Recreation Demand Analysis under Truncation, Overdispersion, and Endogenous Stratification: An Application to Gros Morne National Park

Roberto Martínez-Espiñeira*and Joe Amoako-Tuffour†
Economics, St. Francis Xavier University

November 7, 2005

Abstract

Using on-site survey data from Gros Morne National Park in Newfoundland, this paper estimates and compares several truncated count data models of recreation demand. The model that not only accounts for the truncated and overdispersed nature of the data but also for endogenous stratification due to the oversampling of avid users, while allowing for flexible specification of the overdispersion parameter dominates on the basis of goodness of fit. The results are used to estimate the users' value of access to the park.

Keywords: on-site sampling, endogenous stratification, consumer surplus, count data, overdispersion, recreation demand, travel cost method, truncation.

JEL CODES: Q26 C24

*Corresponding Author: Economics, St. Francis Xavier University, PO Box 5000, Antigonish B2G2W5, Nova Scotia, Canada. Tel: 1-902-867-5433, Fax: 1-902-867-3610. E-mail: rmespi@stfx.ca.

†We would like to thank Joe Hilbe (who helped us refine the STATA code), Jeff Anderson, Danny Major, Gareth Horne, Colleen Kennedy, Ken Kennedy, Dave Lough, John Gibbons, Bob Hicks, Paul Parsons, and the Parks Canada staff at Gros Morne. Courtney Casey, Jarret Hann, Perry Payne, Tracy Shears, and Thomas Khattar did an excellent job administering the survey, and Brian O'Shea provided invaluable research assistance. The survey effort was made possible by funding and/or logistic support provided by Parks Canada, ACOA, the Gros Morne Co-Operating Association, the Viking Trail Tourism Association, and SSHRC funds through a grant of the Centre of Regional Studies at St. Francis Xavier University.

1 Introduction

Planners and administrators of protected areas believe that in order for these areas to be accorded their proper place in regional, national and international economies, their benefits and impacts must be clearly demonstrated. The sound design of policies that relate to national and provincial parks management is based on the knowledge of both the costs and benefits associated with maintaining a park. Since access to this type of recreational areas is only subject to entry fees that clearly underestimate the maximum willingness to pay by most visitors, the true value of parks to the public is unknown and must be estimated using non-market valuation methods.

The travel cost method is the rubric of recreation demand analysis. The number of trips to the site is assumed to be related to travel cost, time, and other demographic or locational variables (Parsons, 2003). A convenient and affordable way to collect data is to conduct an on-site sample of recreationists. However, the nature of the sampling and the nature of the raw data will exhibit a series of features that do not lend themselves to standard demand specification. Many issues must be addressed in moving from the travel cost method to estimation and interpretation of the recreation demand parameters. First, since the dependent variable is a count calculated on the basis of the number of visits to the site,[1] it can only take on values that are nonnegative integers. Second, since all observed visitors have taken at least the current trip, non-visitors are not observed, so the sample is truncated at zero. Third, sampling on site often leads to what is known as choice-based sampling or size-bias, because avid visitors are more likely to end up being sampled than occasional visitors. This means that the data will be endogenously stratified. Finally, the data frequently exhibit overdispersion, which means that the variance is greater than the mean, because a few visitors make many trips and most make only a few.

Truncated count data models are now routinely applied in single-site recreation demand studies after their evolution since the late eighties (Shaw, 1988; Creel and Loomis 1990, 1991; Grogger and Carson, 1991; Hellerstein and Mendelsohn, 1993). A great leap forward was made by Englin and Shonkwiler (1995a) when they developed and empirically applying a truncated, endogenously stratified negative binomial model. However, empirical applications that rigorously correct for all the problems associated with on-site sampling are still relatively few (Ovaskainen et al., 2001; Curtis, 2002; McKean et al., 2003, 2005; Englin et al., 2003).

In this paper truncated count data models are used to estimate the demand curve for trips and consumer surplus per predicted trip using data from an on-site survey of visitors to Gros Morne National Park in Newfoundland. We use the data to attempt a comprehensive treatment of the issues that arise in the estimation of the recreation demand analysis, because they have implications for the measurement of the consumer surplus that users derive from park visits. Standard truncated count data models are compared with models that account not only for the truncated and overdispersed nature of the data but also for endogenous stratification due to the oversampling of avid users resulting form the on-site sampling. We follow the work of Englin and Shonkwiler (1995a) by parameterising the overdispersion parameter of the negative binomial specification so that it varies according to visitor characteristics. This most flexible specification proves to dominate the more restrictive ones often used in previous studies. To our knowledge, there is no published study of the recreational values of National Parks in Atlantic Canada.[2]

The rest of the paper is organized as follows. In the next section, we briefly outline the Travel Cost Method and its application to a single-site sample. This is followed in Section 3 by the methodology of the survey and the data collection procedures. The econometric and estimation issues are dealt with in Section 4, The data description and the choice of variables for the estimated model appear in Section 5. Section 6 includes the discussion of estimation results, followed by the conclusions.

2 The travel cost method

The theoretical starting point in estimating recreational demand is the Travel Cost Method (TCM). The technique is one of several revealed preference methods applied to the valuation of non-marketed goods and services (Braden and Kolstad., 1991; Freeman, 1993, Garrod and Willis, 1999). Examples of the application of the method to value national parks include Beal (1995) and Liston-Heyes and Heyes (1999). The TCM method relies on the assumption that, although access to recreational site has a minimal price or no explicit price, individual's travel costs, including transportation, accommodation, and lost wages, can be used as surrogate prices to approximate the nonexistent prices for their recreational experience. The basic premise is that visitors perceive and respond to changes in travel costs to the site in the same way they would respond to changes in an entry fee, so the number of trips to a recreation site should decrease with increases in distance travelled and other factors increasing the total travel cost. Exploiting this postulated relationship permits the researcher to estimate a true demand relationship. Socioeconomic characteristics of the individuals and information concerning substitute sites and environmental quality indicators can also be included.

Weak separability of recreation demand from non-recreation consumption and weak complementarity (Mäler, 1974) of the marketed goods and services required to get to and to enjoy the site make it possible to estimate a demand curve for individual sites and, from it, a measure of the consumer surplus derived from the site. However, it is clear that the TCM measures only user values of the site. The TCM cannot calculate any type of non-use value (Krutilla, 1967), such as intrinsic value, existence value, option value, or bequest value. The estimates of full-economic value obtained from TCM studies will therefore err on the conservative side and can only be considered as a lower-bound measure of the full benefit of recreational sites.

3 Data collection

Established in 1973, Gros Morne National Park covers 1,805 Km^2 on the Southwestern side of the Great Northern Peninsula in the Canadian province of Newfoundland and Labrador. The park was identified in 1987 as a UNESCO World Heritage Site, due to its rather unique geological features (in particular the Tablelands, the Long Range Mountains and Western Brook Pond), and it is considered one of Canada's most spectacular and unspoiled locations. To borrow Sohngen et al. (2000)'s sentiment about Headlands State Park in Ohio, Gros Morne National Park is simply natural. Gros Morne is a key contributor to the Newfoundland's appeal as an exotic, high quality wilderness area (Locke and Lintner, 1997). Most visitors hike in the park mainly during the peak season of July and August. The hiking experience provided by the varied and attractive scenery is enhanced by the opportunities to encounter wildlife (for example arctic hare, caribou, and moose). Other recreational activities that include angling, swimming, and whale watching contribute to attracting approximately 120,000 visitors to the park annually.

