RURAL ECONOMY

Modeling Congestion as a Form of Interdependence in Random Utility Models

Peter C. Boxall, Grant Hauer and Wiktor L. Adamowicz

Staff Paper 05-01

Staff Paper

Department of Rural Economy
Faculty of Agriculture, Forestry
and Home Economics
University of Alberta
Edmonton, Canada

Modeling Congestion as a Form of Interdependence in Random Utility Models

Peter C. Boxall, Grant Hauer and Wiktor L. Adamowicz

Staff Paper 05-01

The authors are, respectively, Professor, Research Associate and Professor, Department of Rural Economy, University of Alberta, Edmonton, Alberta.

We thank Sean Cash, Assistant Professor, Department of Rural Economy, and the participants of the International Institute of Fisheries Economics and Trade Conference Held at Oregon State University, Corvallis Oregon, July 10-14, 2000 for comments on this paper.

Modeling Congestion as a Form of Interdependence in Random Utility Models

This paper develops a theory of interdependent utility functions in examining congestion in recreation demand equilibria. The notion is examined empirically through the development of congestion forecasting functions which individuals use to sort themselves among a set of recreation sites. These forecasts are used in site choice models estimated on revealed preference information. An interesting outcome of this exercise is that recreation site attribute changes, which are projected to provide positive utility, may in fact generate negative utility if the effect of the change on congestion is jointly considered.

Modeling Congestion as a Form of Interdependence in Random Utility Models

Introduction

Consider the following story. A recreationist wishes to visit a specific park that has recently been improved. However, this individual desires solitude and does not wish to see others during her visit. In considering the trip, she notes that others may be aware of this improvement and thus the solitude she seeks at this park is not likely to exist. As a result she chooses to visit another park where she feels more confident about experiencing the solitude she seeks. In evaluating the alternatives this individual has explicitly considered the actions that other individuals might take. Thus, the utility functions of the individuals making the recreation site choices are interdependent.

Congestion is a critical attribute associated with many forms of recreation. Evaluation of congestion is also important when examining environmental quality changes. An environmental quality improvement at a recreation site will attract more visitors, which in turn increases congestion to levels that may degrade the positive changes. However, what is a congested recreation facility to some may not be to others (e.g. Shelby 1981) and, in fact, may actually be attractive to some others. In other words, the impact of interactions among people probably differs between individuals and this 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 in the future.

One way to understand congestion may be 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 involve heterogeneity in that the critical number for one individual may not be the same for another individual in the group.

Defining a suitable empirical measure of congestion has been difficult. 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 be the most relevant measures of congestion determining recreation site choice decisions, particularly in models that use revealed preference information. Jakus and Shaw (1997) define anticipated congestion as an individual's own estimate of congestion that holds at a site prior to when they actually visit that site. 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 study 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 simple 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.

A Theory of Congestion Equilibria

In this study we examine the choices of congested recreation sites as a game. Recreationists (indexed $n = 1,...,N$) receive utility, U, from visiting sites (indexed $i = 1,...,I$) equal to $U_{ni} = U(X_i, Z_n, C_{-ni})$, where X_i is a vector of the characteristics of site i, Z_n is a vector of individual characteristics, and C_{-ni} is the number of recreationists other than individual n that visit site i. In this framework, individual recreationists react to the actions of other recreationists through the congestion argument C_{-ni}. The choice of recreation sites is denoted by the vector $w_n=(w_{n1},\ldots w_{ni},\ldots w_{nI})$ where the elements represent binary decisions to visit each site in the set of sites.[1] A Nash equilibrium for this situation is defined as a list of recreation site choices and congestion levels $\left(\overline{w}_{ni}, \overline{C}_i\right)_{\forall ni}$ such that for each recreationist n this vector $\overline{w}_n$ solves the following utility maximization problem:

$$\max_{w_{ni}} \sum_{i=1}^{I} w_{ni}\, U\left(X_{ni}, Z_n, \overline{C}_{-ni}\right) \tag{1}$$

subject to

$$\sum_{i=1}^{I} w_{ni} = 1 \qquad \text{for all } n; \tag{2}$$

$$\overline{C}_{-ni} = \sum_{m \neq n} \overline{w}_{mi} \quad \text{for all } n,\ i. \tag{3}$$

Although this is a simple representation of a recreation system with congestion, it is possible to generate a rich set of equilibrium outcomes that provide insights into a variety of congestion situations. Figures 1 and 2 provide summaries of recreation site choice games for the simplest case of two recreationists and two recreation sites. Figure 1 represents cases where preferences for congestion are homogeneous; in other words the two individuals hold identical preferences for congestion. Figure 2 depicts cases where one individual prefers congestion while the other dislikes congestion. Both figures present cases where preferences for site attributes are heterogeneous. The congestion situations are organized as various preferences for congestion in the rows and site preferences in the columns. The rows represent a gradient of congestion preferences from "strongly dislike congestion" in the first row, to "don't care about congestion" in the middle row, to "strongly prefer congestion" in the final row. The critical issue for site preferences is whether the two individuals prefer the same sites or not. The columns represent a gradient of preferences for the two sites from both recreationists "prefer the same site (site 1)", to indifference between the two sites, to both recreationists preferring different sites.

