wilderness trippers (segment 4) would be more negatively affected by higher levels of encounters than the other 3 segments.

A final set of utility parameters result from the alternative specific constants (ASCs) used to identify the 5 parks (brands) in the choice experiment. These parameters are shown in Table 5. Recreationists in segment 1 strongly prefer Quetico followed by the BWCA and Woodland Caribou parks. Segment 2 individuals, the weekend challengers, only prefer the Manitoba parks and would tend to avoid the other four parks, all other things being equal. Segment 3 individuals exhibit negative parameters for all five parks suggesting that they may prefer parks not included in the choice experiment or are more likely to participate in other activities. This seems to be an odd result and will be addressed further below. Finally, individuals in segment 4 exhibit higher utility, all else being equal, for the two Ontario parks. These individuals also exhibit a negative association with the Boundary Waters. Once again, comparison of the alternative specific constants with the single-segment model suggests that the simpler model would not capture sources of heterogeneity associated with the latent segment model.

Since segment membership parameters (λ_s) were jointly estimated with the utility parameters (β_s), one should expect consistent behavioural relationships among the two parameter vectors. These features appear to be present in this dataset. For example, the trip choices of weekend challengers should be positively influenced by recreation areas with higher chances of gaining entry; nature nuts are more likely to avoid areas with high levels of development; and wilderness trippers should seek areas where few other recreationists would be encountered. Thus, in this empirical example, the latent segment model appears to have identified sources of heterogeneity in recreation site choice and to have incorporated this by identifying different utility functions.

The role of sociodemographic characteristics in explaining latent class membership (from Fig. 1) was explored by computing segment membership probabilities for each individual and then regressing these four probabilities against the individuals' characteristics.[5] In this procedure, the method of Bucklin and Gupta (1992) was used in which the probabilities were transformed by the following formula: $\ln(\pi_s / 1-\pi_s)$. It must be recognized that the error terms for the four regressions would be correlated and that seemingly unrelated regression (SUR) methods should be used to improve the efficiency of the parameter estimates. However, this would require the addition of parameter restrictions which requires theoretical justification. This is a topic for future research and thus equation by equation OLS methods were used to provide an illustration of the role of these individual characteristics.

The results of these regressions (Table 6) identify that high levels of experience in

---

[5] An argument can be made for including these characteristics in the membership likelihood function. However, the computational effort required for adding these variables was beyond the scope of the resources available. Thus in this example it is assumed that these individual features are correlated with the variables included in the membership function (Table 4).

wilderness recreation are associated with the escapists, but that low levels of specialization in canoeing are associated with the weekend challengers. Residency in the USA is associated with the weekend challengers. Other factors such as household size, age, education levels also are associated with the various membership probabilities.

In order to complete the application of the framework proposed in Fig.1 for the wilderness recreation data, estimates of park choices were calculated. This required knowledge of the attributes of the parks and placement of individuals in the segments. First, segment membership probabilities were computed for each of the 620 individuals using equation (7). Each individual was assigned to one of the four segments based on the largest probability. This assignment method determined that 41.4% of the sample were members of segment 1, 7.3% were members of segment 2, 0.8% were members of segment 3 and the remaining 50.4% were assigned to segment 4. Thus, escapists and wilderness trippers dominate the sample of wilderness recreationists.

Second, the levels of the attributes of the five parks were determined using individual's perceptions of their attributes as outlined in Figure 1. For this, indicators of perceptions of campsite types, levels of development, and numbers of encounters were utilized from the questionnaire. The questions used to collect this information, and a summary of the results, are shown in Appendix 1. For fees and chances of entry, the objective levels of these were used. For the majority of the 620 recreationists, the objective and perceived levels of these two variables were identical. These indicators provide linkage to the latent perceptions as outlined in the model (Fig.1).

Figure 4 displays the predicted distribution of wilderness recreation choices among the five parks by segment. Individuals in segments 1 and 3 would be more likely to visit Quetico than the other two segments. Those in segment 2 would be more likely to visit the Manitoba Parks. The other parks appear to be less attractive to these members of segment 2, particularly the Boundary Waters and Wabakimi. These findings have implications for identifying the relevant choice sets across the segments, but these results are not as strong as those reported by Swait (1994).

Welfare Measures in the Latent Segment Model

One of the roles of recreation economic choice models is to examine the welfare implications of environmental or management changes. McFadden (1996) outlines the theory required for deriving welfare measures using conditional logit models. In what follows, this theory is applied to the latent segment model in two ways. The first involves the derivation of welfare measures on a segment by segment basis. In this case, the distributional impacts of policies can be understood. However, computing these welfare measures requires that respondents be assigned to a segment. The second way of applying this theory involves, in a sense, correcting the standard aggregate procedure, which assumes homogeneous preferences, for heterogeneity. Using this method, welfare measures are computed segment by segment for each individual, and these are then weighted by the segment membership probabilities and summed to compute a total welfare measure.

McFadden (1996) and Hanemann (1982) show that the expected utility on any given choice occasion is the sum of utility gained from each choice times its respective probability of being chosen. Thus, measuring a change in welfare associated with decreasing some attribute in the indirect utility function involves estimating the amount individuals must be compensated to remain at the same utility level as before the decrease. The following formula from Hanemann (1982) shows this calculation under the assumption of no income effects:

$$CV_n = \frac{1}{\gamma}[\ln(\sum_{k\in C}\exp(\beta X_k^0)) - \ln(\sum_{k\in C}\exp(\beta X_k^1))] \qquad \textbf{(12)}$$

where CV_n is the compensating variation for individual n, γ is the marginal utility of income, βX_k represents the indirect utility function over k choices, the 0 superscript refers to the initial state and the 1 superscript refers to the new state following some change in an attribute in X in at least one of the choices in k. Applying this formula to understand the distribution of welfare effects across segments necessitates the incorporation of segment-specific utility parameters and the assignment of individuals to segments. Hence:

$$CV_{n|s} = \frac{1}{\gamma_s}[\ln(\sum_{k\in C}\exp(\beta_s X_k^0)) - \ln(\sum_{k\in C}\exp(\beta_s X_k^1))]. \qquad \textbf{(13)}$$

Employing this further to generate an aggregate welfare measure, weighted by segment membership, can be calculated by:

$$= \sum_{s=1}^{S} \pi_s \left(\frac{1}{\gamma_s}\left[\ln\left(\sum_{k \in C} \exp(\beta_s X_k^0)\right) - \ln\left(\sum_{k \in C} \exp(\beta_s X_k^1)\right) \right.\right. \quad (14)$$

where π_s refers to the probability of membership is segment s.

The parameters on fees (Table 5) were chosen as the marginal utility of income parameter (γ). This choice was based on the fact that the distances between recreationists' homes and each of the five parks were not significant in explaining park choice in exploratory analyses of the choice experiment data. In turn, this was probably the result of the alternative specific constants in the model confounding the distance parameter. Thus, fees were considered the most appropriate price variable associated with a trip in this sample.

Two welfare simulations were conducted. The first involved the hypothetical closure of Quetico Provincial Park. This scheme, while hypothetical, is not far-fetched as the portions of the park can be closed during severe forest fires and in some cases entry can be completely prevented. The second simulation involved increasing congestion levels at each park, one at a time, and at all parks simultaneously. This scenario is related to the possibility that demand for experiences in these areas is increasing (Boxall et al. 1999) and would result in increasing levels of visitation and encounters between recreation parties in the backcountry. In both scenarios the base levels for the attributes in the utility function involved the actual levels of fees, objective assessments by park managers of the chances of entry, and the modal perceptions of the three wilderness attributes used in estimating the park choices shown in Figure 4.

The welfare impacts of these changes are shown in Table 7 for a representative recreationist in the sample for the single segment model and in each segment for the 4 segment model. In these simulations equation (12) was used for the former and equation (13) for the latter. The results highlight the limitations of single segment models in understanding the distribution of welfare impacts. For example, the closure of Quetico has a larger impact on members of segments 1 and 3, and a relatively minor impact on members of segment 2 in comparison to the single segment case.

Simulated increases in congestion also suggest distributional effects not captured by the simple model. Increasing congestion at individual parks illustrates the effect of segments and substitution among the parks in the choice set. As a result the welfare differences between the two models are not remarkable except for those segments which exhibit preference for the park in which congestion is changed. However, the simulation for all parks highlights the effects of segmentation alone. In this case impacts are estimated at \$-18.44/trip for the simple model, but the latent segmentation model suggests that the negative impacts of this scenario on wilderness trippers would be almost twice as much (\$-30.04/trip). It would be about half as much for escapists (\$-7.27) and nature nuts (\$-8.80).

The weighted welfare measure (equation (14)) was examined by extracting a sub-sample of 17 individuals from the sample who provided complete information on the perceptions of campsite type, development, and congestion for all five parks[6]. The mean welfare loss for the closure of Quetico was estimated to be \$-9.44/trip/person in this sub-sample and the individual welfare measures ranged from \$-20.85 to \$-3.89. The single segment welfare measure estimated the welfare loss for the same group of individuals at \$-8.78/trip and the range was \$-18.59 to \$2.27. Thus, in this empirical examination failure to incorporate heterogeneity in the welfare measure associated with the closure policy would probably underestimate the value of the loss to the wilderness recreationists.

CONCLUSIONS

This paper was motivated by the need to simultaneously incorporate and explain sources of heterogeneity in random utility models. Current approaches in the literature involve simple parameterizations of the scale factor in conditional logit models or the random coefficients logit method proposed by Train (1998) and others. An alternative model proposed here involves the use of latent class analysis in concert with the conventional random utility structure to explain choices. This latent segment model simultaneously groups individuals into relatively homogeneous segments and explains the choice behaviour of the segments. A major advantage of this latent segment approach may be its ability to enrich the traditional economic choice model by including

---

[6] These 17 individuals were chosen because they reported complete information for all of the required explanatory variables. These people did not appear to be a unique group in the sample. The mean (SD) probabilities of membership in each of the 4 segments over these 17 people were: 0.33 (0.11), 0.11 (0.16), 0.18 (0.08), and 0.38 (0.11) respectively. The max/min probabilities for each segment were 0.59/0.14, 0.62/0.001, 0.36/0.06, and 0.57/0.15.

psychological factors. However, this integrated modeling strategy also offers an opportunity to merge various social psychological and economic theories in explaining behaviour.

To illustrate these features, a latent segment model was developed and applied to recreation demand in a set of wilderness parks by a sample of 620 people. The theoretical basis for this involved the incorporation of sociodemographics and latent constructs relating to motivations in describing segment membership. These constructs were integrated using indicators derived from survey responses related to reasons for taking a wilderness trip. The development of the utility function involved recreation site choice attributes which were examined in a choice experiment.

The results from this integrated approach provided a much richer interpretation of wilderness recreation site choice behaviour than a traditional single segment model (which assumed homogeneity of preferences). For these data the latent segment approach suggested that heterogeneity was related to the motivational constructs underlying wilderness trips, sociodemographic characteristics, preferences for specific wilderness parks (holding changes in their characteristics constant), and perceptions of managerial attributes and congestion levels at the five parks. These findings support both economic and social psychological constructs related to wilderness recreation behaviour.

The latent segment model may be at a considerable advantage in adding to understanding the distribution of the effects of management policies among members of a population. To illustrate this, three welfare measures were developed. The first is the form frequently used in the empirical literature (Hanemann 1982) in which homogeneous preferences are assumed. This welfare measure is relevant for the single segment case. The other two welfare measures were variants of this case and explicitly included segment differences. One measure was employed to assess welfare impacts in each segment and the results were used to examine the distributional impacts of policies across segments. The other welfare measures utilized the probability of segment membership and used this probability to adjust the weights of the segment welfare impacts and generated a single welfare measure. The resulting welfare measures from these latter two approaches were quite different than the single segment case.

The empirical application of latent segmentation to the wilderness data suggests that this method holds considerable promise in understanding recreation choice behaviour. The method may be even more useful when applied to other types of recreation data, for example those in which the participants are more heterogeneous than are the wilderness recreationists examined in this study.[7] Regardless of the application, however, the underlying theory which incorporates latent psychographic constructs must be relevant to the activity being studied, and the indicator variables used to describe these constructs have acceptable explanatory power (see Ben-Akiva et al. 1997). The recreation literature abounds with theoretical and empirical studies on attitudes, perceptions, and motivations and would appear to offer fertile ground for further applications of the latent segment approach. For example, the success of the empirical

[7] An example of this heterogeneity may be participation in automobile camping in which equipment preferences, social and environmental settings, and facilities may drive site choice behaviour.

application in this paper was related to prior existence of suitable instruments (see Beard and Ragheb 1983) to measure motivations for taking a trip.

