Valuation of Outdoor Recreation Spaces
using the
Travel Cost Model

by

Irka Poplawetz
3989116

Major Paper presented to the
Department of Economics of the University of Ottawa
in partial fulfillment of the requirements of the M.A. Degree
Supervisor: Professor Kathleen Day
ECO7997

Ottawa, Ontario
January 2007

Abstract

This paper uses the Travel Cost Model to estimate the demand for and use value of natural outdoor areas for recreation in Canada. Estimations using OLS to estimate the linear model and Maximum Likelihood to estimate the semi-log Poisson model of the demand for recreation are compared, using actual travel cost as a proxy for price and using demand shifters in the form of demographic variables. Our results indicate that consumption of active outdoor recreation in natural areas increases as participants' level of education increases, as participants' age decreases, and as participants' income increases. Using the estimated demand equations, estimates of consumer surplus are calculated to find the use value of natural outdoor recreational areas. The linear model and semi-log Poisson models are compared. We conclude that the Poisson model is a more appropriate model to use to estimate the demand for and the valuation of natural outdoor areas for recreation in Canada.

1. Introduction

The correct valuation of outdoor natural areas in Canada is important because of the existence of competing uses for these areas. Such competing uses include residential development, private recreational development (such as a private golf course or a private ski area) and industrial development (including the harvesting of lumber and the mining of minerals). By identifying the value of these natural outdoor areas to individuals who enjoy them for recreational purposes and comparing it to their value if developed will help to ensure that areas with a high value as a recreational space will remain available for such use.

Valuing natural areas is a challenging task because natural outdoor areas are a non-market good. Credible values for the recreational use of areas are necessary in order to justify using them to compare to the values calculated for alternate uses of the areas. This requires that appropriate models and methods of estimation be used to estimate the demand for recreation and to calculate the consumer surplus generated by recreational activity. The existing literature on this topic is abundant and its evolution over the last several decades has validated the work that has been done to discover the best method of valuing natural outdoor recreational areas.

One of the most popular methods of evaluating natural areas is the Travel Cost Model (TCM), which was first suggested by Hotelling (1949). The TCM uses the amount spent to travel to the area of recreation as a proxy for price. Given this proxy for price, one can estimate a demand curve for recreation in natural outdoor areas. The demand curve can then be used to calculate the consumer surplus associated with a given level of consumption of the natural area. Consumer surplus is widely regarded as a reasonable proxy for the true value of a non-market good.

This paper uses the individual Travel Cost Model, based on an off-site survey, to estimate the demand for and the valuation of natural outdoor recreation areas in Canada. The data used are from the *Survey on the Importance of Nature to Canadians During 1996*. This off-site survey has the advantage that non-users of the natural areas are also surveyed and thus avoids selection bias.

To discover the most appropriate models and methods of estimation it is necessary to compare them against each other, and thus this paper will compare the results from estimation of a linear model using OLS to results from estimation of a semi-log Poisson model using Maximum Likelihood. The number of days spent in outdoor areas is the dependent variable, and the explanatory variables include a "price" variable, and income variable and other demographic variables. Both of these demand functions are then used to value the natural outdoor areas by estimating their respective consumer surpluses.

In contrast to previous studies, this paper estimates the aggregate value of the natural outdoor recreation areas in Canada. Instead, previous studies have valued a relatively small area, such as a portion of a state or province, or a geographic region, or a specific feature such as a park or a lake. This paper thereby augments the existing literature by using comparison techniques similar to previous researchers, but by estimating a demand function and consumer surplus for a larger area. The results confirm those of previous studies that suggest that the Poisson model is better than the linear model and that estimates of consumer surplus calculated using the Poisson model are smaller than those calculated using the linear model. The estimates for consumer surplus obtained for the sample used in this paper are $523M and $125M, respectively, for the linear model and the Poisson model.

2. Literature Review

2.1 Overview of valuation of non-market goods

The general problem of valuing natural areas arises because of the difficulty of measuring a resource that is a non-market good and thus has no price. The total value of outdoor recreation and the consumer surplus (the difference between the amount consumers actually spent engaging in outdoor recreation and how much they would be willing to spend) are important values to measure. In order to arrive at credible values, it is necessary to design a model that can be used to quantify the value of these outdoor recreational spaces (Burt and Brewer, 1971). Burt and Brewer suggest that interviewing households directly is the only feasible way to gather data to estimate the demand equation(s). The Travel Cost Model meets this requirement by interviewing individuals and gathering information directly on the quantity and price of outdoor trips, as well as gathering demographic data on each individual.

There are various methods of valuing non-market goods. These methods are generally grouped into two main categories: revealed preference methods and stated preference methods. Revealed preference methods, such as the Travel Cost Model, use data gleaned from surveys designed to record the actual use of the non-market good, based on the behaviour of individuals that reflect utility maximization subject to constraints. Stated preference methods, such as the Contingent Valuation Method, use data based on what respondents indicate they would do with respect to paying for the use of natural outdoor areas.

The Contingent Valuation Method first came into use in the early 1960s when economist Robert K. Davis used questionnaires to estimate the benefits of outdoor recreation in an outdoor recreational area in Maine (Davis, 1963). The defining characteristic of the CVM is its hypothetical nature. The Contingent Valuation Method is an approach to valuation that uses individual responses to contingent circumstances suggested in an artificially structured market (Stoll, Loomis and Bergstrom, 1987). In other words, in contingent valuation surveys respondents do not actually pay the amounts they indicate as their valuation of the good.

However, Boyle (2003) points out that economists generally prefer to observe individuals' valuations of goods and services through their behaviour. These observations reveal the consumers' preferences. The Travel Cost Model uses revealed behaviour by asking respondents to a survey to indicate the amount of travel expenditures actually incurred in their visit(s) to an outdoor recreational area. The Travel Cost Model was first suggested by Harold Hotelling (1949) and the model was initially applied by Samuel Wood and Andrew Trice (1958) and by Marion Clawson (1959).

Most applications of the Travel Cost Model have been to value recreational sites. Boardman et al. (2001) suggest that estimating the demand schedule for a particular site, using the TCM, involves selecting a random sample of households that are surveyed to determine the number of visits they made to the site during a time period, the costs they incurred visiting the site, their incomes and other characteristics that could affect demand. After a functional form has been specified, these data are used to estimate the demand schedule. If the market for visits to a site is over a wide geographic area, then visitors incur different travel costs depending on their travelling expenses. The resulting differences in total cost (the implicit price), and the differences in the number of visits, provide data to estimate a Marshallian demand function (Smith, 1990).

The main assumption of the TCM is that individuals respond to changes in the travel-related part of the cost of a trip to a recreation site in the same way as they would respond to a change in the price of admission to the site (Freeman, 2003a). Thus, in the Travel Cost Model the costs of consuming the services of the environmental asset are used as a proxy for price (Hanley and Spash, 1993). Costs of travelling to and from the site may include the opportunity cost of travel time and on-site time, the cost of vehicles used, accommodations, food and all other expenses incurred while on the trip. These costs are summed and the total is used as an explanatory variable in determining the quantity of visits to the site.

Whether travel costs should include time costs is not clear. Some researchers have included various estimates of time costs while others have assumed that time costs are zero. However, it is becoming evident that many current researchers are including time costs for some individuals, but not others, dependent upon their employment status. The time cost factor is further discussed at several points later in this paper.

There are two variants of the Travel Cost Model: individual and zonal. For the zonal travel cost model, researchers survey actual visitors at a site. The analysis is based on the zone in which the users of the site live. The zones are determined by concentric circles around the site, or by municipal boundaries. The zonal model assumes that households within a zone have similar travel costs, income and other characteristics (Boardman et al., 2001). In contrast, the individual Travel Cost Model undertakes an analysis based on an individual's decisions and data are recorded for each respondent.

The individual Travel Cost Method is favoured over the zonal Travel Cost Method for several reasons, mentioned by Bowker, English and Donovan (1996). First, it better addresses statistical efficiency. Second, it better addresses theoretical efficiency in the modeling of behaviour. Third, it avoids using arbitrarily defined zones, and fourth, it considers the heterogeneity of populations within zones. Parsons (2003) indicates that the zonal model has fallen out of favour because it is not consistent with economic theory. Hanley & Spash (1993) also agree that the zonal model is unlikely to be consistent with consumer theory. Haab and McConnell (2002) summarize the current view of the zonal version of the Travel Cost Model:

> The zonal travel cost model has been supplanted by models of individual behaviour for several reasons. Zonal models require aggregate visitation data, often not available. But a more difficult problem is that zonal models assume that behaviour of individuals within a zone is identical. (Haab and McConnell, 2002, p. 181).

2.2 Data Requirements

The Travel Cost Model uses surveys to gather data. Regarding the data that are required to be collected in the survey and the most feasible way to gather them, Burt and Brewer (1971) suggest that households be directly interviewed in order to gather the following required information:

> (i) number of days the household spent at each category of recreation sites, inclusive of travel time; this information will almost certainly be collected trip by trip and the site classified by the analyst; (ii) expenditures incurred that were specific to each trip (excluding auto expenses which are estimated separately by mileage); (iii) mileage driven on each trip (could

> be estimated by the analyst if each site were identifiable on a map); and
> (iv) family income. (Burt and Brewer, 1971, p.819)

Interviewing households directly can be achieved using various survey methods, such as personal interviews, telephone interviews, or mail-in questionnaires.

Surveys can be administered in two ways: on-site and offsite. On-site surveys gather data only from site users. Off-site surveys gather data from users and non-users. Off-site surveys avoid selection bias and provide information on non-users. The major problems with on-site surveys are that they result in truncation at one trip and they over-sample more frequent users (Parsons, 2003). Parsons adds:

> "Off-site random samples avoid the problem of truncation and endogenous stratification and have the advantage of including non-participants. This allows the researcher to model the decisions to participate in recreation at the site or not." (Parsons, 2003, p. 288)

Where possible, an off-site survey is preferred, in order to include non-users. This greatly increases the cost of the survey because it increases the number of individuals surveyed; however, the information provided is more representative of the population as whole.

2.3 Non-econometric limitations

The Travel Cost Model poses several difficulties to those using it as a method of valuing non-market goods. Stoll, Loomis and Bergstrom (1987) point out that the TCM approach to valuation requires that travel be necessary in order to use the resource, as do Hanley and Spash:

> Only recreational resources which necessitate significant expenditure for their enjoyment can have user values estimated. For example, no estimate could be made for the value of a park in the middle of city which is visited only by pedestrians. (Hanley and Spash, 1993, p.93).