It is useful to start with the case where congestion does not affect preferences and where both individuals are indifferent between the sites. This situation gives rise to multiple equilibria (ME in the third row of Fig. 1). In this case there are 4 possible equilibrium distributions of the recreationists over the 2 sites. However, if indifference among sites is relaxed and if both recreationists prefer site 1, then there is a unique equilibrium (UE) where both visit site 1. If congestion preferences are heterogeneous and recreationist 1 prefers site 1 and recreationist 2 prefers site 2, then there is also a unique equilibrium with each recreationist going to their preferred site.

Based on previous literature (e.g. McConnell and Sutinen 1984) the more common case is that people dislike congestion (i.e $\partial U_i / \partial C_{-ni} < 0, \ \forall i$). Rows one and two in Fig. 1 illustrate this situation. Rows 4 and 5 deal with the opposite case where congestion is attractive for recreationists. This latter case could represent cases where individuals are uncertain of their experience or safety in remote wilderness recreation settings and would prefer to see other recreationists nearby.

Row 2 represents the case where both individuals mildly dislike congestion. If both recreationists prefer the same location (homogeneous site preferences), or prefer different locations (heterogeneous site preferences) then congestion has no effect on the equilibrium site choice outcome relative to the case above. However, if the individuals are indifferent between the locations (i.e. homogeneous site preferences), then there are two equilibria and each recreationist chooses sites to avoid the other. This situation is a pure coordination problem with both individuals indifferent to the actual site chosen, but would prefer that the other individual be at the other site.

If congestion is strongly disliked by both individuals (row 1 in Fig. 1) then this effect creates coordination games with multiple equilibria. For example, if both recreationists prefer site 1 then stronger negative preferences against congestion can overwhelm the preferences for site location such that the presence of someone else at the preferred site creates an incentive to move to the site with less preferred site attributes. In this case, each equilibrium results in one of the recreationists being better off than the other because one individual will have the preferred site in terms of site attributes and the other will not. However, if the recreationists prefer different locations (Fig. 1, col. 3 row 1), then this recreation site choice game becomes an assurance problem with multiple equilibria. One of these equilibria has both individuals at their preferred sites in terms of site attributes. There is no negative effect from congestion because each is on different sites. On the other hand, the other equilibrium is one where both individuals occupy their least preferred site in terms of site attributes. This equilibrium is Pareto inferior to the former equilibrium. Selecting the Pareto superior equilibrium becomes a matter of assurance that there will be no one at the site with the preferred site attributes.

For the opposite case (i.e. individuals prefer congestion, $\partial U_i / \partial C_{-ni} > 0, \; \forall i$) there is a tendency for multiple equilibria to appear especially when preferences for congestion are strong. If the individuals prefer the same location and congestion is strongly preferred then the game becomes an assurance problem. The equilibrium with the preferred site attributes in this case is Pareto superior to the other site. The situation with both individuals on the site with less preferred site attributes is also an equilibrium because of the strong preference for congestion. In the case, however, where they prefer different locations, a strong preference for congestion turns the game into a coordination problem with two equilibria. The attraction for being at the same location overrides site attribute preferences - the individual that is at the site with her preferred site attributes is better off than the other recreationist. Other possible cases under the positive preferences for congestion are illustrated in rows 4 and 5 in figure 1.

To complete the set of possibilities in this simple world figure 2 illustrates cases where preferences for congestion are heterogeneous. If recreationists have different preferences for congestion (i.e. one prefers congestion and the other dislikes it) then recreation site choice equilibria may not exist. If one recreationist strongly prefers congestion and the other does not then a cycling effect occurs where the recreationist that dislikes congestion is always trying to get away from the one who prefers it.

These simple Nash games illustrate the issues that have plagued the study of congestion in many economic situations. Consumers are heterogeneous regarding their preferences for the characteristics of goods and services that they desire and in the congestion case, one consumer's preference is affected by another's. Thus, the central problem for a researcher modelling congestion is that the models should capture the endogeneity inherent in a world of interdependent yet heterogeneous preferences.

Congestion Forecasting

In the model developed above (equation 1), the congestion variable (C_{-ni}) is taken as given by each individual. However, modelling congestion using this approach may not be appropriate for studying recreation demand (Jakus and Shaw 1997). Recreationists can only experience the level of congestion by visiting the site (an *ex post* measure). Thus, it is difficult to incorporate the actual current levels of congestion in an empirical setting because congestion cannot enter the utility function and influence site choice before it is experienced.

This characteristic of congestion, however, could be manifested in the formation of prior expectations or anticipations of congestion levels. Recreation site choices could be made on the basis of prior perceptions or anticipated congestion levels (as well as other choice attributes), not necessarily objective measures of these levels. Empirical support for the hypothesis that recreation site choices could be made on the basis of perceptions of environmental quality has been obtained by Adamowicz et al. (1997). Thus, a more accurate view of the site choice problem in a congested world is to assume that recreationists anticipate congestion prior to selecting trip alternatives.

