The Socio-Economic Marine Research Unit (SEMRU)
National University of Ireland, Galway

Working Paper Series

Working Paper 11-WP-SEMRU-05

Estimating recreation demand with on-site panel data: An application of a latent class truncated and endogenously stratified count data model.

Stephen Hynes and William Greene

SEMRU ***Working Paper Series***

For More Information on the SEMRU Working Paper Series
Email: stephen.hynes@nuigalway.ie, **Web:** www.nuigalway.ie/semru/

Estimating recreation demand with on-site panel data: An application of a latent class truncated and endogenously stratified count data model.

Stephen Hynes[1*] and William Greene[2]

1. Socio-Economic Marine Research Unit, J.E. Cairnes School of Business and Economics, National University of Ireland, Galway
2. Economics Department, Stern School of Business, New York University

Abstract

In this paper, we present an extension of Shaw's (1988) and Englin and Shonkwiler's (1995) count data travel cost models corrected for on-site sampling to a panel data setting. We develop a panel data negative binomial count data model that corrects for endogenous stratification and truncation. We also incorporate a latent class structure into our panel specification which assumes that the observations are drawn from a finite number of segments, where the distributions differ in the intercept and the coefficients of the explanatory variables. Results of this model are compared to some of the more common modelling approached in the literature. The chosen models are applied to revealed and contingent travel data obtained from a survey of visitors to a beach on the outskirts of Galway city in Ireland. The paper argues that count data panel models corrected for on-site sampling may still be inadequate and potentially misleading if the population of interest is heterogeneous with respect to the impact of the chosen explanatory variables.

This work was funded through the Beaufort Marine Research Award, which is carried out under the *Sea Change* Strategy and the Strategy for Science Technology and Innovation (2006-2013), with the support of the Marine Institute, funded under the Marine Research Sub-Programme of the National Development Plan 2007–2013.

* Stephen Hynes is the contact author (e-mail: stephen.hynes@nuigalway.ie), Office: +353 (0)91 493105, Fax: +353 (0)91 524130.

Introduction

The Travel Cost Method (TCM) of non-market valuation, based on the count nature of recreation trips, can only measure values associated with the current use of a recreational site. However an analyst, site manager or policy maker may be more interested in the value to the user of potential changes to the facilities of a site or the value associated with some environmental change at the site. Stated preference techniques such as contingent valuation are particularly suited to estimating the value of these hypothetical changes in non-market goods. An extension to TCM surveys therefore has been to supplement the usual questions related to trips taken with one or more contingent behavior questions where recreationalists are asked to state the number of trips they would take given either changes in site quality or varying percentage changes in trip prices. This revealed and contingent response data can then be used in count data models to estimate the change in welfare associate with the change in the site or environmental attribute (Hanley et al., 2003).

Combining revealed preference information and intended behaviour responses involves obtaining multiple responses from the same individual. As pointed out by Loomis (1997) an individual's multiple responses will likely be correlated due to individual specific but unobservable taste parameters. Standard statistical count models fail to account for this correlation and are therefore inefficient[1]. Panel estimators such as fixed and random effects poison and negative binomial models have been previously employed to account for the possible correlation of multiple responses of the same individual (Greene, 2008).

[1] It has also been previously noted that estimates from TCM models that combine both stated and revealed trip information should result in more efficient parameter estimates as more information on the same set of underlying preferences is employed in constructing the estimates (Hanley et al. 2003).

Endogenous stratification and truncation are two important issues of relevance for on-site collected contingent behaviour data and the associated panel count data models. Truncation refers to the fact that on-site data contains information on active visitors only and is therefore truncated at positive demand for trips to the site (Shaw, 1988 and Englin and Shonkwiler, 1995). Secondly, an on-site survey is subject to the problem of endogenous stratification. Due to the method of data collection the likelihood of being sampled depends on the frequency with which an individual visits the site.

To date, few attempts have been made to account for these on-site sampling issues in panel data count models. Papers that have done so have been Egan and Herriges (2006), Beaumais and Appéré (2010) and Moeltner and Shonkwiler (2010). Egan and Herriges (2006) developed random effects mixed Poisson regression models to jointly model observed and contingent behaviour data and to correct for on-site sampling and Beaumais and Appéré (2010) address on-site sampling issues within the framework of a random-effect Poisson gamma model. Moeltner and Shonkwiler (2010) also employ a Poisson-lognormal model that accounts for on-site sampling In their case the focus is on the estimation of trip demand and economic benefits for visitors to recreation sites when past-season trip information is elicited from respondents intercepted on-site.

However, even these models may be inadequate and potentially misleading if the recreational group of interest is heterogeneous with respect to the impact of explanatory variables. We account for this issue in this paper by extending a panel negative binomial model that accounts for endogenous stratification and truncation to account for unobserved heterogeneity through the use of a latent class modeling framework which assumes that the observations are drawn from a finite number of segments, where the distributions differ in the intercept and the coefficients of the explanatory variables.

In what follows we first discuss (section 2) the use of panel data or contingent behaviour TCM models in the literature with particular regard to studies that used on-site data and review the alternative approaches used to account for endogenous stratification and truncation. Section 3 then presents our extension of Englin and Shonkwiler's specification to a panel data negative binomial count data model that corrects for endogenous stratification and truncation and also allows for unobserved heterogeneity in the population. Section 4 provides a description of the recreational beach site used in the application of our models and includes a brief description of survey design and data collection procedures. Our estimation results are then presented in section 5. Finally, the paper concludes with a discussion of its major findings and their implications for recreational demand modelling.

The Contingent Behaviour Modelling Approach and accounting for On-site Sampling in Panel Data Count Models

There have been several attempts in the literature to combine the TCM revealed preference method and stated preference contingent valuation approaches to non-market valuation in the form of the contingent behaviour model. This is done with the objective of measuring the welfare impact of a hypothetical change in implicit price or in environmental quality (Whitehead et al., 2008b). Usually, this variation in site or environmental quality is obtained through a stated change in hypothetical visits. Examples of the use of the Contingent Behaviour TCM approach in recreational demand modelling include Englin and Cameron (1996), Hanley et al. (2003), Alberini and Longo (2006), Christie et al. (2007), Hynes and Cahill (2008), Martınez-Espineira and Amoako-Tuffour (2008) and Beaumais and Appéré (2010)[2].

While the majority of contingent behaviour studies use panel rather than pooled count data specifications (although studies such as Eiswerth et al (2000) and Hesseln et al. (2004) used pooled poisson specifications only) other approaches have included panel data ordinary least squares models (Englin and Cameron, 1996[3]), binary probit and random effects probit models (Loomis, 1997), panel tobit models (Azevedo et al., 2003) and pooled generalised least squares modelling approaches (Alberini et al., 2007). What does stand out from the literature is the fact that the correction for endogenous stratification and truncation in contingent behaviour models has, until very recently, been largely ignored.

