Resource and Energy Economics 25 (2003) 177–195
RESOURCE and ENERGY ECONOMICS
www.elsevier.com/locate/econbase

Heterogeneous preferences for congestion during a wilderness experience

Peter Boxall [a], Kimberly Rollins [b,*], Jeffrey Englin [b]

[a] *Department of Rural Economy, University of Alberta, Edmonton, Alta., Canada*
[b] *Department of Applied Economics and Statistics, University of Nevada, Mail Stop 204, Reno, NV 89557-0105, USA*

Received 7 December 2000; received in revised form 22 July 2002; accepted 5 August 2002

Abstract

This analysis breaks down the congestion levels experienced during specific parts of a wilderness canoe trip. By explicitly addressing the heterogeneity in preferences for congestion during a trip, we are able to determine the relative value canoeists place on solitude at different points of a trip. Our econometric model utilizes a random effects probit framework to efficiently estimate the welfare impacts of congestion on each trip portion. The welfare effects of congestion levels vary across wilderness areas, parts of a trip and individuals.

JEL classification: Q26

Keywords: Panel estimators; Contingent valuation; Wilderness recreation; Congestion; Heterogeneous preferences

1. Introduction

Welfare reducing congestion effects can result from attempts to accommodate an increasing demand for recreational user days in natural areas. The recognition of the economic costs of congestion of recreational resources dates from Fisher and Krutilla (1972). Cicchetti and Smith (1973, 1976) were among the first to empirically measure costs of congestion, using contingent valuation to estimate the willingness-to-pay by hikers to avoid encounters with others at trailheads and campsites. Since then, the literature has grown substantially to address a number of specific problems, including the definition of congestion, empirical

* Corresponding author. Tel.: +1-775-784-1677/6701; fax: +1-775-784-1342.
E-mail address: krollins@cabnr.unr.edu (K. Rollins).

PII: S0928-7655(02)00025-8

issues related to measuring congestion costs, and the welfare implications of heterogeneous preferences when rationing by price and by quota.

Efforts to develop policy-relevant measures of congestion are typically motivated by the problems of setting optimal user fees and site visitation quotas. Another use for congestion measures that is not well described elsewhere, is to optimize site capacity in circumstances where trips are routed through sites using a number of discrete and overlapping segments. Routes may be developed to accommodate both natural features and management objectives. In these cases, systems of routes can be designed to relieve congestion at key points where users are most sensitive. For this type of problem, it is desirable to determine how marginal costs of encounters vary among different segments within a site for individual users, as well as across users who have different characteristics and preferences. While a number of recent studies treat preferences for congestion as heterogeneous, we are not aware of economic studies that explicitly model preferences as heterogeneous both within users and across users for components of a trip.

Theory and empirical studies indicate that the assumption of homogenous preferences is likely to lead to incorrect conclusions about optimal use levels and incorrect welfare measures. Freeman and Haveman (1977) first considered effects of heterogeneous preferences for congestion among recreational users and concluded that optimal allocation and prices will generally depend upon the distribution of user preferences. In demonstrating that “the socially optimal level of congestion of a facility at any one time will depend on the levels of congestion at other times,” Dorfman (1984) does not model preferences as heterogeneous, but instead considers the good to be a trip at a specific time and date, with other times and dates at the same site modeled as substitute goods. Michael and Reiling (1997) model optimal allocation between weekend and weekday trips as a heterogeneous preference problem, which they test empirically to demonstrate that failure to account for heterogeneous preferences would overestimate congestion costs.

McConnell (1988) shows that, if demand for a congestible recreational good is income elastic, rationing via price among heterogeneous users will increase the demand for site visits among some groups, even while decreasing overall demand. The effect is to make the users more homogenous, favoring higher income user groups. McConnell contrasts this result with the effects of increasing site capacity, which results in unambiguous reductions in congestion costs while not decreasing demand by any one group. He notes that the theoretical results support site managers’ reluctance to use price rationing, in favor of setting quotas or rationing via congestion, due to equity considerations. He concludes that management strategies for congestible facilities must account for interactions among different users with different preferences.

A number of empirical problems have made the practical application of congestion measures somewhat difficult. Due to the complexity of recreational decision-making, the definition and measurement of congestion continues to be controversial to social psychologists and economists (Shelby, 1981; Jakus and Shaw, 1997). Congestion measurements that are relevant to the recreationist may not be equivalent to those that have been developed and measured by the outside observer. A popular approach used in recreational planning is to derive encounter measures that describe the number of other visitors an individual remembers seeing during a trip or at a given location. Crowding is then defined as a negative evaluation of those encounters (Shelby et al., 1989). Most recreation researchers subscribe to an approach that predicts that disutility associated with crowding will increase until a visitor's tolerance limit is reached (Manning, 1995; Vaske et al., 1986; Graefe et al., 1984; Donnelly et al., 2000). Visitors who experience disutility from crowding will be willing to pay to reduce the number of encounters. However, over some initial range of congestion it is possible that a visitor will be indifferent or even value an additional encounter. This is an empirical issue.

Jakus and Shaw (1997) point out that all valuation estimates for reduced congestion currently in the literature are derived from pure stated preference models. They suggest that the apparent void with respect to valuation of congestion costs based on observed trip-taking behavior is because congestion is difficult to measure in a revealed preference model. Past experiences, expectations, self-selection of people among various sites, and mitigating actions on a given trip can confound revealed preference measures. Others (Graefe et al., 1984) note that recreationists may adjust their expectations to account for anticipated special features of a given trip so that their tolerance for congestion changes with other trip attributes. Similarly, unexpected factors that occur in a given trip may tend to increase or reduce the marginal contribution of congestion to the overall level of welfare achieved from the recreational experience. Therefore, it has been suggested that obtaining revealed preference measures of congestion costs from actual choices involves interpreting the behavior of a moving target. We take the view that this is an empirical issue and that an appropriately designed random utility model can condition estimates of the marginal costs of congestion on unobservable expectations and mitigating behavior by using a combination of revealed and stated preference data.

Dorfman (1984, p. 105) notes that the concept of congestion can be ambiguous in an analytical context, because "for the purpose of establishing demand curves, expected, or anticipated congestion is relevant, but for appraising the results in terms of social and consumers' surpluses, the actual level of congestion experienced is appropriate." Jakus and Shaw (1997) define two types of measures for congestion. Ex ante measures are related to the individuals' expectations in advance of taking a trip, and are thus likely based on recollections of past congestion experience from a variety of sites and trips. Ex post measures are, on the other hand, related to perceived feelings of crowding during a given trip, and are thus likely to be affected by on-site experiences, taking into account mitigating behaviors of the individuals during and in preparation for the given trip. They suggest that these are two different measures, and that it is important to define which is being measured, as demand will depend upon the measure.