One way to consider anticipated congestion in the theoretical framework outlined above is to assume that individual recreationists make *forecasts* of congestion levels (C^{a}_{-ni}) at sites before visiting them. In the case of wilderness canoeing, for example, forecasts may be based on past experience, information provided on site attributes by management agencies, information obtained from other canoeists through networking in recreation associations, and perceptions about how other canoeists will react to this information. If individuals had full knowledge of other recreationists' preferences then the forecasts of congestion (anticipated congestion) could be based on the following equation:

$$C^a_{-ni} = \sum_{m \neq n} w^*_{mi}\left(X_m, Z_m, C_{-m}\right) \forall ni \qquad (4)$$

where $w^*_{mi}(\cdot)$ is recreationist *m*'s optimal choice for site *i* (0 or 1) given the attributes of all ($i=1,..,I$) recreation sites for individual *m* (X_m), the individual characteristics of recreationist *m* (Z_m), and congestion levels for all ($i=1,..,I$) sites, C_{-m}. In this forecast model, the right hand side congestion variable (C_{-m}) could be based on previous congestion experiences with the recreation site, or information provided by the management agency and recreation associations about previous congestion levels. Thus, (4) can be considered a forecasting equation that incorporates full knowledge of other individuals' behaviour and feedback responses. This can be written more simply as:

$$C^a_{-ni} = F_{-ni}\left(X, Z, C_{-n}\right) \forall ni \qquad (5)$$

where *X* is a vector of individual site attributes for each individual and each site; *Z* is a vector of characteristics for each recreationist; and C_{-n} is a vector of previous congestion levels for each site.[2]

To sum up, recreation site choice in this framework would involve individuals forecasting congestion levels based on knowledge of how other recreationists select recreation sites. This now assumes that individual recreationists anticipate the choices of other recreationists, but that these anticipations are based on previous levels of congestion.

While (4) and (5) provide a basis for thinking about the congestion forecast this framework probably represents the recreationist as an agent that is more knowledgeable and capable in terms of forecasting the actions of others than is realistic. Hence, an approach based on (5) but more limited in terms of informational requirements may be more valuable in an empirical setting. A simplified version of (5) with a reduced number of arguments is developed in the following equation:

$$C^{a}_{-ni} = F_{-ni}\left(X, Z, C_{-n}\right) = F\left(X^{s}_{i}, Z^{s}_{n}, C_{-n}\right) \quad \forall ni \tag{6}$$

In (6) the latter function *F(.)* includes X^{s}_{i} which is a subset of *X* containing only common site attributes for each site *i,* and Z^{s}_{n} which is a subset of *Z* containing only the characteristics of individual *n* that pertain to the individual's ability to forecast congestion. An example of a characteristic that would fall into this category would be whether the individual is a member of a recreation association as suggested above. Thus, this equation allows for heterogeneity in recreation forecasts, but assumes each agent possesses limited knowledge of other agents' individual attributes, site preferences and forecasting abilities.

In order to examine this theory in an empirical setting, the forecasting function must be linked to a model of recreation site choice behaviour. The random utility model (commonly used in recreation economic research) represents a candidate for this choice process. Random utility theory considers *U* as a random variable where part is known or observable to the investigator and the remainder is not. Thus, $U_{ni}=V_{ni}+ \varepsilon_{ni}$ where $V_{ni}= V(X_i, Z_n, C_{-ni})$ is the former component and ε_{ni} the latter. Congestion in this context can be considered a site attribute, and can be separated from other elements in the vector of site characteristics X_i. The error term ε_{ni} is considered to arise from imperfect knowledge on the part of the researcher.

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, *I*. Thus, the probability, π, of visiting site *i* is:

$$\pi_n(i) = Pr\ \{V_{ni} + \varepsilon_{ni} > V_{nk} + \varepsilon_{nk};\ \forall k \in I\}. \tag{7}$$

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:

$$\pi_n(i) = \frac{e^{\mu U_{ni}}}{\sum_{k \in I} e^{\mu V_{ni}}} \tag{8}$$

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

Typical applications of this model specify a linear indirect utility function, which in the case of congestion may be expressed as:

$$V_{ni} = V(X_i, Z_n, C_{-ni}) = X_i\beta + Z_n\gamma + \delta C_{-ni}\ \forall\ n, i \tag{9}$$

where β, γ, and δ are parameters to be estimated and where C_{-ni} may be replaced by C_{-ni}^{a} in the case of anticipated or forecasted congestion as described above. There are a number of ways to capture heterogeneity in congestion preferences. For example, δC_{-ni}, could be generalized to $\delta(Z_n)C_{-ni}$, where $\delta(Z_n)$ is a function of individual specific variables, such as gender and experience, that may affect congestion preferences. Another approach would be to estimate the choice model with δ as a random parameter that follows some distribution. This latter approach allows for the interpretation of this parameter reflecting heterogeneous congestion preferences in the data that arise from different levels of experience or other characteristics (see Train 1999).