To avoid dealing with the issue of truncation in panel count data specification many studies have discussed their per trip welfare estimates as being representative of their sample only and not of the general population of users (e.g. Hanley et al. 2003; Starbuck et al., 2006; Christie et al., 2007). Indeed, Hanley et al. (2003) note that since they do not adjust their model to take account of those individuals that currently do not make a single trip to the beach site under investigation, the resulting welfare estimates relate only to those who currently visit the beach site. They also point out that this failure to account for truncation will likely underestimate welfare gains to the wider population of all those who could take a trip under the contingent scenario but who currently do not.

The non-correction of contingent behaviour models based on on-site sampled data for endogenous stratification is even more prevalent in the literature than the

---

[2] For an in-depth review of the contingent behaviour modeling literature the interested reader should see Whitehead et al. (2008b).

[3] Englin and Cameron (1996) also applied a fixed effects Poisson model to compare to the fixed effects ordinary least squares model and to test for differences in price elasticities and consumer surplus from separate demand equations estimated with observed number of trips and intended number of trips for three hypothetical cost increases.

non-adjustment for truncation. This may be due to the fact that most modern statistical packages have standard routines to model Poisson and negative binomial count data models adjusted for truncation but in general the researcher herself needs to specifically program the adjustment of the log-likelihood function to correct for endogenous stratification. To get around the fact that no standard program is available in statistical packages to deal with endogenous stratification in panel data count models some studies have simply pooled the revealed and contingent observation points and run endogenously stratified truncated Poisson or negative binomial models which are routinely available.

For example Starbuck et al. (2006) examined the linkages between fire and fuels management activities to changes in forest recreation demand using the contingent behaviour methodology. A pooled endogenously stratified truncated Poisson model was used to estimate consumer surplus and predict changes in recreation visits under three alternative fire and fuels management scenarios. This pooling technique however ignores the fact that there is likely to be substantial correlation between the revealed and contingent behaviour responses from the same individual. It is also worth noting that a simple adjustment used to correct for endogenous stratification in the univariate Poisson model, is to transform the dependent variable (number of trips taken by individual i (Y_i) to equal $Y_i - 1$ (this adjustment is possible assuming a univariate Poisson distribution for the dependent variable and Shaw's (1988) derived on-site sampling distribution). Hesseln et al. (2004) use this adjustment technique in a pooled contingent behaviour model that examined the effects of fire on hiking demand in Montana and Colorado. However, as Egan and Herriges (2006) point out the above technique only applies to the univariate setting and this simple adjustment is not appropriate with use of the panel data specification.

Others have used strategies that avoid the need to account for endogenous stratification at all in a TCM framework. Mendes and Proença (2009) for example applied count-data travel cost methods to a truncated sample of visitors, to the Peneda-Gerês National Park but state that they did not need to account for endogenous stratification as people were intercepted at the park entrance. Presumably this is due to the fact that the respondents might as easily pass by as enter the park. Elsewhere Hynes and Hanley (2006) avoid the need of adjusting their truncated negative binomial TCM for endogenous stratification by combining data from their on-site survey with a non-site based survey - in their case, survey data collected via the internet. In this manner the sample incorporates individuals who visit the recreational site but who have a lower probability of being sampled on-site due to less frequent visits.

Finally, it has been suggested that the issue of endogenous stratification can be dealt with in a panel data count specification by simply applying a sampling weight to observations equal to the inverse of the estimated probability that an individual will visit the site. This reduces the proportional influence on the estimated model of individuals that have a higher probability of being included in the sample because of the on-site sampling design (i.e. those who are more likely of being sampled due to the increased frequency with which they visit the recreational site). Woolridge (2002) demonstrates how this inverse probability weighting recovers the population moments from a selected sample.

As mentioned earlier, only 3 papers to date have produced panel data count model that explicitly correct for both truncation and endogenous stratification. These are Egan and Herriges (2006), Beaumais and Appéré (2010) and Moeltner and Shonkwiler (2010). In the case of Egan and Herriges (2006) the authors develop a multivariate Poisson-log normal model to jointly model revealed and contingent behaviour data and to correct for on-site sampling. They also estimate Winkelmann's (2000) seemingly unrelated negative binomial (SUNB) model, also adjusted for truncation and endogenous stratification. The resulting models are used to analyze survey data collected on-site at Clear Lake in north central Iowa. The authors conclude that there is substantial bias in the results if the sampling procedures are ignored, overstating both the average number of trips to the site (by a factor of 14) and the welfare associated with the recreational opportunities at study site.

Beaumais and Appéré (2010) extend the work of Egan and Herriges (2006) by addressing the on-site sampling issue within the framework of a random-effect Poisson gamma model[4]. Their modelling approach constrains the correlation across counts for the same panel to be positive. This is not *a priori* the case of Egan and Herriges's (2006) multivariate Poisson log-normal specification. Similar to Egan and Herriges, Beaumais and Appéré (2010) and Moeltner and Shonkwiler (2010) also find that correcting for on-site sampling has a significant impact on model parameters and the consumer surplus estimates.

Finally it should be noted that count data panel models that incorporate unobserved heterogeneity have also been previously developed. For example, Wang et al. (1998), in the analysis of patent data, developed Poisson regression models for count data that accommodated heterogeneity arising from a distribution of both the intercept and the coefficients of the explanatory variables. The study assumed that the mixing distribution was discrete, resulting in a finite mixture model formulation.

[4] Beaumais and Appéré (2010) also introduce the concept of a "twin site". In their surveying approach they introduce to the respondent the hypothetical existence of a site strictly identical to the study site with a difference only in the environmental quality of certain attribute and try and establish the maximum distance the respondents would be willing to travel to such an alternative site and the number of extra trips if any that individual would make to such a site. This approach to defining the hypothetical scenario in a contingent behaviour study differs from that usually found in the literature where the change is environmental condition is defined in terms of the study site itself.

Advances in computational capabilities have also meant that statistical packages such as Nlogit (Greene, 2008) and Latent Gold Choice (Vermunt and Magidson 2005) now contain standard commands that allow the researcher to readily incorporate a discrete mixture distribution into panel count models.

As is evident from the previous (non-exhaustive) review of the literature much has been written in term of the issues surrounding on-site sampling issues related to the TCM. To date however no count data model exists for panel data that simultaneously accounts for the on-site sampling issues of endogenous stratification and truncation and the presence of unobserved heterogeneity via slope coefficients for the explanatory variables. The specification of such a model is presented in the following section. In particular we develop a random effects panel data model with a latent class framework that also accounts for truncation and endogenous stratification.