We take this as a starting point, but suggest that while the two measures are not equivalent, they are both necessary. It is possible to define a single measure of congestion costs that combines the two. Indeed, as Dorfman (1984) points out, at the optimum, the two measures would tend to alignment. Observable ex post measures of congestion are conditional on ex ante expectations and the decisions over mitigating factors (such as choice of site, date of trip) are made based on ex ante expectations. This can be modeled using an approach that combines observed and stated preferences, and treating the stated preference responses as conditional on observed behavior.

1.1. Trip routing to increase efficiency

A management option to deal with excess demand for a recreational area is to increase the efficiency with which the site can provide recreation benefits by increasing capacity or optimizing use for an existing quota. Quotas tend to be the most common form of rationing access to parks, where the quota is indicative of marginal congestion costs. The knowledge of which specific points in a wilderness experience congestion is most costly can be crucial to the efficient design of trip routes and quota schemes.

As an example, we consider the three main backcountry canoe areas within Ontario's Provincial Park system (Quetico, Killarney and Algonquin Provincial Parks). In these areas the backcountry permit is a license to a specific route of campsites, or lakes, for specific nights. Within a given park, a variety of lakes each support different numbers of campsites; portages differ in length, level of difficulty, and probability of encounters with other groups; alternative entry points vary in remoteness and congestion during the first and last days of a trip; and waterways differ in width, level of difficulty, current, and other features that would be expected to affect the numbers of other paddlers encountered in a given day of paddling. Daily quotas are specified by lake and/or campsite. The feasible number of routes through each park is constrained by natural features and by the locations of portages and campsites. Each park supports a number of different routes, each being a different combination of entry and exit points, campsites, portage sites, paddling sites, and trip length. While there tend to be bottlenecks at entry/exit points and first portages, congestion at these areas is expected, and anecdotal evidence indicates that marginal congestion costs may not be as great at these points as they are elsewhere within the parks.

Individual user groups choose routes based on their preferences and park features and are issued permits associated with their chosen routes. Users choose routes by booking campsites through the central computerized reservation system for Ontario Provincial Parks. Individuals found camping at a site without a valid license can be charged, unless they can demonstrate that the violation was due to unforeseen circumstances, such as bad weather or accident. Because quotas for Algonquin and Killarney Provincial Parks are fully subscribed throughout most of the season, any attempt by a canoeist to change their route is likely to result in a greater sense of crowding than they would have experienced otherwise, and/or a potential fine. Permits require that users vacate their sites no later than 2:00 p.m. each day. Canoeists tend to respect the route plans and terms of the permits. While people are ticketed each season, there are few complaints about congestion. Managers agree that without a highly efficient system of quotas, the wilderness character would be quickly eroded.

The management feature that makes this possible is the definition of the backcountry permit as a right to a specific route through the site on a specific set of dates, rather than as the right to enter the park. This effectively makes alternative routes (and those on different dates) substitute goods. In order to make such a system most efficient, it would be reasonable to identify where within the park additional encounters create greatest losses, and where reductions in encounters generate the greatest gains. Ontario Provincial Parks are not unique in their experience of increasing demand. Since solitude is an important attribute of wilderness recreation, it would be expected that more precision in predicting and managing for congestion is a desirable avenue for further research. Knowledge of marginal costs of congestion for various parts of a trip could allow managers to more efficiently allocate visitor quotas within parks.

This paper develops and applies a random utility model to value changes in congestion experienced during different segments of a wilderness canoe trip. The model uses multiple stated preference responses from individuals and combines these with revealed preferences to account for the correlation between congestion values and an individual's own preferences, experiences and expectations, and across individuals who engage in similar actual experiences. The analysis undertaken in this paper follows the general econometric insights provided by Loomis (1997), who has suggested random effects probit specifications to incorporate multiple stated preference responses per individual and also revealed preferences into the estimation of welfare effects of changes in the quality of recreation sites.

The model is applied to changing congestion levels for wilderness canoe trips in the three Ontario parks (Killarney, Quetico and Algonquin) that represent Ontario's main wilderness canoeing destinations. All three parks contain formally designated wilderness areas that cater to canoeists, and all use quotas to manage visitation levels to their wilderness zones. If the effects of encounters with other groups are the same for a given individual, regardless of where they occur during a trip, then congestion enters the utility function once and only one measure is needed. If encounters with other groups are evaluated differently, depending upon when and where they occur during a trip, then assuming otherwise would lead to biased estimates of congestion costs and inefficient use of the sites. This paper tests whether preferences are heterogeneous for a given individual and across individuals for different components of a trip.

2. The econometric model

2.1. Behavioral model

The underlying behavioral model is based on the random utility model of consumer behavior. The standard structure of the random utility model represents utility as two parts, consisting of a systematic and a random component. Thus, preferences for congestion can be represented as:

$$U = V(c^j, y, s) + \varepsilon, \quad (1)$$

where V is the systematic portion of utility and ε represents the random portion, y represents income, s is a vector of individual characteristics, and c^j represents congestion at level j. An individual will agree to pay an additional cost, P, to reduce congestion if the value of the change in congestion is greater or equal to the lost income:

$$V_1(c^1, y - P, s) + \varepsilon_1 \geq V_0(c^0, y, s) + \varepsilon_0. \quad (2)$$

Otherwise, the individual will choose to keep the income and live with the old congestion levels. The process, however, is stochastic. The individual's response is a random variable with a probability distribution represented by:

$$\text{Prob(paying)} = \text{Prob}[V_1(c^1, y, -P, s) + \varepsilon_1 \geq V_0(c^0, y, s) + \varepsilon_0], \quad (3)$$

and Prob(not paying) $= 1 -$ Prob(paying). The maximum willingness-to-pay (P^*) for the change in congestion is found by setting $\Delta V = 0$ and solving for P. Hanemann and Kanninen (1999) show that if one assumes that the difference between random terms is generated by a standard normal CDF, the probability that an individual pays to change the congestion level corresponds to a standard probit model.