A major challenge in the use of choice models incorporating psychographic information is out-of-sample prediction. This is a problem because one generally does not know nor can predict the answers to attitudinal questions from those outside of the sample. In recreation contexts involving managed areas like parks, however, there is usually considerable information on the number and types of visitors who visit these areas. In these cases it may be possible to construct attitudinal instruments which, in concert with socioeconomic and experiential information, may be generalized to the recreation population of interest. In essence what is required is reasonable confidence in allocating out-of-sample individuals to segments and then using the segment-specific choice parameters to predict their behaviour.

In the case of broader issues in which prediction to a more general population is of interest, the use of psychographic information may be problematic. Successful out-of-sample prediction in these instances will require the development of attitudinal questions and sufficient understanding of the answers to these before out-of-sample individuals can be allocated to segments with confidence. This represents a considerable challenge to social science research agendas. In the absence of this kind of knowledge analysts may have to rely on the traditional sets of socioeconomic variables to understand membership in segments and their behaviour.

LITERATURE CITED

Adamowicz, W.L., J. Louviere, and M. Williams. 1994. Combining stated and revealed preference methods for valuing environmental amenities. *Journal of Environmental Economics and Management* 26:271-296.

Adamowicz, W.L., J. Swait, P.C. Boxall, J. Louviere, and M. Williams. 1997. Perceptions versus objective measures of environmental quality in combined revealed and stated preference models of environmental valuation. *Journal of Environmental Economics and Management* 32:65-84.

Allenby, G. 1990. Hypothesis testing with scanner data: The advantage of Bayesian methods. *Journal of Marketing Research* 27:379-389.

Beard, J.G. and M.G. Ragheb. 1983. Measuring leisure motivation. *Journal of Leisure Research* 15:219-228.

Ben-Akiva, M. and J. Swait. 1986. The Akaike likelihood ratio index. *Transportation Science* 20:133-136.

Ben-Akiva, M., J. Walker, A.T. Bernardino, D.A. Gopinath, T. Morikawa, and A. Polydoropoulou. 1997. Integration of choice and latent variable models. Paper presented at the 1997 IATBR, University of Texas at Austin.

Boxall, P.C., W.L. Adamowicz, and T. Tomasi. 1996. Nonparametric tests of the traditional travel cost model. *Canadian Journal of Agricultural Economics* 44:183-193.

Boxall, P.C., J. Englin and D.O. Watson. 1999. Valuing backcountry recreation in wilderness parks: A demand systems approach in the Canadian Shield. *Information Report* NOR-X-361, Northern Forestry Centre, Canadian Forest Service, Edmonton, Alberta.

Bucklin, R.E. and S. Gupta. 1992. Brand choice, purchase incidence and segmentation: An integrated modeling approach. *Journal of Marketing Research* 20:201-215.

Cameron, T.C. and J. Englin. 1997. Respondent experience and contingent valuation of environmental goods. *Journal of Environmental Economics and Management* 33:296-313.

Crandall, R. 1980. Motivations for leisure. *Journal of Leisure Research* 12:45-54.

Famulari, M. 1995. A household-based, nonparametric test of demand theory. *Review of*

Economics and Statistics :372-382.

Hanemann, W.M., 1982. Applied welfare analysis with qualitative response models, Working Paper. No. 241. University of California, Berkeley, 26 pp.

Gupta, S. and P.K. Chintagunta. 1994. On using demographic variables to determine segment membership in logit mixture models. *Journal of Marketing Research* 31:128-136.

Kamakura, W. and G. Russell. 1989. A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research* 26:379-390.

Layton, D.F. 1996. Rank-ordered, random coefficients multinomial probit models for stated preference surveys. Paper Presented at the 1996 Association of Environmental and Resource Economists Workshop, June 2-4, Tahoe City, California.

McCutcheon, A.L. 1987. *Latent Class Analysis*. Sage University Papers Series: Quantitative Applications in the Social Sciences Number 07-064. Sage Publications, Newbury Park, California.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. In: P. Zarembka (Editor), *Frontiers in Econometrics*. Academic Press, New York, pp. 105-142.

McFadden, D. 1986. The choice theory approach to market research. *Marketing Science* 5:275-297.

McFadden, D. 1996. Computing willingness to pay in random utility models. *Energy and Resource Economics* (forthcoming).

Morey, E.R., R.D. Rowe, and M. Watson. 1993. A repeated nested-logit model of Atlantic salmon fishing. *American Journal of Agricultural Economics* 75:578-592.

Pollack, R.A. and T.J. Wales. 1992. *Demand System Specification and Estimation*. Oxford Univ. Press, New York.

Salomon, I. and M. Ben-Akiva. 1983. The use of the life-style concept in travel demand models. *Environmental and Planning A* 15:623-638.

Schmidt, P. and R. Strauss 1975. The predictions of occupation using multinomial logit models. *International Economic Review* 16:471-486.

Shocker, A.D., M. Ben-Akiva, B. Boccara, and P. Nedungadi. 1997. Consideration set influences on consumer decision-making and choice: Issues, models, and suggestions. *Marketing Letters* 2:181-197.

Shonkwiler, J.S. and W.D. Shaw. 1997. Shaken, not stirred: A finite mixture approach to analysing income effects in random utility models. Paper Presented at the 1997 Annual Meeting of the American Agricultural Economics Association. August 2-4, Toronto, Ontario.

Swait, J.R. 1994. A structural equation model of latent segmentation and product choice for cross-sectional revealed preference choice data. *Journal of Retailing and Consumer Services* 1:77-89.

Swait, J.R. and J.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.

Titterington, D.M., A.F.M. Smith, and U.E. Makov. 1985. *Statistical Analysis of Finite Mixture Distributions*. John Wiley & Sons, New York, NY.

Train, K.E. 1997. Mixed logit models for recreation demand. In: C. Kling and J. Herriges (editors) *Valuing the Environment using Recreation Demand Models*, Edward Elgave, Lyme New Hampshire. (Forthcoming)

Train, K.E. 1998. Recreation demand models with taste differences over people. *Land Economics* 74(2):230-239.

Wind, Y. 1978. Issue and advances in segmentation research. *Journal of Marketing Research* 15:317-337.

TABLE 1
Attributes and Levels Used in the Park Level Branded Choice Experiment.

Attribute	Level
User fee	None: no fees
	$5.00 per day per person
	$10.00 per day per person
	$15.00 per day per person
Chance of entry due to management restrictions such as quotas	Always get in
	3 in 4
	1 in 2
	1 in 4
Campsite type	Anywhere
	In designated areas only
	In designated areas with fireboxes
	In designated areas with fireboxes, tent pad and pit toilets
Level of development	None: no development in the park and no roads directly to or in the park and no motor boats
	Outposts: unstaffed outpost cabins in places and a road exists to the boundary of the park, but not inside, some motor boats may be present
	Lodges: fishing or hunting lodges present with motor boats and a road goes through the park
	Cottages: some places have cottage sub-divisions and there is a network of roads that allows improved access; motor boats will be present
Encounters with other wilderness groups	None: no other groups will be encountered
	1-3 groups are encountered each day
	4-9 groups are encountered each day
	over 9 groups are encountered each day

TABLE 2
Distribution and Factor Analysis of Attitudinal Statements reflecting Motivations for Wilderness Recreation in the System of Wilderness Parks

Variable	Factor Loadings			
	Factor 1 *(Challenge and freedom)*	Factor 2 *(Nature appreciation)*	Factor 3 *(Social relationships)*	Factor 4 *(Escape from routine)*
To challenge my skills and abilities	0.714	0.153	0.118	0.031
To develop my skills	0.636	0.145	0.206	0.1
To be in charge of a situation	0.635	0.056	0.039	0.19
To feel independent	0.573	0.267	0.091	0.094
To feel free from society's restrictions	0.501	0.071	0.093	0.434
To challenge nature	0.418	0.031	0.162	0.086
To be alone	0.395	0.188	-0.271	0.312
To feel close to nature	0.345	0.669	-0.001	-0.001
To observe the beauty of nature	0.05	0.66	0.014	0.142
To obtain a feeling of harmony with nature	0.329	0.632	0.037	0.011
To find quiet places	0.076	0.579	0.003	0.26
To enjoy the sights, sounds, and smells of nature	0	0.567	0.006	0.103
To be with my friends or family	-0.024	0.023	0.746	0.063
To strengthen relationships with friends or family	0.14	0.12	0.666	0.059
To do things with other people	0.183	-0.109	0.665	0.109
To be with people with similar interests	0.304	0.029	0.533	0.09
To escape from the pressures of work	0.043	0.125	0.153	0.708
To relieve my tensions	0.25	0.132	0.049	0.667
To get away from my everyday routine	0.08	0.101	0.221	0.649
To be away from other people	0.278	0.225	-0.239	0.431
Eigenvalues	4.619	1.989	1.263	1.18

TABLE 3

Latent Segment Model with Factor Scores, Intercepts and Trip Dummy as Segment Membership Variables[1]

Number of Segments	Number of Parameters (P)	Log Likelihood (LL) at Convergence	Log Likelihood evaluated at 0	ρ^2	AIC {-2(LL-P)}	ρbar^2 {1-AIC/2LL(0)}	BIC {LL-(P/2)*ln(N)}
1	16	-7040.49	-8765.30	0.197	14112.98	0.194	-7108.45
2	38	-6931.97	-8765.3	0.209	13939.9	0.205	-7093.38
3	60	-6791.82	-8765.3	0.235	13703.64	0.218	-7046.68
4	82	-6697.57	-8765.3	0.236	13559.14	0.227	-7045.88
5	104	-6648.11	-8765.3	0.242	13504.22	0.230	-7089.87
6	126	-6625.25	-8765.3	0.244	13502.50	0.230	-7160.46

[1] Sample size (N) is 4,892 choices from 620 individuals

TABLE 4
Parameters (standard errors) on the Segment Membership Variables for the Four Segment Model

Variables	Segment 1 (*Escapists*)	Segment 2 (*Weekend Challengers*)	Segment 3 (*Nature Nuts*)	Segment 4 (*Wilderness Trippers*)
Intercept	0	-2.7239 (0.0800)	-0.6293 (0.0870)	0.1167 (0.0400)
Short Trip Dummy	0	5.8356 (0.0934)	-0.2814 (0.0629)	-2.8806 (0.0951)
Factor 1 *(challenge and freedom)*	0	1.5267 (0.0606)	-0.2811 (0.0288)	-0.0617 (0.0134)
Factor 2 *(nature appreciation)*	0	0.0411 (0.0192)	0.7013 (0.0542)	0.4711 (0.0135)
Factor 3 *(social relationships)*	0	0.7737 (0.0102)	-0.6453 (0.0233)	-0.6507 (0.0262)
Factor 4 *(escape from routine)*	0	-0.6127 (0.0229)	-0.0557 (0.0224)	-0.0321 (0.0162)

TABLE 5
Parameters (standard errors) on the Variables for Two Recreation Site Choice Models

Variable	1 Segment Model	4 Segment Model			
		Segment 1 (*Escapists*)	Segment 2 (*Weekend Challengers*)	Segment 3 (*Nature Nuts*)	Segment 4 (*Wilderness Trippers*)
Fee	-0.0566 (.0035)	-0.0655 (0.0081)	-0.0594 (0.0180)	-0.1110 (0.0170)	-0.0943 (0.0107)
Chance of Entry	0.3857 (0.0472)	-0.1025 (0.0175)	1.8264 (0.0205)	-1.2318 (0.1134)	1.3654 (0.0539)
Campsite 1	-0.0036[1] (0.0156)	-0.0819 (0.0358)	0.1301 (0.0201)	1.3067 (0.0469)	-0.2600 (0.0264)
Campsite 2	-0.0175[1] (0.0120)	0.2328 (0.0252)	-0.0498 (0.0216)	-1.2395 (0.0364)	0.0849 (0.0300)
Campsite 3	-0.2290 (0.0138)	-0.2017 (0.0167)	-0.4529 (0.0177)	0.3883 (0.0494)	-0.5016 (0.0198)
Level of development 1	0.2400 (0.0258)	0.1931 (0.0409)	-0.2561 (0.0242)	2.5219 (0.0774)	0.0826 (0.0331)
Level of development 2	-0.4287 (0.0215)	-0.2536 (0.0319)	-0.0974 (0.0291)	-2.5507 (0.0236)	-0.6027 (0.0400)
Level of development 3	-0.6141 (0.0219)	-0.4829 (0.0432)	-0.1854 (0.0294)	-3.3136 (0.0253)	-0.9583 (0.0330)
Level of encounters 1	0.5783 (0.0257)	0.5064 (0.0203)	0.5817 (0.0327)	0.5329 (0.0922)	1.2711 (0.0263)
Level of encounters 2	-0.5123 (0.0257)	0.0322[1] (0.0236)	-0.3195 (0.0145)	-1.0459 (0.0816)	-1.6401 (0.0738)
Level of encounters 3	-1.0189 (0.0277)	-0.6382 (0.0346)	-0.7693 (0.0420)	-0.8151 (0.1861)	-2.2357 (0.0574)
Woodland Caribou ASC	-0.1124 (0.0363)	0.4569 (0.0788)	-0.2084 (0.0184)	-3.4795 (0.0174)	0.4541 (0.0347)
Quetico ASC	0.8561 (0.0317)	2.0588 (0.0244)	-0.9935 (0.0190)	-0.1756[1] (0.1096)	0.9222 (0.0276)
BWCA ASC	0.2206 (0.0256)	1.2308 (0.0536)	-1.8306 (0.0318)	-2.2187 (0.0209)	-1.1667 (0.0166)
Wabakimi ASC	-0.4637 (0.0348)	-0.6305 (0.0132)	-0.8476 (0.0173)	-2.0425 (0.0876)	-0.1331 (0.0316)
Manitoba Parks ASC	-0.4456 (0.0383)	-2.5380 (0.0380)	0.2427 (0.0090)	-2.2124 (0.1302)	-0.2514 (0.0241)

[1] Indicates that the parameter is not significantly different than 0 at the 5% level.