Methodology

In our study of recreational demand at Silverstrand Beach, the variables of interest are a count of beach trip demand in the previous 12 months and a count of potential beach trip demand that the same individuals would make given some hypothetical change in site quality or facilities. In effect, each person i in the data set yields two responses. The first is the number of trips (y_{i1}) they make to the beach under current conditions ($j=1$) and the second observation is how many trips (y_{i2}) the person says they would make if a specified improvement in recreational facilities at the beach occurs under hypothetical conditions ($j=2$). These counts are limited to non-negative integers. In the contingent behaviour modelling framework, we require a panel data modelling approach. The distribution of data on beach trip recreation is also positively skewed towards zero, thus preventing the use of a standard ordinary linear regression model (Cameron and Trivedi 2005).

Following the work of Shaw (1988), Grogger and Carson (1991), Englin and Shonkwiler (1995) and Greene (2008) we assume that, based on such data, a panel data count model of recreational demand can be estimated using a negative binomial distribution for the dependent count variable. As with Englin and Shonkwiler (1995) we also need to adjust our modeling strategy to control for the fact that our data were collected on-site. Unique in the literature we also adjust our random effects panel data negative binomial model corrected for on-site sampling to allow for the mixing of taste intensities over a finite group of taste segments in the population. Unobserved heterogeneity in the distribution of y_{ij} is assumed to impact the mean (and variance) λ_{ij}. The continuous distribution of the heterogeneity is approximated using what Greene (2008) refers to as a finite number of "points of support". The distribution is approximated by estimating the location of the support points and the mass probability in each interval. We interpret this discrete approximation as producing a sorting of individuals into C classes, c= 1,…C. Therefore in what follows we modify our random effects panel data negative binomial model corrected for on-site sampling for a latent sorting of y_{ij} into C classes.

Our starting point for a panel of trip data, where $i=1,\ldots,N$ individuals take y_i trips under site conditions $j=1,\ldots J_i$, is the standard negative binomial model for count data that allows for overdispersion in the responses;

$$P(y\mid\mathbf{x})=\frac{\Gamma(y+1/\alpha)}{\Gamma(1/\alpha)\Gamma(y+1)}\left(\frac{1/\alpha}{\lambda+1/\alpha}\right)^{1/\alpha}\left(\frac{\lambda}{\lambda+1/\alpha}\right)^{y} \qquad (1)$$

where $\lambda = \exp(\boldsymbol{\beta}'\mathbf{x})$ is the conditional mean function and $1/\alpha$ is the overdispersion parameter. (For convenience at this point, observation subscripts are omitted.) The vector $\mathbf{x}$ represents the set of explanatory variables reported for each individual *i*. It is a $k\times1$ vector of observed covariates and $\boldsymbol{\beta}$ is a $k\times1$ vector of unknown slope parameters. The scalar α and the vector $\boldsymbol{\beta}$ are parameters to be estimated from the observed sample. Finally α is a structural parameter to be estimated along with $\boldsymbol{\beta}$. Larger values of α correspond to greater amounts of overdispersion. The model reduces to the Poisson when $\alpha = 0$.

The density that applies to the observations obtained on site was shown by Shaw (1988) to equal:

$$P(y\,|\,\mathbf{x},\ on\ site) = \frac{yP(y\,|\,\mathbf{x})}{\sum_{t=1}^{\infty} P(t\,|\,\mathbf{x})}. \tag{2}$$

For the negative binomial model in particular, the result [see Englin and Shonkwiler, 1995, p. 106, (9)] is

$$P(y\,|\,\mathbf{x},\ on\ site) = \frac{y\Gamma(y+1/\alpha)\alpha^{y}\lambda^{y-1}(1+\alpha y)^{-(y+1/\alpha)}}{\Gamma(1/\alpha)\Gamma(y+1)}, \; y = 1,2,... \tag{3}$$

The second extension in our model is the accommodation of the latent sorting of individuals into *C* groups, or classes. The analyst does not observe directly which class, $c = 1,\ldots, C$, generated observation $y_{ij}\,|\,c$ and class membership must be estimated. The latent class model, in generic form, conditioned on the particular class can therefore be written as:

$$P(y|\mathbf{x},\ on\ site,\ class = c) = F(y|\mathbf{x}, \boldsymbol{\beta}_c, \alpha_c). \tag{4}$$

It should be noted that there is a separate dispersion parameter in each class as well. The unconditional *prior* probabilities attached to the latent classes are given by:

$$\pi_c = \text{Prob}(class = c) = \frac{\exp(\tau_c)}{\Sigma_{q=1}^{C} \exp(\tau_q)}. \tag{5}$$

The logit formulation for the probabilities is a convenient parameterization that allows the prior class probabilities to be constrained to the unit interval and to sum to one. The normalization $\tau_C = 0$ is imposed because only C-1 parameters are needed, with the adding up restriction, to specify the C probabilities. With this structure, there is a one to one correspondence between the set of parameters, $(\tau_1,\ldots,\tau_{C-1},0)$ and the set of class probabilities, $(\pi_1,\ldots,\pi_{C-1},1- \Sigma_{c=1}^{C=1} \pi_c)$. For an individual observation, the unconditional probability is averaged over the classes,

$$P(y \mid \mathbf{x},\ on\ site) = \sum_{c=1}^{C} \pi_c P(y \mid \mathbf{x},\ on\ site,\ class = c). \tag{6}$$

The probability $P(y|\mathbf{x},\ on\ site)$ is the term that enters the log likelihood that is maximized to obtain the estimates of $\boldsymbol{\theta} = [(\boldsymbol{\beta}_1,\alpha_1),(\boldsymbol{\beta}_2,\alpha_2),...,(\boldsymbol{\beta}_C,\alpha_C),(\tau_1,...,\tau_C)]$. The log likelihood for the observed sample is therefore

$$\log L = \sum_{i=1}^{N} \log\left\{\sum_{c=1}^{C} \pi_c P(y_i \mid \mathbf{x}_i,(\boldsymbol{\beta}_c,\alpha_c),\ on\ site,\ class = c)\right\} \tag{7}$$

where π_c is given in (5) and $P(y_i|\mathbf{x}_i,(\boldsymbol{\beta}_c,\alpha_c)$ on site, class = c) is given in (3) with $\lambda_i = \exp(\boldsymbol{\beta}_c'\mathbf{x}_i)$.