The foregoing model assumes that an individual experiences a constant level of congestion throughout the trip. In reality, a wilderness canoe trip consists of a number of different activities with different congestion levels among the activities. Congestion levels may change systematically with the different activities that an individual participates in during a trip. If this were the case, the estimation of the welfare costs of congestion would need to account for that systematic variation. There is also the possibility that preferences for congestion are different across different activities engaged in during a trip. It has been well established elsewhere that welfare impact of congestion is a function of, but not equivalent to, numbers of encounters with others. Thus, even if congestion levels remain the same over the different types of activities in a trip, there is no reason to assume that the change in welfare induced by a marginal change in congestion would be similar across activities. We denote congestion as c_t for a trip that consists of four main types of activities, $t = 1, 2, 3, 4$ (camping, paddling, portaging, and activities during the first and last days of the trip). The utility function can be represented by:

$$U = V(c_1, c_2, c_3, c_4, y, s) + \varepsilon. \tag{4}$$

Given the utility function in Eq. (4), we would like to determine how changes in congestion levels during each of the four different activities affects welfare, and whether those individual valuations are statistically different. One can envision a valuation procedure where congestion is changed during one activity, holding congestion constant during the other activities.[1] Modeling this process requires replication of (2) for each activity. In the case where there are four activities within the trip, the resulting model would be a panel with the following structure, in which '*' denotes a congestion level held constant:

$$\begin{aligned}
&V_1(c_1^*, c_2^*, c_3^*, c_4^*, y - P_5, s) + \varepsilon_5 \geq V_0(c_1^*, c_2^*, c_3^*, c_4^*, y, s) + \varepsilon_0,\\
&V_1(c_1^1, c_2^*, c_3^*, c_4^*, y - P_1, s) + \varepsilon_1 \geq V_0(c_1^0, c_2^*, c_3^*, c_4^*, y, s) + \varepsilon_0,\\
&V_1(c_1^*, c_2^1, c_3^*, c_4^*, y - P_2, s) + \varepsilon_2 \geq V_0(c_1^*, c_2^0, c_3^*, c_4^*, y, s) + \varepsilon_0,\\
&V_1(c_1^*, c_2^*, c_3^1, c_4^*, y - P_3, s) + \varepsilon_3 \geq V_0(c_1^*, c_2^*, c_3^0, c_4^*, y, s) + \varepsilon_0,\\
&V_1(c_1^*, c_2^*, c_3^*, c_4^1, y - P_4, s) + \varepsilon_4 \geq V_0(c_1^*, c_2^*, c_3^*, c_4^0, y, s) + \varepsilon_0.
\end{aligned} \tag{5}$$

Such a panel would consist of five observations per individual. The first, corresponding to the first line in Eq. (5), would include levels of congestion during each of the four activities actually experienced by an individual on a trip, and a 'yes' or 'no' response to whether they would take the same trip if the cost to them was increased by P_5. This observation establishes willingness-to-pay for existing levels of congestion. The second observation would include

[1] The assumption that the number of encounters with other groups is independent across activities during a given trip may be quite strong for recreational experiences that are characterized by large numbers of visitors, where congestion levels may be correlated across activities. However, for the backcountry experiences that are considered here, it is not unreasonable to assume that the number of encounters with other groups while paddling is independent of that experienced while camping, or portaging. The mean numbers of encounters per day actually experienced in the backcountry canoe areas studied, given in Table 2, vary from 1.4 to 5.8. At these levels, a marginal change in encounters per day while portaging is not likely to have a significant impact on the probability of being the sole party using a backcountry camping area overnight. If the model were applied to an activity for which congestion is not independent among activities, then combinations of changes in congestion for other activities would need to be included as additional equations in (5).

the individual's 'yes' or 'no' response to whether they would take a trip with a specified change in congestion in activity c_1 and a corresponding change in the amount that the trip would cost, P_1. The levels of congestion in the other three activities remain unchanged. The following three observations per individual refer to changes in congestion for one activity at a time, with the others held constant, and changes in trip cost variables. The resulting four error terms, ε_1 through ε_4, are then systematically correlated with changing congestion levels for each activity.

A vector of variables, s, includes information specific to the respondent and their overall trip, such as experience, education, expectations about the trip, where the trip was taken (one of three backcountry recreation areas), the dates of the trip, and the duration of the trip. Variables in s will not vary within any one individual's observations. The variation across individuals and trips provides a second error component. A sixth observation per individual could include no change in congestion, no change in P, and a 'yes' response to the binary response. This last observation corresponds to the actual behavior of each respondent, and provides as a point of reference that we observe the trip at a given cost and actual congestion levels.

Using the binary process outlined above to model willingness-to-pay among the activities, a trip is specified as a series of T activities. Each canoeist is presented with a series of T congestion levels along with respective changes in income. Let Y_{it} represent the 'yes' or 'no' answers to whether respondents would take the trip in response to changes in cost and levels of congestion at the tth activity. The result is a set of continuous random variables represented by $Y_{i1}, Y_{i2}, \ldots, Y_{iT}$. If one assumes that each Y_{it} is independently distributed $\mathrm{IN}(0, \sigma^2)$, one can pool the observations across activities and use the standard probit model to estimate the parameters.

However, there can be activity-specific responses to changing congestion levels, or a cumulative overall response to congestion levels as a trip proceeds. Modeling these features of a wilderness experience requires distinguishing among three sources of variation in congestion responses across individuals: (i) purely random factors that arise independently in each trip activity; (ii) random factors, including variables unobservable from the researcher's perspective, that are correlated across the activities during a trip; and (iii) deterministic variables such as income or levels of wilderness experience that affect the probability of paying to change congestion levels.

This correlation pattern is captured in a random effects model (after Heckman and Willis, 1975):

$$Y_{it}^* = \alpha + \gamma c^j + X_{it}\beta + \mu_i + \varepsilon_{it}, \tag{6}$$

where Y_{it}^* is an unobserved latent variable, c^j represents the level of congestion, X_{it} is a $1 \times k$ vector of exogenous variables, α and γ are coefficients, and β is a $k \times 1$ vector of coefficients. The other two terms represent error components that are mutually independent. The first, μ_i, represents an unobservable characteristic specific to individual i that does not vary among the t observations from i and is $\mathrm{IN}(0, \sigma_\mu^2)$. These effects include the trip specific conditions (which cannot be identified separately from the individual effects). The second, ε_{it}, is $\mathrm{IN}(0, \sigma_\varepsilon^2)$ and is a component that varies among individuals and across the t observations from each individual. The observed random variable, Y_{it} is defined by: $Y_{it} = 0$ if $Y_{it}^* = 0$ and $Y_{it} = 1$ if $Y_{it}^* \neq 0$.