TABLE 6
Results of Posterior Analysis of the Characteristics of Segment Membership Probabilities using OLS.

Variable	Segment 1 (*Escapists*)	Segment 2 (*Weekend Challengers*)	Segment 3 (*Nature Nuts*)	Segment 4 (*Wilderness Trippers*)
Intercept	-1.5630[1]	-0.3945	-3.0117[1]	-3.4880[1]
Years of experience	-0.0107[1]	0.0226[1]	-0.0074[2]	-0.0185[1]
Level of specialization	0.006	-0.4578[1]	0.19861[1]	0.4353[1]
Canadian	0.0055	-0.3648[2]	0.0486	0.2312
Household size	0.0141[1]	0.2201[1]	-0.0827[1]	-0.1115[1]
Years of age	0.0122[1]	-0.0190[2]	0.0085[2]	0.0161[1]
Level of education	0.0294	-0.1161[1]	0.0575[1]	0.1021[1]
Income	0.0177	-0.0633[2]	0.0104	0.005

[1] Coefficient is significant at the 5% level or better
[2] Coefficient is significant at the 10% level or better

TABLE 7

Compensating Variation ($/trip) for Some Hypothetical Changes in Recreation Quality for a Representative Individual in the Sample[1]

Change	One Segment Model	Four Segment Model			
		Segment 1 (*Escapists*)	Segment 2 (*Weekend Challengers*)	Segment 3 (*Nature Nuts*)	Segment 4 (*Wilderness Trippers*)
Closure of Quetico Provincial Park	-8.98	-18.44	-1.63	-12.49	-7.25
Increase congestion by 1 level:					
At Woodland Caribou Park	-1.15	-0.55	-1.48	-0.06	-1.14
At Quetico Park	-5.43	-4.47	-0.93	-6.76	-6.66
At BWCA	-0.43	-1.26	-0.12	-0.18	-0.01
At Wabakimi Park	-0.71	-0.16	-0.68	-0.19	-0.51
At Nopiming/Atikaki Parks	-5.24	-0.05	-8.42	-0.83	-4.17
At all parks	-18.44	-7.27	-14.54	-8.8	-30.04

[1] These estimates used the modal perception of campsite type, development, and number of encounters as the base case

Structural Model of Choice and Latent Segmentation

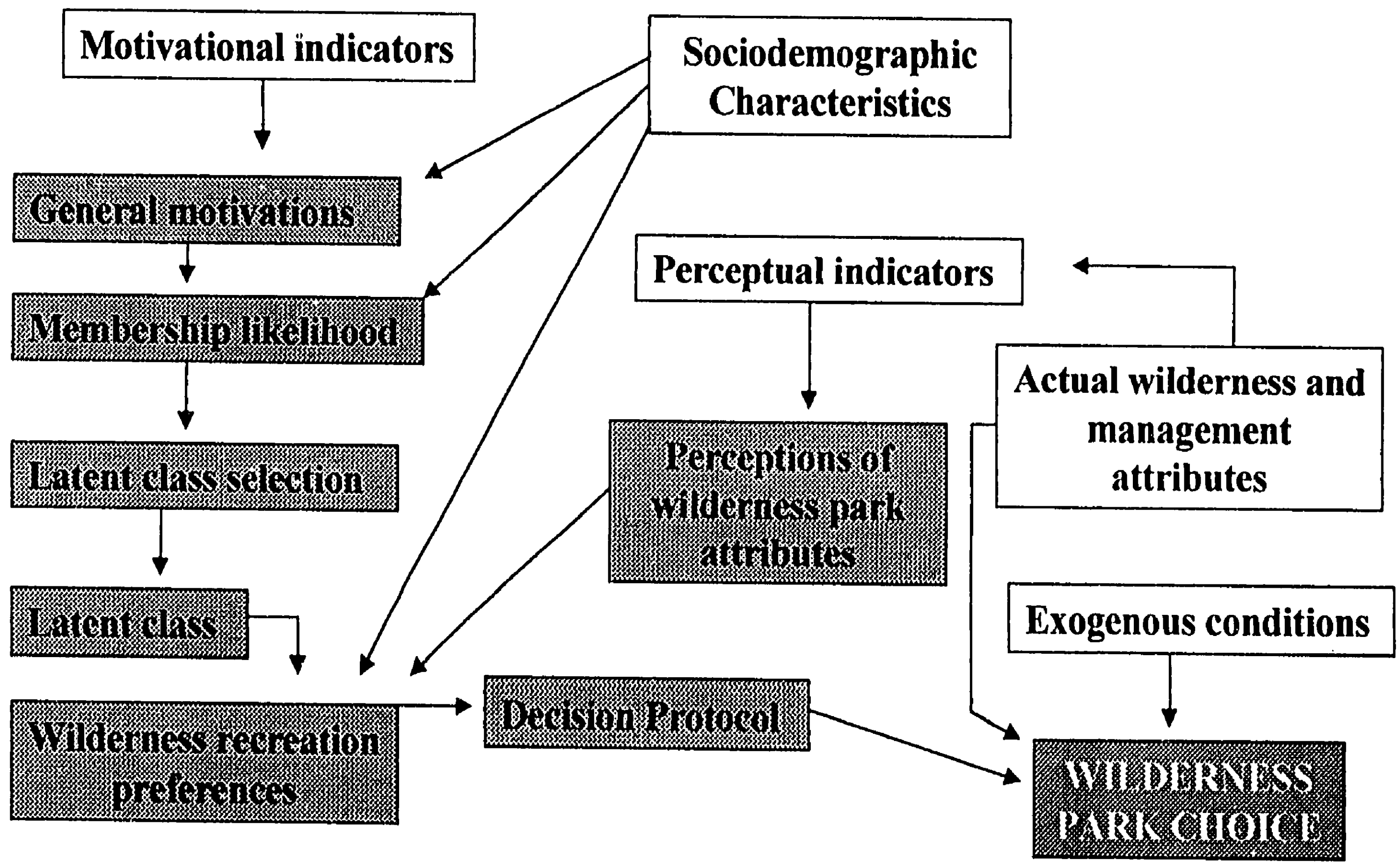


Figure 1. A path diagram outlining the application of the latent segmentation choice model to backcountry recreation in a set of wilderness parks in the Canadian Shield region . Shaded boxes refer to the latent constructs utilized in the model.

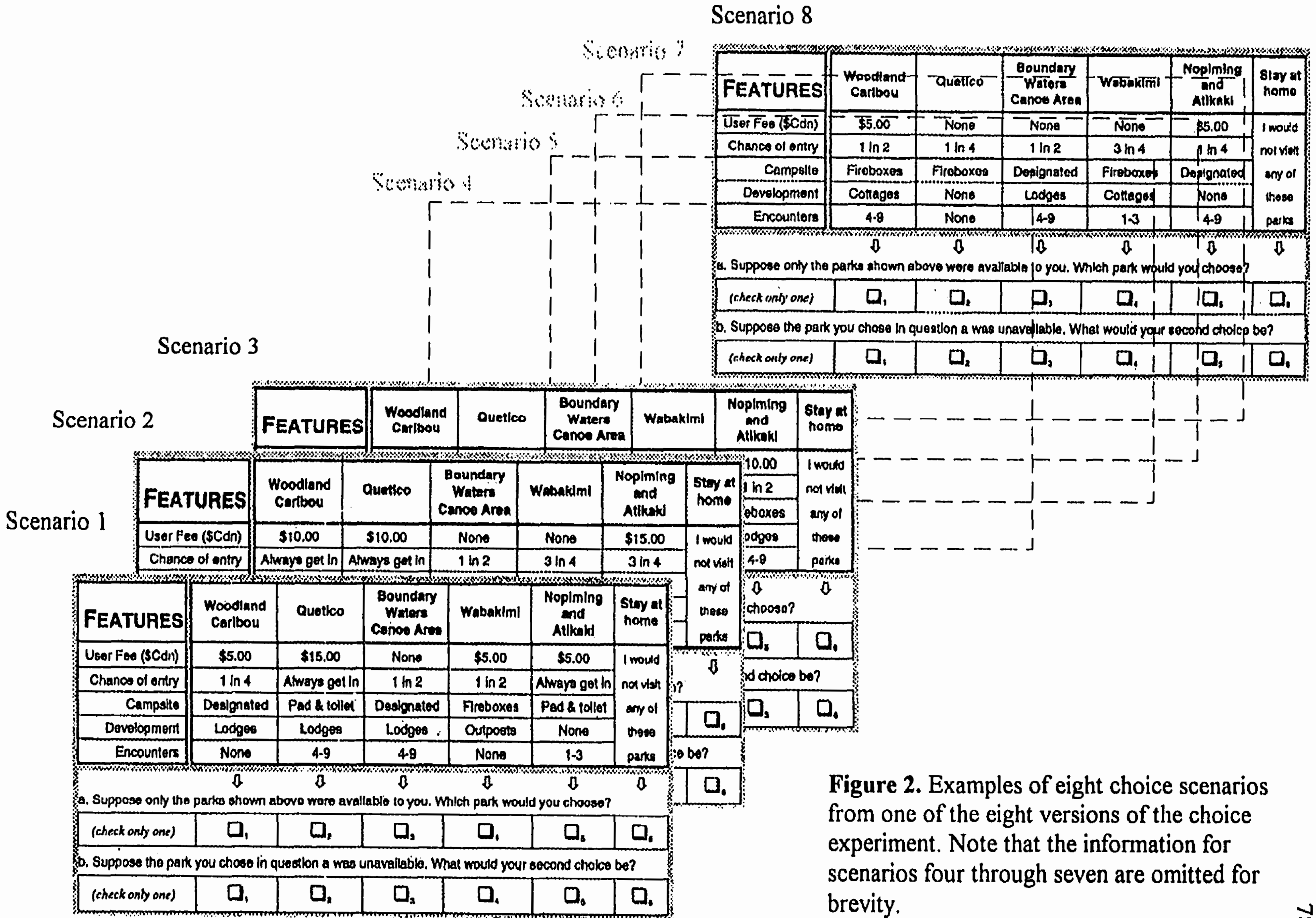


Figure 2. Examples of eight choice scenarios from one of the eight versions of the choice experiment. Note that the information for scenarios four through seven are omitted for brevity.