Individuals are observed more than once in the sample. We make the usual assumption that conditional on the class membership, which does not change for the person, the trip choices are made independently. There is correlation induced across choices in that the observed variables, x_i are correlated across visits and, as well, since the class membership is fixed, the individuals preferences, embodied in β_c are also common across visits. However, we have not assumed that there are unobserved factors that are omitted from the model and which are correlated across visits. With these assumptions, the joint probability of the T_i trip choices by individual i is given by

$$P(y_{i1},...,y_{iT_i} \mid \mathbf{x}_{i1},...,\mathbf{x}_{iT_i}, \boldsymbol{\beta}_c, \alpha_c, \textit{on site, class} = c) = \prod_{t=1}^{T_i} P(y_{it} \mid \mathbf{x}_{it}, \boldsymbol{\beta}_c, \alpha_c, \textit{on site, class} = c) \quad (8)$$

The log likelihood for the panel of data is obtained by inserting the joint probability in (8) in the log likelihood in (7);

$$\log L = \sum_{i=1}^{N} \log\left\{\sum_{c=1}^{C} \pi_c \prod_{t=1}^{T_i} P(y_{it} \mid \mathbf{x}_{it}, \boldsymbol{\beta}_c, \alpha_c, \textit{on site, class} = c)\right\} \quad (9)$$

The function in (9) is maximized with respect to $\boldsymbol{\theta} = [(\boldsymbol{\beta}_1, \alpha_1), (\boldsymbol{\beta}_2, \alpha_2), ..., (\boldsymbol{\beta}_C, \alpha_C), (\tau_1, ..., \tau_C)]$.

Finally, it should be noted that the approach of adjusting for truncation and endogenous stratification in both the observed and contingent observations distribution is different from that in Egan and Herriges (2006) and Beaumais and Appéré (2010) where the observed behavior data are assumed truncated to zero and endogenously stratified but the contingent behavior data are not. Thus the on-site sampling correction is only specified through observed data in their case. Even though our second observation for each person is the hypothetical number of trips they would make under changed site conditions, we argue that the problem of endogenous stratification and truncation still holds. The respondent is still someone who has a higher likelihood of being included in the sample due to their frequency of use.

Also, given that the contingent behaviour question is commonly set up such that respondents are asked how many more trips (if any) they would make to the site given an improvement in facilities (and therefore y_2 cannot be less than y_1) truncation still exists in the second period as we are still only dealing with individuals who will use the facility at least once. Interestingly, Moeltner and Shonkwiler (2010) showed that on-site sampling issues persist even for past season trip reports if the respondent is intercepted on-site this season. The authors labelled this effect "avidity carryover". They found that for their sample of lake visitors relatively stronger preference or "avidity" for the interview site carries over across seasons. We argue that a similar effect could apply to hypothetical trip reports, if we interpret them as "future season trips". If that is indeed the case then this again implies that the contingent behavior data as well as the observed behavior data should be assumed truncated.

For consumer utility maximization subject to an income constraint, and where the number of trips are a nonnegative integer, we follow Hellerstein and Mendelsohn (1993) and calculate the expected value of consumer surplus, $E(CS_{ij})$ from our count models as $E(CS_{ij}) = E(y_{ij}|x_i)/\beta_{pi} = \hat{\lambda}_{ij}/(\beta_{pi})$ where Y_{ij} is the number of trips to the beach for individual *i* under conditions *j*, and λ_{ij} is some underlying rate at which the number of trips occur, such that one would expect some number of trips in a particular year, i.e. λ_{ij} is the mean of the random variable Y_{ij}. $ß_{pi}$ is the individual specific price (*i.e.* travel cost) coefficient. The per-trip $E(CS_{ij})$ is simply equal to $1/\text{-}ß_{pi}$. The change in the consumer surplus resulting from an improvement in the coastal amenities is then given by

$$\Delta E(CS_i) = \Delta E(y_{ij}|x_i)/\beta_{pi} = (\hat{\lambda}_i^* - \hat{\lambda}_i)/\beta_{pi} \qquad (10)$$

where $\hat{\lambda}_i$ is the expected number of trips before any improvements are made to the coastal amenities *(j = 1)* and $\hat{\lambda}_i^*$ is the expected number of trips after improvements are made to the coastal amenities *(j = 1)*. This suggests that the change in consumer surplus for individual *i* can be calculated by dividing the change in the predicted number of trips to the beach site by the coefficient of the travel cost variable. It is important to state that the relevant comparison in welfare terms is between the number of predicted trips at the current level of coastal amenity provision at the beach site and the predicted number of trips at the improved level. Also, one cannot disaggregate benefit estimates into additional utility from those who take no extra trips to the beach and additional utility from those who visit most frequently. The beach travel cost study and the on-site collected dataset employed are described in the next section prior to the presentation of model results and welfare estimates.

Data and Study Background

The application of our model is to a data set generated from a survey that examined the possible welfare impact associated with the development of a coastal trail that connects two beach areas along the Galway Bay coastline in the west of Ireland. The data was generated from an on-site survey of visitors to Silverstrand beach approximately 7km outside of Galway city which is accessible by public road only. The beach was awarded a blue flag status in 2009 and is therefore required to comply with certain standards in terms of lifeguard safety and patrol as well as high water quality. The beach itself is only 300m long and has only limited facilities in the form of parking, benches, picnic tables and toilet facilities. Nevertheless it is a popular destination, particularly in the summer months for outdoor enthusiasts and is used heavily by the local urban community of Galway city and surrounding area as a recreational amenity. The beach caters for a wide range of uses including walking, swimming, sun-bathing, bird watching, kayaking and kite surfing. The beach was of interest as it is a site where potential exists to add recreational value through the establishment of a walking trail that would link it to another area of beach currently cut off by a small area of farmland.

Coastal based recreational activities have been recognized in Ireland as having the potential to deliver substantial economic benefits to rural areas through locally run tourism activities (Vaughan et al., 2000). Failte Ireland (2008) reported that holidaymakers do not visit Ireland for the typical beach holiday, but rather seek out soft adventure activities such as walking, kayaking, etc along the coast. It has also been noted that one of the best means for improving the value of coastal resources, such as beaches, is through the provision of walking trails. These not only provide a valuable source of recreation to the public but also provide increased access to the coastline. However some of the best coastal walking areas in Ireland can only be accessed through private farm land and under Irish law access to privately owned land, for the purpose of recreation, is at the discretion of the landowner. A variety of issues such as potential interference with agricultural activities, insurance liability and potential invasion of privacy have been reported by landowners as reasons why they may be unwilling to permit public access to their farmland for walking related activities (Buckley et al., 2009).

Silverstrand beach was chosen as a site to investigate the issue of coastal access as a strip of privately owned agricultural land which has a cliff face at the waters edge prevents the access of recreationalists to a much larger area of beach and access along the shore to the nearby Salthill beach and promenade. If recreations could freely cross this section of agricultural land it would open up a coastal walk of over 4 miles. At present users of Silverstrand have no right to cross the private farmland to access the additional beach area. With this in mind respondents were asked a contingent behaviour question in relation to how their usage of the beach facility would change if the length of beach at their disposal was increased through the opening up of a cliff walk that would give them access to an additional 1km of beach) and also access along the shore to Salthill beach and promenade. The features of the new walking trail were pointed out to respondents on a map as well as information on how the new walking trail would also open up access to the nearby Salthill beach as well. Respondents were informed how walkers would be granted formal right of way access along the walk, a marked path with a fence to separate the walk from the farm land and cliff edge and informational plaques detailing the surrounding countryside. They were also informed how all facilities would be built with material that blends in with the coastal amenity.