Let $\sigma^2 = \sigma_\varepsilon^2 + \sigma_\mu^2$, $\rho = \sigma_\mu^2/\sigma^2$ and impose the normalization that $\sigma^2 = 1$. Then if $Y_i = [Y_{i1}, Y_{i2}, \dots, Y_{iT}]$ is the observed sequence of choices for i, and defining $\upsilon_i = \mu_i/\sigma_\mu$:

$$P(Y_i) = \int_{-\infty}^{\infty} \prod_{t=1}^{T} \Phi \left[\left\{ \left(\frac{X_{it}\beta}{\sigma_\varepsilon} \right) + \mu_i \left(\frac{\rho}{1-\rho} \right)^{1/2} \right\} (2Y_{it} - 1) \right] f(\upsilon_i)\, \mathrm{d}\upsilon_i, \quad (7)$$

where $\Phi[\cdot]$ is the normal cumulative distribution function. In this expression, ρ represents the correlation coefficient between responses from i in any two trip activities, and allows the measurement of the proportion of total variance explained by systematic correlated components. If no correlation is present, $\rho = 0$ and the information can be pooled across activities and the model parameters estimated using the standard probit model. If $\rho \neq 0$, then use of the standard probit results in biased standard errors of the coefficients (Guilkey and Murphy, 1993) and one must consider the random effects in the estimation.

The likelihood function for the random effects probit for an observed sample of N Y_{it}'s is simply:

$$L = \prod_{i=1}^{N} P(Y_i), \quad (8)$$

where $P(Y_i)$ represents Eq. (7). As numerous authors note, the estimation of parameters is difficult due to the computation of the joint probabilities of a T-variate normal distribution, which involves evaluating T-dimensional integrals. However, by conditioning on the individual effects, this problem can be reduced to a single integral involving the product of a standard normal density and the difference of two normal cumulative density functions and solved using Gaussian quadrature procedures (Butler and Moffitt, 1982).

3. The data

The three parks under examination present different management challenges. The parks vary by the degree of crowding, proximity to urban areas, size, number of canoe routes, amount of advance time required for reservation booking, and landscape features. Killarney is the smallest (485 km^2), has a low interior camping quota and therefore often has long trip reservation lists. Killarney currently operates at 95% capacity all season, with a 5% vacancy rate to allow for unforeseen circumstances and natural recovery of vegetation at the sites. Some of the harder to access core areas may have an 80% occupancy rate at times. Killarney has a total of 171 campsites, with 2–14 sites at each lake and a quota of 45,000 camper nights between May and November.

Quetico (4655 km^2) offers the greatest degree of solitude and virtually no waiting time due to its size and remote location. Canoeists generally go to Quetico to take longer trips, and are attracted by its remoteness. This park lies on the US border across from the Boundary Waters Canoe Area, and 83% of the backcountry permits issued during 1993 were to US citizens.

Algonquin Park (7725 km^2) is situated within a half-day drive to the large metropolitan centers of Ottawa and Toronto. As a result, Algonquin has the largest quotas since its use is under the greatest pressure due to its proximity to large urban centers. While large, only a fraction of its area is designated as wilderness zoned backcountry. The rest is more heavily used front country with multiple uses, including commercial forestry. Canoeists destined for the interior typically meet large numbers of day users on the first and last days of an interior trip. There is no theoretical reason to presume that the marginal cost of congestion is precisely the same among different parks, different routes within parks, or that the impacts of congestion on the value of a trip are homogenous over all activities during a trip.

During the months of June–October 1993, trained park staff handed out questionnaires to backcountry canoeists taking trips of 2 or more days in the three parks, as the participants purchased their backcountry camping permits at the start of their trips. Canoeists returned the questionnaires by mail. Those who did not respond within 3 weeks of receiving their questionnaires received reminder postcards, and those who had not responded within another 2 weeks received follow-up mailings with duplicate survey materials. A total of three follow-up mail contacts were issued to non-respondents. As a final prompt, non-respondents were contacted by phone and asked to return their questionnaires. Approximately 50 participants were not contacted either because their backcountry permit form was not returned by the park, the address on the form was illegible or repeated attempts to locate them at the given address were unsuccessful.

As questionnaires were returned over the season, responses to dichotomous choice contingent valuation questions were analyzed, allowing for subsequent offer amounts on the remaining surveys to be updated for an increasingly more efficient bid design. Details of the survey design and effects of updating are given in Rollins et al. (1997). The sampling design stratified by park, park entry points, and the month of trip. A response rate of 88% resulted in 2434 usable surveys. The item non-response rate for the contingent valuation questions was less than 1.5%. An example of a contingent valuation question included in the questionnaire is shown in the Appendix A.

Table 1 summarizes trip lengths and expenditures per trip over the three parks. Respondents reported mean total trip-related expenditures of US$ 208 and 210 per trip for Algonquin and Killarney. Since Quetico is a more remote location, trips there were longer and total trip costs were higher (US$ 740). The mean willingness-to-pay (WTP) per day for a canoeing trip, over and above current expenses, estimated by Rollins et al. (1997) were US$ 67.37, 65.82 and 66.76 for Algonquin, Quetico and Killarney Parks. While WTP per trip day was similar for the three parks, due to the differences in average trip lengths the mean surplus per trip for each of the three parks was US$ 303.16, 514.74 and 320.45, respectively.

The questionnaire was designed to elicit valuation of congestion as a function of numbers of groups actually encountered over the course of the trip. Respondents were asked to report the number of encounters during each of four different activities during their trip: (1) the first and last day of the trip; (2) while paddling; (3) while portaging; and (4) while

Table 1
Respondents' trip length and expenditures by park

Park	Mean trip length (days)	Mean total trip expenditures (US$)
Quetico	7.7	740
Killarney	4.8	210
Algonquin	4.5	208

Table 2
Mean number of groups encountered per day by trip activities

Canoe trip segments	Quetico	Killarney	Algonquin
First/last days	5.6 (4.0) [1299]	4.3 (3.2) [468]	9.5 (9.2) [619]
While paddling	2.8 (1.8) [1303]	3.0 (2.0) [472]	5.8 (5.3) [622]
At portages	2.2 (1.7) [1275]	2.0 (1.9) [471]	3.9 (4.0) [598]
At campsites	1.5 (1.5) [1294]	1.4 (1.7) [464]	2.7 (3.7) [616]

The values given in parentheses indicate standard errors, and the values in square brackets indicate the number of observations.

camping. Table 2, which summarizes the means and standard errors for the encounter variables by park, illustrates the wide range of experiences in the revealed part of the data. Canoeists experienced a variety of congestion conditions among the three parks and within trip activities. Paddlers in Algonquin Park encountered more groups of canoeists at all trip activities than those in the other parks. Respondents in all three parks reported that the average size of their own group was about four people (two canoes), the same as the average size of the other groups they encountered. In most cases, there was little variation in group size, since wilderness backcountry permits limit the number of users per permit. Special permits for large groups are typically limited to a restricted set of routes through designated areas where large groups will not disturb other users. Since there was little variation in the numbers of people per group, our analysis uses the numbers of groups to describe congestion, rather than numbers of people.