The primary data used in this study come from an on-site survey of visitors conducted between June and September 2004. Using the 'next available vehicle' methodology at the entries of the park and intercepting on-foot visitors at a series of hotspots within the park, a team of interviewers randomly sampled visitors daily (except Sundays). Interviewers were distributed across the park according to a careful sampling plan (developed by Parks Canada) ensuring that visitors from all origins and using different facilities had the same likelihood of being interviewed.[3] Visitors were briefly interviewed (mainly about party size and place of residence) and asked to take with them a questionnaire and mail it back after finishing their visit to the Park. A total of 3140 questionnaires were administered with 1213 returned, giving a response rate of 0.386. Note that the format of the survey prevented the use of reminders, since interviewers only asked about zipcodes and postcodes, rather than actual names and addresses.

The questionnaire included among others questions on the main reasons for the trip, number of times the respondent had visited the park in the previous five years, home location, duration of visit, attractions visited, income, travel cost, size and age composition of travel party, distance to substitute sites, and other sites visited on the same holiday.

Briefly, 18% of the visitors were over 65, 58% were between 35 and 64 years, 14% in the range of 17 to 34 years and 10.25% were under 16 years. By origin, 41% came from Newfoundland and the other Atlantic provinces, 42% from outside the Atlantic provinces of Canada, 13% from the USA, and 4% from other countries. Most visitors (83%) were from within Canada. By the highest level of education, 46% of visitors identified themselves as having college/University degree with an additional 34% having a graduate degree, professional certificate of diploma. The mean income of respondents was $90,000 (in 2004 Canadian dollars) with a standard deviation of $44,500. About 43% of the respondents earned above the mean income and only 15% earned less than one standard deviation below the mean income. Travel cost models assume that trips are for a single purpose only. The majority of visitors (64%) intended this to be a single purpose –vacation or pleasure- trip and about 65% of respondents indicated that Gros Morne National Park either was or played a major influence in their decision to visit the island.[4]

4 Econometric Methods

Count data models have become the standard in single-site recreation demand models (Creel and Loomis, 1990; Englin and Shonkwiler, 1995; Gurmu and Trivedi, 1996; Shrestha et al., 2002).[5] Regression models for counts differ from the classical regression model in that the response variable is discrete with a distribution that places probability mass at nonnegative integer values only. Count data distributions are also characterized by a concentration of values on a few small discrete values (such as 0, 1 and 2), skewness to the left, and intrinsic heteroskedasticity with variance increasing with the mean (Cameron and Trivedi, 1998 and 2001).

4.1 Poisson

Hellerstein and Mendelsohn (1993) provide a theoretical basis for the use of count data to model recreational demand. On any choice occasion, the decision whether to take a trip or not can be modelled with a binomial distribution. As the number of choices increases this asymptotically converges to a Poisson distribution. The density of this distribution for the count (y) is given by:

$$\Pr[Y = y] = \frac{e^{-\mu}\mu^{y}}{y!}, \qquad y = 0, 1, 2, ... \tag{1}$$

where μ is the intensity or rate parameter. The first two moments of this distribution equal each other ($E[Y] = \mu = V[Y]$), a property known as *equidispersion.* This model can be extended to a regression framework by parameterizing the relation between the mean parameter μ and a set of regressors x. An exponential mean parametrization is commonly used:

$$\mu_i = \exp(x'\beta), \qquad i = 1, ..., n \tag{2}$$

where x is the matrix of k regressors and β is a conformable matrix of coefficients to be estimated. Since $V[y_i|x_i] = exp(x_i'\beta)$, the Poisson regression is intrinsically heteroskedastic. Given (1) and (2), the Poisson regression model can be estimated, under the assumption that $(y_i|x_i)$ are independent, by maximum likelihood.

4.2 Negative Binomial

Data on the number of trips are often overdispersed relative to the Poisson distribution. That is, the variance is larger than the mean for the data, because a few respondents make a large number of trips while most respondents make only a few. This makes the Poisson model overly restrictive. Overdispersion has qualitatively similar consequences to heteroskedasticity in the linear regression model. Therefore, as long as the conditional mean is correctly specified, the Poisson maximum likelihood estimator with overdispersion is still consistent, but it underestimates the standard errors and inflates the t-statistics in the usual maximum-likelihood output.

For cases where the overdispersion problem is serious, a widely-used alternative is the negative binomial model. This is commonly obtained by adding an additional parameter that reflects the unobserved heterogeneity that the Poisson fails to capture. Let the distribution of a random count y be Poisson, conditional on the parameter λ, so that $f(y|\lambda) = exp(-\lambda)\lambda y/y!$. Suppose now that the parameter λ is random, rather than being a completely deterministic function of the regressors x. In particular, let $\lambda = \mu\nu$, where μ is a deterministic function of x, say $\mu = exp(x'\beta)$, letting $\nu > 0$ be independently and identically distributed with density $g(\nu|\alpha)$, where α is denoted the overdispersion parameter. This is an example of unobserved heterogeneity, as different observations may have different λ (heterogeneity) but part of this difference is due to a random (unobserved) component ν, which would not be captured by the Poisson regression model.

If $f(y|\lambda)$ is the Poisson density and $g(\nu)$, $\nu > 0$, is assumed to be the gamma density with $E[\nu] = 1$ and $V[\nu] = \alpha$, we obtain the negative binomial density:

$$h(y|\mu, \alpha) = \frac{\Gamma(\alpha^{-1} + y)}{\Gamma(\alpha^{-1})\Gamma(y+1)} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu}\right)^{\alpha^{-1}} \left(\frac{\mu}{\mu + \alpha^{-1}}\right)^{y} \qquad \alpha > 0 \tag{3}$$

where $\Gamma(\cdot)$ is the gamma function. The parameter α determines the degree of dispersion in the predictions. Special cases of the negative binomial include the Poisson ($\alpha = 0$) and the geometric ($\alpha = 1$). A likelihood-ratio test based on the parameter α can be employed to test the hypothesis of no overdispersion.[6]

4.3 Truncation

An additional feature of the distribution of $persontrip$ is that it is truncated at zero, since the data collection was done on-site. Failing to account for truncation leads to estimates that are biased and inconsistent because the conditional mean is misspecified (Shaw, 1988; Creel and Loomis, 1990; Grogger and Carson, 1991; Yen and Adamowicz, 1993; Englin and Shonkwiler, 1995). The density of the Poisson distribution truncated at zero for the count (y) is given by:

$$\Pr[Y = y|Y > 0] = \frac{e^{-\mu}\mu^{y}}{y!} \cdot \left[\frac{1}{1 - e^{-\mu}}\right], \qquad y = 1, 2, ... \tag{4}$$

The standard Poisson model is unbiased even with overdispersion but this is not the case with the truncated version of Poisson. If there is overdispersion, the truncated Poisson model yields inconsistent and biased estimates (Grogger and Carson 1991). In that case, the truncated negative binomial is in order. The density of the negative binomial distribution truncated at zero for the count (y) is given by:

$$\Pr[Y = y|Y > 0] = \frac{\Gamma(y + \alpha^{-1})}{\Gamma(y + 1)\Gamma(\alpha^{-1})}(\alpha\mu)^y(1 + \alpha\mu)^{-(y+\alpha^{-1})} \cdot \left[\frac{1}{1 - (1 + \alpha\mu)^{-\alpha^{-1}}}\right] \tag{5}$$

Examples of applications of this model include Bowker et al. (1996); Liston-Heyes and Heyes (1999); Zawacki et al (2000); and Shrestha et al. (2002). Yen and Adamowicz (1993) compare welfare measures obtained from truncated and untruncated regressions.