As part of the study, 146 personal interviews were carried out at the beach site. The questionnaire was piloted over a 2 week period in June 2009. This was followed by the main survey which took place at Silverstrand during the months of July and August 2009. Due to the non-response to certain questions in the main survey, 18 surveys were not deemed usable in the final analysis which resulted in a final sample 126 individual responses being used for model estimation. The on-site interviews were conducted on both week days and weekends, during all daylight hours. The questionnaire solicited information on trips taken to the beach, activities undertaken, personal demographics, income, employment status, education, social relations and obligation free time. Each interview took approximately 20 minutes.

Respondents were provided with background information on the study and were then asked to outline how they used the beach for recreation. Next, they were presented with information on how the beach (where they were sampled) might be improved for recreation. Respondents were then presented with the contingent behaviour scenario (as shown in Figure 1) and asked to identify the extent to which their number of planned trips to the beach in the next 12 months would change if the stated change was made.

In particular, they were asked if the changes described on the card were implemented at the beach resource, would they change the number of trips they would take to the site over the next 12 months. This was followed up with an option of choosing 1. *no change in number of trips taken*, 2. *more trips* or 3. *fewer trips*. Finally the respondent was asked to state the increased (or decreased) number of trip if they had chosen option 2 (or 3) [5]. Thus, 2 observations for trips taken were collected from each respondent; the actual number of trips taken in the previous 12 months and the contingent number of trips that would be taken if the walking trail was put in place. This resulted in a panel data set of 256 observations. Finally, attitudinal data was also collected from the respondents.

Each respondent's travel cost was computed following the standard approach in the literature by considering the direct costs and the opportunity cost of travel. For each respondent *i* and each scenario *j*, the travel cost was calculated as:

$$TC_{ij} = \left(\frac{Dis\tan ce_{ij} \times CostperKM}{Groupsize_i} \right) + (0.25 \times \left(\frac{Income_i}{2000} \right))$$

where $Dis\tan ce_{ij}$ is the round-trip distance from the respondent's home to the site. $CostperKM$ is the average petrol cost per mile (the Automobile Association of Ireland's calculation of €0.224/mile obtained from http://www.aaireland.ie/infodesk/cost_of_motoring.asp was used) and $Groupsize_i$ is the number of people that travelled to the site in the respondent's vehicle. Following Shaw and Feather (1999), the opportunity cost of travel time is included in the travel cost calculation as a proportion (0.25) of the hourly wage, where the hourly wage rate

[5] As is often the case in contingent behavior studies of this type no respondent chose option number 3.

was taken as the respondents reported income divided by 2000, based on a 40 hour week for 50 weeks in a year. No allowance for on-site time was made in the travel cost calculation[6].

Relaxing/Sun bathing was highlighted as the main activity of 35% of all respondents in the survey followed by entertaining children (21%), swimming (13%), walking (11%) and other water sports (6%). Also, it is notable that 49% of respondents were male, 57% were in full-time employment and 63% had been educated up to degree level. Mean annual visits to the beach where each respondent was sampled were 11.76 (range 1-60). The day of the survey was the first ever visit to the beach for 7% of the sample and respondents spend on average 2 hours 31 minutes on site. A visit to the beach was the main purpose of the day's journey for 61% of the sample, and participants in the survey used the beach resource for, on average, 4.1 different recreational activities. Mean one-way distance travelled was 24 miles and respondents to the survey tended to be at the beach in groups of, on average, 2.2 persons (range 1 to 13). Further summary statistics associated with the sample are presented in table 1.

Results

Given the contingent behaviour scenarios described in Figure 2 and the model specifications described in section 3 we present here the results of 2 models. Table 2 then presents the results of both a random effects negative binomial model corrected for on-site sampling and a random effects latent class negative binomial panel model also corrected for endogenous stratification and truncation. Although not presented

[6] An in-depth discussion of the many issues that surround the calculation of the travel cost variable is beyond the scope of the article but for a good over view of the treatment of time and the specification of the travel cost variable in recreation demand models the interested reader is advised to see Amoako-Tuffour and Martınez-Espineira (2008) and Hynes et al. (2009).

here, both pooled versions of the Poisson and negative binomial model were also initially fitted as were random effects Poisson and negative binomial models uncorrected for on-site sampling[7]. While the results for these models are not presented we do, for the purpose of comparison estimate and present the mean consumer surplus per trip estimates and the change in consumer per trip estimates as a result of the new coastal walking trail for all models in table 3.

In all models, the average number of trips undertaken by individual *i* under (the real or contingent) scenario *j* is assumed to be a function of the travel cost to the site, the travel cost to the respondent's next preferred substitute site, whether the respondent participates in a water sport while on-site, is a member of a recreation or environmental organisation, income, age, income whether the visit to the beach is by chance due to the respondent being in the area for other business and a 'Contingent Behaviour' variable, which indicates whether the visits we are explaining are actual, with current facilities, or hypothetical, with improved facilities. A further description of each of the independent variables is given in table 1.

The model in the first column of table 2 is the random effects panel negative binomial accounting for on-site sampling (henceforth referred to as the NB corrected model) while the second and third columns present the results of the negative binomial panel model that allows for unobserved heterogeneous with respect to the impact of explanatory variables on the number of trips taken as well as accounting for the issue of on-site sampling (henceforth referred to as the latent class corrected NB model). The travel cost coefficients in both models are significant at the 5% level and have the expected negative signs. This indicates that, on average, as the cost of travelling to the beach site decreases, the number of trips made to the site increases.

[7] Whether a panel specification was preferred to a pooled specification was tested, and the Likelihood Ratio test statistic confirmed the need for a panel rather than a pooled regression.

The ‘travel cost to the nearest substitute site’ and the ‘Incidental visit to the beach’ variable are also significant and have the *a piori* expected signs.

The one major difference between both models in terms of the estimated coefficients is that the ‘contingent behaviour’ variable is insignificant in the NB corrected model. This finding would appear to suggest that the hypothetical trail that facilitates access to a further area of beach does not have a statistically significant effect on the number of planned trips to the site. Once we account for the unobserved heterogeneity in our sample however the ‘contingent behaviour’ variable in our latent class corrected NB model is significant (at the 90% level in class 1 and at the 99 % level in class 2). In fact, all variables bar being a member of a recreation or environmental group are now significant at the 95% level in at least one of the two class segments.