This problem involves weak complementarity (Mäler, 1974): the assumption that a good is valued only if there is a marketed good that is also consumed. Thus, if the purchase of the marketed good (e.g. gasoline, accommodations) is zero, then the marginal demand price of the outdoor recreation area is also zero (Hanley and Spash, 1993).

Another problem with the TCM is the difficulty of valuing travel time to and from the site visited for outdoor recreation. If included, the value of the time spent must be estimated, usually based on a particular hourly wage. Van der Straaten (2000) points out that in all the discussions on valuing travel time, the fact that the value of travel or visiting time is only relevant for consumers who have earned income is not taken into account. Randall (1994) has argued that the difficulty of observing the monetary value of the time component of the cost of a trip is a fundamental limitation of the Travel Cost Model. Burt and Brewer (1971) also address the issue of whether or not to include travel time in the calculation of travel costs and point to the difficulty of deciding whether travel time is a cost of the trip to the recreational area, or whether it is part of the recreational experience. If part of the travel time has some utility, then travel cost is overestimated if time is given a value. Alternatively, if time is not included as a cost, the cost variable is biased downwards when there is disutility to the travelling portion of the trip. In practice, researchers decide on an individual basis whether or not to include travel time as a cost. Hanley and Spash (1993) give time a zero value, as do Seller et al. (1985), while Freeman (2003b), Parsons (2003) and McKean (2005) include the wage rate as the opportunity cost of time.

Bowker, English and Donovan (1996) provide some insight into the sensitivity of valuations to the treatment of travel time. They found that the biggest difference in estimates of consumer surplus results from the different values used for time costs. This study underlines the critical role of the value placed on time in the Travel Cost Model. The time issue in still unresolved in the TCM literature (McKean, Johnson and Walsh, 1995), however it is evident that most current researchers are including time costs in their studies.

A third problem with the TCM is that it cannot measure the value of non-use of the resource. Natural outdoor areas provide use benefits and non-use benefits. Use benefits include rivalrous consumption (such as hunting), direct non-rivalrous consumption (such as hiking), and indirect non-rivalrous consumption (such as watching a movie about hiking). Boardman, Greenberg, Vining and Weimer (2001) also mention the category of non-use benefits that include pure existence value (valuing the natural order) and altruistic existence value (such as valuing other people's use or non-use value of wilderness). Because non-use values do not involve any expenditure, the TCM is unable to measure them and is restricted to measuring use value only. A consequence of this short-coming is that the model will underestimate the value of a natural area because it is unable to include the wide range of non-use values (Hanley and Spash, 1993). The interpretation of the value calculated using the Travel Cost Model, then, is specifically the use value of the area.

Yet another problem with using the Travel Cost Model is that some trips are multi-purpose trips (e.g. travel to spend a day hiking combined with business in a nearby city). This difficulty is dealt with by ensuring that respondents identify the primary purpose of the trip. It is assumed that each trip to the site is for the sole purpose of visiting the site. Parsons (2003) suggests that if the purpose of the trip is to engage in other activities while on a trip, then at least part of the travel cost would be a joint cost that cannot be distributed among different purposes. If another destination can be identified as the primary purpose of the trip, then the relevant cost of the visit to the recreation site is the additional cost of adding the site visit to the trip, given the trip to the main destination. Hanley and Spash (1993) include only the data from respondents for whom the trip to an area was the primary purpose. In fact, Haab and McConnell (2002) list, as a condition for calculating welfare measures, that the trips included in the data be single-purpose trips taken to the recreation site. Including only those trips for which recreation in the natural area is the primary purpose would lead to underestimation of the use-value of a site because some secondary-purpose trips to the site are not taken into consideration in the valuation.

Another limitation of the TCM is that it provides an estimate of the value of the entire area, but not of its specific features. The TCM does not provide the needed information to value changes in specific features of a site, as discussed by Brown and Mendelsohn (1984). In order to value changes in features of a site, a valuation of the site would need to be done before the changes and again after the changes in order to place a value on the change in features of the site.

The TCM also may encounter difficulties with poor respondent recall of amount spent on travel, if there are no researcher-imposed mileage rates to calculate travel costs based on distances. However, gathering this cost information from the respondent provides more variability in trip cost data. In addition to recall bias, respondents may differ in what they perceive as trip costs (Bowker et al., 1996).

2.4 Econometric implementation

The general estimating equation in the Travel Cost Model has a dependent variable for number of days or number of trips and several independent variables, one of which is the total cost-per-day or total cost-per-trip. The other independent variables are demographic variables such as income, level of education and age. A typical TCM estimating equation is

$$(1) \qquad visits_i = \beta_0 + \beta_{tc}tc_i + \boldsymbol{\beta}'\mathbf{x} + \varepsilon_i,$$

where the variable *tc* (total cost) is the proxy for price; $\mathbf{x}$ is the vector of demographic variables, such as income, education, age and marital status; and ε_i is the error term .

Early studies of the Travel Cost Model, such as that of Burt and Brewer (1971), used OLS to estimate the demand equation, even though they found that their data contained a large proportion of zero values. More recent researchers have criticized the OLS technique. Censoring and truncation were found to be problems, followed by the recognition that the data in such studies are count data.

A truncated sample is one in which observations with values of the dependent variable above or below predetermined bounds are deleted from the sample. A censored sample is one in which observations with values of the dependent variable above or below predetermined bounds are altered to equal the boundary values when they are recorded. Truncated and censored samples both make OLS an inconsistent estimator of a regression equation's slopes and intercept. OLS inconsistently estimates linear regression models when the data are from truncated or censored samples. For on-site surveys truncation is the problem in that the complete distribution of the dependent variable may not be observable (we do not see any of the data for individuals who choose not to visit the site). For off-site surveys (they include individuals who never visit the site) censoring is the problem because we see a large part of the distribution massed at zero.

In order to consistently estimate regression models with data from truncated and censored samples, Maximum Likelihood methods can be used. One such ML method is the Tobit model. It has been widely used in studies of demand for goods and services using individual data where there are many zero observations. The Tobit model has been used to address the issue of truncated or censored data in estimating the demand for recreation in natural areas and other environmental amenities (Smith and Desvousges, 1985 and Choe, Whittington and Lauria, 1996). The Tobit model is appealing because it allows for many observations massed at zero. However, shortcomings of the Tobit model when applied to the Travel Cost Model have caused it to fall out of favour. Maddala (2002) points out that the Tobit model is only applicable when the dependent variable can take on negative values but they are just not observed because of censoring. In the case of estimating travel demand, the dependent variable can never take on negative values, thereby making the Tobit model inappropriate.

Another development in the literature has been the recognition that the data for the dependent variable involved in applying the Travel Cost Model are count data. It is more reasonable to treat the dependent variable as an integer rather than a continuous variable because the number of days or number of visits is measured in integer values. The implication this has for estimation is that a Poisson or Negative Binomial model is appropriate. Both the Poisson and Negative Binomial models are estimated using Maximum Likelihood techniques.

In most of the recent applications the model is estimated as a count data model. The dependent variable (quantity of outdoor recreation demanded) is a non-negative integer and the data set is comprised of a large proportion of zero and small values, making it suitable for a count model such as the Poisson regression. Parsons (2003) suggests a Poisson model of recreation demand such as the following:

(2) $ln(\lambda_i) = \alpha_{tc}tc_i + \boldsymbol{\alpha}'\mathbf{x},$

where λ_i is the expected number of trips, tc_i is travel costs, and x is a vector of demographic variables. In the Poisson model we have

(3) $E(visits_i) = \lambda_i$

where $\lambda_i = exp(\alpha_{tc}tc_i + \boldsymbol{\alpha}'\mathbf{x})$ and the derivative, for any variable, x_k, is $\alpha_k exp(\alpha_{tc}tc_i + \boldsymbol{\alpha}'\mathbf{x})$, or $\lambda_i\ \alpha_k$.

In the linear model we have

(4) $E(visits_i) = \beta_0 + \beta_{tc}tc_i + \boldsymbol{\beta}'\mathbf{x}$

and the derivative, for any variable, x_k, is β_k. Thus in the linear model we can look directly to the coefficient of any variable and its value will tell us the effect a one-unit change in that variable will have on the dependent variable.

The Negative Binomial model is also used to estimate travel cost models and may be preferable to the Poisson model. In the Poisson model the mean and the variance are both constrained to be equal to λ. But this is an unreasonable constraint that is relaxed when the Negative Binomial model is used (Greene, 1997). However, the Negative Binomial model is generally more difficult to estimate than the Poisson model. There is a consensus in the current literature that for count data, the model used should be the Poisson or the Negative Binomial.

Another econometric issue that arises with the Travel Cost Model is that the travel cost variable may be endogenous, such as when individuals decide to reside in an area because it is close to the recreational site (thereby reducing their travel costs), or when individuals make equipment purchases based on their individual tastes (thereby causing their costs to increase or decrease based solely on their choice of quality of equipment). There does not appear to be a remedy for this issue and it does not appear to be frequently addressed in the literature.

2.5 Measuring consumer surplus

Generally, consumer surplus per-person-per-day is estimated to be the reciprocal of the absolute value of the coefficient on the travel cost variable. In order to compute the total consumer surplus for an area, this per-person-per-day value must be multiplied by the total number of days in the sample and then factored up to be representative of the entire population. When using the Poisson distribution, for individual i, the surplus is

$$(5) \qquad CSP_i = \lambda_i / -\alpha_{tc} = -exp(\alpha_{tc} tc_i + \boldsymbol{\alpha}'\mathbf{x})/\alpha_{tc}$$

where CSP_i is the consumer surplus for individual i, λ_i is the expected number of visits spent in outdoor recreation for individual i, and α_{tc} is the coefficient on the travel cost variable (Parsons, 2003). Once the parameters of the model are estimated, the above equation is used to calculate the surplus value for each individual in the sample and is then aggregated over the population of users to arrive at a total access value.