Participants answered a series of contingent valuation (CV) questions that changed the number of encounters in each of the four trip activities. Each respondent was asked four single-bounded dichotomous choice contingent valuation questions about their willingness-to-pay for marginal changes from actual encounters at the four different activities during each trip. One of two versions of the CV questions was randomly assigned to respondents. The first version asked whether the respondent would be willing to take the same trip if their trip costs increased by a given dollar amount and if they were to experience half as many encounters with other groups per day. The second version specified that they would experience twice as many encounters per day. Therefore, for each of the four activities, the hypothetical good is the change in the number of groups encountered, which might be an increase or decrease. The change in congestion is calculated as the actual number of encounters for the activity times either 0.5 or 2.0, depending on whether the version was a half as many or twice as many encounter change from that experienced on the trip. The variables representing these four congestion goods are labeled in Table 3 as Cong first/last, Cong portaging, Cong paddling and Cong camping.

The combination of the valuation of the actual trip under existing crowding conditions, the variation in actual crowding conditions experienced, and the valuation of marginal changes from actual conditions results in an experimental design that allows for efficient panel estimators to be employed, and the impact of the hypothetical nature of the contingent valuation surveys to be quantified. The panel consists of six observations per individual as described above. The first observation corresponds to the actual trip taken, where the congestion at each trip activity is what was actually experienced, the change in congestion is zero and the

Table 3
Parameter estimates for three random effects probit models

Variables	Model 1	Model 2	Model 3		
			Algonquin (base)	Interaction term for Quetico	Interaction term for Killarney
Constant	2.4565[a] (0.2141)	2.6080[a] (0.2299)	2.8145[a] (0.2291)		
Fee	−0.0120[a] (0.0003)	−0.0123[a] (0.0003)	−0.0151[a] (0.0005)	0.0048[a] (0.0005)	−0.0011[a] (0.0006)
Cong first/last	0.0106[a] (0.0042)	0.0473[a] (0.0119)	0.0432[a] (0.0119)		
Cong paddling	−0.0258[a] (0.0102)	−0.0429[a] (0.0165)	−0.0338[a] (0.0110)		
Cong portaging	−0.0222[b] (0.0121)	−0.0092 (0.0237)	−0.0387[a] (0.0135)		0.2673[a] (0.0817)
Cong camping	−0.1647[a] (0.0100)	−0.3150[a] (0.0245)	−0.2152[a] (0.0299)	−0.1750[a] (0.0276)	−0.1861[b] (0.0768)
(Cong first/last)^2		−0.0007[a] (0.0002)	−0.0006[a] (0.0002)		
(Cong paddling)^2		0.0003 (0.0004)			
(Cong portaging)^2		−0.0008 (0.0009)			−0.0149[a] (0.0068)
(Cong camping)^2		0.0071[a] (0.0009)	0.0047[a] (0.0011)		0.0087[b] (0.0053)
Duration	0.1189[a] (0.0128)	0.1193[a] (0.0133)	0.0970[a] (0.0145)		
Age	0.0133[a] (0.0048)	0.0127[a] (0.0049)	0.0112[a] (0.0049)		
Income (US$ 1000)	0.0052[a] (0.0008)	0.0053[a] (0.0008)	0.0049[a] (0.0008)		
Version dummy	−1.1521[a] (0.1031)	−1.1622[a] (0.1072)	−1.1512[a] (0.1073)		
ρ	0.8739 (0.0249)	0.8821 (0.0251)	0.8786 (0.0258)		
Log likelihood	−4413.87	−4376.20	−4312.89		
Percentage success	77.2	77.0	77.5		

Model 3 presents the differences between parks, using Algonquin as the base and the parameter estimates for the other two parks as deviations from that of Algonquin, where those differences are significant.

[a] Significant at 1% or better.

[b] Significant at 10% or better.

change in willingness-to-pay is zero. For this observation the dependent variable is 'yes' for all individuals. The second observation corresponds to the response to a dichotomous choice contingent valuation question about whether the respondent would take the same trip, with the same levels of congestion, but at a higher cost. For this observation, the change in congestion is zero, and the dependent variable is the standard 'yes' or 'no' response. The last four observations correspond to the responses to the respondents' responses to whether they would take the trip under hypothetical conditions of halving or doubling congestion levels at each of the four trip activities, for a specified cost increase. So for these, the change in congestion is either twice or half of actual reported congestion levels. A total of 2110 surveys were used (324 had missing values for one or more variables, and so were omitted from the analysis) for a total of 12,660 observations in the panel.

Three random effects probit models were estimated to test a series of hypotheses. The first model is a random effects probit that treats congestion effects for each of the four activities as linear. In this model, responses are pooled across the three parks, so that parameters are constrained to be equal across the three parks. This model allows us to test whether marginal congestion costs differ across trip segments. That is, a given user values congestion differently depending upon where in the trip it occurs. The null hypothesis is: $\text{Cong first/last} = \text{Cong portaging} = \text{Cong paddling} = \text{Cong camping} = 0$. We also hypothesize that ρ is significantly different from zero due to the systematic differences in preferences among different individuals. That is, preferences are hypothesized to be heterogeneous among users.

The second model is identical to the first, but treats marginal congestion effects for each activity as non-linear. That is, the null hypotheses are $(\text{Cong}\, c_i)^2 = 0$, for $i = 1, \ldots, 4$, where Cong c_i refers to congestion levels for each of the four activities. We also use this model to test the null hypothesis that congestion costs are the same across the three parks. A maximum likelihood ratio test would indicate whether the pooled model performs just as well as three individual models, where a separate one is estimated for each park. Failure to reject the null hypothesis would suggest that preferences for congestion differ by park. If participants have different expectations for the three parks, and park features can mitigate for differences in expected crowding, then we would expect to see evidence of this in terms of differences between parks in preferences for congestion.