4.4 Endogenous stratification

Finally, since the data have been obtained on-site, the sample is endogenously stratified. This is because a visitors' likelihood of being sampled is positively related to the number of trips they made to the site. That is, frequent visitors are more likely to be sampled. This problem (sometimes referred to as choice-based sampling) was first addressed by Shaw (1988), while Englin and Shonkwiler (1995) extended their analysis with an application of the truncated and endogenously stratified negative binomial model.[7]

If the assumption of equidispersion holds, standard regression packages can be used to estimated a Poisson model adjusted for both truncation and endogenous stratification. In this case (Shaw 1988) show that:

$$\Pr[Y = y|Y > 0] = \frac{e^{-\mu}\mu^{y-1}}{(y-1)!}, \qquad y = 1, 2, ... \tag{6}$$

so it suffices with regressing $persontrip^* = persontrip - 1$ with a conventional Poisson regression model (Haab and McConnell, 2002, p. 174-181).

This model has been used in several applied studies (Bin et al, 2005; Hagerty and Moeltner, 2005) under the, sometimes untested (Fix and Loomis, 1997; Hesseln et al., 2003; Loomis, 2003), assumption that overdispersion is not significant.

For the case where overdispersion is significant, the density of the negative binomial distribution truncated at zero and adjusted for endogenous stratification for the count (y) was derived by Englin and Shonkwiler (1995) as:

$$\Pr[Y = y|Y > 0] = \frac{\Gamma(y_i + \alpha_i^{-1})}{\Gamma(y_i + 1)\Gamma(\alpha_i^{-1})} \alpha_i^{y_i} \mu_i^{y_i - 1} (1 + \alpha_i \mu_i)^{-(y_i + \alpha_i^{-1})} \tag{7}$$

Unfortunately, this expression cannot be manipulated into an easily estimable form, so it needs to be programmed as a maximum likelihood routine, with the associated increase in computational burden. Englin and Shonkwiler (1995) provide an empirical application of this specification. Englin et al. (2003) and Ovaskainen et al. (2001) also used this model and found that correcting for choice-based sampling on top of zero-truncation does not make much difference in estimates.[8]

However, these applications are based on a variation of (7) that restricts α to a common value for all observations (so $\alpha_i = \alpha$). To the authors' knowledge, only (Englin and Shonkwiler 1995a) parameterized α (as $\alpha_i = \frac{\alpha_0}{\lambda_i}$).[9] In this paper we use the more flexible approach that allows the overdispersion parameter to vary[10] according to the characteristics of the visitor and compare it with the more restrictive approach.[11]

5 Model specification and variable definitions

Within the framework of the individual Travel Cost Method, the single-site demand function is

$$Y_i = f(TC_i, S_i, D_i, I_i, V_i) \tag{8}$$

where TC_i is travel cost, inclusive of the cost of travel time and S_i, information on substitutes sites. D_i represents demographic characteristics of the respondent and the visitor party. I_i is a measure of income. V_i are features of the current visit to the park and i indexes the individual respondents.

The dependent variable in our study is defined as $persontrip$, calculated as the product of current *partysize* times the number of times the respondent visited Gros Morne during the past five years (including the current trip). This type of variable was proposed by Bowker at al. (1996) to circumvent the problem of lack-of-dispersion endemic to individual Travel Cost Method models (Ward and Loomis, 1986). Bhat (2003) also used this format for the Florida Keys because, as in the case of Gros Morne, group travel by car is very common in the Florida Keys (Leeworthy and Bowker, 1997).

The independent variables in Expression 8 were constructed on the basis of answers to the questionnaire.[12] These include:

Home location (postcode for Canadian residents, zipcode for US residents and country for residents of other countries). This allowed us to calculate flying distances to the airport of entry or driving distances to Gros Morne (measured as driving distances to Rocky Harbour). We trimmed off from the sample 12 respondents living further than 7500 Km away from the Gros Morne, because long haul travellers are often not well described by the recreational demand model applicable to visitors from closer areas. Bowker at al. (1996) and Bin et al. (2005), for example, also trim the sample at 1000 miles or about 1700 km and Beal (1995) discard all overseas visitors to Carnarvon National Park in Queensland. In particular, long haul travellers are much more likely to visit the park as part of a multipurpose trip.

Travelcost (in CAN$ 1000 per year). Following the general approach commonly taken in the literature (Hesseln et al., 2003; Englin et al., 2003), the travel cost is calculated as the number of round-trip kilometers from the visitor's residence to the park times 0.35 $CAN/Km if the visitor entered Newfoundland by ferry. If the visitor entered through any of the airports in Newfoundland we assume that she took a flight from her residence and we value the cost of flying at $CAN/Km 0.20 if the one-way distance is less than 4000 Km and 0.10 $CAN/Km if the one way distance exceeds 4000 (a similar calculation was done by Bhat, 2003). Unfortunately, we only had information about the point of entry in Newfoundland, not the modes of transportation used during the whole trip. It is likely that some visitors flew from their destination to Halifax in Nova Scotia or one of the main hubs in central Canada (Montreal, Toronto, or Ottawa) before renting a car to drive through the Maritime Provinces. We treated those visitors as having driven all the way from home to the park. However, since the daily cost of renting a car is much higher than the opportunity cost of using a private car, the finally calculated driving cost per kilometre should average to more or less the same as the cost per kilometre worked out assuming that they just drove all the way from home in their own car. Those visitors who entered through St. John's airport had to drive 700 Km to Gros Morne. This was accounted for when calculating *travelcost*. This variable is then divided by *partysize* before adding it to the estimated cost of travel time (see definition of *traveltimecost* below) to compute the full cost of travelling to the park (Cesario, 1976). Due to the high collinearity between the two measurers, it was not possible to enter them separately in the model. Multicollinearity may result in wrong signs and/or implausible results (Smith et al., 1983; Earnhart, 2004). Englin and Shonkwiler (1995) faced the same problem of multicollinearity and could not independently estimate the effect of the cost of travel time. Other studies (e. g. Fix and Loomis, 1998) use reported travel costs. To avoid survey overburden and avoid response and recall bias, we did not ask respondents to calculate their travel costs themselves. Bowker et al. (1996) report models using both approaches to variable travel cost calculation, finding no appreciable differences.

Traveltimecost The valuation of travel time is a thorny issue in travel cost method studies (Englin and Shonkwiler, 1995b; Feather and Shaw, 1999; Zawacki et al., 2000; Hesseln et al., 2003; McKean et al., 2003). We used the product of round trip time times a fixed fraction of the wage rate to proxy the opportunity cost of time. Cesario (1976) used 0.43 as the relevant fraction, Zawacki et al. (2000) and Bowker et al. (1996) use 0, 0.25, and 0.5 as wage multipliers. Liston-Heyes and Heyes (1999) and Hagerty and Moeltner (2005) use 1/3 of the wage. Sohngen et al. (2000) and Sarker and Surry (1998) use 0.3. In the present study the fraction used was 0.3 and the wage rate was roughly approximated as the ratio of the annual income divided by 1080 hours of work per annum (Sohngen et al., 2000; Bin et al., 2005). Travel time was calculated from the estimated travel distance to the Park by assuming a driving average speed of 80 Km/hour and a flying[13] average speed of 600 Km/hour.[14]

CTC: the proxy of the travel cost that acted as a price in Expression 8, was the sum of $travelcost$ and $travetimecost$. It was measured in CAN$ 1000.