An alternative method of estimating aggregate surplus is to compute an average per-trip-per-person value and then multiply this by an estimate of the total number of trips taken to the site. This gives us an average per-trip value in a Poisson model, according to Parsons (2003) as

$$(6) \qquad CSP_i = \frac{\lambda_i/\text{-}\alpha_{tc}}{\lambda_i} = 1/\text{-}\alpha_{tc}.$$

This applies to on-site and off-site models alike. To arrive at an aggregate value for the site, we multiply the average per-trip value by the number of trips taken to the site during the relevant time period.[1]

For a linear model, Adamowicz et al. (1989) provide the following formula for the consumer surplus of one individual:

$$(7) \qquad CSL_i = visits_i^2/\text{-}2\beta_{tc},$$

where CSL_i is the consumer surplus for individual i, *visits* is the quantity of recreation consumed and β_{tc} is the coefficient on the travel cost variable.

In a TCM analysis the choice of functional form can affect the value for consumers' surplus per visit generated because various functional forms will result in varying values for the coefficient of the travel cost variable. Smith and Desvousges (1985) point out that the progressive refinement in the generalized travel cost model illustrates how sensitive the consumer surplus estimates can be to the modeling and estimation judgements made in deriving them. Adamowicz et al. (1989) also focus on the variability of welfare measures across functional forms, pointing out that this is a result of the coefficient of the price variable being in the denominator of the consumer surplus equation. Englin and Shonkwiler (1995) reflect this, too, in saying that the consumer surplus per trip is inversely proportional to the responsiveness of the model to price changes, determined by the coefficient on the price variable. Haab and McConnell (2002)

[1] Some studies use the actual days consumed, while others use the predicted number of days.

also stress the great sensitivity of welfare estimates from the travel cost model to the actual specification estimated. Including non-participants in the survey (Yen and Adamowicz, 1993) by using an off-site survey (using an untruncated model) causes estimation to result in lower consumer surplus estimates.

2.6 Review of case studies

It is interesting to review some of the wide range of applications of the Travel Cost Model in the literature over time. These applications vary with respect to type of survey (on-site versus off-site), whether time costs are taken into consideration, the model and estimation method used (OLS, Tobit, Poisson, Negative Binomial) the particular area being valued and various other features of the study. This review of case studies will follow the development of the TCM from 1971 to the present.

Burt and Brewer (1971) use generalized least squares to estimate a linear functional form, using an on-site survey of water-oriented outdoor recreation in Missouri. They do not include time costs. The coefficient on the price variable is -0.0222, with annual net benefits of the three lakes in the study calculated at 8.5 million dollars.

Smith and Kopp (1980) also do an on-site survey, using zero time costs, with a log-linear functional form for the demand model, in a zonal TCM study of the Ventana area of the US (includes parts of California, Oregon, Nevada and Arizona) in 1972. The coefficient on the log of travel cost variable is -1.587 and the resulting consumer surplus is $14.40 per trip.

Seller, Stoll and Chavas (1985) did an off-site survey (but excluded observations in which the dependent variable is zero). They do not include time costs and use a linear functional form as did Burt and Brewer. They chose four lakes in East Texas to be the focus of their recreational boating study. They estimated the coefficients by SUR (seemingly unrelated regression) for a system of four demand equations (one for each lake).

Smith and Desvousges (1985) report that they previously did an OLS study to estimate demand functions for 33 water-based recreation sites and wanted to redo the study using other methods of estimation. They used the Poisson model to reflect the discrete probability structure. Their Poisson results differ substantially from the original OLS estimates: the Poisson estimates of the travel cost parameter are larger in absolute magnitude implying more elastic site demand. The newer estimates imply substantially smaller site value per unit of use, ranging from \$3.58 to \$105.26 for the Poisson estimates in comparison to \$39.00 to \$400.00 for the OLS model (all in 1977 dollars).

Smith (1988) compared off-site surveys using OLS, Tobit and Poisson estimations to value water quality of the Monongahela River area, specifically for North Park Lake. For the linear specification of OLS the consumer surplus was \$1.67 per trip; for the linear specification of the Tobit model the consumer surplus was \$0.23 per trip; for the semi-log specification of the Poisson model the consumer surplus was \$0.38 per trip (all in 1981 dollars).

Adamowicz, Fletcher and Graham-Tomasi (1989) estimate several functional forms of demand functions for recreation. The data were collected by mail survey for a bighorn and Rocky Mountain sheep hunting site in Alberta. In their linear model the slope coefficient is -0.0025 with consumer surplus of \$1,367 for an individual with a mean number of visits of 2.61.

Hellerstein (1991) uses 1980 permit data from the Boundary Waters Canoe Area in northern Minnesota to compare various count data models. For comparative purposes a semi-log model is also estimated, which requires that occurrences of zero visits be dropped. The coefficients are -0.0208, -0.0158 and -0.0111, respectively for Poisson ML, Negatie Binonmial ML and semi-log OLS. He shows that the count models outperform the drop-zero semi-log model.

Bowker, English and Donovan (1996) are the first of several researchers reviewed who pay particular attention to time costs in their estimation of the per-trip consumer surplus for guided white-water rafting on the Chatooga River, using guide records (thus, it is an on-site survey). They estimate cases using zero time costs, time cost valued at 25% of the household wage rate, and time cost valued at 50% of the household wage rate. The estimated demand model was a Negative Binomial model in order to address the integer nature of trips and to correct for zero truncation. For the zero-time-cost case the coefficient on the travel cost variable is -0.0072, for the 25%-of-wage-rate case the coefficient is -0.0052 and for the 50%-of-wage-rate case the coefficient is -0.0035. The corresponding consumer surplus per trip is \$139.56, \$192.99 and \$286.22. This case demonstrates how the valuation can differ greatly based only on a change in the value placed on the opportunity cost of time.

Choe, Whittington and Lauria (1996) also include the opportunity costs of travel time in their study, and use the linear OLS and Tobit models to estimate demand, both before and after a public health advisory, in order to value water quality improvement in Davos, Philippines. The valuation is in pesos per month. The OLS before and after coefficients were -0.104 and -0.029; the Tobit before and after were -0.201 and -0.144.

The estimated mean WTP for water quality improvement (in pesos per month) is 51 pesos in the Tobit model and 36 pesos in the model estimated using OLS.

McKean, Walsh and Johnson (1996) continued the trend to including time as a cost, including both the opportunity cost of travel time and on-site time. They used an on-site survey, using Poisson and Negative Binomial models, in an application of the TCM to a sample of 161 visitors to Blue Mesa reservoir in 1986. The coefficients of the Poisson and Negative Binomial were identical, although t-values were slightly lower for the Negative Binomial regression. For the Poisson model they estimate a coefficient on the travel cost variable of -0.01445, with a corresponding consumer surplus of $69.20 per trip. They report that consumer surplus decreases when time costs are excluded: the coefficient is -0.022505 with a corresponding consumer surplus of $44.43 per trip.

Englin, Boxall and Watson (1998) also include time costs and use a count model in their study. The model is applied to individual wilderness recreation trips in a system of four Canadian wilderness parks in 1993-1994. The value of travel time was based on an average speed of 80 km/hour and the opportunity cost value at one-third of the average wage rate. They report that trips to the most remote of the four parks were estimated to be worth over $700 per day.

Parsons (2003) also includes time costs, using an imputed wage to value the opportunity cost of time. A Poisson model is used to estimate demand for two parks in Ohio. The survey is conducted on-site in 1997; random beach users were given a survey and asked to return it by mail. Four different functional forms were estimated: two continuous (linear and log-linear) and two count (Poisson and Negative Binomial). The Poisson model parameter estimates are reported for each of the two parks, with coefficients -0.040 and -0.026. The per-person-per-trip consumer surplus values are $25 and $38. The total access value of the beaches at the parks was reported to be $5.6 million each.

McKean (2005) includes travel time and uses the Negative Binomial method to estimate demand for non-angler recreation at the impounded Snake River in eastern Washington. A question in their survey asked if income was foregone while travelling or recreating. The estimated coefficient on the travel cost variable was -0.04056. Consumer surplus was the reciprocal, or $24.65. Average recreationist trips per year were 7.36. Total surplus is the product of average annual trips and surplus per trip or $181.42 per recreationist per trip. It was estimated that 52,984 unique non-angler recreationists visit the reservoirs per year. They applied their estimated consumer-surplus-per-person to all non-angler recreationists, with a resulting total consumer surplus of $181.42 x 52,984 = $9.61 million.

These case studies demonstrate that most modern valuations using a travel cost model include time costs and use a count data model to estimate the demand curve and to calculate consumer surplus, or use value. They also reveal that researchers who have compared a linear model and a Poisson model have found that the Poisson model is a better model to use when estimating a demand for recreation equation and when calculating consumer surplus from that demand equation.

3. Econometric Model

Two econometric models will be estimated in this paper. The first model is

(8) $$DAYS_i = \beta_0 + \beta_C COSTPERD_i + \mathbf{x}'_i\boldsymbol{\beta} + \varepsilon_i,$$

where $DAYS_i$ is the number of days individual i spent recreating in natural areas outdoors in 1996, $COSTPERD_i$ is the cost per day individual i incurred on such recreation and **x** is a vector including all other explanatory variables: gender, urban/rural residence, level of education, household size, age, marital status and income. Linear models are simple to estimate and have been frequently used in the literature.[2]

In this equation, the coefficient of *COSTPERD* is expected to be negative because we expect the demand curve for outdoor recreation in natural areas to be downward-sloping; that is, as the cost-per-day of travel increases we expect that the number of days of travel for outdoor recreation in natural areas decreases.

Because gender may affect the quantity of outdoor recreation, the vector **x** includes a dummy variable indicating whether the individual is male or female. However, the effect of gender may not be large, given that many vacationers travel with a person of the opposite gender.

The place of residence (urban or rural) of an individual may affect the quantity of outdoor recreation they engage in. Whether urban or rural dwellers engage more in outdoor recreation is difficult to anticipate. Rural inhabitants may have easier (in terms of convenience and time) access to outdoor recreational activities than their urban counterparts. Urban dwellers may be more likely to take advantage of engaging in these outdoor activities to escape the city.

The level of education a respondent has attained is an important factor in the model and we expect that the level of education attained by a respondent will strongly influence the number of days they engage in outdoor recreation. For this group of

---

[2] For example, see Burt and Brewer (1971); Seller, Stoll and Chavas (1985); and Adamowicz, Fletcher and Graham-Tomasci (1989)

variables we expect that as we progress from lower levels of education to higher levels of education the coefficient will increase, because individuals with increasing levels of education are more aware of the health and well-being benefits of outdoor activity in natural areas.