If anticipated congestion (see equation 6) is used in equation (9), then equations (8), (9) and (6) can be thought of as a simple two equation recursive simultaneous equation system. This may be written:

$$U_{ni} = X_i\beta + Z_n\gamma + \delta C^a_{-ni} + \varepsilon_{ni} \tag{10}$$

$$C^a_{-ni} = X^s_i\theta + Z^s_n\alpha + \rho C_{-n} + \nu_{ni} \tag{11}$$

where β, γ, δ, θ, α, and ρ are parameter vectors to be estimated and X^s_i and Z^s_n are site attributes and individual characteristics that are not necessarily the same as in X_i and Z_n. Empirical implementation of this anticipated congestion model requires information on anticipated congestion levels at recreation sites, information on trips to these sites, the development of an anticipated congestion function (which is linear in equation 11), and the use of this function in a two-stage instrumental variables estimation procedure. In this case, equation (11) would be estimated first, then the forecasted values of C^a_{-ni} are used as instruments in the estimation of equation (10).

The next section describes a specific application of these approaches to modelling congestion in a wilderness recreation setting.

Data and Econometric Analysis

We chose to examine congestion in a Canadian wilderness canoeing application. Congestion is important in wilderness canoeing because solitude and remoteness are significant influences on the experience. Furthermore, the expenses involved in accessing the remote areas in which this activity takes place are quite high and most individuals take few trips each year. Thus, selecting congested sites can be costly and as a result recreationists examine existing information sources and rely on past experiences in selecting wilderness sites. Our study focuses on visitors to a system of wilderness parks located in eastern Manitoba and Northwestern Ontario.

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% was from Wabakimi. This distribution was selected because it approximately represented the levels of visitation across the five parks (see Boxall et al. 1999).

The Congestion-Forecast Model

A questionnaire was developed that gathered information about opinions of wilderness management, levels of past visitation to the 5 parks[3] and an additional park, the Boundary Waters Canoe Area (BWCA), descriptions of a typical wilderness trip, and socio-demographic characteristics. The questionnaire was mailed to respondents during November of 1996 and after two follow-ups and adjustment for non-deliverables, an 80% response rate was achieved. Respondents took 1,723 trips to the 5 parks during 1995 and 1996. The most frequently visited parks were Quetico and the Boundary Waters Canoe Area (BWCA).

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 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.

Completed responses to the congestion question were pooled (N=1,297) and these formed dependent variables of the anticipated congestion model represented by equation (11). A number of individual-specific variables were used as explanatory variables in the congestion-forecast model. These variables included: years of experience in wilderness trips in the region, membership in conservation or recreation organizations, the typical trip length, gender, income, education, and household size.[4] 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 would have experienced the increasing visitation levels over time. Similarly, those who were members of wilderness or recreation oriented 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. Thus, this variable was expected to 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 parks 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 canoe routes one is able to select, and may affect the spatial distribution 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 effect of individual respondent and park characteristics on forecasts of congestion levels. The estimation results are shown in table 1.

Being male and preferring long trips is inversely related to increasing congestion perceptions. 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 connection between the highly specialized recreationist (likely male, experienced, takes long trips and is a member of an organization) visiting places where they do not expect to see high numbers of other individuals. Finally, as expected, high levels of perceived human development at these 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 mediated by relationships between visitation levels and park characteristics, however. In this data the size of the wilderness area had a negative effect, while the number of roads accessing a wilderness area had a positive effect on congestion forecasts. These relationships support the hypothesized connections between park size, access, and congestion levels.

Park Choice Models

Park choice models were estimated using the revealed preference information collected in the survey. Park choice was modelled as a function of travel costs, perceived chances of entry, 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, corresponding to the actual and anticipated congestion models described in the theory section. The first used the congestion level by park reported by each respondent from the questionnaire[5]. The model using this variable corresponds with the actual congestion recreation site choice game described in the theory section and is termed the reported congestion (RC) model here, 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 of congestion from the congestion-forecast model described above and this choice model is labelled the instrumental variable (IV) congestion model.

The parameter estimates (table 2) 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 the results are noteworthy. First, higher congestion forecasts exhibit a significant negative effect on park choice. However, the negative effect of congestion is much more pronounced in the IV model than the RC model. This suggests that using predictions from the congestion forecast function as instruments for the anticipated congestion model may be a more powerful way of predicting recreationists' responses to congestion than using reported congestion. The instrumental variable approach is revealing that congestion has a greater effect on park choice than could be understood with the other more typical modelling approach.