With respect to the definition and testing of hypothesis on the number of classes to include in the latent class corrected NB model the conventional specification tests used for maximum likelihood estimates are not valid as they do not satisfy the regularity conditions for a limiting chi-square distribution under the null (Hynes et al., 2008). Therefore, in order to decide the number of classes, we used the information criteria statistics first developed by Hurvich and Tsai (1989). We report the Akaike information criterion (AIC), the Baysian information criterion (BIC) and the Hannan Quinn statistic for all models in tables 2 and 3. In terms of the latent class corrected NB model no one number of classes minimize each of the measures. The 3 class specification has the lowest score on 2 of the criteria while the 2 class specification is lowest for the BIC. As Scarpa and Thiene (2005) point out these statistics provide guidance on the number of latent classes to choose but this decision also requires the discretion of the researcher. We hence choose only to report in table 3 the latent class corrected NB model estimates for the 2 class model even though two of the information criteria statistic were lower for the 3 and 4-class models. We reject the 4 class model as one of its classes has a complete set of insignificant parameter estimates and also both the 3 and 4 class models displayed a high number of insignificant parameter estimates in at least one of their other classes.

As can be seen from table 2, the two-class model specification allocated 22% of respondents to class one and 78% to class two. Importantly, the travel cost coefficients in both classes are negative and significant at the 5% level and, as mentioned above, the contingent behaviour variable is also significant in both classes. It is also interesting to note that the income coefficient is now significant for the smaller group of recreationists likely to be represented by class 1. This coefficient was insignificant in all earlier versions of the contingent behaviour model. Only by allowing for taste heterogeneity in the sample do we pick up in the importance of this characteristic for a certain portion of recreationalists using the site. It should also be noted that for this smaller segment participation in water sports has no influence on the number of trips made to the site whereas it has for class 2. The travel cost variable would appear to have more or less the same influence on both which would suggest that both classes exhibit ‘price’ sensitivity to the same degree.

Finally, it should be noted that that the latent class corrected NB model had a lower log likelihood value (in absolute terms) and a lower score on all of the information criteria statistics than the NB corrected model indicating that the latent class structure provides a better fit for our on-site sampled data that when we assume a homogenous mean influence of the explanatory variables amongst our beach recreationists.

Following Beaumais and Appéré (2010) we also carried out a Vuong test (Vuong, 1989) to examine if the on-site sampling correction to the negative binomial specification was appropriate. Previously Greene (1994) adapted the Vuong test to examine the appropriateness of a zero-inflated negative binomial versus a standard negative binomial model. The Vuong statistic has a limiting distribution that is normal with large positive values favoring the corrected model and with large negative values favoring the standard panel version of the negative binomial model unadjusted to account for on-site sampling. Values close to zero in absolute value favor neither model. The calculated Vuong statistic of 9.54 results in a clear rejection of the null hypothesis that not accounting for on-site sampling has no effect on the means or the variances in the negative binomial panel specification of the contingent behaviour model (i.e. that the models are indistinguishable).

Estimating the welfare effects of changes in the quality or supply of site facilities or environmental goods is the main objective of most contingent behaviour studies. We therefore consider the implications for welfare measures of controlling for on-site sampling and unobserved heterogeneity. In particular we compare the consumer surplus (CS) per trip (real behaviour), the estimates of the change in number of trips taken and the change in total CS per recreationalist as a result of the hypothetical extension to the beach being provided through the creation of an adjoining walking trail, across the alternative model specifications. Table 4 also reports the estimates for the basic (unadjusted for on-site sampling issues) pooled and panel Poisson and Negative Binomial specifications even though the model parameter estimates are not presented for these models.

The panel negative binomial models accounting for truncation and endogenous stratification result in lower mean CS estimates and lower predicted trips taken than the basic pooled and panel Poisson and NB models. The distribution of CS estimates for the latent class corrected NB model varies across classes, with each class having a specific CS per trip estimate. The class weighted population estimate of per-trip consumer surplus for the latent class corrected NB model is estimated with 95% confidence to be between €16.93 and €27.21. With a mean CS per revealed trip estimate of €21.67 and €15.67 for class 1 and 2 respectively this model provides the most conservative mean CS estimates across all the reported models.

While nothing in the construction of the latent class model assures that the consumer surplus measures in a two class model will bracket the result from a one class model (the NB corrected model) it is still interesting to note that the CS estimate in the NB corrected model does not fall between the 2 class estimates of the latent class corrected NB model. This may be an indication that the one class model is forcing an overestimate of the consumer surplus measure and that that controlling for heterogeneous in the population with respect to the impact of the chosen explanatory variables provides more reliable CS estimates.

To estimate the recreation benefits from the access improvements and the addition of the walking trail and additional beach area, the steps outlined in the methodology section were followed. To calculate the proportional change in recreationalist welfare from implementation of the coastal walking trail, we first take into account the stated change in trips to the beach site if the trail were to be put in place. Such a facility improvement would increase visits by an estimated 3.32 trips per year under the NB corrected model. This is the lowest predicted change in trips across all model specifications.

Even though the latent class corrected NB model provides the lowest mean CS per trip estimates it predicts the second largest change in the number of trips taken per individual as a result of the beach site changes being implemented (6.04 additional trips per person per annum). However, the relative low CS per trip estimate for the latent class corrected NB model means that the estimated total increase in consumer surplus from the beach facility improvements per person per year (the class weighted estimate) is only €0.82 higher than the estimate associated with the NB corrected model (€102.26 and €101.44 respectively). The panel negative binomial model that does not account for truncation and endogenous stratification produces estimates for the change in CS per person per year that are approximately 65% larger the models that do account for on-site sampling while the pooled unadjusted models (which is still a popular approach in the literature) provide estimates that are over 300% higher.

Discussion and Conclusions

In this paper, we presented an extension to Shaw's (1988) and Englin and Shonkwiler's (1995) count data models corrected for on site sampling to a panel data setting. We contrasted a number of modelling techniques; namely, random effects panel Poisson and Negative Binomial specifications that did not account for the on-site sampling issues of truncation and endogenous stratification, a panel negative binomial model that did account for the aforementioned on-site sampling issues and finally, and uniquely in the literature a latent class random effects panel data negative binomial model corrected for on-site sampling but at the same time allowing for the mixing of taste intensities over a finite group of taste segments in the population. The chosen models were applied to revealed and contingent travel data obtained from a survey of visitors to a beach on the outskirts of Galway city in Ireland. While the estimated models did not provide coefficient estimates that were dramatically different from one another, it was still clear that that the failure to correct for on-site sampling results in biases in the estimated average number of both observed and contingent trips to the beach.