Household size is also a variable that will probably influence the quantity of outdoor recreation that an individual engages in. However, the signs of the coefficients of this group of variables are not as easy to anticipate as, say, those of the education group of variables. On the one hand, a single person may find it easier to schedule a trip in the outdoors; however, larger households may find that it is more economical to take trips because they can spread travelling costs over a number of people. Furthermore, trips to engage in outdoor activities may be the only affordable type of vacation for a large household.

The age of a respondent is another demographic factor that we expect will strongly influence the quantity of outdoor recreation consumed because age typically determines a person's lifestyle. We expect the coefficients of the age group variables to decrease as age increases, as outdoor recreational activities become increasingly less frequent as children enter into families, and as individuals encounter health problems and enter into old age when physical mobility becomes a factor.

Marital status is included as an explanatory variable because it is a factor that influences lifestyle and thereby influences opportunities to engage in outdoor recreation. We expect that single individuals have the greatest opportunity to participate due to fewer conflicting responsibilities. However, we expect that married people and those living in common law relationships would participate because they can go as a couple or as a family. It is difficult to say what the net effect of these opposing effects is. We expect that widows and widowers would participate less due to lack of a convenient companion with whom to travel. Similarly, we expect that divorced or separated individuals would engage less in outdoor recreational activities, because of lack of a travelling companion.

Finally, dummy variables for income classes are also expected to strongly influence the quantity of outdoor recreation consumed. Regarding the coefficients on these variables, we expect that as income increases, participation in recreational outdoor activities will increase because of the ability to afford such trips.

To summarize, we expect the coefficient on the travel cost variable to be negative, and the education, age and income variables are expected to strongly influence the quantity of days engaged in outdoor recreation, causing quantity to increase as education and income increase and as age decreases. Variables that may have a weaker effect are gender, urban versus rural residence, household size and marital status.

Equation (8) will be estimated using OLS. Diagnostic tests will be applied to the OLS estimates. These include tests for heteroscedasticity, specification error, and normality of the errors. To test for heteroscedasticity we exclude the computation of the White test statistics for heteroscedasticity because we have dummy variables in our model. To test for specification error we use Ramsey's Reset Test. To test for normality of the errors we use the Jarque-Bera test for normality of the residuals.

The second model is a Poisson regression model. We chose to estimate it because the data we have are count data, and previous researchers have successfully used a Poisson model to estimate models with count data. This Poisson model is

$$(9) \qquad ln[E(DAYS_i)] = exp(\alpha_0 + \alpha_c COSTPERD_i + \mathbf{x}'_i\boldsymbol{\alpha}),$$

where the variables are the same as in equation (1). Expectations regarding the signs of the coefficients are the same as for the linear model.

Consumer surplus will be calculated using each method of estimation. The formulas used are based on those suggested by Adamowicz et al. (1989). The consumer surplus formula for the linear functional form is

$$(10) \quad CS_L = DAYS_i^2 / -2\beta_c,$$

where $DAYS_i$ is the number of days of outdoor recreation consumed by individual i. The consumer surplus formula for the semi-log functional form is

$$(11) \quad CS_P = DAYS_i / -\alpha_c.$$

4. Data

The data used for this paper are from Statistics Canada's *Survey on the Importance of Nature to Canadians during 1996*. This survey was conducted by Statistics Canada from February to June 1997. The survey was administered to a sub-sample of the households in the Labour Force Survey. As a result, the data are accurate and reliable, with a resulting high quality of the sampling variables. The respondents to the survey are Canadians 15 years of age and over.[3]

Several of the many sections in the survey are used in this paper in order to study trips taken during 1996 for the main reason of participating in one or more of 17 specified outdoor activities: sightseeing in natural areas; photographing natural areas; gathering nuts, berries or firewood; picnicking; camping; swimming/beach activity; canoeing/kayaking/sailing; power boating; hiking/backpacking; climbing; horseback

[3] Nunavut was not in existence at the time the survey took place. Residents of the Northwest Territories were not surveyed due to the high cost of surveying that territory.

riding; cycling; off-road vehicle use; downhill skiing; cross-country skiing/snowshoeing; snowmobiling; and relaxing in an outdoor setting. The survey specifically included only those trips for which the primary reason was to take part in one or more of these 17 activities. This was done in order to avoid the difficulty of valuing multi-purpose trips.

The data in the *Survey on the Importance of Nature to Canadians during 1996* provide most of the information generally required for conducting a valuation using the Travel Cost Model. The data include the costs incurred by individuals to engage in outdoor recreation (including costs of travel, accommodation and equipment). The data include the number of days respondents spent engaged in outdoor activity in 1996. The data also include demographic information usually included in the TCM.

The data used in the model include all outdoor recreation areas in Canada that were enjoyed by the respondents in the survey. The TCM can be used to value a single site or a cluster of sites. Recent developments in the use of the model study regions including a group of sites, rather than for a single site (Seller et al., 1985). Thus the use of the *Survey on the Importance of Nature to Canadians during 1996* can be regarded as an extension of this approach to the case where the group of sites surveyed is extremely large.

The survey was conducted off-site, thereby including both non-users and users of the outdoor recreational areas. As a result, the data are free of selection bias and include useful information on individuals who do not participate in outdoor recreational activities. However, because the dependent variable includes zeros (those respondents who engaged in zero days of outdoor recreation during 1996) a semi-log demand model cannot be estimated using OLS unless the zero values are dropped. This disadvantage is outweighed by the advantage of having information on non-users. As a result the data are superior to data collected in an on-site survey.

Unfortunately, the sample includes numerous missing values as a result of the reluctance of some individuals to answer some of the survey questions. The variable suffering from the most missing values is the income variable. In order to deal with this problem of missing values for various variables, any observation that had one or more missing values was excluded from the data. The data set originally included 60,789 observations. A total of 12,899 observations were excluded due to missing values for one or more variables. The remaining 47,890 observations were used for estimation.

In addition to missing values, there are also other sampling errors in the data. For example, the respondent may have given incorrect responses due to recall error or misunderstanding of the interviewer's question. (The interviews were conducted by telephone.) Travel distances and mileage rates were not used; instead, the actual cost of travel recalled by respondents was used. This may actually be a more accurate measure because the perception of the travel cost recalled by the respondent is used and it may be the perception of costs that influence an individual in making consumption decisions. A shortcoming of the data is that time costs are not included. Although many valuations using the TCM do not include time costs (they assume time costs are zero) most modern valuations using the model do consider the opportunity cost of time.

Non-use value cannot be measured using the data in this survey because information is collected based on revealed behaviour only. In addition, there are some observations for which time was spent engaging in outdoor activities with no incursion of cost (perhaps an individual walked to the recreational area from their home and no expenditure was necessary for the trip). Neither of these issues can be considered shortcomings of the data, but rather features of the Travel Cost Model.

The variables used in the estimation are in Table 1. The dependent variable (*DAYS*) is the number of days that each individual spent engaged in outdoor recreational activities during 1996. This dependent variable takes on non-negative integer values, thus it is not distributed continuously; rather, it is a count variable and is censored (non-negative). The dependent variable was chosen to be the number of days spent engaged in outdoor recreation in the year, rather than the number of trips taken, because the decision unit in the model must be quantities of equal length. The data set reports the number of trips and the length, in days, of each trip; the total number of days per year per respondent is also included in the data set.

The variable *COSTPERD* is the variable that will serve as a proxy for the per-day price of engaging in outdoor recreational activities. This "price" variable has been chosen to reflect the choice of independent variable (*DAYS*). It is calculated by dividing the total annual cost of outdoor recreation incurred by the respondent by the number of days experienced by the respondent. The total annual cost (and each of its component costs) was provided in the data set.

The dependent variable, *DAYS*, and the "price" variable, *COSTPERD*, are both quantitative variables. The remaining explanatory variables are dummy variables constructed from the categorical variables available in the survey microdata. (These categorical variables are listed in Table 2.) These dummy variables include the income variables and other variables representing demographic demand shifters: sex, urban/rural residence, education, household size, age, marital status and income. *SEX* has been transformed into the dummy variable *MALE*, with female as the reference category. The urban/rural residence variable *UR* has been transformed into the dummy variable *RURAL*; the reference category is urban. *EDUC* has been transformed into the following dummy variables, the reference category being elementary education: high school education (*EDUCHS*), college education (*EDUCCOL*) and university education (*EDUCUNI*). *HHS* has been transformed into the following dummy variables, the reference category being a household size of one: two in household (*HHS2*), three in household (*HHS3*), four in household (*HHS5*), and five or more in household (*HHS5PLUS*). *AGE* has been transformed into the following dummy variables for teens (*AGETEEN*), individuals in their 20s (*AGE20S*), individuals in their 30s (*AGE30S*), individuals in their 40s (*AGE40S*), individuals in their 50s (*AGE50S*), and individuals in their 60s (*AGE60S*); the reference category is 70 years or older. The reference category for *MARST* is single, with the following dummy variables: married or common law (*MARRCOMM*), widowed (*WIDOW*), divorced or separated (*DIVSEP*). Finally, the *INCOME* group of variables has zero income as the reference category and the following dummy variables: income less than ten thousand (*INLT10*), income between ten thousand and thirty thousand (*IN10TO30*), income between thirty thousand and fifty thousand (*IN30TO50*), and income greater than fifty thousand (*INGT50*).

The "zero income" category of income may include individuals from any age group, any household size, of any educational background, and of either gender. For example, individuals with no income during the period of the survey may have been parents staying home with young children, teenagers and post-secondary students who have no part-time jobs and individuals who have, for one reason or another, decided to take time away from paid employment and have no other source of income. Sources of funds for these individuals may include support from a family member or a friend or savings, or perhaps they are dissaving.

Unfortunately, the data set did not include a wealth variable or a measure of family income. Such variables would have been useful in helping to explain decisions made by individuals with no personal income. In order to discover the importance of the omission of these variables the estimations were also run with the zero-income observations omitted.

Zones of origin (residence) are not included in the estimation because the individual Travel Cost Model is being used, in which data on each individual respondent are collected. Regardless, the data in the survey do not provide the information required to do a use a zonal TCM.