If we are unable to reject the hypothesis that the three parks differ, then a third pooled model should be generated that would include additional variables as interaction terms for each activity and park, where the differences between parks are significant. For example, park-specific interaction terms for Killarney would be constructed so as to take on the value of zero for observations associated with Quetico or Algonquin. Thus, K_{fee} and $K_{\text{Cong camping}}$ are interaction terms that, if significant, would indicate that the parameter values for the variables for the fee amount and congestion while camping at Killarney are different from the other parks. Note that one park serves as a base, and the interactions are interpreted as differences from the base. Thus, the null hypothesis that preferences regarding congestion while camping do not differ among parks is represented by $A_{\text{Cong camping}} = Q_{\text{Cong camping}} = K_{\text{Cong camping}}$, where A, Q and K represent interaction terms for each park.[2]

[2] Table 3 results for Model 3 uses Algonquin as the base and presents significant park-specific interaction terms for Quetico and Killarney in separate columns, as an alternative to using A, K or Q to precede each variable name.

A dummy variable is used to differentiate between survey versions, with 1 indicating versions where congestion is increased and 0 indicating a decrease in congestion. We would expect the sign on this variable to be negative, given that an increase in congestion represents a loss from the status quo.

Others have demonstrated that users with heterogeneous preferences self-select by day of week (Michael and Reiling). The data do include the beginning and ending dates for each trip. However, for the three parks, quotas do not vary by weekday and quotas are also heavily subscribed. Hence, day of week indicates less about preferences for solitude than does the duration of the trip, since longer trips allow canoeists to access more remote areas. A variable for trip duration is instead included for this reason. We would hypothesize that the sign on this variable would indicate that groups who are more likely to place higher values on reducing congestion take longer trips.

4. Results and discussion

Table 3 provides the econometric results. The standard probit models are not reported because ρ is always highly significant in the random effects models. In all three models the version dummy variable is negative and significant, indicating that increases in congestion are more costly than decreases are valuable. This finding persists across the three models and is consistent with a non-linear marginal willingness-to-pay.

The first model presented in Table 3 treats congestion at the four activities as linear, and the responses are restricted to be the same at each park. Older and wealthier people are more willing to pay to reduce congestion than are younger or poorer people. These demographic effects persist through all specifications. The scale of the parameters is also robust across specifications.

The second model incorporates non-linear congestion effects at each activity during the trip by including quadratic terms for each of the congestion levels. Two of the activities' quadratic terms, camping and first/last day, are significant at conventional levels. First/last day congestion willingness-to-pay increases at a decreasing rate, while camping congestion decreases at a decreasing rate. The other two essentially exhibit linear willingness-to-pay. A log likelihood ratio test for Model 2 (pooled) versus unpooled models for each park results in a χ^2 value of 166.48, indicating that park differences are highly significant.[3] Model 3 allows examination of the robustness of these findings on a park-by-park basis.

Model 3 allows for non-linear responses to congestion and also allows each park to have its own willingness-to-pay function. In this analysis Algonquin is the base park and the parameters associated with Quetico and Killarney measure the change from Algonquin. The first point is that the fee coefficient is statistically different at each park. Ceteris paribus, reducing congestion will be most valuable at Quetico and least valuable at Killarney. However, the impact on willingness-to-pay is also a function of the congestion coefficients. Thus, the actual economic effects of congestion are complex.

[3] The log likelihood ratio test for the pooled (Model 2) versus the unpooled models for Algonquin (LL = 1224.53), Quetico (LL = 2224.51) and Killarney (LL = 854.21), indicates that park differences are highly significant.

Congestion is modeled using quadratic functions for all four of the trip activities. Many of the coefficients are significant individually. The remaining coefficients, while not significant at conventional levels, are significant when they are tested jointly. For example, none of the paddling coefficients are significant individually, but a likelihood ratio test of the hypothesis that they are jointly different from zero has a χ^2 value of 21.9 with six degrees of freedom. Thus, the null hypothesis that they can be left out is rejected.

Following Hanemann (1989) the median and unrestricted mean WTP for congestion level c_0 in a trip segment in Algonquin Park would be:

$$\mathrm{WTP} = \frac{\alpha + \beta_1 c_0 + \beta_2 c_0^2}{\beta_{\mathrm{fee}}},$$

where β_1 is the Algonquin parameter on congestion at the segment of interest, β_2 the parameter on congestion at that segment squared, β_{fee} the Algonquin coefficient on the fee amount offered to respondents, and α represents all of the other Algonquin parameters in the model multiplied by their respective values for the other trip segments and individual characteristics (at the means of the sample). The marginal value of a change in congestion from c_0 to c_1 at that segment would be the difference in the respective WTP amounts. This would be obtained by:

$$\mathrm{WTP}(c_0) - \mathrm{WTP}(c_1) = \frac{\beta_1(c_0 - c_1) + \beta_2(c_0^2 - c_1^2)}{\beta_{\mathrm{fee}}}. \tag{9}$$

To calculate marginal WTP at the same segment in the other two parks, Eq. (9) must include the Algonquin parameters jointly with the shift congestion and bid coefficients for the relevant park.

The marginal WTP curves for congestion on the first and last days, portaging, camping and paddling for each park are provided in Fig. 1 for congestion levels of 1–10 groups. The graphs tell an interesting story of homogeneity and heterogeneity of preferences among the parks. First, it is important to recall that Killarney is a small park, Quetico is an isolated wilderness park, and Algonquin, which has large tracts of wilderness zones, is easily accessible to a large population.

In all three parks, meeting people on the first day in or the last day of the trip is unambiguously a good. The park where it is most valuable is the wilderness park, Quetico. It seems that there is a strong collegial effect of meeting groups during this trip activity at all the parks. Spirits among canoeists are high beginning a trip and finishing a trip. There is a great deal of story-telling and information traded amongst users at these times. The information from exiting paddlers can include details that are not provided on maps, or timely information about the conditions of portages or campsites to avoid. Upon exiting, many canoeists look forward to sharing tales on their last day, with others who are also exiting, or canoeists starting their trips.

Congestion experienced while camping, however, is exactly the opposite. Other canoeists sharing a campsite are unambiguously a bad, regardless of the park. The effect in the wilderness park, Quetico is the most pronounced where congestion at campsites appears to erode the value of the experience. The effects are less pronounced in Killarney and Algonquin but they are clear nevertheless. This is a straightforward result to interpret.