Unlike previous attempts at accounting for on-site sampling in contingent behaviour models we assumed that, similar to the observed trip counts, the contingent behaviour trip responses are also truncated and endogenously stratified. In their modelling approach Egan and Herriges (2006) and Beaumais and Appéré (2010) assumed observed behaviour data were truncated to zero, but contingent behavior data was not. We argue that the contingent observation is likely to still suffer from endogenous stratification as the respondent being interviewed on-site is still someone who has a higher likelihood of being included in the sample due to their frequency of use. Also, given that the contingent behaviour question is usually set up such that respondents are asked about a counterfactual situation where there have been changes to the current site; truncation still exists in the second period as we are still only dealing with individuals who will use the facility at least once. Thus unlike previous studies, in our modelling approach the on-site sampling correction is specified through both the observed and contingent behaviour data.

While Egan and Herriges (2006), Beaumais and Appéré (2010) and Moeltner and Shonkwiler (2010) have previously developed count data panel models corrected for on-site sampling their approaches may still be inadequate and potentially misleading if the population of interest is heterogeneous with respect to the impact of the chosen explanatory variables. The error term added to the parameterized mean function of the Poisson models used by the aforementioned authors can be interpreted as capturing unobserved heterogeneity. However what was still missing in the literature up until this paper was an on-site corrected count data model that captures unobserved heterogeneity via slope coefficients for explanatory variables. Our proposed methodology accounts for heterogeneity in both the underlying mean number of trips taken and the regression coefficients. That is, our model assumes that the observations are drawn from a finite number of segments, where the distributions differ in the intercept and the coefficients of the explanatory variables. Within each class the population interest is homogenous with respect to the impact of explanatory variables but this assumption is relaxed across classes.

We would contend that the use of latent class modeling approach is particularly relevant for on-site sampled recreationalists. Users of a recreational site such as a beach or a forest park tend to be diverse and have different reasons for wanting to visit such sites. In the discrete choice recreational demand literature this has been a well recognized fact since Train (1998) and now the publication of almost all work involving the estimation of destination choice random utility models involves modelling the site choice decision for recreationists allowing for the mixing of taste intensities either over a finite group of taste segments (the latent class approach) or over continuous value distributions (random parameter logit approach) [8]. This recognized heterogeneity across recreational groups using a site such as a beach (and indeed even within particular recreational groups) has not been given the same treatment in count data travel cost models of recreation demand as it has in the discrete choice literature. This paper fills that gap in the literature.

[8] Hynes et al. (2008) highlight the fact that there are different types of boaters within a population of kayakers using a random utility site choice latent class modelling framework while Scarpa and Thiene (2005) do the same for rock climbers. An early paper by Morey (1981) developed a model of skier behaviour implicitly taking into account whether the skier was a novice, intermediate or of advanced level. The results of that study indicated that the number of days spent at a particular skiing site depended significantly on the individual's skiing ability.

The latent class corrected NB model facilitated a much deeper analysis of the factors driving the decision to make a particular number of trips to the beach site. It also highlighted the fact that there are distinct segments of the population who make that decision based on different influences. For instance in one segment, being wealthier has a significant (and negative) influence on the number of trips taken while participating in a water sport at the beach site did not. In the second segment income had no significant influence while participating in a water sport at the beach site was highly significant.

The latent class approach also generates additional information which is potentially very useful to recreational site managers, simply by identifying groups of users with particular demands. Planners and policy makers may be concerned with how changes to coastal sites will affect visitor numbers or the utility of the individuals that visit the sites. Being able to identify different segments of users within a count data modelling framework will allow such managers to better allocate resources between policy issues such as beach congestion, beach access, coastal access such as roads and trails and beach developments and facilities. In our empirical investigation for example the results of the latent class corrected NB model would suggest that policies impacting on water sport participation would have an impact on a larger group of beachgoers.

Given the relatively small sample size it would be wise to take a cautious view as to how representative the estimated welfare results are of the population of beach users in the west of Ireland. Nevertheless the estimated models still demonstrate how controlling for on-site sampling and unobserved heterogeneity can have a significant impact on predicted trips taken and on welfare estimation. Also, it should be noted that the modelling framework proposed here relies heavily on the assumption that the non-site participants we are trying to account for with the on-site sampling corrections are the same as the users that are actually sampled but are simply coming from a different portion of the same underlying distribution. If non-site participants are fundamentally different from the sampled site users, the correction will be inappropriate. This assumption is not unique to our case but is something that is relevant to all procedures in the recreational demand literature that have attempted to account for on-site sampling issues.

It is important to state that while the focus of the paper was on a model of contingent behaviour the developed modelling framework is just as applicable to cases where data has been collected on-site in relation to trips taken by the same individuals over repeat time periods or on an individual's trip activity to alternative sites over a fixed period. Finally, an area for future research is to compare the welfare impacts derived using the latent class specification developed here to a count data model where the unobserved heterogeneity of the population with respect to the explanatory variables is specified as continuous rather that over finite segments (i.e. specifying the slopes as random coefficients). This would allow for a broader discussion of how unobserved heterogeneity could be best captured in on-site panel count data models.

References

Amoako-Tuffour, J. and R. Martınez-Espineira. 2008. "Leisure and the opportunity cost of travel time in recreation demand analysis: A re-examination". MPRA working paper 8573.

Azevedo, C., Herriges, J. and C. Kling. 2003. "Combining revealed and stated preferences: consistency tests and their interpretations". *American Journal of Agricultural Economics*, 85: 525-537.

Beaumais, O. and G. Appéré. 2010. “Recreational shellfish harvesting and health risks: A pseudo-panel approach combining revealed and stated preference data with correction for on-site sampling”. *Ecological Economics* 69 (12): 2315–2322.

Buckley, C., Hynes, S., van Rensburg, T.M. and E. Doherty. 2009. Walking in the Irish countryside – Landowner preferences and attitudes to improved public access provision. *Journal of Environmental Planning and Management,* 52(8): 1053-1070.

Christie, M., Hanley, N. and S. Hynes. 2007. “Valuing enhancements to forest recreation using choice experiments and contingent behaviour methods”. *Journal of Forest Economics*, 13 (2): 75-102

Egan, K. and J. Herriges. 2006. “Multivariate count data regression models with individual panel data from an on-site sample”. *Journal of Environmental Economics and Management*, 52: 567–581.

Englin, J. and T.A. Cameron. 1996. “Augmenting travel cost models with contingent behaviour data”. *Environmental & Resource Economics*, 7: 133–147.

Englin, J. E. and J. S. Shonkweiler. 1995. “Estimating social welfare using count data models: An application to long-run recreation demand under conditions of endogenous stratification and truncation”. *The Review of Economics and Statistics*, 77(1): 104-112.