The descriptive statistics for the model variables are shown in Table 3[4]. The mean and standard deviation of the *DAYS* variable are 7.48 and 17.44, respectively. The standard deviation for *DAYS* is relatively large and demonstrates the wide variation in the number of days of outdoor activity experienced among the respondents. The mean and standard deviation of *COSTPERD* are $50.51 and 176.35. The standard deviation for *COSTPERD* is large and demonstrates the wide variation in the data.

The remainder of the variables in Table 3 are dummy variables. The *MALE* variable has a mean of 0.47, indicating that 47 per cent of the sample are male and 53 per cent are female. The *RURAL* variable has a mean of 0.26, indicating that 26 per cent of the sample live in a rural area and 74 per cent live in an urban area. The education group

[4] These means are all calculated based on the 47,890 observations that are used in the estimation. They reflect the observations for which none of the variables are missing.

of variables indicate that the means for *EDUCHS*, *EDUCCOL* and *EDUCUNI* are 0.39, 0.35 and 0.14, respectively. Thus, 12 per cent of the sample have an elementary school education, 39 per cent have a high school education, 35 per cent have a college education and 14% have a university education. For the household size group of variables, we see that *HHS2*, *HHS3*, *HHS4* and *HHS5PLUS* have respective means of 0.30, 0.20, 0.24 and 0.16. Thus, 10 per cent of the sample live alone, 30 per cent live in a household of two, 20 per cent live in a household of three, 24 per cent live in a household of four and 16 per cent live in a household of five or more.

For the age group of variables the means are 0.10 (*AGETEEN*), 0.17 (*AGE20S*), 0.23 (*AGE30S*), 0.22 (*AGE40S*), 0.07 (*AGE50S*) and 0.11 (*AGE60S*). Thus we know that 10 per cent of the sample are between the ages of 15 and 19, 17 per cent are in their 20s, 23 per cent are in their 30s, 22 per cent are in their 40s, 7 per cent are in their 50s, 11 per cent are in their 60s and 10 per cent are 70 years of age or older.

For the marital status group of variables the means for *MARRCOMM*, *WIDOW* and *DIVSEP* are 0.62, 0.06, and 0.06, respectively. Thus 26 per cent of the sample are single (never married), 62 per cent are either married or living in a common law relationship, six per cent are either widows or widowers and six percent are either divorced or separated. For the income group of variables the means are 0.24 (*INLT10*), 0.38 (*IN10TO30*), 0.17 (*IN30TO50*) and 0.09 (*INGT50*). Thus we know that 12 per cent of the sample have zero income, 24 per cent have income of less than ten thousand dollars, 38 per cent have income from ten thousand to just under thirty thousand, 17 per cent have income between thirty thousand and fifty thousand and nine per cent have income of fifty thousand dollars or more.

5. Results

The results of the diagnostic tests performed on the OLS model are shown in Table 4. They indicate that the errors are not normal, the model is mis-specified and there is heteroscedasticity. The Jarque-Bera normality test has a p-value of 0.0000, so we can reject the null hypothesis that the errors are normally distributed. Likewise the goodness-of-fit test for normality of the residuals has a χ^2 value of 80215.7777 with 34 degrees of freedom and a p-value of 0.0000. We conclude that the errors are not normally distributed. The Ramsey Reset specification tests shown in Table 4 also clearly indicate that the model is mis-specified, with all p-values equal to 0.0000. The results for the tests of heteroscedasticity are shown in Table 4. We can conclude from the p-value = 0.0000 results that the model suffers from heteroscedasticity.

The results of the diagnostic tests suggest that the linear model does not have normally distributed errors and that it suffers from specification error and heteroscedasticity. This leads us to believe that the linear model is not appropriate for the study. The comparison of the parameter estimates below will help us to further consider the appropriateness of the linear model. These results are consistent with the view that a simple linear model estimated using OLS is not appropriate. The result that the errors are not normal suggests that some other model, such as the Tobit or Poisson model, would be more appropriate. Wooldridge (2003) suggests that the RESET test primarily detects errors in functional form; since the functional form of the two models is so different, this result also supports the hypothesis that the linear model is inappropriate.

Because heteroscedasticity was found, the model was corrected for this problem using White's heteroscedasticity-consistent covariance matrix estimator. The standard errors that were reported have been corrected for this problem. Without correction for the heteroscedasticity the t-statistics may not have been reliable.

The estimation of the linear model results in an R^2 of 0.0307 and an adjusted R^2 of 0.0302. The F-statistic is 442.287 in the analysis of variance (from zero). The R^2 value indicates the proportion of variance in the dependent variable that is explained by the amount of variance in the explanatory variables. The values of R^2 for the linear model is small, telling us that there is a great deal that is not explained by the model. For example, there may be other variables, not included in the model, that would help to explain the number of days of outdoor recreation demanded and thereby increase the value of R^2. The F-statistic is for a test of the null hypothesis that all the slope coefficients are jointly zero. Given the large value of the F-statistic and its p-value of 0.0000 we see that we can reject the null hypothesis and conclude that the slope coefficients are not jointly zero and that our slope estimates are good. The p-value of 0.0000 indicates that there is approximately a zero probability that all the slope coefficients are jointly zero.

The estimation of the Poisson model results in an $R^2(d)$[5] of 0.07483. The two measures of R^2 for the linear model and the Poisson model are not comparable. For the Poisson model we can also calculate the likelihood ratio test statistic to test the null hypothesis that all the slope coefficients are zero. The log of the likelihood function for the unrestricted model is –499244. The log of the likelihood function for the restricted model is –679796. The likelihood ratio statistic is

$$LR_{model} = -2[\ln L_r - \ln L_u] = -2[-679796 - (-499244)] = 361{,}104$$

[5] This measure is described in Greene (1993).

and is asymptotically distributed as a χ^2 random variable with degrees of freedom equal to the number of parameter restrictions in the restricted model. Given the value of the test statistic we conclude that we have enough evidence to reject the null hypothesis that the slope coefficients are equal to zero.

The parameter estimates for both models are summarized in Table 5. Numbers in brackets below the coefficient estimates are standard errors. Coefficients that are significant at the 1% level are indicated by "**" and coefficients that are significant at the 5% level are by indicated "*".

The estimated constants for both models are relatively small and significant at the 1% level, but the constant is negative for the linear model and positive for the Poisson model. The negative constant in the OLS estimation suggests that there is a problem with the model because when it is considered with the negative coefficient on the price variable the result is a negative number of days of outdoor recreation for the reference individual (the individual for whom all of the dummy variables are equal to zero). The negative constant is problematic in and of itself, since it implies negative demand when the price is zero, which does not make sense.

The coefficient of the "price" variable, *COSTPERD*, is negative and significant at the 1% level for both estimations. This result is consistent with our prior expectations and implies a downward-sloping demand curve. The elasticity value indicated by the coefficient of the *COSTPERD* variable is –0.0202 in the linear model and –0.0193 in the Poisson model. They can both be considered to be approximately equal to –0.02. This indicates that for every 1% increase in the cost per day of recreation, the number of days of recreation decreases by 0.02%. The demand for recreation in outdoor areas is price inelastic.

In the linear model the coefficient of the *COSTPERD* variable tells us that, ceteris paribus, for each dollar decrease in the price of a day of outdoor recreational activity, the number of days spent engaged in outdoor recreational activity increases by approximately .002 days. Or, put differently, for every $100 decrease in the price of outdoor recreational activity, the number of days spent engaged in outdoor recreational activity increases by approximately 0.2 days. This gives us some insight into the relative importance of the cost of engaging in outdoor activity in determining how many days are spent in the activity. "Price" is not a major factor in determining the demand for recreation so we must look to the other explanatory variables to explain changes in demand.

Given the reference category for each dummy variable in the model, we deem the reference individual to be a woman who lives in an urban area, has 0-8 years of schooling, lives alone, is 70 years of age or older, has never been married and has no income. Using the average cost per day of outdoor recreation ($50.51) we calculate, when using the linear model, that this reference person consumes –1.98 days of outdoor recreation per year. The calculations for the effect of each of the dummy variables relative to the reference person are shown in Table 6.

When we look at the coefficient of the *MALE* dummy variable (female is the reference category) we see that, ceteris paribus, males spend more time in the outdoors than females. Compared to the reference category a male with otherwise identical characteristics would spend –1.78 days engaged in outdoor recreation. Using the OLS estimation, males can be expected to spend an extra 0.2 days per year engaged in outdoor recreational activities, compared with a female with otherwise identical characteristics. Generally, men are expected to spend more time on recreational activities because of fewer time constraints than women and perhaps a higher inclination to engage in an active lifestyle. However, the coefficient is not significant at the 5% level. The variable has no effect at the 5% level.

The coefficient of the *RURAL* dummy variable (reference category is urban) is significant at the 1% level for the OLS estimation. It shows that rural inhabitants consume marginally more days of outdoor recreational activity than an otherwise identical individual, consuming –1.37 days, or 0.61 days more than the reference individual.

When we look at the group of dummy variables for level of education (0-8 years of schooling is the reference category) we very clearly see that as the level of education increases, the number of days spent engaged in active outdoor recreation increases. Relative to the reference individual, a person who has a high school education would consume 1.72 days more, or –0.26 days. An individual with a college education, but otherwise identical to the reference person would consume 3.04 more days, or 1.06 days. An individual with a university education, but otherwise identical to the reference individual would consume 4.94 more days, or 2.96 days. For the OLS estimation, all three education dummy coefficients are significant at the 1% level of significance. The number of days of outdoor active recreation consumed increases for every incremental increase in education level.

For the coefficients describing household size (household size of one is the reference category) there is not a clear pattern. For the OLS estimation, only two of the four dummy coefficients are significant and only at the 5% level. If the reference individual was in a household size of two, 0.72 more days of outdoor recreation would be consumed, or –1.26 days in total. Increasing household size to three increases days consumed by 0.33 days to –1.65 days. Increasing household size to four increases days consumed by 0.9 to –1.08. Increasing household size to five increases days consumed by 0.67 to –1.31.