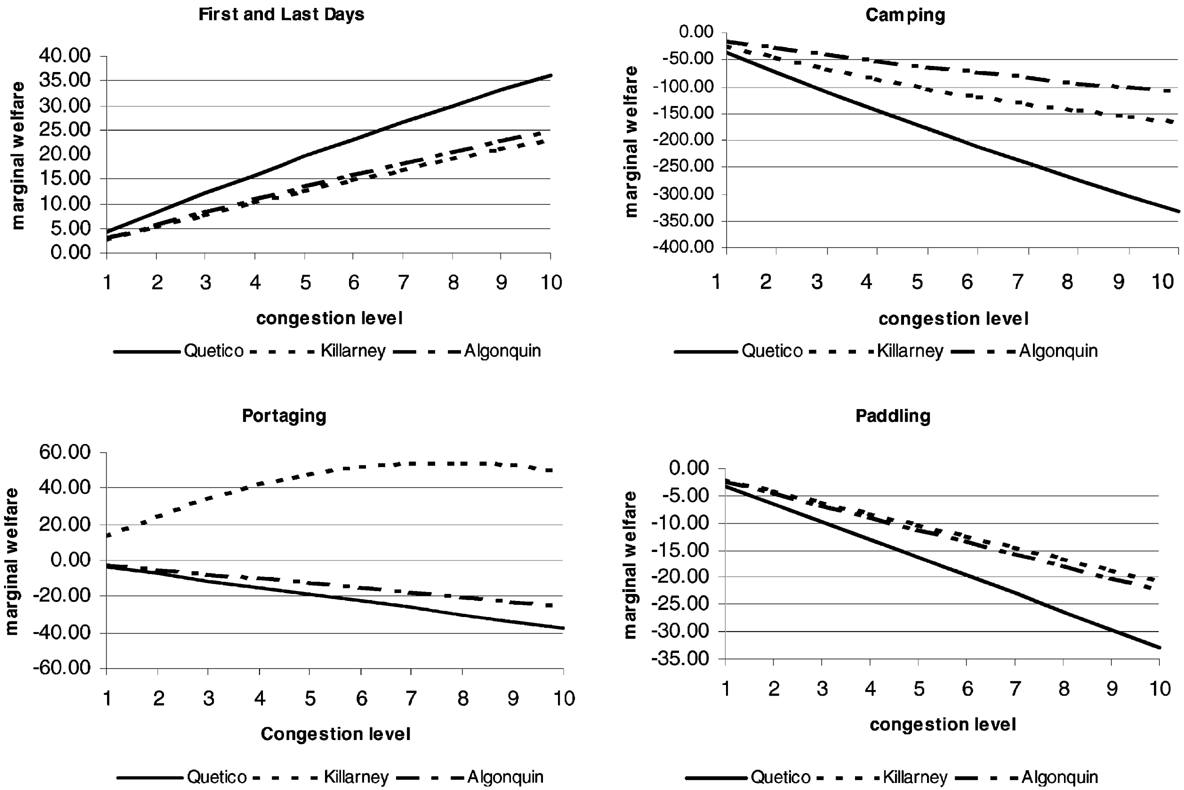

Fig. 1. The marginal welfare effects of increasing congestion during various segments of a wilderness trip.

Campsites are constrained. There are tent sites of varying quality around any site, as well as differences among the quality of sites. The displeasure or pleasure of interacting aside, some places are simply better places to spend the night than others and ones chance of getting a superior campsite diminishes with competition. A poor nights sleep or lack of collegiality around the fire resulting from a poor site diminishes the value of the trip.

People self-select among the three parks based on the types of wilderness experiences that they expect from each, and according to their own preferences and constraints. Thus, Algonquin visitors are more tolerant of congestion at campsites than are visitors to Quetico, who visit Quetico in part because they want and expect a much more isolated environment. All three parks experience excess demand. Based on raw visitor numbers Algonquin might appear to have a more significant congestion problem at campsites, and Quetico the least. However, increasing the quota at Quetico would be more costly in terms of congestion than an increase in the quota at Algonquin.

Congestion at portages differs among the parks. Moderate increases in encounters to about seven groups, provides a good at Killarney. Additional congestion provides disamenity values, which eventually become unambiguously negative. Sharing portages is simply bad at both Quetico and Algonquin. There are at least two reasons for this effect. Killarney has some extremely steep terrain among its portages. The area is hilly with rocky outcrops, and it can be difficult to plan a trip route that avoids steep and difficult portages. Meeting another group at the right spot can be viewed as good fortune for both groups since canoeists often help each other over the worst parts of portages. But, as the number of encounters overall increases beyond a useful level, the negative effect of congestion overtakes the usefulness of assistance.

5. Summary and conclusions

Wilderness canoeists experience a wide range of congestion levels, depending upon specific circumstances such as the wilderness area chosen, the time of season, the days of the week, the route chosen, specific points within a given trip, and other random events beyond the control of the canoeist. Therefore, responses to questions about willingness-to-pay to achieve hypothetical changes in the numbers of encounters with other groups could be expected to depend upon the actual numbers of encounters experienced during a trip, and at what point in a trip those encounters occurred. Few studies examine whether encounter norms differ during the different components of a recreation experience. This study takes advantage of the considerable heterogeneity in the actual congestion levels experienced by individuals in the revealed part of the data, to condition values for hypothetical changes in congestion.

Individuals' responses to the contingent valuation questions were analyzed using a random effects probit panel estimator. The random effects estimator, which allows for an individual level error term, was the most efficient estimator and its use had important implications for the welfare analysis. Both linear and quadratic specifications of congestion variables by trip activities were analyzed. Not only were the non-linear terms important in the model, but the non-linear effects differed among and within the parks. The marginal value of an encounter during the first or last day was either zero or slightly positive, while an encounter during the paddling or portaging activities was significantly negative. However, the most significant negative values were associated with encounters at the campsite. These were estimated to be about four times greater than while paddling or portaging. Important differences were also found among the three parks.

These results suggest that park visitors self-select according to individual preferences and expectations over park attributes in relation to congestion. This information is important for optimal management of public lands subject to entry quota restrictions. The existing park reservation system and route recommendations for the three Ontario parks in the study do not currently consider marginal economic values of crowding. Routing plans and quota systems could incorporate valuations of encounters by trip activities, so that routes, defined as combinations of trip activities, could be developed to optimize recreational use value over a season.

Ontario Provincial Parks managers are able to set differential user fees for each park commensurate with the value of the experience and the costs of provision. At this time, the fees are typically set according to a cost recovery mission. However, the costs of congestion have not before been empirically determined. Estimates of the marginal value of a unit of congestion for trip activities at each park could allow for more finely tuned management and generate greater net recreational benefits.