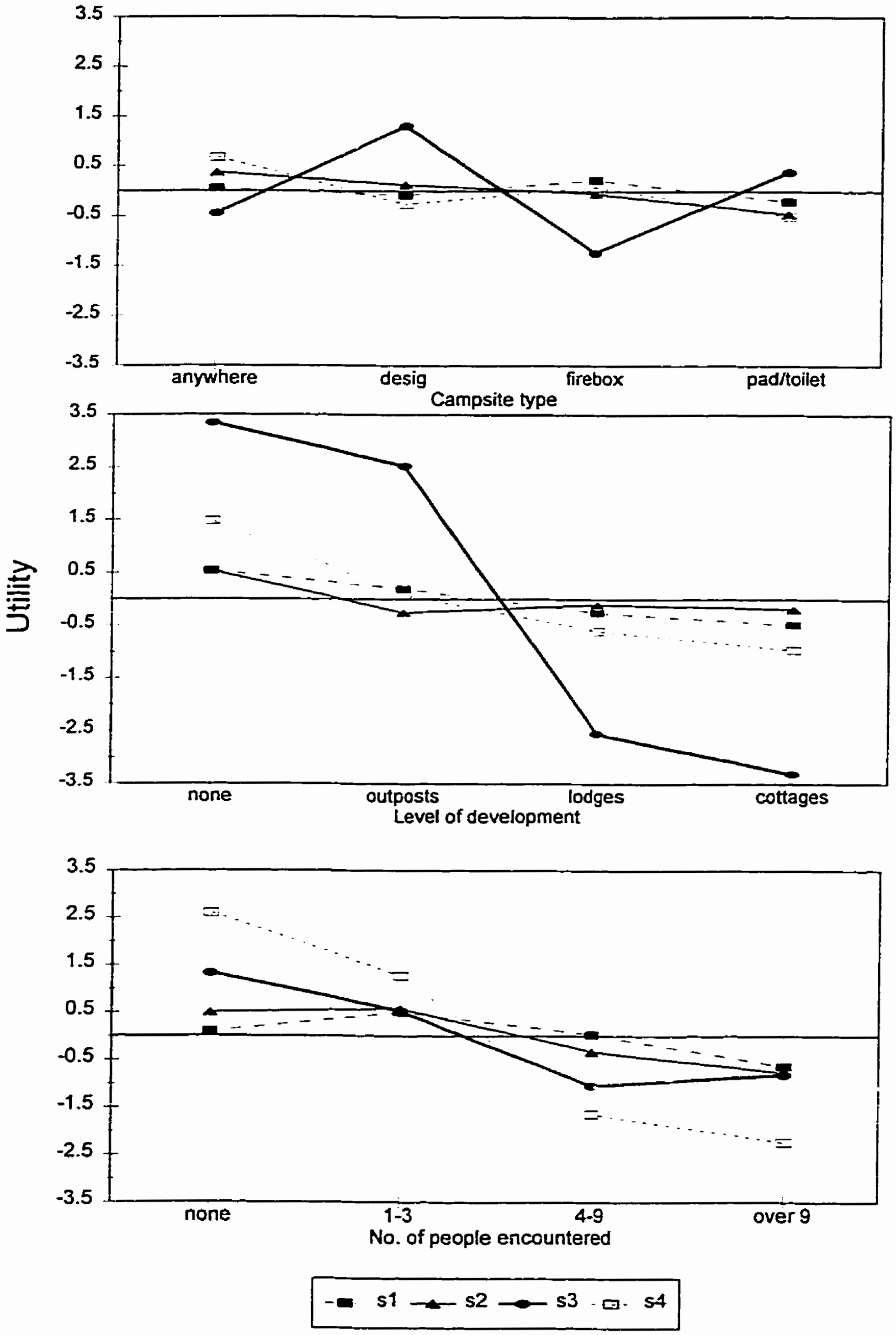


Figure 3. Plots of parameters associated with four levels of three attributes in the utility function over wilderness park choice. Note that the parameters on these levels were effects coded in the development of the model.

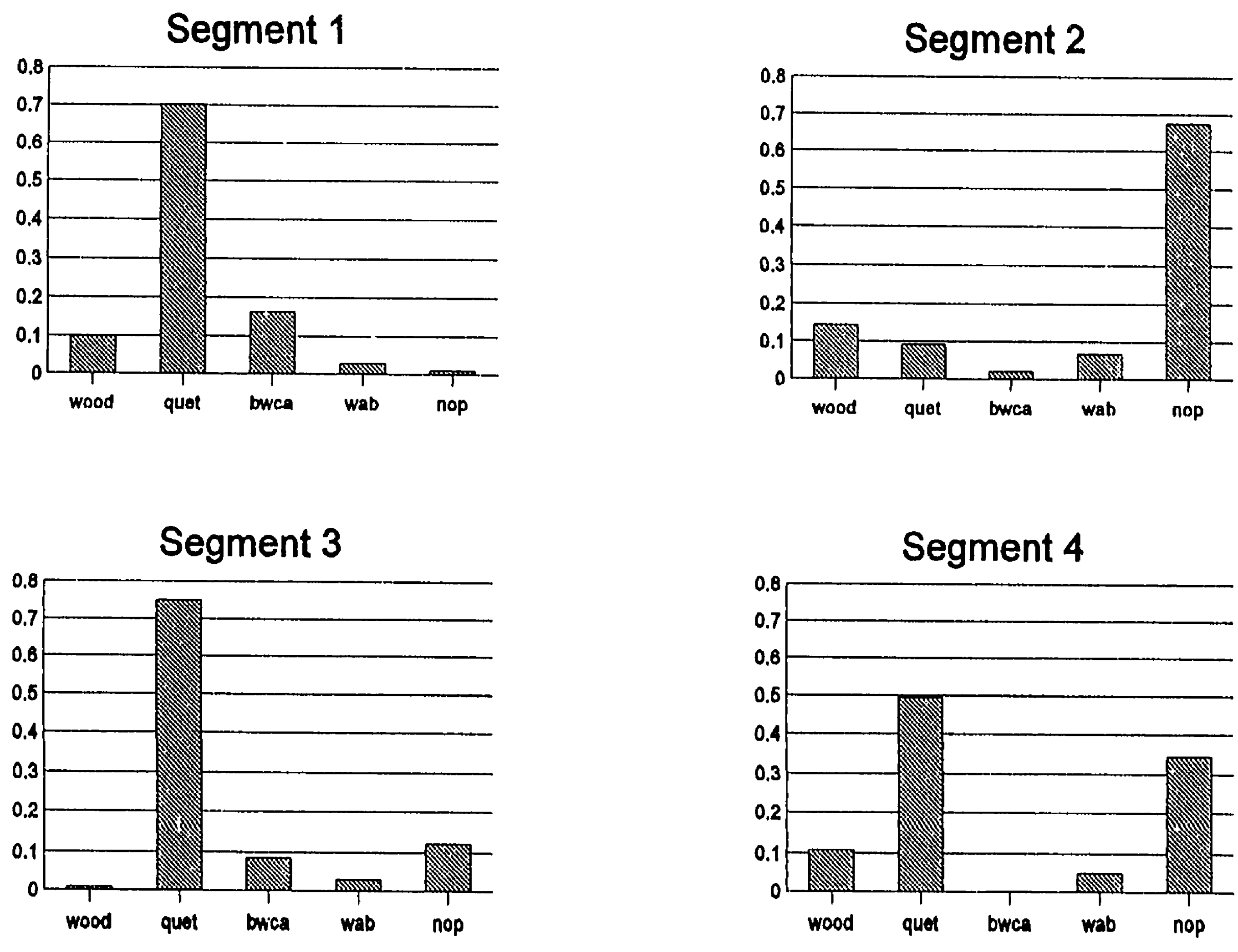


Figure 4. The estimated probability distribution of trips among the five parks for each of the four segments in the latent segment model.

CHAPTER 4

Incorporating Endogenous Perceptions of Environmental Attributes in RUMs: The Case of Congestion

INTRODUCTION

There are many occasions when individuals behave in ways that are contingent on one another. For example, it is commonly observed that individuals who belong to an identified group behave in similar ways. Manski (1995) describes three hypotheses that attempt to explain this observation. The first he calls *correlated effects* in which observed behaviour is related to the similarity of the institutional environment(s) to which group members are subjected. These effects are nonsocial phenomena, and for an environmental economist could represent the kinds of relationships between environment quality and recreation site choice behaviour.

Manski's second hypothesis is labelled *contextual effects*. (This was termed exogenous effects in an earlier paper Manski (1993)). In this situation individual behaviour is related in some way to exogenous characteristics of the group. These effects represent social phenomena related to group membership and the behavioural outcomes are related to the type of group of which one is a member. These types of effects are perhaps illustrated by the latent class analysis conducted in the previous chapter of this thesis.

In this chapter we are interested in Manski's third hypothesis which he calls *endogenous effects*. These are cases in which the propensity of an individual to behave in a certain way is dependent on the prevalence of that behaviour among members of a group. This effect describes endogenous feedback in which individuals interacting with each other reciprocally influence or mutually adjust their behaviour (Erbring and Young 1979). This concept has been of interest to sociologists and social psychologists who typically examine *reinforcing* endogenous effects such as conformity, herd behaviour, band wagon effects, etc. (e.g. Granovetter 1978). In this chapter, however, the interest is in examining *attenuating* endogenous effects in which interactions between individuals cause mutual avoidance or in some way may have a negative influence on behaviour.

One can imagine many examples of this phenomenon. For example, recreational fishing may be influenced by the chance to catch lots of fish; yet this ability may be compromised by the number of other fishers harvesting fish. This example implies that exploring the relationship between the behaviour of individual fishers who comprise a recreational fishing group and the characteristics of that group would provide information on the levels of fishing at a set of sites. Thus, fishing site choice behaviour is related not only to the ability of the individual to catch fish, but also to the numbers of other anglers catching fish at the sites and the relative ability of the aggregate group of fishers to catch fish. As Schelling (1978:14) points out:

> "situations in which people's behavior or peoples' choices depend on the behavior or the choices of other people are the ones that usually don't permit any simple summation or extrapolation to the aggregates. To make that connection we usually have to look at the *system of interaction* between individuals and their environment, that is, between individuals and other individuals or between individuals and the collectivity."

The system of interaction described above could have a number of characteristics, one of which is the concept of a critical mass or a critical number (Schelling 1978:94).

This is described by the observation that an individual's behaviour is dependent upon how *many* others are behaving in a particular way or how *much* they are behaving that way. This system of interactions can also involve heterogeneity in that the critical number for one individual may not be the same for another individual in the group. One class of interactions that typifies this concept of critical numbers with attenuating endogenous effects is the concept of congestion.

The Case of Congestion

Congestion is an example of an externality that occurs when an economic agent engages in an activity that imposes negative effects on the welfare of some other agent. Congestion is a critical attribute associated with many forms of recreation: with greater levels of the number of recreationists anticipated, fewer will wish to participate. However, what is a congested recreation facility to some may not be to others and, in fact, may actually be attractive to some others. In other words, the degree of interaction among people could differ between individuals and plays a role in the effect of the interaction. Thus, congestion could be characterised by considerable heterogeneity in terms of the critical number of people an individual sees or interacts with before they decide not to participate.

This concept is best illustrated by Schelling's (1978) diagrammatic model of critical mass using the North American dataset on wilderness use described in the last chapter. This sample of 716 wilderness recreationists were asked about their critical levels of encounters with other recreationists. In the first question they were asked how many encounters would represent an ideal trip, how many an acceptable trip, and how many they would tolerate before not visiting the area again. In the top panel of Figure 1 the percent distributions of answers to these questions are plotted against discrete categories of expected levels of encounters with other recreationists. What is striking about this figure is the variation among individuals in the various levels of encounters, particularly in the relationship examining the expected encounter level at which recreationists would not visit the area again. This latter relationship portrays the idea of a individual specific threshold beyond which an individual would not participate further.

This threshold concept is examined further in the bottom panel of Figure 1 which illustrates Schelling's hypothesis on congestion. This graph plots the expected proportion of the sample participating with increasing levels of expected encounters. This relationship incorporates the heterogeneity in thresholds in the sample and allows one to predict visitation levels as expected encounters increase. Can this notion have significance in examining and incorporating congestion in recreation demand models? Answering this question requires a model which incorporates expectations of congestion levels, heterogeneity in these expectations, and the incorporation of some of the consequences of congestion.

Cicchetti and Smith (1973) conducted the first comprehensive examination of congestion in a wilderness recreation setting. Using a mail survey, they assessed the response of recreationists to hypothetical variations in encounters with other recreationists. They found clear evidence of willingness to pay by individuals to avoid encounters with others at trail heads and camp sites. While the economic significance of congestion is important in resource allocation where those who impose the costs are not required to pay for the full consequences of their actions, for most wilderness recreation settings this is problematic because the facilities are usually unpriced (Cicchetti and Smith, 1976).

A cost of congestion, however, is avoidance of the site where this occurred and subsequent visitation of some other site further away. Here, the costs imposed on an affected individual are the additional travel costs required to visit substitute sites. It is reasonable to hypothesize that recreationists anticipate congestion levels at sites and that these predictions enter the decision to visit them. Thus, anticipated congestion levels may be an important feature of recreation site choice. In fact, this has been demonstrated empirically by Menz and Mullen (1981) who found a strong relationship between predicted congestion levels and willingness to travel.

Defining a suitable measure of congestion has been difficult for economists, however. Early research on this issue used the density of recreationists at sites or the number of interactions with other recreationists as measures of congestion or "crowding" (e.g. Cicchetti and Smith 1973; McConnell 1977; Deyak and Smith 1978). Shelby (1980), however, showed that neither of these measures affected recreationists' assessments of satisfaction nor did they represent suitable assessments of crowding.

Jakus and Shaw (1997) suggest that *ex ante* assessments of congestion may the most relevant measures of congestion determining recreation site choice decisions, particularly in models that use revealed preference information. Jakus and Shaw (1997) define four possible measures of congestion: actual, expected, perceived, and anticipated. They define actual congestion as the level determined by someone (for example a researcher) outside of the sample so that the level is exogenous to those in the sample. Expected congestion is defined as the average congestion level. Perceived congestion is interpreted as an *ex post* measure involving individuals' understanding of congestion once at the site and uncertainty about congestion has been eliminated. Finally, anticipated congestion is defined as an individual's own estimate of congestion that holds at a site prior to when they actually visit that site.