The coefficients on the age variables are relatively large and show a distinct decrease in consumption of days of outdoor activity as one moves from one age category to the next. The reference category is the oldest age category, 70 years and older. For the OLS estimation the coefficients for every age category dummy variable are significant at the 1% level. A person who is in the 15 to19 age group, but is otherwise identical to the reference individual, will consume 7.49 more days of outdoor recreation, consuming 5.51 days. A person in their 20s will consume 4.96 more days than the otherwise-identical reference individual, consuming 2.98 days in total. A person in their 30s will consume 4.68 more days, or 2.7 days. A person in their 40s will consume 3.55 days more than the reference individual, or 1.57 days. A person in their 50s will consume 2.38 more days, or 0.4 days. A person in their 60s, will consume 1.76 more days than the otherwise-identical reference individual, consuming –0.22 days of outdoor recreation.

The effect of marital status, as for household size, is not as clear. The reference category is a single (never married) person. In the OLS estimation, the only variable that has a significant coefficient is *MARRCOMM*, and its coefficient is significant at the 1%

level. The OLS regression ranks (statistically insignificantly) marital status in the following order of increasing consumption: single, widow(er), divorced or separated, married or living in a common-law relationship. Relative to the reference person, a person who is married or living in a common-law relationship will consumer 0.27 more days of outdoor recreation per year, or –1.71 days. A widow, with otherwise identical characteristics as the reference person, will consume 0.32 more days, or –1.66 days. If the person is divorced or separated, but otherwise identical, they will consume 0.69 more days, or –1.29 days per year.

The income coefficients are, as for the education and age coefficients, significant at the 1% level for all four dummy variables in the OLS estimation. The reference category is no income. We see clearly that as income increases from one category to the next, the number of days of outdoor recreational activity consumed increases. If the reference person had an income of less than $10k she would consume 1.16 more days of outdoor recreation, or –0.82 days. If the reference person had an income of $10-30k, she would consume 1.44 more days, or –0.54. If she had an income of $30-50k she would consume 2.94 more days, or 0.96 days. If the reference person had an income greater than $50k, she would consume 4.73 more days, or 2.75 days of outdoor recreation in the year. These are the results that we intuitively expect.

Overall, we see that in the OLS estimation some of the coefficients are not significant, some are significant at the 95% level, and some are significant at the 1% level. The strongest conclusions we can draw are for the education, age, and income coefficients. The OLS estimation has coefficients significant at the 1% level for each coefficient in every category for these three characteristics. We can clearly state that consumption of active outdoor recreational activity increases as education increases, as age decreases, and as income increases.

We also notice in Table 6 that in the OLS estimation the quantity of outdoor recreation consumed in some cases is negative. This is clearly a problem as mentioned above when discussing the negative constant in the estimation. This provides us with more evidence to consider the linear model inappropriate.

For the Poisson estimation we notice that every coefficient is significant at the 1% level. To interpret the value of each of the coefficients we note that when a dummy variable is turned on (as its value changes from 0 to 1), the value of λ_i in the Poisson model will also change. In the Poisson model, using the average cost per day of outdoor recreation ($50.51) we calculate that this reference person consumes 1.08 days of outdoor recreation per year. The calculations for the effect of each of the dummy variables relative to the reference person are shown in Table 7.

In the Poisson estimation, if the person were male, but otherwise identical to the reference person, 0.05 more days would be consumed, or 1.13 days in total. If the reference person were a rural dweller, all other characteristics held equal, she would consume 0.10 more days of outdoor recreation for the year, or 1.18 days.

If the reference individual had, instead of 0-8 years of schooling, a high school education, she would demand 0.70 more days, or 1.78 days of outdoor recreation. Alternatively, if she had a college education, she would consume 1.04 more days, or 2.12 days. If she had a university education, she would consume 1.51 more days, or 2.59 days.

The reference individual lives in a household of one. If, instead, she lived in a household of two, she would consume 0.15 more days, or 1.23 days of outdoor recreation.

If she lived in a household of three, rather than alone, her days of outdoor recreation would increase by 0.08 days to 1.16 days. If she lived in a household of four her consumption would increase by 0.16 days to 1.24 days. If she lived in a household of five or more her consumption of outdoor recreation in the year would increase by 0.14 days to 1.22 days.

It is evident that age strongly influences the decision to participate in activities in outdoor areas. If the reference individual were aged 15-19 she would consume 2.9 more days, or 3.98 days. If she were in her 20s, she would consume 1.81 more days, or 2.89 days. If she were in her 30s she would consume 1.37 more days, or 2.45 days. If she were in her 50s she would consume 1.01 more days, or 2.10 days. If she were in her 60s she would consume 0.78 more days, or 1.86 days of outdoor recreation.

If the reference individual were married or living in a common-law relationship, rather than being single, she would consume 0.12 more days of outdoor recreation, or 1.20 days in the year. Alternatively, if she were a widow she would consume 0.20 more days, or 0.89 days. If she were divorced or separated, she would consume 0.9 more days, or 1.17 days of outdoor recreation in the year.

The level of income strongly affects the number of days a person engages in outdoor recreational activity. If the reference person's income were less than $10k, rather, than zero, she would increase her outdoor recreation by 0.18 days per year to 1.26 days. If, instead, her income were $10-30k the quantity would increase by 0.26 days to 1.34 days. If her income were $30-50k the quantity would increase by 0.55 days to 1.63 days. If her income were more than $50k, rather than zero, she would engage in 0.86 more days of outdoor recreation, or 194 days of outdoor recreation in the year.

The general results of the Poisson model are similar to the OLS results in that we can see that education, age and income strongly influence the amount of outdoor recreation that an individual engages in. More outdoor recreation is demanded as a person attains more education, as they earn more income, and as their age decreases. However, the results differ in how many days of outdoor recreation are predicted as the dummy variables are turned on and off. The OLS results have many negative values for the estimated value of the dependent variable (*DAYS*), while all of the Poisson results for the dependent variable are positive. This is an importance difference, because the number of days should be non-negative.

Both models were also estimated excluding those individuals with zero income. There was a small effect on the coefficient of the *COSTPERD* variable of excluding these individuals. The coefficients of the *COSTPERD* variable in the estimations omitting individuals with no income were -.0029931 and -.0028669, respectively, for the linear and Poisson models, as compared to -.0029906 and -.0028628, respectively. However, after rounding to three decimal places the difference is no longer apparent.

From the results of the estimation of the demand for outdoor recreation in natural areas we can calculate the use value of natural areas. A different formula is used for the different models estimated. To calculate the aggregate use value using the OLS model we use formula (10):

$$CS_L = DAYS^2 / -2\beta_c ,$$

where β_c = -.0029906. The predicted value of *DAYS* for each individual in the sample was calculated. Each of these values was then squared and then the squared values were summed, arriving at a value of 3,129,532 days. This sum of squared predicted numbers of days was then divided by $-2\beta_{c,}$, where β_c has a value of -.003. This results in a use value for the sample of $523M. The population size (15 years and older) in 1996 was 23,582,516 so we factor the sample value up by 492.43 to arrive at a total use value for the population of $258B.[6]

The estimate for the population is not accurate because ideally the sample survey weights should be used to produce an aggregate measure, but they were not available.[7] As an alternative, we have assumed that each individual in the sample represents the same number of individuals in the population. It is difficult to say whether this will result in an overestimate or an underestimate of the true population consumer surplus because we do not know whether this procedure gives a higher-than-actual weight to heavy consumers or to light consumers of outdoor recreation.

If, instead of using the predicted values of days for calculation of consumer surplus, we used the value of days for the average individual, the value of consumer surplus would differ from the value above. Using the average number of days (7.48 days), the consumer surplus per person for the year is $9354. For the sample the consumer surplus is $448M and for the population the consumer surplus is $221B. These consumer surplus values calculated using the number of days consumed by the average individual are summarized in Table 8.

To calculate the aggregate use value using the Poisson model we begin by using the consumer surplus formula for the semi-log functional form in equation (11):

$$CS_P = DAYS / -\alpha_c.$$

[6] This inflation factor is calculated by dividing the population size by the sample size.
[7] The sample survey weights do exist, but were not retrieved with the data used for this paper.

We calculate the per-person-per-day value as *1/-α_c*, where α_c is the coefficient on the travel cost variable *COSTPERD.* In the Poisson estimation the coefficient on the price variable is -.0028628, resulting in a per-person-per-day use value of $349.31. To calculate the use value of the sample (47,890) we multiply by the predicted number of days spent in the outdoors by the respondents in 1996 (358,426) and get the result: $125M. To estimate the use value for the entire population (23,582,516) we factor up by 492.43 and get the result $62B. (As for the consumer surplus calculation for the linear model, the estimate of the value for the entire population is not accurate due to the lack of the sample survey weights.)

If, instead of using the predicted values of days for calculation of consumer surplus, we used the value of days for the average actual individual, the value of consumer surplus for the sample would be $125,128,299 (compared to $125,201,786 using the predicted values) and the value of consumer surplus for the population would be $61,616,912,040 (compared to $61,653,115,510 using the predicted values). When taking these values to the nearest million and nearest billion (see Table 8) there is no difference in the consumer surplus calculated when using predicted values and average actual values for the Poisson model.

It is immediately apparent that the use values calculated from the OLS estimation are much larger than the use values calculated using the Poisson estimation. This was expected, given that the price coefficients are similar in value, but the formula for the consumer surplus calculations are quite different. Previous researchers also found that OLS estimates result in larger consumer surplus values than do Poisson estimates. For example, Zeimer, Musser and Hill (1980) found that in their study of warm-water fishing in Georgia their value for consumer surplus in the linear model was $1,004,321 compared to $335,926 for the semi-log model. Smith and Desvousges (1985) found that for the OLS estimates the values per unit of use ranged from $39.00 to $400.00 while the ML estimates ranged from $3.58 to $105.26 (both in 1977 dollars). In a study of North Park Lake in the area of the Monongahela River in Pennsylvania, Smith (1988) calculated that, using the OLS method for a linear specification, the per-trip value was $1.67, versus a value of $0.38 using the Poisson method for a semi-log specification (both in 1981 dollars).[8] All of these studies found that the consumer surplus calculated using an OLS model was much greater than the consumer surplus calculated using a Poisson model. This is the same result as was found in our study.

6. Conclusion

The objective of this paper was to use the Travel Cost Model to estimate the demand for and use value of natural outdoor areas for recreation in Canada. Two alternative models of demand were estimated: a linear model and a semi-log Poisson model. Quantity demanded was measured as the number of days spent in a year engaged in outdoor recreational activities (17 specific activities that take place in natural outdoor areas were included). The independent variables included a "price" variable measuring the price per day of the activity and dummy variables for gender, rural/urban residence, level of education, household size, age, marital status and income. Both the linear and the semi-log Poisson estimations produced a negative (although small) coefficient for price; both were statistically significant at the 1% level. Level of education, age, and income had the largest effect on quantity of outdoor recreational activities consumed.