Second, while the size of the park and the number of roads accessing a park has positive effects on choice in both models, there are differences in the sizes and the statistical significance of the parameters. In the RC model the park size parameter is highly significant and about six times larger than in the IV model. Roads, however, are not statistically significant in the RC model, but are in the IV model 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 IV models. In the latter, congestion and roads have been estimated with much greater precision than in the former. 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. In the RC model the greater precision may be spurious due to the incorrect specification of congestion. In the IV model the instrument variable has successfully identified the endogenous congestion condition in the choice models. This is further supported by the observation that the value of the log likelihood at convergence and the ρ^2 statistic for the IV model are larger than those for the RC model.

The IV model described above incorporates heterogeneity in terms of forecasting ability through the incorporation of individual specific parameters in the ordered logit model (table 1). A substantive issue of interest in the congestion game, however, was heterogeneity in preferences for congestion. As described above this form of heterogeneity should be incorporated in the park choice model. We chose to re-estimate the IV model while specifying the congestion parameter as a normally distributed random parameter using the procedures described by Train (1999).[6]

The results for this second IV model are shown in the fourth column of table 2. The mean congestion parameter is strongly negative and the SD parameter is large, signifying considerable heterogeneity in congestion preferences. The mean and variance parameters from this model suggest that while the majority of individuals in our sample consider congestion a negative attribute of a wilderness experience, about 6% of the sample would fall into the positive region of the distribution of congestion preferences. These individuals tend to hold weak positive preferences for congestion. However, the majority of the sample will avoid congestion and, given similar preferences for the other attributes, will fall into the categories in the upper left hand portion of the conceptual models in figure 1.

Welfare Implications

A question that remains with the endogenous congestion condition is the effect it would have on economic welfare associated with environmental quality changes. To examine this issue a policy simulation was imposed on one of the five parks in the system. The policy involved increasing road access to Quetico Provincial Park. While this increased road access is hypothetical at present, it is plausible given possible expansion of forest harvesting in the region 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 their boundaries or in the park itself (e.g. Nopiming Provincial Park).

The welfare implications of road access expansion were examined using Hanemann's (1982) formula for estimating compensating variation in conditional logit models. For the IV 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 remained constant and only the roads variable was changed in the choice model.

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 (figure 3 bottom). 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 percentage of the sample forecasts congestion at the highest level (Fig. 3).

The welfare implications of this expansion of access are also shown in figure 3. Note that an additional road at Quetico would generate benefits valued at over $200/trip. More than one additional road, however, would generate substantial 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 threshold 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 appears to generate additional benefits through their impact on site choice utility.

Linking the Theory with the models

In section 2 we presented a theory of congestion that results in many possible equilibrium situations depending on the heterogeneity of preferences over recreation site attributes, the degree of disutility toward congestion and heterogeneity of preferences for congestion. In our sample 71% of the respondents went to Quetico Provincial Park or the Boundary Waters Canoe Area. In addition, some of the parameter estimates together with the data for this park system (e.g. Boxall et al. 1999) suggest that these parks are "signature" parks, highly desired by most of the canoeists in our sample. Since most people that use this park system are from the north Midwestern states and northern Ontario, the BWCA and Quetico are closer than the other parks in the system. Hence, the negative parameter on travel cost would indicate a preference for Quetico and BWCA. Roads are also an important factor in park determining park choice in our model. Quetico has more roads that can be used as entry points than the other Canadian parks in the system. Therefore, the positive and highly significant coefficient on roads would also indicate preference for Quetico. The negative coefficient on the ASC representing Manitoba Parks in the system tends to rule them out as a preferred destination based on site attributes alone. Hence, based on site attributes alone, Quetico can be viewed as the most desired of the Canadian parks for most people in the sample. In terms of the theory presented in figure 1, the relevant case appears to be the first column where people prefer the same location.

The random parameters model suggested that 94% of the sample dislikes congestion. If we relate this to the theory, the relevant cases are in the upper left corner of figure 1. The stylized facts for this sample are that people tend to prefer the same location, namely Quetico, and that they mildly or strongly dislike congestion. If people mildly disliked congestion everyone would go to Quetico. If there are people in the sample who strongly dislike congestion then a coordination problem ensues and in equilibrium more people would go to the other parks in the system.

Historically, the situation appears to fit the upper left area of figure 1, probably that characterized as mildly averse to congestion. As demand for wilderness canoeing increased the number of visitors to the BWCA and Quetico increased. This has led to more and more people to seek alternatives and go to the other parks in the system to avoid congestion in these areas. Empirical data support this as during the period 1988-2000 trips to Woodland Caribou Provincial Park, which is considerably farther north from the traditional market areas than the BWCA and Quetico, experienced a 938% increase in backcountry registrants (Engel Consulting Group 2002).

Figure 3 can be interpreted in a way that also supports this. When the number of roads accessing Quetico increases, the welfare per trip increases because access is a positive attribute. While increasing the number of roads from 3 to 4 increases congestion, it does so only slightly and therefore welfare increases. When the number of roads increases to 5, congestion increases significantly as a result of the access roads. At this point those with strong negative preferences for congestion are pushed toward the other parks.