Time spent on the site ($daysatGM$): The expected effect of this variable was uncertain *a priori*. although Shrestha et al. (2002) and Creel and Loomis (1990) find that the longer the duration of the trip the less the trips taken and Bell and Leeworthy (1990) also find that people living far away make fewer trips but stay longer at the site.

Distance to substitute sites ($DSUB$): If a person lives near a substitute recreational site, the number of trips to the site analyzed will likely decrease.[15] We followed Bowker et al. (1996) and used a dummy (*substitute*) that takes the value of one if the respondent suggested an alternative site or the distance to it. Liston-Heyes and Heyes (1999) chose not to include the distance to substitute sites and describe the difficulties involved in introducing this variable in the demand model. McKean et al. (2003) also found the effect of this variable non-significant and many respondents failing to provide a value for it.

Other sites visited during the current trip: We asked respondents if they had visited other national parks in the Atlantic region, as in Liston-Heyes and Heyes (1999). The final model included dummies for Terra Nova National Park ($TerraNova$) and Cape Breton Highlands ($Highlands$) National Parks.

Reasons for visiting Newfoundland and Labrador and the relative influence of Gros Morne in the decision to visit this province: This helped us screen out those visitors from outside the province whose decision to visit Newfoundland and Labrador had little to do with their visit to Gros Morne. Similar variables were also used by Beal (1995); Sohngen (1998); and Liston-Heyes and Heyes (1999).

Number of people in the visitor group sharing expenses in the current trip ($partysize$) as in Liston-Heyes and Heyes (1999) and Hesseln et al. (2003) and age composition of the visitor group in the current trip (Siderelis and Moore, 1995). The proportion of party members under sixteen ($propou16$) was used in the final models.

$Income$ (in $CAN1000). Often the influence of income is found to be weak in travel cost studies. Many found it negative or non-significant (Creel and Loomis, 1990; Sohngen et al., 2000; Loomis, 2003). Liston-Heyes and Heyes (1999) find visits an inferior good. Bin et al. (2005) find a significant positive effect of income on the number of trips to North Carolina Beaches. Given the remoteness of Gros Morne, we expected income to exert a positive effect on the number of visits, even though residents of Newfoundland, whose average income is relatively low, would have of course visited very often.

Level of educational attainment (*educat*). The sign of this effect was expected to be positive *a priori*, although Shrestha et al. (2002) find a negative effect on fishing trips

Different aspects of their experience during the current trip were considered, including an estimate of combined total spending in the Gros Morne area per member of the visiting party ($expenses$, in thousands of $CAN). The respondents were also asked about having visited certain parts of the park and used certain facilities there (several dummies were used in the final models: *camped*, about use of campgrounds; *dine* about use of fine dining outlets, and *pool*, about using the park swimming pool); the type of accommodation used; the level of satisfaction with the services and facilities ($satisfied$, which takes the value of one if the visit met or exceeded expectations); and source of information about the Park.

Visitors were asked about the time of decision to visit the park and the degree of influence of different activities (hiking, backpacking) within and different features (the fact that it is a World Heritage site, etc.) of the park in the decision to make the visit. The variable *camping* (about the influence of camping) was kept in the final model. Visitors ranked (from 1 to 5) the importance they attached to different features and facilities available in Gros Morne. Included in the final models were the importance given to the availability of accommodation at 3.5 rating or below (*valueaccomm*); campgrounds with 30 amp outlets or more ($valuecamp30$); importance given to the fine dining outlets in the park (*valuedine*); and importance given to the park's swimming pool (*valuepool*).

[INSERT TABLE 1 ABOUT HERE]

6 Results

A great proportion of questionnaires were discarded due to item non-response, out of the 1213 completed. Additionally, only those visitors whose residence was no more than 7500 Km away from Gros Morne were used for the analysis.[16] And only those visitors who planned the visit to Gros Morne 'before leaving home' were included in the analysis, because visitors who planned the visit to the park after leaving home would clearly be multisite travellers. We also screened off those visitors from outside Newfoundland for whom Gros Morne did not strongly influence their coming to Newfoundland.[17] Finally, the sample was also cleaned off of 3 outliers that would drive the deviance residuals in a plain Poisson model beyond the value of 6. Further inspection revealed that those observations corresponded to visitor who had visited Gros Morne more than 12 times in the past five years. The final sample contained 658 observations. Summary descriptives of the variables used by the demand models are shown in Table 1.

6.1 Model selection

On the basis of the signs and magnitudes of the coefficients, the different models appear highly robust. There are no sign changes across specifications and only the statistical significance and the goodness of fit differ. The robustness of results confirms that the need for corrections due to the sampling characteristics of the data largely improves the efficiency and consistency of the estimates.

Table 2 shows that the econometric specification that best fits the data is the one that accounts not only for the truncated nature of the data, but also for the endogenous stratification resulting from the fact that an on-site sampling was used. Model $TSNBIN3$ also accounts for the overdispersed nature of the data, and it allows the overdispersion parameter α to vary across visitors according to characteristics of the visitor group.

The high significance of the overdispersion parameter α in models $TRNBIN0$ (a truncated negative binomial) and $TSNBIN$ (a truncated negative binomial that corrects for endogenous stratification) confirms that overdispersion is a problem. Therefore the models based on the Poisson distribution, $TRPOIS0$ (which accounts for truncation) and $TSPOI$ (which also corrects for endogenous stratification) are overly restrictive. In fact, a likelihood-ratio test of $\alpha = 0$ based on $TRNBIN0$ results in a $\overline{\chi}^2(01) = 134.52$ with $Prob >= \overline{\chi}^2 = 0.000$.[18]

Therefore the choice of the best model rests among the models based on the negative binomial specification, which correct the overestimation of t-ratios and the underestimation of consumer surplus in the Poisson. Among these ($TRNBIN0$, $TSNBIN$, and $TSNBIN3$), it can be seen that under the models that account for endogenous stratification the log-likelihood increases and the size of the coefficient of the price proxy variable (CTC) also rises. The latter means that the consumer surplus measures become smaller. The correction for endogenous stratification results in an increase in the absolute size of the price coefficient and, as a consequence, in an appropriate reduction of the estimates of consumer surplus.

The chosen model $TSNBIN3$ dominates $TSNBIN$, by allowing the overdispersion parameter α to vary according to *income* and a proxy for the age composition of the visitor group[19] ($propu16$). A Likelihood-ratio test yields the value of $\chi^2(2) = 29.06$ with $Prob > \chi^2 = 0.0000$. The coefficients of both covariates in the α equation are highly significant, again revealing that using the same overdispersion parameter for all observations would be overrestrictive.