While their descriptions of the first three of these terms is generally not congruent with the recreation literature, their concept of anticipated congestion clearly captures the important role *ex ante* assessments of congestion may play in recreation demand. Jakus and Shaw (1997) suggest that in order for economists to conduct empirical analyses of congestion effects, individual-specific demand modeling should be undertaken in which anticipated congestion levels vary with the individual. An implication of anticipated congestion playing a role in recreation site choice is the fact that the demand model must incorporate possible endogenous effects.

The objective of this chapter is to incorporate anticipated congestion in an individual-specific demand framework using random utility theory. In this process the notion of interdependent utility functions will be utilized and implemented in a relatively econometric structure. The model will be illustrated in an empirical application to wilderness recreation in which congestion plays a key role in recreation site choice behaviour.

THEORY

A recreationist (indexed n) receives utility, U, from visiting a site (indexed i) equal to $U_{ni}=U(X_{ni},Z_n)$, where X_i is a vector of characteristics of site i, and Z_n is a vector of individual characteristics. Random utility theory considers U as a random variable where part is known or observable to the investigator and the remainder is not. Thus, $U_{ni}=V_{ni}+\epsilon_{ni}$ where $V_{ni}= f(X_i,Z_n)$ is the former component and ϵ_{ni} the latter. ϵ is considered an error term related to researcher error or to randomness (hence the term random utility) of the individual.

The probability that site i will be visited by n is equal to the probability that the utility gained from visiting i is greater than or equal to the utilities of choosing any other site in some finite set of available sites, C_n . Thus, the probability, π, of visiting site i is:

$$\pi_n(i) = \Pr \{V_{ni} + \epsilon_{ni} \geq V_{nk} + \epsilon_{nk}; \forall k \in C_n\}. \quad \textbf{(1)}$$

The conditional logit model, developed by McFadden (1974), can be used to estimate these probabilities if the ϵ's are assumed to be independently distributed Type-I Extreme Value variates. McFadden (1974) shows that this assumption allows the choice probabilities to take the form:

$$\pi_n(i) = \frac{\exp \mu(V_{ni})}{\sum \exp \mu(V_{nk})}. \quad \textbf{(2)}$$

where μ is a scale parameter that is typically assumed to equal 1.

Congestion in this theoretical context can be considered a site attribute, and can be separated from other elements in the vector of site characteristics, X_i. The utility function now can be represented by:

$$V_{ni} = f(X_i, Z_n, c_i) , \quad (3)$$

where c_i represents congestion at site i. Based on discussion in the recreation literature (e.g. McConnell and Sutinen 1984) and the information displayed in Fig. 1, one can assume that $\partial V_i/\partial c_i < 0$ for the majority of individuals. However, as pointed out by Jakus and Shaw (1997), utilizing this structure using revealed preference data is problematic because recreationists can only experience the level of congestion by visiting the site (an *ex post* measure). Thus, it is difficult for researchers to incorporate current levels of congestion because congestion cannot enter the utility function and influence site choice before it is experienced. An approach that properly incorporates congestion should consider its uncertain nature and individual variability in the disutility (or indeed for some utility) that it may provide recreationists.

The uncertain nature of congestion could be manifested in the formation of prior expectations or anticipations of congestion levels and choices could be made on the basis of these prior perceptions of congestion and other choice attributes, not necessarily objective measures of them. This idea is considered by Schelling (1978) in his discussion of congestion. Empirical support for the hypothesis that recreation site choice could be made on the basis of perceptions of environmental quality has been obtained by Adamowicz et al. (1997). One way to consider anticipated congestion in the theoretical framework outlined above is to assume that individual recreationists make *forecasts* of congestion levels at sites before visiting them. An individual's forecast of congestion at site i is based on their view of the probabilities of other recreationists visiting the site. Considering person n's utility function and expanding congestion in (3) yields:

$$V_{ni} = f(X_i, Z_n, (-n \cdot Pr_{-ni})), \quad (4)$$

where -n represents the number of other recreationists (not including n) who would visit i, and Pr_{-ni} portrays the probability of these other people visiting site i. Thus, the congestion level in this expression is represented by the number of people who may visit this site multiplied by their probabilities of choosing this site. However, the probabilities of other recreationists visiting i are related to their indirect utility functions as explained above, making congestion endogenous. Incorporating this idea into (4) produces:

$$V_{ni} = f(X_i, Z_n, G(-n, V_{-ni}(X_i, Z_{-n}, c_i))) . \quad (5)$$

where congestion is captured in the function $G(\cdot)$ which provides n's estimate of -n, in turn depending on $V_{-ni}(\cdot)$.

This way of thinking about congestion results in interdependent utility functions in that n's choice of a particular recreation site is dependent on -ns' choice of that same site. Each recreationist in the market does not have information about how others choose sites but may anticipate other recreationists' choices.[1] Since each person selects a site to maximize his or her utility, given the other recreationists' utility maximizing behaviour,

---

[1] Note also that this formulation is quite different from other examinations of interdependent preferences in which the mean or some other measure of the demand of the population or reference group enters demand analysis (e.g. Pollak 1976).

the behaviour associated with recreation site choice is analogous to the behaviour hypothesized in a Nash game. In this case, the arguments of the game are the forecasts of others visitation levels which are taken as given. Thus, recreationists predict V_{-ni}, take this as given, and act on their best response to the resulting expected structure of the market.

An outcome of this framework is that environmental quality changes at site i appear not only in n's indirect utility function, but also in the functions of the other recreationists. Thus, site quality improvements may have a positive effect on the recreation benefits provided by a site on n, but may have an opposite effect through their impact on congestion levels by making the site attractive to other recreationists. To illustrate this (5) can be rewritten as:

$$V_{ni} = g(X_i) + h(Z_n) + k(V_{ni}), \qquad \textbf{(6)}$$

and examining the effect on n of a change in quality at site i yields:

$$\frac{\partial V_{ni}}{\partial X_i} = g'(X_i) + k'(V_{-ni}),$$
$$\text{where } k'(V_{-ni}) = \frac{\partial V_{ni}}{\partial V_{-ni}} \cdot \frac{\partial V_{-ni}}{\partial X_i}. \qquad \textbf{(7)}$$

The latter term in (7) represents the reaction function of n to the other recreationists following a quality change at site i. Therefore, measuring the benefits of environmental quality changes must include both of these forces.

APPLYING THE ENDOGENOUS CONGESTION FRAMEWORK

Given the concepts outlined above, recreation site choice would involve individuals forecasting congestion levels based on their view of the probabilities of other recreationists visiting the sites of interest. Modelling this theoretical structure of recreation choices in the presence of congestion as a Nash equilibrium could involve representing the visitation reaction functions depicted by the term $k'(V_{-ni})$ by some instrument such as congestion forecasts, $F(c_i)$. Thus, the indirect utility function of n for site i could be represented by:

$$V_{ni} = f(X_i, Z_n, F(c_i)) \tag{8}$$

Empirical implementation of this idea requires explicit estimation of anticipated congestion levels at some set of sites. For example, forecasts of congestion at site i may be some function of the qualities or characteristics of site i, the experience levels, and socioeconomic characteristics of the individual. Now the utility function can be portrayed by:

$$V_{ni} = f(X_i, Z_n, F(c_i));\ \text{where}\ F(c_i) = f(X_i^s, Z_n^s) \tag{9}$$

where the superscript s refers to a subset of the vector of site or individual characteristics. Empirical implementation of this theoretical structure requires information on anticipated congestion levels at some set of recreation areas and information on trips to these same areas, the development of an anticipated congestion function, and the use of this function in an instrumental variables procedure with the site choice model. Note that now congestion is not only a site attribute, but is also an individual attribute. Thus, n's forecast of congestion is unique to n. This framework now allows for the existence and inclusion of Schelling's (1977) concept of the critical number, in this case the critical number of recreationists.

Data and Econometric Analysis

During 1995, a sample of 1000 visitors to Nopiming and Atikaki Provincial Parks in Manitoba, and Woodland Caribou, Quetico, and Wabakimi Provincial Parks in Ontario was drawn from park registrations or on-site registrations administered by the Canadian Forest Service. About 71% of individuals in this sample were from Quetico, about 18% from Woodland Caribou, 10% from both Manitoba parks, and about 1% were from Wabakimi. This distribution was selected because it approximately represented the levels of visitation across the five parks (see Boxall et al. 1999).

A questionnaire was developed that gathered information about opinions of wilderness management, levels of past visitation to the 5 parks[2] and an additional park, the Boundary Waters Canoe Area (BWCA), descriptions of a typical wilderness trip, and socio-demographic characteristics. One section of the questionnaire solicited perceptions of congestion at each of the five parks. These perceptions were solicited using the following question: "***In planning your last trip*** *to wilderness parks or areas, what were your perceptions of existing park conditions and management?*" A table was presented to

---

[2] In this analysis Nopiming and Atikaki parks were combined into an eastern Manitoba Parks unit.

respondents and they were asked to indicate the number of expected encounters per day with other wilderness visitors in each park by checking one of four levels: none, 1-3 groups, 4-9 groups, or over 9 groups. The questionnaire was mailed to respondents during 1996 and after two follow-ups and adjustment for non-deliverables, an 80% response rate was achieved.

Table 1 provides information on the responses to the anticipated congestion question. Woodland Caribou and Wabakimi parks have the lowest reported anticipated congestion levels while the BWCA had the highest. Respondents took 1,723 trips to the 5 parks during 1995 and 1996. The most frequently visited parks were Quetico and the BWCA.

The responses to the congestion question were pooled (N=1,297) and these formed dependent variables of the anticipated congestion model ($F(c_i)$ in equation (9)). A number of individual-specific variables formed the Z_n vector in the congestion forecast model. These variables included: years of experience in wilderness trips in the region, membership in a conservation or recreation organization, the typical trip length, gender, income, education, and household size.[3] Since the five parks represent an increasingly highly sought wilderness experience (Boxall et al. 1999), and that in at least one park (the BWCA) visitors were increasingly "feeling crowded" (Cole et al. 1995), the years of experience variable was expected to have a positive effect on increasing congestion forecasts. The rationale here was that individuals visiting the area many times in the past

---

[3] It is recognized that this is a limited set of variables and that others, such as attitudes towards crowding, may be better explanators of congestion forecasts.

would have experienced the increasing visitation levels over time. Similarly, those who belonged to organizations would have more information on visitation levels and the increasing use of the parks over time. Thus, the effect of membership was also expected to be positive.

However, individuals who typically take short trips were thought to take more of them with families or other types of social groups. This characteristic suggests that they may not have experienced the increasing use of backcountry areas and may not be as sensitive to congestion as those taking longer trips to congestion levels. Thus, it was thought that this variable would probably have a negative effect on congestion forecasts. Similarly, the household size variable was hypothesized to have a negative effect on congestion forecasts due to the fact that families with many children would not have the time or background to have experienced the increasing visitation levels. The signs of the other individual-specific variables were uncertain.

Finally, the perceived level of development at a park was thought to influence anticipated congestion. In this case the development category reported by each respondent from a park for which a congestion forecast was received was used. It was hypothesized that forecasted congestion would be greater if an individual thought that the level of development was greater. Thus, the parameter on development was expected to be positive.

For the X_i vector, there were few choices relating specifically to each of the limited set of park and due to the diversity of routes in each park an individual was able to choose. However, the size of the park is probably representative of the number of routes one is able to take, and may affect the spacing of recreationists such that their chances of encountering each other are reduced. Thus, park size was expected to have a negative effect on congestion forecasts. In addition, the degree of access of the routes at each of the five parks varies and was thought to play an important role in determining congestion levels. This variable was expected to have a positive impact on congestion forecasts; greater accessibility would mean more visitors.

Since the dependent variable was discrete, but ordered, ordered logit models were used to determine the effects of individual respondent and park characteristics on forecasts of congestion levels. Two models were estimated (Table 2). The first included only the individual-specific variables, and the second included the park characteristics along with these variables.

In both models, being a male and preferring long trips is inversely related to increasing congestion perceptions (Table 2). As expected, high levels of wilderness recreation experience and membership in a conservation or recreation organization have positive effects on the levels of anticipated congestion. These relationships point to a possible connection between the highly specialized recreationist (likely male, experienced, takes long trips and is a member) probably visiting places where they do not expect to see high numbers of other individuals. Finally, as expected, high levels of perceived human development at wilderness parks have a significant positive effect on congestion levels. All other individual-specific variables were statistically insignificant in explaining congestion forecasts.

These individual effects on congestion are probably mediated by relationships between visitation levels and park characteristics, however. In this wilderness data the size of the wilderness area has a negative effect, while the number of roads accessing a wilderness area has a positive effect on congestion forecasts. These relationships support the hypothesized connections between park size, access, and congestion levels. It is noteworthy that the inclusion of these park level attributes in the congestion forecast function made some of the individual-specific variables statistically significant, improved the value of the log likelihood function, and increased the model's explanatory power.