It is worth noting again that the past efforts to use revealed preferences for estimating congestion values have not been successful, presumably due to the variety of factors that influence behavioral decisions before the trip and during the trip, which in turn affect welfare losses associated with congestion (Jakus and Shaw, 1997; Graefe et al., 1984). The panel model presented in this paper combines revealed preferences (actual trips and encounters experienced) with stated preference data (willingness-to-pay for doubling or halving the numbers of encounters experienced), to take into account a number of these factors. We feel that the approach is promising for further work in this area.

Future research could incorporate into the experimental design and data collection more information about the nature of encounters, such as whether encounters were with other backcountry canoeists, or with recreational users of other types, such as motor boats in some areas of Quetico, or hikers in some areas of Killarney. For example, a variable that was not considered in the data collection for this study, was whether encounters were with groups traveling in the same or in opposite directions along the respondents' canoe routes. This variable would reasonably affect marginal values for encounters, and can also be manipulated by wilderness managers in designing routes.

Furthermore, our panel approach did not capture the actual temporal nature of the activities experienced during a trip. For example, one could capture the actual sequence of activities involved at a particular route in a park, and the panel structure could be designed to examine this temporal effect. In other words, there may be different correlations between responses to changes in congestion at two specific trip activities. Understanding these would provide more information to managers on the effects of congestion, and allow the model to be applied to recreational experiences that have levels of congestion that are correlated between activities during a given trip. However, this would involve more complex survey instruments that would have to be custom-designed for each potential route through a park.

Appendix A. Survey questions

Q-18 a) On average, how crowded did you feel **while paddling**? (*circle one number that best represents your feelings on the scale*).

1	2	3	4	5	6	7	8	9
not at all crowded		slightly crowded			moderately crowded		extremely crowded	

b) On average, how many other groups per day did you encounter **while paddling**? (*fill in blank*)

NUMBER OF GROUPS PER DAY: _____

[The above was repeated for the three other activities, camping, portaging and first/last day.]

Q-24 In question Q-7 you told us what it cost you to take this trip. Suppose that canoeing and camping conditions next year are roughly similar to those for the trip on which you received this questionnaire, with one exception:

- **Your costs next year will be $ 160 higher than what you paid this year.**

Under these conditions, would you still go on this canoe trip? (*circle one number*)

1 No
2 Yes, I would still go on this trip under these conditions

Q-25 Now suppose that canoeing and camping conditions next year are the same as for the trip on which you received this questionnaire with the following two exceptions:

- **First, you experience twice as many encounters while paddling.**
- **Second, your trip expenses will be $ 45 higher than what you paid for the trip on which you received this questionnaire.**

Under these conditions, would you still go on this canoe trip? (*circle one number*)

1 No
2 Yes, I would still go on this trip under these conditions

The above was repeated for the three other activities, camping, portaging and first/last day.
Dollar amounts used as bid values varied according to the design described in Rollins et al, 1997.

References

Butler, J.S., Moffitt, R., 1982. A computationally efficient quadrature procedure for the one factor multinomial probit model. Econometrica 50, 761–764.

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

Cicchetti, C.J., Smith, V.K., 1976. The Costs of Congestion: An Econometric Analysis of Wilderness Recreation. Ballinger, Cambridge, MA.

Donnelly, M.P., Vaske, J.J., Whittaker, D., Shelby, B., 2000. Toward an understanding of norm prevalence: a comparative analysis of 20 years of research. Environmental Management 25, 403–414.

Dorfman, R., 1984. On optimal congestion. Journal of Environmental Economics and Management 11, 91–106.

Fisher, A.C., Krutilla, J., 1972. Determination of optimal capacity of resource-based recreational facilities. Natural Resources Journal 12, 417–444.

Freeman Jr., A.M., Haveman, R.H., 1977. Congestion, quality deterioration and heterogeneous tastes. Journal of Public Economics 8, 225–232.

Graefe, A.R., Vaske, J.J., Kuss, F.R., 1984. Social carrying capacity: an integration and synthesis of 20 years of research. Leisure Sciences 6, 395–431.

Guilkey, D.K., Murphy, J.L., 1993. Estimation and testing in the random effects probit model. Journal of Econometrics 59, 301–317.

Hanemann, W.M., 1989. Welfare evaluations in contingent valuation experiments with discrete choice data: reply. American Journal of Agricultural Economics 71, 1057–1061.

Hanemann, W.M., Kanninen, B., 1999. The statistical analysis of discrete-response CV data. In: Bateman, I.J., Willis, K.G. (Eds.), Valuing Environmental Preferences. Oxford University Press, Oxford, pp. 302–441.

Heckman, J.J., Willis, R.J., 1975. Estimation of a stochastic model of reproduction: an econometric approach. In: Terleckyj, N.E. (Ed.), Household Production and Consumption. National Bureau of Economic Research, New York, pp. 99–138.

Jakus, P., Shaw, W.D., 1997. Congestion at recreation areas: empirical evidence on perceptions, mitigating behavior and management preferences. Journal of Environmental Management 50, 389–401.

Loomis, J.B., 1997. Panel estimators to combine revealed and stated preference dichotomous choice data. Journal of Agricultural and Resource Economics 22, 233–245.

Manning, R., 1995. Studies in Outdoor Recreation: Search and Research for Satisfaction, 2nd ed. Oregon State University Press, Corvallis, 374 pp.

McConnell, K.E., 1988. Heterogeneous preferences for congestion. Journal of Environmental Economics and Management 5, 251–258.

Michael, J.A., Reiling, S.D., 1997. The role of expectations and heterogeneous preferences for congestion in the valuation of recreational benefits. Agricultural and Resource Economics Review 27, 166–173.

Rollins, K., Wistowsky, W., Jay, M., 1997. Wilderness canoeing in Ontario: using cumulative results to update dichotomous choice contingent valuation offer amounts. Canadian Journal of Agricultural Economics 45, 1–16.

Shelby, B., 1981. Crowding models for backcountry recreation. Land Economics 56, 43–55.

Shelby, B., Vaske, J.J., Heberlein, T.A., 1989. Comparative analysis of crowding in multiple locations: results from 15 years of research. Leisure Sciences 11, 269–291.

Vaske, J.J., Shelby, B., Graefe, A.R., Heberlein, T.A., 1986. Backcountry encounter norms: theory, method and empirical evidence. Journal of Leisure Research 18, 137–187.