At the bottom of Table 2 pseudo-R^2 values are reported. These are calculated as $R^2 = 1 - \frac{\ln L}{\ln L_0}$ where $\ln L$ is the log-likelihood of the full model and $\ln L_0$ is the log-likelihood of a model with the restriction $\beta = 0$. This measure of fit is not comparable to ordinary least squares R^2, but still provides an indication of the improvement of the fit of the model over a restricted model with only a constant term. Following Ovaskainen et al. (2001), for a common point of reference and for the overall rather than incremental fit for the negative binomial, the restricted log-likelihood of the relevant Poisson (with restrictions $\beta = 0$, $\alpha = 0$) was used as the restricted log-likelihood for both the Poisson and negative binomial models when computing the pseudo-R^2. That is, the full $TRNBIN0$ is compared with a restricted $TRPOIS0$ and the full $TSPOIS$, $TSNBIN$, and $TSNBIN3$ are each compared to a restricted $TSPOIS$. This measure confirms that the most flexible model works best. The discussion below therefore corresponds to Model $TSNBIN3$.

[INSERT TABLE 2 ABOUT HERE]

CTC presents the expected negative sign, which yields a negatively sloped demand curve for *persontrips*. This means that the further away a visitor lives the fewer the visits to the park in the past five years and/or the smaller the visitor party in the current trip. The size of the coefficient is remarkably similar across all the negative binomial models, although it rises in absolute value, as expected, when the specification adjusts for oversampling of avid users. The effect of the variable *income* appears significant at the 10% level and has a positive sign. Often income is found to be non-significant in travel cost studies. It is likely that the remote location of Gros Morne makes the visit expensive enough for many visitors for visits to be a normal good. Bin et al (2005) find a significant positive effect of income on the number of trips to North Carolina Beaches. The variable *educat* presents positive although non-significant sign, which suggests that that income and education are perhaps too collinear to allow for independent estimation of the effect of education.The variable *expenses* presents the expected negative sign: those who tend to spend more on a visit to the park, tend to make fewer trips.

The length of the visit (*daysatGM*) has a significant and positive sign. Bowker et al. (1996) also find a positive sign for time spent at the site. However, this result is at odds with previous findings. Shrestha et al. (2002) and Creel and Loomis (1990) find that the longer the duration of the trip the fewer the trips taken and Bell and Leeworthy (1990) also find that people living far away make fewer trips but longer stays. The fact that the length of stay appears positively correlated with the frequency of visits may be associated with the remote geographical location of Gros Morne and the type of recreational activities that it offers.

The ordinal variable *satisfied* presents a negative sign suggesting that those who were not satisfied with their current trip take less frequent trips. The binary variable *substitute* has a non-significant positive sign. In theory, we would have expected that those visitors who came up with a next best alternative to Gros Morne would visit this park less frequently (e.g. Parsons, 2003). However, it is also possible that avid recreatonists have a more readily available mental list of recreational destinations than those who travel less frequently. In fact, one should also consider the possibility that those visitors most interested in outdoor recreation might have actually chosen a residence location close to a recreational site.[20] It is also likely that respondents failed to successfully come up with a valid substitute for Gros Morne (and that explains why there was a great number of item nonresponses for this variable),[21] since this park offers a rather unique combination of features. The fact that nearly 92% of the respondents made it a point to visit Gros Morne before leaving home suggests for many the single-minded purpose of the trip and the irrelevance of alternative sites closes to home in the decision making. Betz et al (2003) also find the effect of this variable nonsignificant.

Visitors were asked about whether they had visited a series of alternative recreational sites in Atlantic Canada. The variable *TerraNova*, enters the final model with a negative sign. It makes sense that those coming to Gros Morne for the first time during the current trip from outside Newfoundland were more likely to take advantage of the trip to also visit Terra Nova National Park. More experienced and knowledgeable visitors were perhaps less likely to visit Terra Nova, since Gros Morne remains the clearly preferred choice among most people who have experienced both sites. *Highlands* has also a negative sign for mainly the same reason. Additionally, informal conversations with visitors during the sampling process revealed that it was very common for visitors from the US and Canadians from west of New Brunswick to drive through Cape Breton if they were visiting the Maritime Provinces for the first time. More frequent visitors of Gros Morne would then be more likely to drive or fly directly to Newfoundland.

The variable *museums* and *pool* have a significant positive sign, suggesting that those who visited the museums and/or the swimming pool in the park had already visited Gros Morne more often in the past. This makes intuitive sense in a destination based primarily on outdoor recreation. On the other hand, *dine* has a negative sign, suggesting that those people who had visited Gros Morne more often before did not use fine dining restaurant so much as those who had visited less often. It makes sense that avid outdoor recreatonists focus on the features of the area other than its restaurants and they also stay longer in the park (as suggested by the sign of *daysatGM*), which probably results in their using self-catering facilities or cheaper food outlets.

The binary variable *camped* presents a somewhat surprising negative sign. It was expected that those who had used the campgrounds would be more frequent visitors to the park. However, since accommodation is quite affordable in Newfoundland relative to the rest of North America and Gros Morne is quite a remote location for most visitors, the effect of *camped* might be confounded with the influence of *income* in the sense that those who camp visit less often, because the trip to the park is the main component of the costs of the visit. Those who can afford several trips to Gros Morne can afford the accommodation there too. Moreover, it is likely that at Gros Morne camping sites are not so much a means of affordable accommodation as an integral part of the overall wilderness experience. High income visitors who seldom travel a long distances to Gros Morne may prefer to camp during their stay, as part of a complete 'wilderness experience'. However, the stronger the influence of *camping* activities in decision to visit the park, the more the expected previous trips to Gros Morne.

The negative sign on the *valueaccomm* variable suggests that those who appreciated more the availability of budget accommodation visited less frequently. As explained above this probably shows that differences in income are at play. Those for whom finding affordable accommodation was a concern probably could not afford to travel to the site that often. Those who valued the pool more (*valuepool*) visited more often and those who valued campgrounds with outlets over 30 amps (*valuecamp*30) visited more frequently. The variable *valuedine* had a non-significant positive sign.

The mean variance inflation factors ($VIFs$) for the independent variables in the model was 1.46 when using CTC (combined total travel cost). Therefore, there were no serious problems of multicollinearity among the variables.

6.2 Welfare calculations

We use the results in Table 2 to calculate welfare measures in terms of the consumer surplus users derive from having access to the park. In addition, since the final model accounts for the truncated and endogenous stratification of the data, welfare measures for a given population could be calculated, provided population values for the parameters in the demand equations were available (Englin and Shonkwiler, 1995). However, truncated individual models of recreation demand can be used to extrapolate welfare measures to nonvisitors only under the assumption that these have the same demand functions as visitors (Hellerstein, 1991). Since it is not clear that this is the case, and it is unclear how the relevant population should be defined in this case, let alone how values for most of the demand parameters could be obtained, the calculations in this section refer only to users.[22]

We only use the estimated coefficients on the combination of *travelcost* and *traveltimecost* (that is CTC) to calculate welfare measures, because *expenses* are mainly endogenous, a choice of the user. It is true that expenses include some component of user fees, but these are usually relatively small compared with the full cost of the visit. In any event, the welfare measures considered can be seen as a conservative lower bound for the full benefit derived by users.

In all the count data models reported in Table 2 the consumer surplus per visit can be calculated as $-1/\beta_{CTC}$ (Creel and Loomis, 1990). If this expression is multiplied by predicted $persontrip$, we obtain the predicted CS per five year period for the typical visitor group in the sample results (Englin et al., 2003). This is the correct measure for policy analysis if it is assumed that the dominant source of error in the analysis is measurement error (Bockstael and Strand 1987, Haab and McConnell 2002, p. 162). Predicted mean *persontrip* can be calculated by aggregating over all visitors and calculating the average count.[23] If instead the error were expected to be mainly specification error, -1/β_{TC} should be multiplied by the sample average.