Park choice models were estimated using the revealed preference information summarized in Table 1. In these models, park choice was modelled as a function of travel costs, perceived chances of entry, perceived development levels, the size of the park, the number of roads accessing the park, and an alternative specific constant. Congestion was included in these models in two different ways. The first used the congestion level by park reported by each respondent from the questionnaire.[4] The models using this variable are termed *reported* congestion (RC) models since respondents reported their prior congestion forecast. This way of including perceptions of attributes has been suggested in other studies of recreation choice behaviour in the literature (e.g. Adamowicz et al. 1997). The second approach used predictions from the congestion forecast models described above and these will be called the *anticipated* congestion (AC) models. For the latter models, two versions were estimated. The first involved the use of forecasts from model 1 and the second from model 2 in Table 2.

---

[3] If this information was missing for a park the modal perception level calculated over the sample was used for a respondent.

Five models are reported in Table 3 and two comparisons were examined. The first comparison involved contrasting two RC specifications with similar specifications of the AC model in which the congestion forecasts were generated using model 2 in Table 2. In this case one RC and one AC model included the perceived development variable in the choice model. In the second set this variable is removed. This use of the perceived development variable was an attempt to examine whether the AC models were over-specified since the congestion forecast function included this variable as well as the roads and park size variables. Thus, removing one of these three allowed examination of the sensitivity of the econometric results to these specifications. The second comparison involved selection of the appropriate specification of the congestion forecast in the AC choice models

The parameter estimates (Table 3) suggest that the variables generally perform as expected in each of the models. For example, travel costs are negative and significant, higher chances of entry to a park are a positive influence on park choice, higher congestion levels are a negative influence on choice, and park size and the number of roads accessing a park have a positive effect on choice. The signs of these variables are consistent across the model, but the magnitude and statistical significance of the effect of these features are different.

A number of findings are noteworthy. First, perceived development has a negative, but insignificant effect on park choice in the one RC model (RC1). This variable is positive and significant in the AC model in which it was included (AC3). Its inclusion, however, affected the size and significance of the park size variable. In the

AC2 model the park size parameter was 10 times larger and statistically significant, while in the AC3 model this parameter was reduced in size and insignificant. Perceivded development did not have the same effect on the park size variable in the RC1 model. These results suggest that including perceived development along with the roads and park size variable in the AC choice model resulted in over-specification. This finding, in addition to the fact that park size is an objective feature subject to policy manipulation, suggests that AC2 is probably the appropriate model to use.[5]

Second, higher congestion levels had a significant negative effect on park choice in four of the models. The one model where it was not statistically significant is the AC1 model which utilized forecasts from model 1 in Table 2. This finding, along with the comparably lower predictive power of the congestion forecast function, suggests that this particular specification of the congestion instrument is not suitable in the park choice model.

However, the negative effect of congestion is much more pronounced in the AC models than the RC models. This further suggests that using predictions from the congestion forecast function as instruments rather than using the reported congestion forecasts, is adding significant information to the analysis of park choice behaviour. The instrumental variables approach is revealing that congestion has a greater affect on park choice than could be understood with the other more typical modelling approach.

---

[5] In addition, it is difficult to imagine a policy that would work to change perceptions of development rather than objective (or actual) measures of development. Adamowicz et al. (1997) discuss this and other issues associated with the use of perceptions of environmental quality changes.

Third, while the size of the park and the number of roads accessing a park have positive effects on choice in both the RC and AC models, there are differences in the sizes and the statistical significance of the parameters. In the RC models the park size parameters are highly significant and are about six times larger than in the AC models. Roads, however, are not statistically significant in the RC models, but are in the AC models and the parameter is quite large signifying that roads have a large influence in determining park choice. Once again this effect has been uncovered as a result of using the congestion forecasts.

Finally, there are other important differences between the RC and AC models. In the AC2 model, congestion and roads have been estimated with much greater precision than in the RC2 model. On the other hand the travel cost, chances of entry, park size and the alternative specific constant were estimated with less precision. This effect is particularly pronounced with the park size variable. These findings are probably the result of the instrumental variable approach used in estimating the AC models and suggests that the instrument (in this case the $F(C_i)$ function) has successfully identified the endogenous congestion condition in the choice models. This is further supported by the observation that the value of the log likelihoods at convergence and ρ^2 statistics for the two AC models are larger than their RC counterparts.

Welfare Implications

A question that remains with the endogenous congestion condition is the effect this condition would have on welfare measures associated with environmental quality changes. To examine this issue a series of policy simulations were conducted on two of the five parks. The policy involved increasing road access to Wabakimi and Quetico Provincial Parks. While these increases are hypothetical at present, they are plausible given possible expansion of forest harvesting and the need for increasing access for logging trucks and other equipment to remote areas. For most of the 5 parks examined in this study, industrial forestry is occurring near these parks, and in some cases (e.g. Woodland Caribou Park), harvesting takes place right up to the boundary of these areas.

An issue with road expansion is its impact on wilderness canoeists. For example, Wabakimi and Quetico Parks are thought to be elements in a system of wilderness parks in the Canadian Shield (Boxall et al. 1999; Englin et al. 1998). This system is currently dominated by the BWCA and a major factor in the spatial distribution of recreation demand is congestion. Since the BWCA is close to the major population centre in the region (Minneapolis), and visitors are increasingly reporting crowding concerns (Cole et al. 1995), congestion is a key issue in its management. This is supported by the high levels of congestion perceived at the BWCA by sampled individuals in this study (Table 1). Congestion may in fact be causing recreationists from this market to visit other parks in this system. The expansion of road networks may therefore have impacts on the distribution of trips taken by individuals in the market among the five wilderness areas.

The welfare implications of road access expansion were examined using Hanemann's (1982) formula for estimating compensating variation in conditional logit models. For the AC model this involved estimating the change in congestion forecasts through adjusting the numbers of roads accessing each of the parks and then incorporating these new forecasts in the park choice model. However, the roads variable must also be modified in the choice model holding all of the other variables (except congestion) at their original values. For the RC model, congestion remains constant and only the roads variable was changed in the choice model. For each simulation the AC2 in Table 3 was used. The results are portrayed graphically in Figure 2 for Quetico and Figure 3 for Wabakimi.

For the current access level at Quetico (3 roads) the ordered logit model predicts that a majority of respondents (575 of 580 individuals), forecast an encounter level of 1-3 groups per day (bottom panel, Fig. 2). However, increasing road access at this park would change this forecast. With six roads for example, every individual in the sample forecasts congestion to be 4-9 groups/day and beyond this road access level, an increasing proportion of the sample forecasts congestion at the highest level.

The welfare implications of this expansion of access are shown in the top panel of Figure 2. Note that an additional road at Quetico would generate benefits valued at over $200/trip. More than one additional road, however, generates dis-benefits. At five roads, this drop is pronounced and is congruent with a major shift in congestion forecasts. These findings support Schelling's (1978) notion of thresholds. These effects are not picked up by the RC model in which increasing road access does not feed back on congestion with the result that each additional road generates additional benefits through their impact on site choice utility.

For Wabakimi Park the findings are somewhat similar. When the current road access level of 0 is expanded, increasing per trip benefits result with a maximum at two roads. Beyond this level, however, the benefits decline and in fact become negative with six roads. These findings are driven by the congestion forecasts which suggest pronounced changes at 4 roads and at 6 roads. Once again these findings support the notion of threshold effects.

When comparing the road access changes at the two parks one is struck with the differences in the scale of the welfare measures. The values for Quetico differ from those at Wabakimi by a factor of 10. This could be the result of the relative importance of these two parks in the recreation demand system. Both the BWCA and Quetico are major destinations for the majority of wilderness canoeists in the market. Wabakimi Park is more remote and is not as well known.

Having demonstrated that congestion forecasts have a significant influence on recreation site choice in this data, it would be instructive to explore the distributional consequences of these effects. This is possible because a significant source of individual heterogeneity in the modelling framework results from the congestion forecast model. Recall that the ordered logit model was estimated with four individual characteristics, one individual level perceptual variable, and two recreation destination characteristics (Table 2). One would predict that the AC model, which incorporates these variables through the congestion forecast, may exhibit more variation in the welfare impacts among the 580 individuals in the sample than the RC model. This prediction is supported by comparing the mean, standard deviation, and the range in the individual welfare measures which are displayed in Table 4. For each road added to either Quetico or Wabakimi parks the variation in the welfare measures is far more pronounced in the AC models than the RC models. Since these effects are the result of the distribution of the welfare measures models. Since these effects are the result of the distribution of the welfare measures among the sample, the variation exhibited in the AC model results from the characteristics of the individual recreationists included in the ordered logit model. This suggests another way to incorporate heterogeneity of preferences in random utility models as well as to understand the distributional consequences of resource management policies.

DISCUSSION

The theoretical framework introduced in this chapter offers a viable solution to incorporating strategic behaviour in economic choice models. The notion of interdependent utility functions has a great deal to offer economists in studying a wide variety of issues, not only congestion. Part of the appeal of this framework is the notion of incorporating endogeneity through forecasts of other individuals' behaviour. These other individuals may be competitors, in which case the interactions are attenuating, or they may be facilitators in which case the interactions are reinforcing. This proposed framework may be a better basis for conducting policy analysis, especially if the analysis is capable of exploring heterogeneity in these types of interactions. This, in particular, is the case Schelling (1978) alludes to where "simple summations or extrapolations to aggregates" are not possible. In the example examined in this present study, the aggregate assessments of welfare masked considerable variation in the impacts of hypothetical policies.

This is not to say that the analysis presented in this chapter cannot be improved. For example, the econometric analysis could be improved to incorporate more fully the statistical linkages between the congestion forecast model and the economic site choice model. For example the standard errors in the choice model should be adjusted using corrections to the variance-covariance matrix in the manner suggested by Murphy and Topel (1985). Alternatively, the parameters of the ordered logit and conditional logit models could be estimated simultaneously using full information maximum likelihood techniques. These procedures will solve potential statistical issues resulting from the generated regressor problem. However, these techniques were beyond the scope of the current analysis.

Another limitation involved the discrete nature of the forecast function. This was restrictive in the number of categories (four) included in the data collection effort. The information displayed in Figure 1 initially had more categories, but these had to be collapsed to the four used in the ordered logit analysis for space considerations in the questionnaire. This discrete approach likely masked considerable additional variation among individuals in the data. While a continuous forecast function may be a better method to examine the issues raised in this chapter, this requires a different way to collect data and poses a considerable challenge for future studies addressing this issue. However, the result of using continuous data may point to the presence of considerably more heterogeneity in preferences, and might uncover more useful information on thresholds in behaviour.

An immediate extension on this present work might be to link the information portrayed in Figure 1 into the two-stage model process. One approach may be to introduce a threshold effect (through the threshold outlined in Fig 1) exogenously and then instead of impacting the utility function through the congestion variable, this threshold could operate to change the choice set in which the unattractive alternative is removed. It remains to be seen, however, if this approach would provide more useful information to decision makers.

Another extension to this analysis includes understanding and incorporating heterogeneity in terms of the types of interactions with people. In the context of wilderness canoeing, this may involve encounters between groups at entry points, while paddling on the water, at portages, and at campsites. In these instances, the critical numbers may differ at each type of location. Furthermore, in recreation contexts the types of people one encounters may have a profound influence on satisfaction levels and consequently may feed back on choice behaviour. For example, I tend to avoid camping at managed sites on long weekends to avoid encountering other campers more interested in drinking and making noise. I perceive that on long weekends these types of people are more likely to be encountered than during the week or on regular weekends.

These types of interactions will be complex to examine and incorporate in choice models. Some of these interactions may be more appropriately modelled using different types of equilibriums. For example, one could envision a leader-follower model in which large recreation groups involved with organizations such as Outward Bound or the Scouts Canada decide to visit wilderness areas and thereby affect the behaviour of the more solitary and smaller groups of recreationists.

Regardless of the improvements and enhancements one could make to the analysis reported here, I believe that the interdependent utility approach used by economists offers considerable promise in addressing the issues raised by Schelling (1978), Manski (1995), and others.

LITERATURE CITED

Adamowicz, W.L., J. Swait, P.C. Boxall, J. Louviere, and M. Williams. 1997. Perceptions versus objective measures of environmental quality in combined revealed and stated preference models of environmental valuation. *Journal of Environmental Economics and Management* 32:65-84.

Boxall, P.C., D.O. Watson, and J. Englin. 1996. Backcountry recreationists' valuation of forest and park management features in wilderness parks of the western Canadian Shield. *Canadian Journal of Forest Research* 26:982-990.