[8] More recent studies do not do a comparison of OLS and Poisson.

The OLS estimation produced a negative constant, thereby producing several negative values for the dependent variable. This caused us to doubt the appropriateness of the OLS model for the data used in this study. In addition, our diagnostics confirmed our doubts about the linear model. We therefore conclude that for the estimation of a demand function for natural outdoor recreational areas the use of a Poisson model is more appropriate.

From the demand estimations we calculated consumer surplus values. For the OLS estimation we calculated an aggregate value of $258B (1996 dollars). For the Poisson model we calculated an aggregate value of $62B (1996 dollars). When compared to the literature our study confirms the result that other studies have found: using the OLS method to estimate the linear model results in higher use values than does the ML method to estimate the semi-log Poisson model.

It is difficult to say whether the consumer surplus value we have calculated is relatively large or small because there are no other studies that valued the same area. However, we can compare it to the amount that was spent by the sample of 47,890 individuals. Our sample spent a total of $14,425,905 on travel costs of outdoor recreation in 1996. For the entire population this would aggregate to approximately $7B. This value is much smaller than the consumer surplus estimate of $62B.

Another implication of the results is that population aging is likely to decrease the demand for recreation in natural areas, because we found that as individuals get older they participate less in outdoor recreational activities. Indeed, recently the *Ottawa Citizen* (9 December, 2006) reported that the number of visits to national parks has declined. The article further suggests that this decline is due to two other factors in addition to population aging: the lack of participation of young people (due to the popularity of electronic entertainment), and the low rate of participation by immigrants. Unfortunately, since our paper did not use panel data nor did it include an indicator for immigrant status, we cannot see the effects that the emergence of electronic entertainment has had on participation in outdoor recreation in natural areas.

The results of this paper do have policy implications because it is important to decide which of the two techniques of estimation should be used for the valuation of natural areas using the TCM. The technique that gives us the most reliable results should be chosen. Given the results we obtained we concluded that the Poisson model was the better model so this is the valuation method one should use. This conclusion is the same conclusion that other researchers have arrived at.

This study has several strengths and weaknesses. The data used in this study were generally of a high quality and were from a large sample. However, the data have some limitations: missing observations, potential for recall error of travel costs, and no data from which we were able to calculate time costs. The inability to include travel time in our model caused the consumer surpluses to be lower than they would be if a value for time were included in the travel cost variable.

Future research on this topic should study the Negative Binomial model for count data in order to compare results to those generated in this paper. The Negative Binomial model is a count data model similar to the Poisson model, but the mean of the distribution is not restricted to be equal to the standard deviation of the distribution. Another promising model for estimating demand for outdoor recreational activity is the Random Utility Model. Parsons (2003) and Haab and McConnell (2002) describe the RUM model as one that is useful for valuing site access and changes in site quality. It enables a researcher to value quality changes in the site(s). The RUM model considers a person's choice of a site for a trip. The dependent variable is the site chosen, rather than the quantity demanded. This choice of a site is based on the "price" of the site (determined by travel costs) and the site's characteristics. The concept of the model is based on the theory that an individual chooses the site that gives him/her the highest utility.

Values of the recreational use of natural areas are often taken for granted. They are not missed until they are no longer available, after the area has been given up for an alternate, ostensibly more productive, use. It is important to bring attention to the demand for and valuation of natural outdoor areas. It is this clear and concise documentation of the demand for and valuation of natural areas' value in recreational use that enables us to compare it against the valuation of alternate uses of the area.

BIBLIOGRAPHY

Adamowicz, Wiktor L., Jerald J. Fletcher and Theodore Graham-Tomasi (1989) "Functional Form and the Statistical Properties of Welfare Measures" *American Journal of Agricultural Economics* Vol. 71, No. 2, pp. 414-421.

Bergstrom, John C., John R. Stoll, John P. Titre and Vernon L. Wright (1990) "Economic Value of Wetlands-Based Recreation", *Ecological Economics, 2, p.129-147* (Amsterdam: Elsevier Science Publishers B.V.).

Boardman, Anthony E., David H. Greenberg, Aidan R. Vining, and David L. Weimer (2001) *Cost-Benefit Analysis, Concepts and Practice* (Upper Saddle River: Prentice Hall, Inc.).

Bowker, J.M., Donald B.K. English and Jason A. Donovan (1996) "Toward a Value for Guided Rafting on Southern Rivers" *Journal of Agricultural and Applied Economics* 28(2).

Boyle, Kevin J. (2003) "Introduction to Revealed Preference Methods" *A Primer on Nonmarket Valuation* edited by Patricia A. Champ, Kevin J. Boyle and Thomas C. Brown (Dordrecht, Netherlands: Kluwer Academic Press).

Brown, Gardner Jr. and Robert Mendelsohn (1984) "The Hedonic Travel Cost Model" *Review of Economics and Statistics* 66, no. 3.

Burt, Oscar R. and Durward Brewer (1971) "Estimation of Net Social Benefits from Outdoor Recreation" in *Econometrica* Vol. 39, No. 5, pp. 813-827.

Choe, KyeongAe, Dale Whittington and Donald T. Lauria (1996) "The Economic Benefits of Surface Water Quality Improvements in Developing Countries: A Case Study of Davao, Philippines" *Land Economics* Vol. 72, No. 4, pp. 519-537.

Clawson, Marion (1959) *Methods of Measuring the Demand for and Value of Outdoor Recreation.* Washington: Resources for the Future.

Davis, Robert K (1963) "The Value of Outdoor Recreation: An Economic Study of the Maine Woods" (Ph.D. dissertation, Harvard University).

Englin, Jeffrey, Peter Boxall and David Watson (1998) "Modeling Recreation Demand in a Poisson System of Equations: An Analysis of the Impact of International Exchange Rates" *American Journal of Agricultural Economics* Vol. 80, No. 2, pp. 255-263.

Englin, J. and J.S. Shonkwiler (1995) "Estimating Social Welfare Using Count Data Models: An Application to Long-Run Recreation Demand Under Conditions of Endogenous Stratification and Truncation" *Review of Economics and Statistics* 77 pp. 104-112.

Freeman, A. Myrick III (2003a) *The Measurement of Environmental and Resource Values, Theory and Methods, 2nd edition* (Washington D.C.: Resources for the Future).

Freeman, A. Myrick III (2003b) in *A Primer on Nonmarket Valuation*, Chapter 8, edited by Patricia A. Champ, Kevin J. Boyle and Thomas C. Brown (Dordrecht, Netherlands: Kluwer Academic Press).

Greene, W. H. (1997) *Econometric Analysis* (3rd edition) Upper Saddle River, NJ: Prentice Hall).

Haab, Timothy C. and Kenneth E. McConnell (2002) *Valuing Environmental and Natural Resources: The Econometrics of Non-Market Valuation* (Northampton, MA, USA: Edward Elger).

Hanley, Nick and Clive L. Spash (1993) *Cost-Benefit Analysis and the Environment* (Aldershot: Edward Elgar Publishing Limited).

Hellerstein, D. (1991) "Using Count Data Models in Travel Cost Analysis with Aggregate Data" *American Journal of Agricultural Economics* Vol. 73, No. 3, pp. 860-866.

Hellerstein, Daniel and Robert Mendelsohn (1993) "A Theoretical Foundation for Count Data Models" *American Journal of Agricultural Economics*, Vol. 75, No. 3 (Aug., 1993), pp.604-611.

Hotelling, Harold (1949) "Letter to National Park Service" *An Economic Study of the Monetary Evaluation of Recreation in the National Parks*, U.S Department of the Interior, National Park Service and Recreational Planning Division.

Maddala, G. S. (2002) *Introduction to Econometrics,* Third Edition (Toronto: John Wiley & Sons, Ltd).

Mäler, Karl-Göran (1974) *Environmental Economics: A Theoretical Inquiry*. Baltimore: John Hopkins University.

McKean, John, R. (2005) "Willingness to Pay for Non Angler Recreation at the Lower Snake River Reservoirs" *Journal of Leisure Research*, Second Quarter 2005.

McKean, J.R., D.M. Johnson and R.G. Walsh (1995) "Valuing Time in Travel Cost Demand Analysis: An Empirical Investigation" *Land Economics* 71(1), February 1995, pp. 96-105.

McKean, John R., Richard G. Walsh and Donn M. Johnson (1996) "Closely Related Good Prices in the Travel Cost Model" *American Journal of Agricultural Economics* Vol. 78, No. 3, pp. 640-646.

National Park Service (1995) *Economic Impacts of Protecting Rivers, Trails and Greenway Corridors – A Resource Book* www.nps.gov/pwro/rtcaa/econ.

Ottawa Citizen. 9 December 2006.

Parsons, George R. (2003) "The Travel Cost Model" *A Primer on Nonmarket Valuation* edited by Patricia A. Champ, Kevin J. Boyle and Thomas C. Brown (Dordrecht, Netherlands: Kluwer Academic Press).

Randall, Alan (1994) "A Difficulty with the Travel Cost Method" *Land Economics* 70(1): pp. 88-96.

Seller, Christine, John R. Stoll and Jean-Paul Chavas (1985) "Validation of Empirical Measures of Welfare Change: A Comparison of Non-market Techniques" *Land Economics* Vol. 61, No. 2, pp. 156-175.

Smith, V. Kerry (1988) "Selection and Recreation Demand" *American Journal of Agricultural Economics* Vol. 70, No. 1, pp. 29-36.

Smith, V. Kerry (1990) "Can We Measure the Economic Value of Environmental Amenities" *Southern Economic Journal* Vol. 56, No. 4, pp.865-878.

Smith, V. Kerry and William H. Desvousges (1985) "The Generalized Travel Cost Model and Water Quality Benefits: A Reconsideration" *Southern Economic Journal* Vol. 52, No. 2, pp. 371-381.

Smith, V. Kerry and William H. Desvousges (1986) *Measuring Water Quality Benefits* (Boston: Kluwer Nijhoff Publishing).

Smith, V. Kerry and Raymond J. Kopp (1980) "The Spatial Limits of the Travel Cost Recreational Demand Model" *Land Economics* Vol. 56, No. 1, pp. 64-72.