As shown in Table 3 the value of consumer surplus per *persontrip* under $TSPOI$ would be \$3, 301, accounting for the fact that all the cost variables were measured in thousands or dollars. The extrapolation to the consumer surplus per visiting group for the five years would accordingly be $(3.081 + 1) \cdot \$3, 301 = \$13, 472$.[24]

Consumer surplus per *persontrip* is $\$4,021$ under $TSNBIN3$. The consumer surplus per visiting group for the five years based on the predicted number of *persontrip*, rather than the observed number, can be calculated following the method in (Englin and Shonkwiler 1995a), whereby $E(persontrip_i/x_i) = \lambda_i + 1 + \alpha_i\lambda_i = 3.26$. Marshallian surplus is $\$13,100$ per *partysize* during the five previous years. Dividing by the sample mean of *partysize* times five (years) we obtain a consumer surplus of $13,100/(2.56535 \cdot 5) = \$1,022$ per year per typical party member.

Note that, as expected, by increasing the coefficient for CTC, the correction for endogenous stratification yields a smaller consumer surplus estimate per *persontrip* (under $TSNBIN$ or $TSNBIN3$) than under $TRNBIN0$. In particular $TRNBIN0$ would yield $\$4,176$ while $TSNBIN3$ yields $\$4,020$.

7 Conclusions, limitations, and suggestions for further research

We have used on-site survey data from Gros Morne National Park in Newfoundland to estimate and compare a set of truncated count data models of recreation demand. The paper's main focus was on the performance of different estimators. Our results confirm, in line with earlier works based on recreational sites in the US and Europe that a model that, at the cost of some extra computational burden, corrects simultaneously for overdispersion, truncation, and endogenous stratification dominates more restrictive models in terms of goodness of fit. Moreover, the results show that the endogenously stratified negative binomial model that allows for flexible specification of the overdispersion parameter (as a function of characteristics of the visitor groups) dominates the model with restricted overdispersion parameter.

The theoretical implications for the estimation of consumer surplus are as expected. Not correcting for overdispersion (by relying on Poisson estimates) substantially understates true consumer surplus, while accounting for endogenous stratification (both under Poisson and negative binomial specifications) appropriately adjusts the consumer surplus downwards. Furthermore, we find that a restricted truncated and endogenously stratified negative binomial model slightly understates consumer surplus relative to the most flexible model. For Gros Morne National Park, the preferred model yields a value of consumer surplus within a range that confirms that the recreational amenities of the park are highly valued by visitors. We also find that visits to the park are a normal good.

As in every application of the travel cost method substantial research judgement has been necessary in some aspects of the analysis and some simplifications have been made. Although the focus of this contribution lies on the comparisons of the different econometric specifications, there are several qualifications to our results and further research is necessary to fully examine the robustness of the welfare values derived from Gros Morne National Park as the basis for park management decisions. For example future work is warranted that look at the sensitivity of the results to alternative ways of approaching the measurement of travel cost and the estimation of the opportunity cost of travel time, to the stratification of observations according to the regional origin of visitors and the length of their stay. These issues deserve further attention but are outside the scope of this paper.

References

Beal, D. (1995). A travel cost analysis of the value of Carnarvon Gorge National Park for recreational use. *Review of Marketing and Agricultural Economics 63*(2), 292–303.

Bell, F. W. and V. R. Leeworthy (1990). Recreational demand by tourists for saltwater beach days. *Journal of Environmental Economics and Management 18*(3), 189–205.

Betz, C. J., J. C. Bergstrom, and J. M. Bowker (2003). A contingent trip model for estimating rail-trail demand. *Journal of Environmental Planning and Management 46*(1), 79–96.

Bhat, M. G. (2003). Application of non-market valuation to the Florida Keys Marine Reserve Management. *Journal of Environmental Management 67*(4), 315–325.

Bin, O., C. E. Landry, C. Ellis, and H. Vogelsong (2005). Some consumer surplus estimates for North Carolina beaches. *Marine Resource Economics 20*(2), 145–161.

Bockstael, N. E. and I. E. Strand (1987). The effect of common sources of regression error on benefit estimates. *Land Economics 63*, 11–20.

Bowker, J. M., D. B. K. English, and J. A. Donovan (1996). Toward a value for guided rafting on southern rivers. *Journal of Agricultural and Applied Economics 28*(2), 423–432.

Braden, J. B. and C. Kolstad (1991). *Measuring the Demand for Environmental Quality.* Amsterdam: Elsevier.

Cameron, A. C. and P. K. Trivedi (1990). Regression-based tests for overdispersion in the poisson model. *Journal of Econometrics 46*(3), 347–364.

Cameron, A. C. and P. K. Trivedi (2001). Essentials of count data regression. In B. H. Baltagi (Ed.), *A Companion to Theoretical Econometrics*, pp. 331–348. Oxford, U.K.: Blackwell.

Cameron, C. and P. K. Trivedi (1998). *Regression Analysis of Count Data.* Cambridge: Cambridge University Press.

Cesario, F. (1976). Value of time in recreation benefit studies. *Land Economics 52*, 32–41.

Creel, M. and J. B. Loomis (1990). Theoretical and empirical advantages of truncated count data estimators for analysis of deer hunting in California. *American Journal of Agricultural Economics 72*, 434–441.

Creel, M. and J. B. Loomis (1991). Confidence intervals for welfare measures with application to a problem of truncated counts. *The Review of Economics and Statistics 73*(2), 370–373.

Curtis, J. A. (2002). Estimating the demand for salmon angling in Ireland. *The Economic and Social Review 33*(3), 319–332.

Dobbs, I. M. (1993). Adjusting for sample selection bias in the individual travel cost method. *Journal of Agricultural Economics 44*, 335–342.

D.W. Knight Associates (2005). Gros Morne National Park visitor assessment 2004. Technical report, Parks Canada.

Earnhart, D. (2004). Time is money: Improved valuation of time and transportation costs. *Environmental and Resurce Economics 29*(2), 159–190.

Englin, J. and J. Shonkwiler (1995a). Estimating social welfare using count data models: An application under conditions of endogenous stratification and truncation. *Review of Economics and Statistics 77*, 104–112.

Englin, J. and J. S. Shonkwiler (1995b). Modeling recreation demand in the presence of unobservable travel costs: Toward a travel price model. *Journal of Environmental Economics and Management 29*(3), 368–377.

Englin, J. E., T. P. Holmes, and E. O. Sills (2003). Estimating forest recreation demand using count data models. In E. O. Sills (Ed.), *Forests in a Market Economy*, Chapter 19, pp. 341–359. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Feather, P. and W. D. Shaw (1999). Estimating the cost of leisure time for recreation demand models. *Journal of Environmental Economics and Management 38*(1), 49–65.

Fix, P. and J. Loomis (1997). The economic benefits of mountain biking at one of its meccas: An application of the travel cost method to mountain biking in Moab, Utah. *Journal of Leisure Research 29*(3), 342–352.

Fix, P. and J. Loomis (1998). Comparing the economic value of mountain biking estimated using revealed and stated preference. *Journal of Environmental Planning and Management 41*(2), 227– 236.