This finding is consistent with our conceptual model involving Nash games. If the number of agents playing the game outlined in the first column and second row of figure 1 was increased, we would see a similar effect. A single agent would clearly go to their preferred site and be unaffected by congestion. Two agents respond as presented in figure 1 with an equilibrium outcome being both agents selecting the same site (but suffering from the congestion costs). Adding a third agent results in a situation with multiple equilbria – one with all agents at site one and one with 2 agents at site one and one indifferent between site 1 and site 2. A model with 4 agents generates a single equilibrium with 3 agents at site 1 and one agent at site 2 (proofs and further details available upon request). This is the pattern we see in the welfare measures in figure 3 and in the actual visitation data. With the addition of roads the welfare measures initially increase but as congestion levels rise, welfare begins to fall. Many individuals will continue to visit the popular site, but will suffer from congestion costs while a few individual will be pushed to alternate, less congested sites.

Conclusions

The theoretical framework introduced in this paper offers a viable solution to incorporating interdependent behaviour in economic choice models. Part of the appeal of this framework is the notion of formally incorporating endogeneity through forecasts of other individuals' behaviour. The negative coefficients on the congestion variables in all of the models suggests that for the majority of recreationists in the case examined here these other individuals are competitors, in which case the interactions are attenuating. For a small group of people, however, additional visitors may be facilitators in which case the interactions are reinforcing.

According to the theory summarized in figure 1 this type of congestion preference will tend to disperse recreationists. For example, recreationists that prefer the same types of site attributes will tend to choose different sites if congestion overcomes previously preferred sites.

The assurance problem that results for recreationists with different site preferences suggests that it is possible they will collectively choose to disperse in distributions over the landscape if congestion is perceived to be too great in the areas they would prefer to go based on site attributes alone. In this case, anticipated congestion will prevent recreationists from switching to more desirable recreation destinations unless they could be assured that those creating the congestion would move when the change in site choice was made.

Extensions to this research include joint estimation of the congestion and visitation models and estimation using better data on recreationist perceptions of congestion. It is also possible that individuals may forecast the presence of different types of agents at the sites (not simply treating all agents as identical), or they may have site specific tolerances for congestion. These issues, as well as various others relating to the theoretical structure provide interesting avenues for future research.

Endnotes

[1] Note that in this model recreation site choices are discrete. This makes the transition from the theory to a random utility model relatively easy. However, it is also possible to think of the recreationist site choice game as one where recreationists choose mixed strategies. In this case, the w_{ni}'s in equations 1 through 4 can be replaced with the probabilities of site choice, π_{ni}

[2] An equilibrium for this case would be defined as a list of recreation site choices, actual congestion levels, and anticipated congestion levels, $\left(\overline{w}_{ni}, \overline{C}_{-ni}, \overline{C}^{a}_{-ni}\right)_{\forall ni}$ such that equation 1, with C^{a}_{-ni} replacing C_{-ni}, is maximized, equations 2 and 3 are satisfied and in addition $\overline{C}^{a}_{-ni} = \sum_{m \neq n} \overline{w}_{mi}$ (i.e., anticipated congestion equals actual congestion).

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

[4] 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.

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

[6] To examine heterogeneity in preferences for site attributes we also estimated random parameters for the other attribute parameters. The coefficients of the SD for all of them (except congestion) were not statistically significant. Thus, in our sample the respondents held similar preferences for the site attributes but heterogeneous preferences for congestion.

References

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

Boxall, P.C., J. Englin and D.O. Watson, "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, 1999.

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

Cole, D.N., A.E. Watson, and J.W. Roggenbuck, 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, 1995.

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

Engel Consulting Group (in association with K. Rollins, L. Johnson and W. Wistowski), Study of the economic and social benefits of the nine Ontario Living Legacy Signature Sites: Analysis of growth and demand for recreational uses associated with parks. Report prepared for Ontario Parks, Peterborough, Ontario, 2002.

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

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

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

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

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

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

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

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

Table 1. Parameter Estimates for an Ordered Logit Model Explaining Reported Forecasted Congestion at Five Wilderness Parks in Eastern Manitoba and Northwestern Ontario.

Variable	Parameters (*t*-statistics)
Constant	2.2632 (5.028)
Gender (male)	-0.4268 (-1.999)
Years of experience in backcountry areas in the study area (years)	0.0122 (2.328)
Typical trip length (days)	-0.0569 (-3.354)
Member of a conservation or recreation organization	0.2457 (2.094)
Perceived level of development	0.1470 (1.825)
Size of park	-0.0584 (-6.997)
Number of roads accessing park	0.9007 (22.373)
μ_1	3.4918 (26.891)
μ_2	5.8445 (33.935)
Log likelihood	-1216.28
% correct predictions	60.8

Table 2. Parameter Estimates for Choice Models Explaining Wilderness Park Choice Among Five Areas in Eastern Manitoba and Northwestern Ontario.