Boxall, P.C., J. Englin and D.O. Watson. 1999. Valuing backcountry recreation in wilderness parks: A demand systems approach in the Canadian Shield. *Information Report* NOR-X-361, Northern Forestry Centre, Canadian Forest Service, Edmonton, Alberta.

Cicchetti, C. J. and V.K. Smith. 1973. Congestion, quality deterioration and optimal use: Wilderness recreation in the Spanish Peaks Primitive Area. *Social Science Research* 2:15-30.

Cicchetti, C. J. and V.K. Smith. 1976. *The Costs of Congestion: An Econometric Analysis of Wilderness Recreation.* Ballinger Publ. Co., Cambridge, Mass.

Cole, D.N., A.E. Watson, and J.W. Roggenbuck. 1995. Trends in wilderness visitors and visits: Boundary Waters Canoe Area, Shining Rock, and Desolation Wildernesses. Research Paper INT-RP-483, Intermountain Research Station, US Department of Agriculture, Forest Service, Ogden, Utah.

Deyak, T.A. and V.K. Smith. 1978. Congestion and participation in outdoor recreation: A household production function approach. *Journal of Environmental Economics and Management* 5:63-80.

Duinker, P.N. 1998. Public participation's promising progress: Advances in forest decision-making in Canada. *Commonwealth Forestry Review* 77:107-112.

Englin, J., P.C. Boxall, and D.O. Watson. 1998. Modeling recreation demand in a Poisson system of equations: An analysis of the impact of international exchange rates.

American Journal of Agricultural Economics 80(2): 255-263.

Erbring, L. and A.A. Young. 1979. Individuals and social structure: contextual effects as endogenous feedback. *Sociological Methods and Research* 7:396-430.

Granovetter, M. 1978. Threshold models of collective behavior. *American Journal of Sociology* 83:1420-1443.

Hanemann, W.M., 1982. Applied welfare analysis with qualitative response models, Working Pap. No. 241. Univ. California, Berkeley, 26 pp.

Jakus, P. and W.D. Shaw. 1997. Congestion at recreation areas: Empirical evidence on perceptions, mitigating behaviour and management preferences. *Journal of Environmental Management* 50:389-401.

Manski, C.F. 1993. Identification of endogenous social effects: The reflection problem. *Review of Economic Studies* 60:531-542.

Manski, C.F. 1995. *Identification Problems in the Social Sciences*. Harvard University Press, Cambridge, Mass.

McConnell, K.E. 1977. Congestion and willingness to pay: a study of beach use. *Land Economics* 53:185-195.

McConnell, K.E. and J.G. Sutinen. 1984. An analysis of congested recreation facilities. In: V.K. Smith and Witte (Editors), *Advances in Applied Micro-Economics*, Volume 3, JAI Press, Norwich, CT, pp 9-36

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. In: P. Zarembka (Editor), *Frontiers in Econometrics*. Academic Press, New York, NY, pp. 105-142.

Menz, F.C. and J.K. Mullen. 1981. Expected encounters and the demand for outdoor recreation. *Land Economics* 57:33-40.

Murphy, K.M., and R.H. Topel. 1985. Estimation and inference in two-step econometric models. *Journal of Business and Economic Statistics* 3:370-379.

Pollak, R.A. 1976. Interdependent preferences. *American Economic Review* 78:745-763.

Schelling, T.C. 1978. *Micromotives and Macrobehaviour*. W.W. Norton and Co., New York, NY.

Shelby. B. 1981. Crowding models for backcountry recreation. *Land Economics* 56:43-55.

Smith, V.K. 1981. Congestion travel cost recreational demand models and benefit evaluation. *Journal of Environmental Economics and Management* 8:92-96.

TABLE 1
Expected congestion levels and trip activity to five wilderness parks in the Canadian Shield

Expected number of groups seen per day	Percent of number of respondents				
	Woodland Caribou	Quetico	BWCA	Wabakimi	Eastern Manitoba
0	47.0	4.4	1.5	38.8	18.5
1-3	50.0	70.8	21.2	51.0	62.0
4-9	1.8	21.3	50.4	10.2	16.7
over 9	1.2	3.4	26.9	0	2.8
Number of respondents reporting congestion information	166	614	458	49	108
Number of trips taken during 1995 and 1996	199	792	494	16	222

TABLE 2
Parameter estimates for an ordered logit model explaining perceived congestion at five wilderness park areas in eastern Manitoba and northwestern Ontario.

Variable	Parameter (t statistics)	
	Model 1	Model 2
Constant	2.5351 (9.196)	2.2632 (5.028)
Gender (male)	-0.3803 (-1.971)	-0.4268 (-1.999)
Years of experience in backcountry areas in the study area	0.0081 (1.615)	0.0122 (2.328)
Typical trip length (days)	-0.0412 (-2.819)	-0.0569 (-3.354)
Member of a conservation or recreation organization	0.1390 (1.286)	0.2457 (2.094)
Perceived level of development	0.1139 (1.541)	0.1470 (1.825)
Size of park		-0.0584 (-6.997)
Number of roads accessing park		0.9007 (22.373)
μ_1	2.6386 (27.829)	3.4918 (26.891)
μ_2	4.3020 (35.640)	5.8445 (33.935)
Log likelihood	-1516.38	-1216.28
% correct predictions	50.9	60.8

TABLE 3

Parameter estimates for a conditional logit choice model explaining wilderness park choice among five areas in eastern Manitoba and northwestern Ontario.

Variables	Parameters (*t* statistics)				
	Reported congestion models		Anticipated congestion models		
	RC1	RC2	AC1[1]	AC2[2]	AC3[2]
Travel cost	-0.00405 (-9.221)	-0.00415 (-9.553)	-0.00425 (-9.828)	-0.00221 (-5.153)	-0.00219 (-5.091)
Perceived chances of entry	0.46766 (11.317)	0.48283 (12.049)	0.47834 (11.950)	0.18530 (4.048)	0.18590 (4.010)
Perceived levels of development	-0.08094 (-1.437)				0.35276 (5.875)
Congestion	-0.11229 (-1.653)	-0.12246 (-1.815)	-11.00700 (-0.072)	-3.22100 (-13.893)	-3.91880 (-14.748)
Park size	0.06260 (11.190)	0.06186 (11.222)		0.01571 (1.994)	0.00159 (0.194)
Roads	0.01937 (0.705)	0.02592 (0.956)		0.93324 (13.139)	1.14900 (14.136)
ASC - Manitoba Parks	-1.07710 (-9.728)	-1.03280 (-9.772)	-1.05250 (-9.975)	-1.08730 (-8.668)	-0.92037 (-7.194)
Log Likelihood at convergence	-2310.97	-2312.01	-2311.48	-2204.4	-2187.5
ρ^2	0.164	0.164	0.164	0.203	0.209

[1] Anticipated congestion levels in this model were estimated from Model 1 in Table 2.
[2] Anticipated congestion levels in this model were estimated from Model 2 in Table 2.

TABLE 4

The magnitude and distribution of the welfare effects of increasing road access at Quetico and Wabakimi Provincial parks in Ontario on wilderness recreationists.

Road expansion increase from base level	Quetico Park							
	Reported congestion model				Anticipated congestion model			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Base level	0				0			
1	2.21	0.65	0.1	4.27	226.22	78.92	-375.53	292.6
2	4.45	1.3	0.2	8.58	-129.65	196.32	-283.57	639.55
3	6.73	1.95	0.3	12.92	-77.15	21.74	-426.5	19.15
4	9.04	2.61	0.41	17.29	-153.25	165.25	-398.8	147.27
5	11.39	3.27	0.52	21.7	-212.48	35.31	-333.37	153.62
6	13.78	3.93	0.63	26.14	-134.25	18.82	-200.08	-32.17
	Wabakimi Park							
1	0.17	0.06	0.01	0.45	8.1	3.01	-42.72	12.77
2	0.35	0.13	0.01	0.91	28.25	6.19	-34.98	43.72
3	0.53	0.19	0.02	1.38	21.18	35.93	-15.87	113.36
4	0.71	0.26	0.03	1.86	3.55	2.22	-44.47	21.38
5	0.9	0.33	0.04	2.36	8.55	11.07	-39.35	26.45
6	1.1	0.4	0.04	2.86	-3.08	2.3	-26.6	42.66

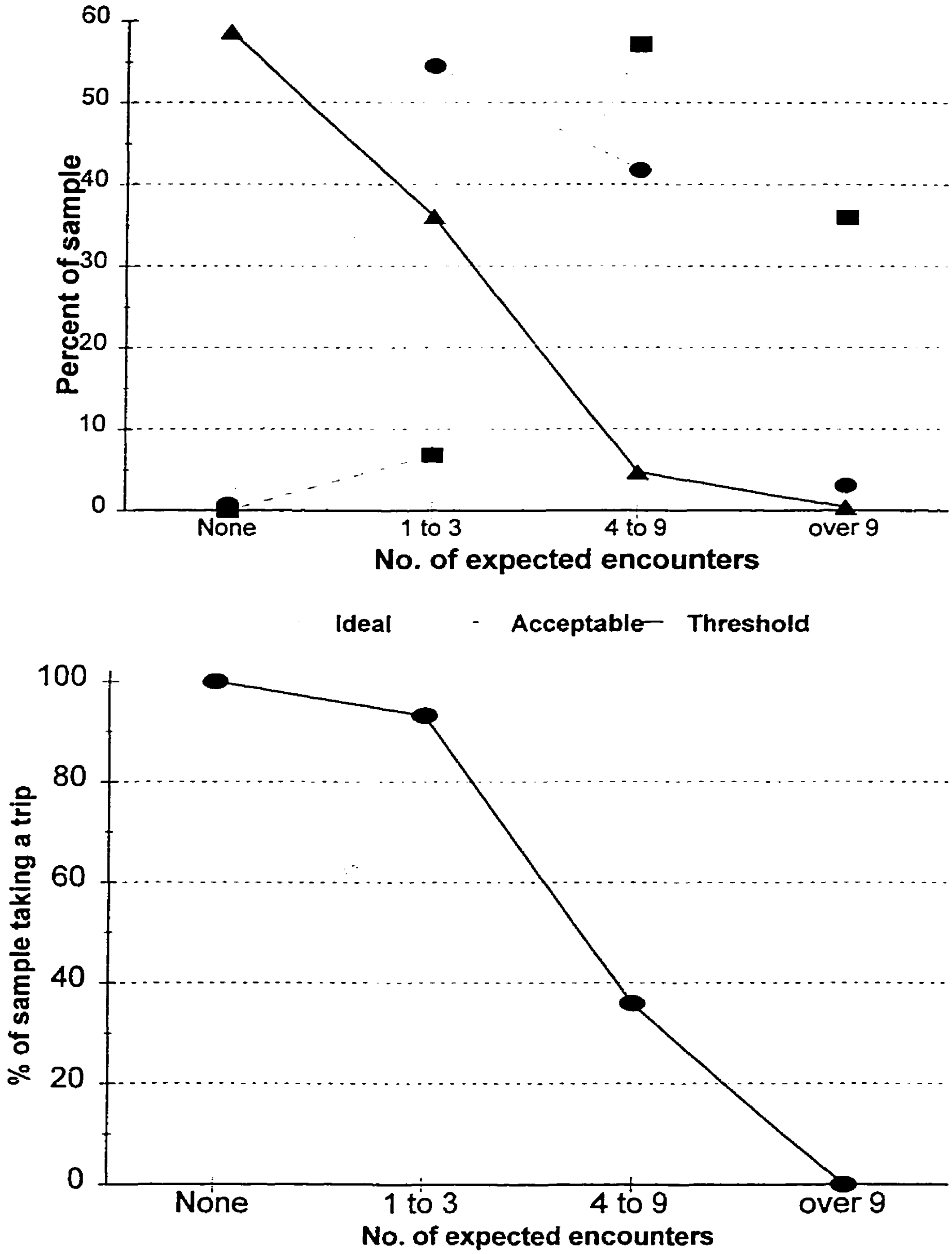
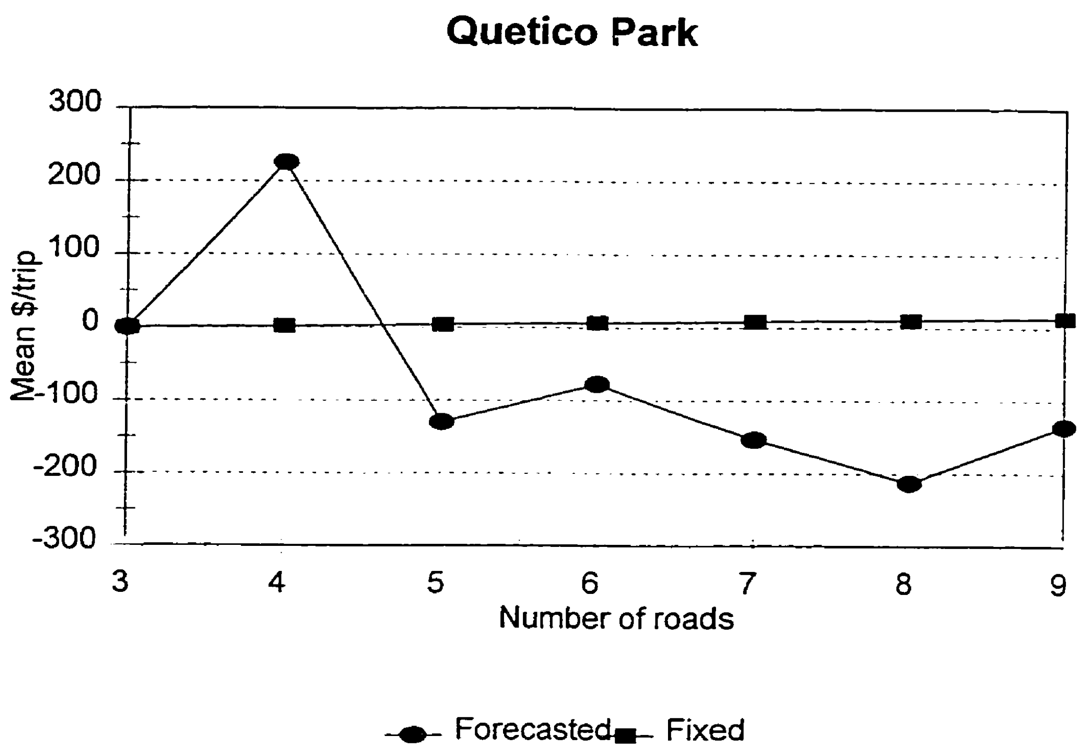


Figure 1. Responses of wilderness recreationists to various levels of expected encounters with other recreationists and the percentage of them taking a trip under these conditions.