Freeman III, A. M. (1993). *The Measurement of Environmental and Resource Values: Theory and Methods*. Washington D.C.: Resources for the Future.

Garrod, G. D. and K. G. Willis (1999). *Economic Valuation of the Environment*. Cheltenham: Edward Elgar.

Grogger, J. T. and R. T. Carson (1991). Models for truncated counts. *Journal of Applied Econometrics 6*(3), 225–238.

Gurmu, S. and P. Trivedi (1996). Excess zeros in count models for recreational trips. *Journal of Business and Economic Statistics 14*, 469–477.

Haab, T. and K. McConnell (2002). *Valuing Environmental and Natural Resources: Econometrics of Non-Market Valuation*. Cheltenham, UK: Edward Elgar.

Hagerty, D. and K. Moeltner (2005). Specification of driving costs in models of recreation demand. *Land Economics 81*(1), 127–143.

Hellerstein, D. and R. Mendelsohn (1993). A theoretical foundation for count data models. *American Journal of Agricultural Economics 75*(3), 604–611.

Hellerstein, D. M. (1991). Using count data models in travel cost analysis with aggregate data. *American Journal of Agricultural Economics 73*, 860–866.

Hesseln, H., J. B. Loomis, A. González-Cabán, and S. Alexander (2003). Wildfire effects on hiking and biking demand in New Mexico: A travel cost study. *Journal of Environmental Management, 69*(4), 359–368.

Krutilla, J. V. (1967). Conservation reconsidered. *The American Economic Review 57*(4), 777–786.

Leeworthy, V. R. and J. M. Bowker (1997). Nonmarket economic user values of the Florida Keys/Key West. Technical report, National Oceanic and Atmospheric Administration, Strategic Environmental Assessments Division, Silver Spring, MD.

Liston-Heyes, C. and A. Heyes (1999). Recreational benefits from the dartmoor national park. *Journal of Environmental Management 55*(2), 69–80.

Locke, W. and A. M. Lintner (1997). Benefits of protected areas: The Gros Morne National Park case study. Technical report, Parks Canada.

Loomis, J. (2003). Travel cost demand model based river recreation benefit estimates with on-site and household surveys: Comparative results and a correction procedure. *Water Resources Research 39*(4), 1105.

Mäler, K. (1974). *Environmental Economics: A Theoretical Inquiry. Resources for the Future.* Baltimore, MD: Johns Hopkins University Press.

McKean, J. R., D. Johnson, and R. L. J. R G. Taylor (2005). Willingness to pay for non angler recreation at the Lower Snake River reservoirs. *Journal of Leisure Research 37*(2), 178–191.

McKean, J. R., D. Johnson, and R. G. Taylor (2003). Measuring demand for flat water recreation using a Two-Stage/Disequilibrium travel cost model with adjustment for overdispersion and self-selection. *Water Resources Research 39*(4), 1107.

Moeltner, K. and J. S. Shonkwiler (2005). Correcting for size biased sampling in random utility models. *American Journal of Agricultural Economics 87*(2), 327–339.

Ovaskainen, V., J. Mikkola, and E. Pouta (2001). Estimating recreation demand with on-site data: An application of truncated and endogenously stratified count data models. *Journal of Forest Economics 7*(2), 125–144.

Parks Canada (2004a). Gros Morne National Park traffic study and attendance formula 2004. Technical report, Management Planning and Social Services Section, Atlantic Service Centre.

Parks Canada (2004b). Gros Morne National Park visitor study 2004. Technical report, Management Planning and Social Services Section, Atlantic Service Centre.

Parsons, G. (1991). A note on choice of residential location in travel cost demand models. *Land Economics 67*(3), 360–364.

Parsons, G. R. (2003). The travel cost model. In P. A. Champ, K. J. Boyle, and T. C. Brown (Eds.), *A Primer on Nonmarket Valuation*, Chapter 9. London: Kluwer Academic Publishing.

Randall, A. (1994). A difficulty with the travel cost method. *Land Economics 70*(1), 88–96.

Sarker, R. and Y. Surry (1998). Economic value of big game hunting: The case of moose hunting in Ontario. *Journal of Forest Economics 4*(1), 29–60.

Shaw, D. (1988). On-site sample regression: Problems of non-negative integers, truncation, and endogenous stratification. *Journal of Econometrics 37*, 211–223.

Shonkwiler, J. and W. Shaw (1996). Hurdle count-data models in recreation demand analysis. *Journal of Agricultural and Resource Economics 21*(2), 210–219.

Shrestha, K. R., A. F. Seidl, and A. S. Moraes (2002). Value of recreational fishing in the Brazilian Pantanal: A travel cost analysis using count data models. *Ecological Economics 42*(1-2), 289–299.

Siderelis, C. and R. L. Moore (1995). Outdoor recreation net benefits of rail-trails. *Journal of Leisure Research 27*(4), 344–359.

Smith, V. K., W. Desvousges, and M. McGivney (1983). The opportunity cost of travel time in recreation demand models. *Land Economics 59*(3), 259–278.

Sohngen, B., F. Lichtkoppler, and M. Bielen (2000). The value of day trips to Lake Erie beaches. Technical Report TB-039, Ohio Sea Grant Extension, Columbus OH.

Statacorp (2003). *Stata Statistical Software: Release 8.1.* College Station, Texas.

Ward, F. A. and J. B. Loomis (1986). The travel cost demand model as an environmental policy assessment tool: A review of literature. *Western Journal of Agricultural Economics 11*(2), 164–178.

Yen, S. T. and W. L. Adamowicz (1993). Statistical properties of welfare measures from count-data models of recreation demand. *Review of Agricultural Economics 15*, 203–215.

Zawacki, W. T. A. M. and J. M. Bowker (2000). A travel cost analysis of nonconsumptive wildlife-associated recreation in the United States. *Forest Science 46*(4), 496–506.

Notes

[1]Note that in the present study the dependent variable (denoted *persontrip* below) is the product of trips made times the group size in the current trip, but still a count.

[2]An unpublished report (Locke and Lintner, 1997) addresses the calculation of the economic benefits of Gros Morne, but using a different approach.

[3]A representative number of sample days were selected throughout the season to cover both the peak and shoulder season and weekend and weekdays within these seasons. In addition the Southern and Northern exit points have been represented in proportion to estimated use. In order to avoid any hourly bias the interview team was directed not to take lunch and break times at the same periods every day.

[4]For further details about the survey effort, the questionnaire, and the data see Parks Canada (2004a, and 2004b) and D. W. Knight Associates (2005).

[5]Englin et al. (2003) summarize the history of the application of count data models to recreation demand analysis.

[6]See Cameron and Trivedi (1990) or Cameron and Trivedi (2001, p. 336) for details.

[7]Dobbs (1993) consider the problem in a continuous context, while Moeltner and Shonkwiler (2005) extend the correction to the case of count data random utility models.

[8]Ovaskainen et al. (2001) also mention that Dobbs (1993) did not find that endogenous stratification made a big difference in slope coefficients. Shrestha et al.(2002) argue that since they used a one-time survey instead of annual vistor-data they do not anticipate bias in their results.

[9]Ovaskainen et al. (2001) did also try this specification but their fixing α for all observations at a value previously estimated using a nonlinear squares regression yielded better results in their study. McKean et al. (2003) appear to have allowed α to vary as a function of a randomly generated parameter, not related to visitor characteristics.