Variables	Reported congestion model	IV model of forecasted congestion	
Travel cost	-0.00415	-0.00221	-0.00179
	(-9.553)	(-5.153)	(-3.208)
Perceived chances of entry	0.48283	0.18530	0.01370
	(12.049)	(4.048)	(2.144)
Congestion	-0.12246	-3.22100	-4.76326
	(-1.815)	(-13.893)	(-6.387)
Congestion σ			3.00232
			(2.585)
Park size	0.06186	0.01571	0.00950
	(11.222)	(1.994)	(1.069)
Roads	0.02592	0.93324	1.08689
	(0.956)	(13.139)	(11.875)
ASC - Manitoba Parks	-1.03280	-1.08730	-1.10934
	(-9.772)	(-8.668)	(-8.407)
Log Likelihood at convergence	-2312.01	-2204.4	-2189.6
ρ^2	0.164	0.203	

Figure 1. Equilibria[a] for Various Recreationist Preferences for Sites and Congestion when I=2 and N=2.[b]

Preferences for Congestion	Prefer Same Location (site 1)		Indifferent Between Locations		Prefer Different Locations	
Both recreationists strongly dislike congestion	-1,-1	1,0	-1,-1	1,1	-1,-2	1,1
	0,1	-2,-2	1,1	-1,-1	0,0	-2,-1
	CP[c]		PCP		AP	
Both recreationists mildly dislike congestion	.5,.5	1,0	.5,.5	1,1	.5,-.5	1,1
	0,1	-.5,-.5	1,1	.5,.5	0,0	-.5,.5
	UE		PCP		UE	
Don't care about congestion	1,1	1,0	1,1	1,1	1,0	1,1
	0,1	0,0	1,1	1,1	0,0	0,1
	UE		ME		UE	
Both recreationists mildly prefer congestion	1.5,1.5	1,0	1.5,1.5	1,1	1.5,.5	1,1
	0,1	.5,.5	1,1	1.5,1.5	0,0	.5,1.5
	UE		PCP		UE	
Both recreationists strongly prefer congestion	3,3	1,0	3,3	1,1	3,2	1,1
	0,1	2,2	1,1	3,3	0,0	2,3
	AP		PCP		CP	

[a] Shaded areas represent equilibria. Player 1 payoffs are the numbers on the left of each cell and player two payoffs are the numbers on the right. The site 1 choice for player 1 is the top row and the site 2 choice is the bottom row. The site 1 choice for player 2 is the left column and the site 2 choice is the right column. Individual 1 prefers site 1and individual 2 prefers site 2.

[b] The payoffs in the table are generated using the following utility function: $U_{in} = 1 + \delta C_{-in}$ for preferred sites and $U_{in} = 0 + \delta C_{-in}$ for the less preferred sites. Variable C_{-in} is equal to 1 if the other individual also chooses site *i* and 0 otherwise. The parameter δ represents preferences for congestion. Above, δ=0 if the individuals don't care about congestion; δ=-0.5(0.5) if congestion is mildly disliked (preferred); and δ = -3 if congestion is strongly disliked (preferred).

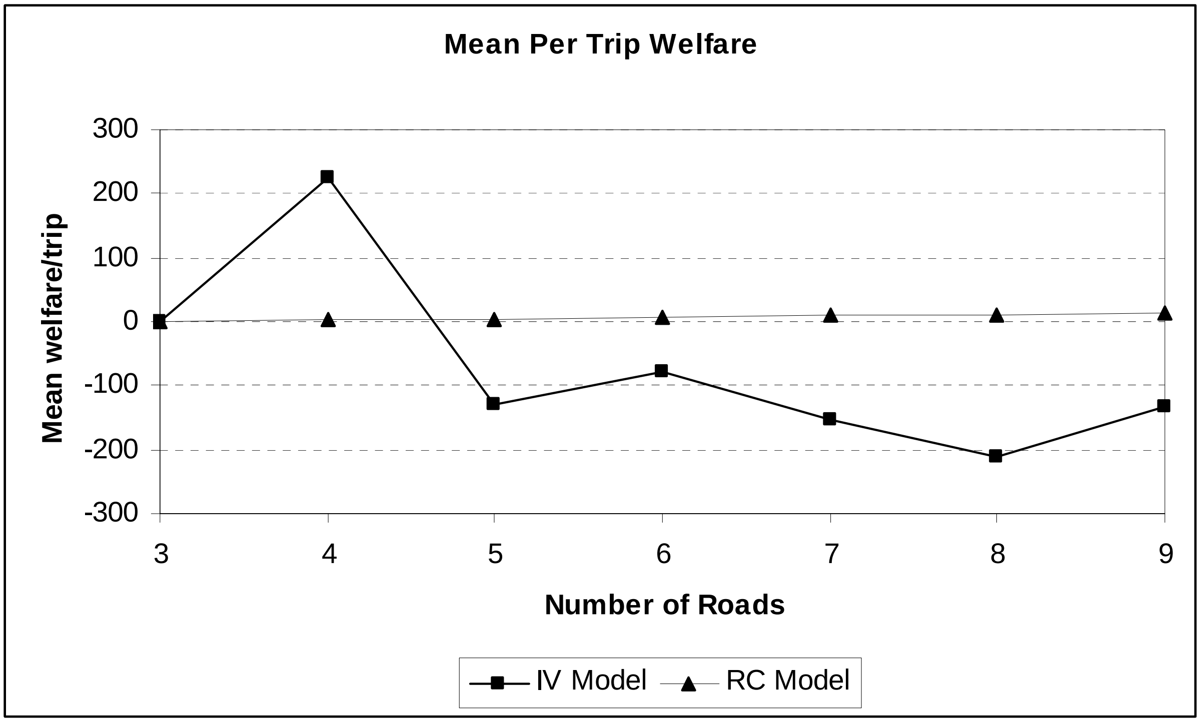
[c] CP refers to a coordination problem, PCP pure coordination problem, AP an assurance problem, UE a game with a unique equilibrium, ME a game with multiple equilibria.