Stoll, John R., John B. Loomis and John C. Bergstrom (1987) "A Framework for Identifying Economic Benefits and Beneficiaries of Outdoor Recreation" in *Policy Studies Review, Vol. 7, No. 2, p. 443-452.*

Van der Straaten (2000) "The Economic Value of Nature" in H. Briassoulis and J. van der Straaten (eds.) *Tourism and the Environment, Second Revised Edition* (Dordrecht: Kluwers Academic Publishers).

Wood, Samuel E. and Andrew H. Trice (1958) "Measurement of Recreation Benefits" *Land Economics,* February, pp. 195-207.

Wooldridge, Jeffrey M. (2003) *Introductory Econometrics: A Modern Approach*, 2nd ed. (Mason, Ohio: South-Western).

Yen, S.T. and W.L. Adamowicz (1993) "Statistical Properties of Welfare Measure from Count-Data Models of Recreation Demand" *Review of Agricultural Economics.*

Ziemer, Rod F., Wesley N. Musser and R. Carter Hill (1980) "Recreation Demand Equations: Functional Form and Consumer Surplus" *American Journal of Agricultural Economics* Vol. 62, No. 1.

Table 1. Variables used in estimation

Variable	Type of variable	Description of variable
DAYS	quantitative (dependent)	number of days spent engaged in outdoor activity
COSTPERD	quantitative	cost per day engaged in outdoor activity
MALE	dummy reference category: female	respondent is male
RURAL	dummy reference category: urban	respondent lives in a rural area
EDUCHS *EDUCCOL* *EDUCUNI*	dummy reference category: 0-8 years schooling	respondent has high-school education respondent has college education respondent has university education
HHS2 *HHS3* *HHS4* *HHS5PLUS*	dummy reference category: household size of 1	respondent's household size is 2 respondent's household size is 3 respondent's household size is 4 respondent's household size is 5 or more
AGETEEN *AGE20s* *AGE30s* *AGE40s* *AGE50s* *AGE60s*	dummy reference category: 70 years and over	respondent is a teenager (aged 15 or over) respondent is in their 20s respondent is in their 30s respondent is in their 40s respondent is in their 50s respondent is in their 60s
MARRCOMM *WIDOW* *DIVSEP*	dummy reference category: single	respondent is married or common law respondent is a widow or widower respondent is divorced or separated
INLT10 *IN10TO30* *IN30TO50* *INGT50*	dummy reference category: no income	respondent's income is less than $10k respondent's income is $10-30k respondent's income is $30-50k respondent's income is over $50k

Table 2. Values for categorical variables

Variable	Categories	Category values
UR	rural	1
	urban	2
HHS	household size is 1	1
	household size is 2	2
	household size is 3	3
	household size is 4	4
	household size is 5 or more	5
AGE	15-19	1 and 2
	20s	3 and 4
	30s	5 and 6
	40s	7 and 8
	50s	9 and 10
	60s	11 and 12
SEX	male	1
	female	2
MARST	married or common law	1
	single	2
	widow or widower	3
	divorced or separated	4
EDUC	elementary	0
	high school	2 and 3
	college	4 and 5
	university	6
INCOME	0 income	1
	less than $10k	2 and 3
	$10k-30k	4 and 5
	$30k-50k	6 and 7
	over $50k	8

Table 3. Descriptive statistics

Name	N	Mean	Standard Deviation
DAYS	47890	7.48	17.44
COSTPERD	47890	50.51	176.35
MALE	47890	0.47	0.50
RURAL	47890	0.26	0.44
EDUCHS	47890	0.39	0.49
EDUCCOL	47890	0.35	0.48
EDUCUNI	47890	0.14	0.35
HHS2	47890	0.30	0.46
HHS3	47890	0.20	0.40
HHS4	47890	0.24	0.43
HHS5PLUS	47890	0.16	0.36
AGETEEN	47890	0.10	0.30
AGE20S	47890	0.17	0.37
AGE30S	47890	0.23	0.42
AGE40S	47890	0.22	0.41
AGE50S	47890	0.07	0.25
AGE60S	47890	0.11	0.32
MARRCOMM	47890	0.62	0.48
WIDOW	47890	0.06	0.23
DIVSEP	47890	0.06	0.23
INLT10	47890	0.24	0.43
IN10TO30	47890	0.38	0.49
IN30TO50	47890	0.17	0.38
INGT50	47890	0.09	0.29

Table 4. Results of OLS diagnostic tests

			d.f.	p-value
Jarque-Bera normality test				
χ^2 = *****			2	0.0000
Goodness-of-fit test for normality of residuals (60 groups)				
χ^2 = 80215.7777			34	0.0000
Ramsey reset specification tests using powers of $\hat{y}$				
RESET(2) = 57.268			1, ****	0.0000
RESET(3) = 28.723			2, ****	0.0000
RESET(4) = 25.926			3, ****	0.0000
Tests for heteroscedasticity				
		χ^2 test statistic		
e^2 on $\hat{y}$		87.018	1	0.00000
e^2 on $\hat{y}^2$		81.664	1	0.00000
e^2 on log($\hat{y}^{2)}$		69.258	1	0.00000
e^2 on lag(e^2)	ARCH test	15.658	1	0.00008
log(e^2) on x	Harvey test	10066.985	23	0.00000
|e| on x |	Glejser test	3391.303	23	0.00000

Table 5. Coefficient estimates for OLS and Poisson

Variable	OLS coefficient estimate	Poisson coefficient estimate
Constant	-1.832**	0.231**
	(0.4305)	(0.0150)
COSTPERD	-0.003**	-0.003**
	(0.0011)	(0.0000)
MALE	0.203	0.043**
	(0.1676)	(0.0036)
RURAL	0.616**	0.087**
	(0.1887)	(0.0039)
EDUCHS	1.723**	0.495**
	(0.2048)	(0.0085)
EDUCCOL	3.048**	0.673**
	(0.2263)	(0.0086)
EDUCUNI	4.942**	0.871**
	(0.3413)	(0.0092)
HHS2	0.719*	0.128**
	(0.3175)	(0.0085)
HHS3	0.335	0.072**
	(0.3486)	(0.0086)
HHS4	0.899*	0.139**
	(0.3558)	(0.0086)
HHS5PLUS	0.674	0.120**
	(0.3759)	(0.0090)
AGETEEN	7.498**	1.303**
	(0.4563)	(0.0130)
AGE20S	4.966**	0.981**
	(0.3234)	(0.0112)
AGE30S	4.686**	0.943**
	(0.2963)	(0.0108)
AGE40S	3.553**	0.818**
	(0.2837)	(0.0108)
AGE 50S	2.385**	0.661**
	(0.3580)	(0.0121)
AGE60S	1.760**	0.542**
	(0.2638)	(0.0112)
MARRCOMM	0.937**	0.105**
	(0.2731)	(0.0057)
WIDOW	0.324	-0.200**
	(0.3614)	(0.0148)
DIVSEP	0.690	0.079**
	(0.4191)	(0.0090)
INLT10	1.164**	0.150**
	(0.2724)	(0.0065)
IN10TO30	1.446**	0.216**
	(0.2726)	(0.0067)
IN30TO50	2.945**	0.410**
	(0.3287)	(0.0074)
INGT50	4.73**	0.585**
	(0.4288)	(0.0081)

Numbers in brackets below the coefficient estimates are standard errors. Coefficients that are significant at the 1% level are indicated by "**" and coefficients that are significant at the 5% level are by indicated "*".

Table 6. Comparison of effect of dummy variables for OLS model

Reference person is a woman who lives in an urban area, has 0-8 years of schooling, lives alone, is 70 years of age or older, has never been married and has no income. The reference person consumes –1.98 days.

Average cost per day of outdoor recreation is $50.51.

Dummy variable	*DAYS*	Difference from reference
MALE	-1.78	+ 0.20
RURAL	-1.37	+ 0.61
HS	-0.26	+ 1.72
COL	1.06	+ 3.04
UNI	2.96	+ 4.94
HHS2	-1.26	+ 0.72
HHS3	-1.65	+ 0.33
HHS4	-1.08	+ 0.90
HHS5PLUS	-1.31	+ 0.67
AGETEEN	5.51	+ 7.49
AGE20S	2.98	+ 4.96
AGE30S	2.70	+ 4.68
AGE40S	1.57	+ 3.55
AGE50S	0.40	+ 2.38
AGE60S	-0.22	+ 1.76
MARRCOMM	-1.71	+ 0.27
WIDOW	-1.66	+ 0.32
DIVSEP	-1.29	+ 0.69
INCLT10	-0.82	+ 1.16
IN10TO30	-0.54	+ 1.44
IN30TO50	0.96	+ 2.94
INGT50	2.75	+ 4.73

Table 7. Comparison of effect of dummy variables for Poisson model

Reference person is a woman who lives in an urban area, has 0-8 years of schooling, lives alone, is 70 years of age or older, has never been married and has no income. The reference person consumes 1.08 days.

Average cost per day of outdoor recreation is $50.51.

Dummy variable	*DAYS*	Difference from reference
MALE	1.13	+ 0.05
RURAL	1.18	+ 0.10
HS	1.78	+ 0.70
COL	2.12	+ 1.04
UNI	2.59	+ 1.51
HHS2	1.23	+ 0.15
HHS3	1.16	+ 0.08
HHS4	1.24	+ 0.16
HHS5PLUS	1.22	+ 0.14
AGETEEN	3.98	+ 2.90
AGE20S	2.89	+ 1.81
AGE30S	2.78	+ 1.70
AGE40S	2.45	+ 1.37
AGE50S	2.10	+ 1.01
AGE60S	1.86	+ 0.78
MARRCOMM	1.20	+ 0.12
WIDOW	0.89	- 0.20
DIVSEP	1.17	+ 0.09
INLT10	1.26	+ 0.18
IC10TO30	1.34	+ 0.26
IN30TO50	1.63	+ 0.55
INGT50	1.94	+ 0.86

Table 8. Consumer surplus values

(all values in 1996 dollars)	**linear, using predicted days consumed**	**semi-log, using predicted days consumed**	**linear, using average actual days consumed**	**semi-log, using average actual days consumed**
use value for sample	523M	125M	448M	125B
use value for population, 15 years and older	258B	62B	221B	62B