[10]This analysis was done using the maximum likelihood programming feature in STATA 8.1. We adapted the code for LIMDEP 7 provided in McKean et al. (2003). Our STATA code is available upon request.

[11]We thank Jeff Englin for very useful suggestions on which covariates to use to estimate α.

[12]The full text of the four-page 27-question survey is vailable upon request.

[13]For those whose point of entry was one of Newfoundland's airports.

[14]Fix and Loomis (1997) use individual data, so they have enough independent variation to include travel time as its own variable in the model. This eliminated for them the need to arbitrarily choose the fraction of the wage used to translate time into money. We did not have that variation, because we did not ask specifically about travel time (we had to infer it from the distance travelled), so we would have perfect collinearity with driving costs. Bowker et al. (1996) divide the calculation by the number of members in the visiting group. Although this would increase the log-likelihood in all our regressions and slightly improve the fit of predictions, we cannot see theoretically how the members of the visitor group can share time costs in the way gas expenses and accommodation expenses can be shared. Liston-Heyes and Heyes (1999) do not divide either travel time costs by the size of the visitor group.

[15]We faced the problem that many people did not mention a site, leaving many missing values for that variable.Perhaps some of those meant that there was no close substitute of Gros Morne they could think of. Therefore, there may be some rationale for substituting the missing values with a very large number of Km as suggested by Parsons (2003). This approach did not work well in our case.

[16]This keeps in the sample visitors from anywhere in Canada and the mainland United States as well as Western Europe, where most of the rest of foreign visitors came from.

[17]On a scale of 0 (no influence) to 10 (primary reason) we only kept those visitors who indicated a value of at least 3, excluding about 19% of the 1213 original observations.

[18]The equivalent likelihood-ratio test between a plain negative binomial and a plain Poisson (not reported but available upon request) yields $\overline{\chi}^2(01) = 40.81$ with $Prob >= \overline{\chi}^2 = 0.000$.

[19]We thank Jeff Englin for pointing us towards the use of age composition of the visitor group to estimate α.

[20]This problem of endogeneity when referred to the travel cost to the site valued is, of course, weel known and represents one of the most untractable shortcomings of the travel cost method (Parsons, 1991; Randall, 1994).

[21]This problem of item nonresponse forced us to use a dummy variable for substitutes rather than the distance to the substitute, as originally intended.

[22]See Shonkwiler and Shaw (1996) for a discussion of whether or not welfare measures calculated from a sample of users can be extrapolated to the general population. The manipulations that one would need to apply to estimated

individual welfare measures to correct the on-site bias when calculating population-wide measures are available in Parsons (2003).

[23]This can be simply done using the conventional commands *predict* and *summarize* in STATA (Statacorp, 2003). However, a more complex procedure, described below is needed for $TSNBIN$ and $TSNBIN3$.

[24]This calculation reflects that TSPOI is based on the transformation $persontrip - 1$ of the dependent variable.

Variable	Obs	Mean	Std. Dev.	Min	Max
camped	658	0.389058	0.487907	0	1
camping	658	3.542553	3.797642	0	10
daysatGM	658	3.978967	2.772878	0.5	40
dine	658	0.443769	0.497206	0	1
educat	658	4.132219	1.083638	1	6
expenses	658	0.284015	0.527075	0	12
Highlands	658	0.218845	0.413778	0	1
income	658	87.80775	43.79637	20	160
museums	658	0.369301	0.482983	0	1
persontrip	658	3.276596	2.617021	1	25
pool	658	0.089666	0.285919	0	1
propu16	658	0.068133	0.173547	0	1
satisfied	658	2.528875	0.529138	1	3
substitute	658	0.659575	0.474213	0	1
TerraNova	658	0.288754	0.453528	0	1
travelcost	658	0.672834	0.511961	0.0036767	4.813977
traveltimecost	658	0.557698	0.660024	0.001844	4.730241
valueaccomm	658	3.358663	1.536643	1	5
valuecamp30	658	2.369301	1.599686	1	5
valuedine	658	2.975684	1.339944	1	5
valuepool	658	2.018237	1.265261	1	5

Table 1: Summary descriptives of the variables used in the econometric model.

.

Variable	TRPOIS0	TRNBIN0	TSPOI	TSNBIN	TSNBIN3
persontrip					
CTC	-0.25070***	-0.23949***	-0.30293***	-0.26905***	-0.24878***
income	0.00081	0.00122	0.00098	0.00150*	0.00300**
educat	0.00423	0.00133	0.00536	0.00076	0.01100
expenses	-0.19160***	-0.19379***	-0.23270***	-0.22201***	-0.20113***
daysatGM	0.04173***	0.04435***	0.05053***	0.05119***	0.05389***
satisfied	-0.23030***	-0.23247***	-0.27945***	-0.26559***	-0.25892***
substitute	0.10309**	0.09789	0.12440**	0.10967	0.11161
TerraNova	-0.13186**	-0.13583*	-0.16018**	-0.15561**	-0.17800**
Highlands	-0.12958*	-0.14172*	-0.15843**	-0.16462*	-0.12826
museums	0.14907***	0.12262*	0.18155***	0.13288*	0.13261**
dine	-0.15295***	-0.14644**	-0.18688***	-0.16557**	-0.20378**
pool	0.32573***	0.33850***	0.39043***	0.38836***	0.27526**
camped	-0.37916***	-0.37670***	-0.45999***	-0.42744***	-0.42545***
valuedine	0.05455**	0.04651	0.06619***	0.05058	0.05517
valuepool	0.06566***	0.07326**	0.08048***	0.08595***	0.08056**
valuecamp30	0.04255***	0.04283**	0.05148***	0.04870**	0.04307*
valueaccomm	-0.08746***	-0.09345***	-0.10639***	-0.10791***	-0.09960***
camping	0.02334***	0.02135*	0.02829***	0.02366*	0.02074*
constant	1.67751***	1.63531***	1.47125***	1.15491***	0.93112***
α					
income					-0.01011**
propu16					2.6852***
constant		0.16396***		0.32968***	0.58698
log-likelihood	-1256	-1217	-1259	-1192	-1177
N	658	658	658	658	658
pseudo-R^2	0.104	0.132	0.124	0.170	0.181
χ^2	291.1***	168.5***	355.7***	190.2***	150.5***
AIC	3.876	3.761	3.885	3.683	3.644
	legend: * p<.1; **p<.05; *** p<.01				

Table 2: Estimation results of the different truncated count data models. Dependent Variable is $persontrip$.

	TRPOIS0	TRNBIN0	TSPOI	TSNBIN	TSNBIN3
$\widehat{\beta}_{CTC}$	-0.2507	-0.23949	-0.30293	-0.26905	-0.24878
CS/persontrip$^{(a)}$	$3, 989	$4,176	$3,301	$3, 717	$4,020
CS/group for five years	$11,285	$11,580	$13,472	$12,130	$13,100
CS/individual$^{(b)}$ for five years	$880	$903	$1051	$946	$1022
Expected *persontrip*	2.83	2.77	4.08	3.26$^{(c)}$	3.26$^{(c)}$

(a) CS/persontrip = $\$1000(1/\beta_{CTC})$
(b) Based on an avegage partysize of 2.5635
(c) Based on $E(persontrip_i/x_i) = \lambda_i + 1 + \alpha_i\lambda_i$

Table 3: Estimated welfare measures.