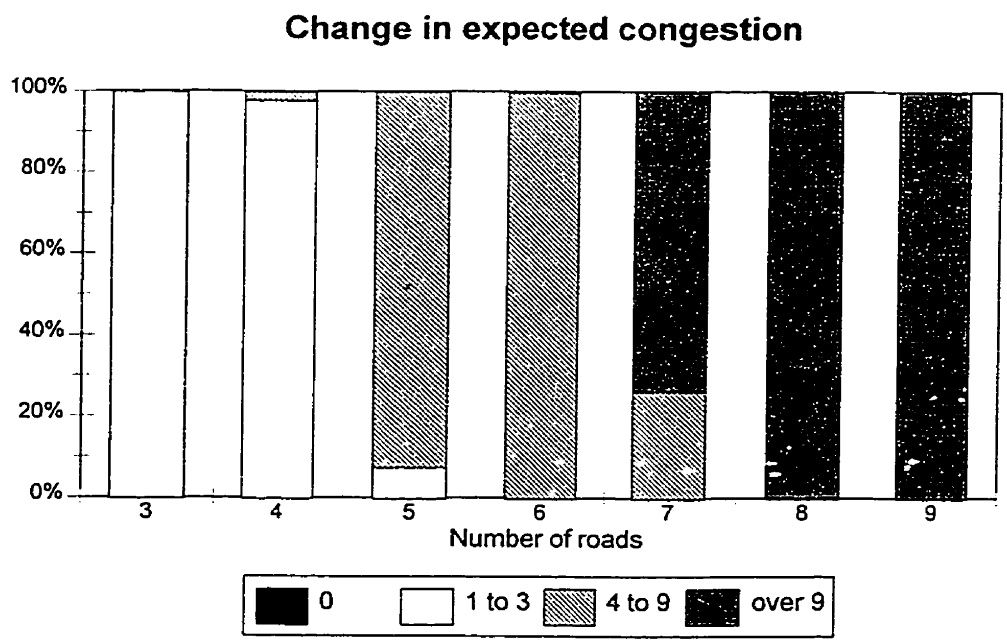
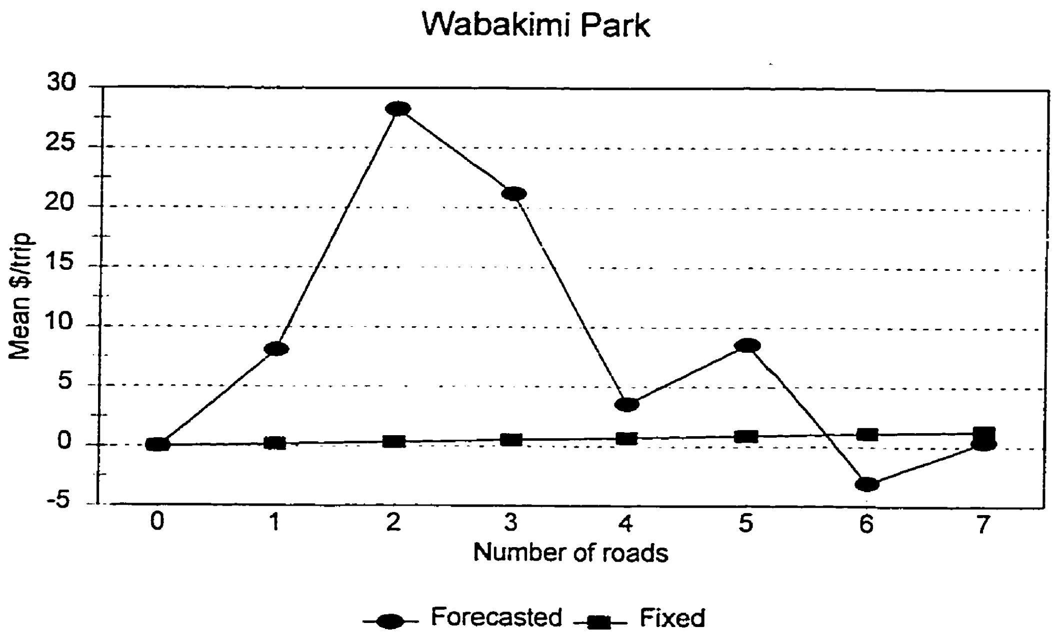


Figure 2. The effects of changing road access at Quetico Provincial Park on welfare and expected congestion of wilderness recreationists.



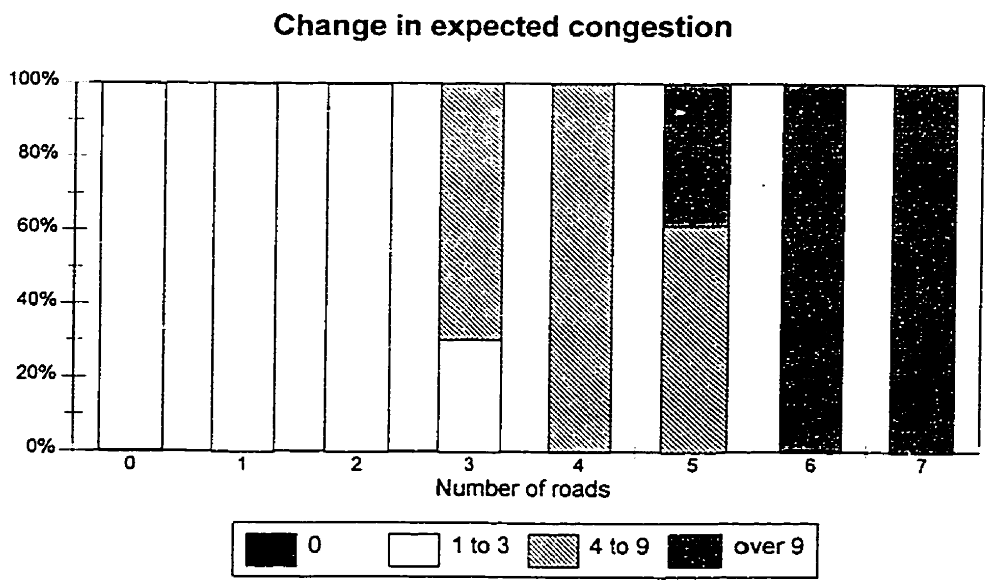


Figure 3. The effects of changing road access at Wabakimi Provincial Park on welfare and expected congestion of wilderness recreationists. Note that in this simulation the size of the park was held at 1550 km^2.

CHAPTER 5

Conclusion

The random utility model has been used by environmental economists to understand the contributions of elements of the environment or an experience to utility. The advantages of using this model are its ability to portray economic choice behaviour in a plausible and useful configuration, and that the model can be developed using either (and both) revealed and stated preference data. However, these applications have been relatively simple in that the behaviour assumed to underpin the analysis is rudimentary. There are a number of instances where more complex behavioural underpinnings would be desirable. The research reported in this thesis attempted to examine some different approaches to understand how and why people make choices.

In the first paper, presented in Chapter 2, the objective was to understand the demand to view aboriginal artifacts in wilderness areas. While this seems a relatively simple question, the problem was that the analysis used data from areas where no such artifacts were known to exist. Thus, the approach used to examine this issue was a combined revealed and stated preference analysis. The results suggested that wilderness users would switch their site choices to areas with artifacts and that consequently, this would generate large recreation benefits. However, with increased levels of visitation comes an increased risk of vandalism, which in turn affects the level of benefits to recreationists and also to aboriginal peoples. This research pointed to the need to understand and incorporate feedback mechanisms in understanding human behaviour. In this case, an attribute would generate utility, which in turn generates higher visitation; but this higher visitation produces the potential for a negative externality. Formally incorporating this effect in the model would probably affect the visitation forecasts and benefits resulting from discovery of artifacts.

In the second paper, the assumption of homogeneity of preferences across a sample of wilderness users was examined. In developing the choice model an integrated approach was used which involved specifying the utility function and a segment membership function. In this analysis, social psychological methods were used to explain segment membership in a latent class framework. The behavioural data used in this analysis came from a choice experiment, a stated preference method. The resulting analysis provided a much richer interpretation of wilderness recreation site choice behaviour than the traditional single segment model (which assumed homogeneity of preferences). For these data the latent segment approach suggested that preference heterogeneity was related to the motivational constructs underlying wilderness trips, sociodemographic characteristics, preferences for specific wilderness parks (holding changes in their characteristics constant), and perceptions of managerial attributes and congestion levels at the five parks. These findings supported both the economic and social psychological constructs related to wilderness recreation behaviour reported in the literature.

The third paper introduced a theoretical framework which offers a viable solution to incorporating strategic behaviour in economic choice models. In this framework, the notion of interdependent utility functions was used to examine the issue of recreational congestion. The analytical structure also involved the notion of an equilibrium which incorporates endogeneity through forecasts of individuals' behaviour. For an individual recreationist other individuals may be competitors, in which case the interactions are attenuating, or they may be facilitators, in which case the interactions are reinforcing. This equilibrium may be a better basis for conducting policy analysis, especially if the analysis is also capable of exploring heterogeneity in these types of interactions. In the empirical case examined, the aggregate assessments of welfare masked considerable individual variation in the impacts of hypothetical policies.

While the research reported in the three chapters provides useful extensions of the RUM, the empirical applications of the models are also of considerable interest. In each case the RUM was used to examine issues relating to wilderness recreation in the Canadian Shield region of Manitoba and Ontario. This area does have a number of signature parks; one in particular being Quetico Provincial Park in Ontario. Congestion in the backcountry areas of these parks is an important management concern, as is the preservation of aboriginal and cultural artifacts. Thus, the empirical applications reported in this thesis should be of considerable interest to wilderness park managers.

However, the development of economic models to facilitate the management of wilderness areas should ideally incorporate *each* of the approaches examined in the three chapters. In other words, the extensions described above should represent components of a broader strategy to examine economic choice behaviour. To illustrate this, the research reported in the second and third chapters could be linked with the analysis in the first chapter in examining the economics of vandalism. For example, the feedback mechanism alluded to in the study of pictographs is the subject of the chapter on congestion. In the pictograph case, as with the congestion case, the interactions between individuals in the case of vandalism are attenuating. Instead of modeling congestion forecasts, one could model forecasts of vandalism. The heterogeneity in responding to these negative interactions could be examined using the latent class methods described in Chapter 3. One could devise a measurement scale that evaluates attitudes to vandalism and uses this as the basis for determining the latent classes. The latent classes would then be the basis for determining differential feedback responses to vandalism. The result of using these techniques in examining the vandalism of pictographs would be better understanding of the demand to see rock paintings and improved knowledge of the potential for and the distributional impacts of vandalism.

A more complete analysis of wilderness recreation incorporating these modeling strategies, however, will have to wait until opportunities to collect more appropriate data arise. This will require the development of scales to assess forecasts of behaviours of interest, attitudinal instruments, and additional surveys of recreationists. Until such opportunities occur one hopes that the questions and issues raised in this thesis will at least stimulate discussion about the appropriate management of public goods.