Figure 2. Equilibria[a] with Heterogeneous Preferences for Congestion when I=2 and N=2.

Preferences for Congestion	Prefer Same Location		Indifferent Between Locations		Prefer Different Locations	
Player 1 mildly prefers congestion Player 2 mildly dislikes congestion	1.5,.5	1,0	3,-1	1,1	1.5,-.5	1,1
	0,1	.5,-.5	1,1	3,-1	0,0	.5,.5
	UE[b]		NoE		UE	
Player 1 strongly prefers congestion Player 2 strongly dislikes congestion	3,-1	1,0	3,-1	1,1	3,-2	1,1
	0,1	2,-2	1,1	3,-1	0,0	2,-1
	NoE		NoE		NoE	

[a] Shaded areas represent equilibria. Player 1 payoffs are the numbers on the left of each cell and player two payoffs are the numbers on the right. The site 1 choice for player 1 is the top row and the site 2 choice is the bottom row. The site 1 choice for player 2 is the left column and the site 2 choice is the right column.

[b] UE a game with a unique equilibrium, NoE refers to a game with no equilibrium outcome.



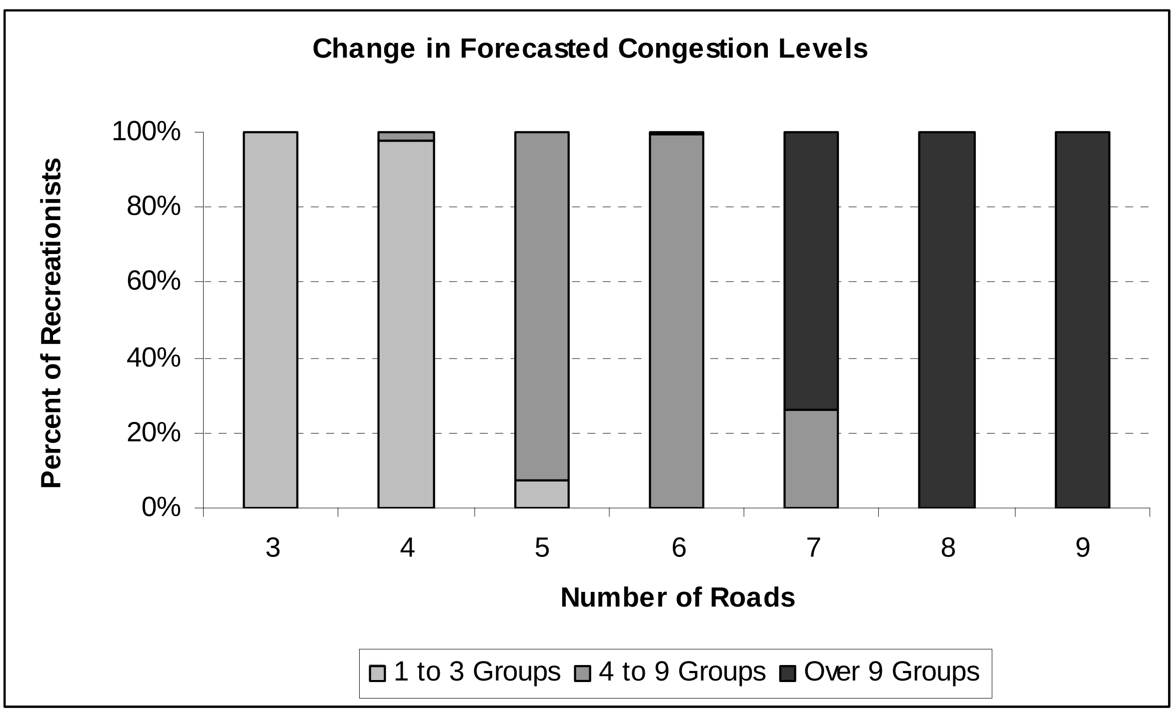


Figure 3. The Effects of Changing Road Access at Quetico Provincial Park on the Per Trip Welfare (top) and on the Percent Distribution of Forecasted Congestion Levels in Terms of Number of Other Recreation Groups (bottom).