Failte Ireland, 2008. *Hiking / walking 2007* [online]. Available from: http://www.failteireland.ie/getdoc/9c7a56d2-299d-460d-943f-b34639f0f07f/Hiking-2007 [Accessed 6th January 2010]

Greene, W. H. 1994. Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models. Working paper, Stern School of Business, NYU EC-94-10.

Greene, W. H. 2007. *LIMDEP 9.0 Econometric Modeling Guide*, Volume 2, Econometric Software, Inc. New York

Greene, W. H., 2008. *Econometric Analysis*, 6th ed., New Jersey: Prentice Hall

Grogger, J. and R. Carson. 1991. "Models for truncated counts". *Journal of Applied Econometrics*, 6: 225-238.

Haab, T. and K McConnell. 2002. *Valuing environmental and natural resources: the econometrics of non-market valuation*. Edward Elgar, Northampton, MA

Hanley, N., Bell, D. and B. Alvarez-Farizo. 2003. "Valuing the benefits of coastal water quality improvements using contingent and real behaviour". *Environmental and Resource Economics*, 24(3): 273-285.

Hesseln, H., Loomis, J. and A. González-Cabánc. 2004. "Comparing the economic effects of fire on hiking demand in Montana and Colorado". *Journal of Forest Economics*, 10 (1): 21-35 .

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

Hurvich, M. and C. Tsai. 1989. "Regression and time series model selection in small samples". *Biometrika,* 76: 297-307.

Hynes, S. and B. Cahill. 2007. "Valuing the benefits to the local community of supplying recreational facilities in community owned forests: an application of the contingent behaviour method". *Small-scale Forestry*, 6(3), 219-231.

Hynes, S., Hanley, N. and C. O'Donoghue. 2009. "Alternative treatments of the cost of time in recreational demand models: an application to white water rafting in Ireland". *Journal of Environmental Management*, 90(2), 1014-1021.

Hynes, S. and N. Hanley. 2006. "Preservation versus Development on Irish Rivers: Whitewater Kayaking and Hydro Power in Ireland". *Land Use Policy*, 23: 170 - 180.

Loomis, J. 1997. “Panel estimators to combine revealed and stated preference dichotomous choice data”. *Journal of Agricultural and Resource Economics*, 22 (2): 233–245.

Martınez-Espineira, R. and J. Amoako-Tuffour. 2008. “Recreation demand analysis under truncation, overdispersion, and endogenous stratification: An application to Gros Morne National Park”. *Journal of Environmental Management*, 88(4): 1320-32.

Mendes, I. and I. Proença. 2009. “Measuring the social recreation per-day net benefit of wildlife amenities of a national park: a count-data travel cost approach”. Instituto Superior de Economia e Gestão, Lisbon - DE Working Paper 35 (ISSN No. 0874-4548).

Morey, E., 1981. “The demand for site-specific recreational activities: a characteristics approach”. *Journal of Environmental Economics and Management*, 8 (4): 245-271.

Moeltner, K. and Shonkwiler, J. S. (2010). Intercept and recall: Examining avidity carryover in on-site collected trip reports, Transportation Research Part D: Transport and Environment 15(7): 418-427.

Scarpa, R. and M. Thiene. 2005. “Destination choice models for rock-climbing in the North-Eastern Alps: a latent-class approach based on intensity of participation”. *Land Economics* 81:426-444.

Shaw, D., 1988. On-Site Sample’s Regression: Problems of Non-Negative Integers, Truncation, and Endogenous Selection, Journal of Econometrics, 37: 211-223.

Shaw, D. and Feather, P., 1999. “Possibilities for including the opportunity cost of time in recreation demand systems”. *Land Economics*, 75(4): 592-602.

Starbuck, M., Berrens, R. and McKee, M., (2006) Simulating Economic Impacts from Hazardous Fuels Treatment and Forest Restoration Management Activities. *Forest Policy and Economics*, 2006(8): 52-66.

Vaughan, D., Farr, H. and R. Slee. 2000. "Economic benefits of visitor spending". *Leisure Studies*, 19(2), 95–118.

Vermunt J. and J. Magidson. 2005. *Technical Guide for Latent GOLD Choice 4.0: Basic and Advanced.* Belmont Massachusetts: Statistical Innovations Inc.

Vuong, Q. 1989. "Likelihood ratio tests for model selection and non-nested hypotheses". *Econometrica*, 57: 307–333.

Wang, P., Cockburn, I. and L., Puterman. 1998. "Analysis of patent data—a mixed poisson-regression-model approach". *Journal of Business and Economic Statistics*, 16 (1), 27–41.

Whitehead, J.C., Dumas, C.F., Herstine, J., Hill, J. and B. Buerger. 2008a. "Valuing beach access and width with revealed and stated preference data". *Marine Resource Economics*, 23(2): 119-135.

Whitehead, J., Pattanayak, S., Van Houtven, G., and B. Gelso. 2008b. "Combining revealed and stated preference data to estimate the non-market value of ecological services: An assessment of the state of the science". *Journal of Economic Surveys*, 22: 872-908.

Whitehead, J.C., Phaneuf, D.J., Dumas, C.F., Herstine, J., Hill, J. and B. Buerger. 2010. "Convergent validity of revealed and stated recreation behavior with quality change: a comparison of multiple and single site demands". *Environmental and Resource Economics*, 45: 91–112.

Winkelmann, R. *Econometric Analysis of Count Data*. Springer-Verlag, New York, 2000.

Woolridge, J. 2002. "Inverse probability weighted M-estimators for sample selection, attrition, and stratification". *Portuguese Economic Journal*, 1(2): 117–139

Figure 1. Scenario examined in contingent behaviour study

Suppose that **NEXT YEAR** a new **WALKING PATH** was built connecting to this beach resource.

The path would consist of:

- An approx 2km round trip walk along the cliffs to the end of the spit at Rusheen Bay
- Walkers would be granted formal right of way along the walk (currently people walk along the cliff but are not supposed to as it is privately owned farm land),
- A marked path with a fence to separate the walk from the farm land and cliff edge
- Informational plaques detailing the surrounding countryside.

All facilities would be built with material that blends in with the coastal amenity.

How would these new facilities affect your use of THIS BEACH?

Table 1. Summary Statistics

Variable Name	*Description*	*Mean*	*Standard Deviation*
Actual trips	Number of trips respondent actually took to the beach in last 12 months	11.76	14.9
Hypothetical trips	Number of trips respondent would take in next 12 months if scenario implemented	17.31	19.23
Age	Age	41.06	13.68
Income	Gross annual income (€)	51,551	29,334
Incidental Visit to Beach	Dummy indicating whether trip to beach occurred by chance as happened to be in the area anyway (1) or was a planned trip to the beach (0)	0.39	0.49
Member of Recreation or Environmental Organisation	Dummy variable Indicating whether the respondent is an active member of a recreational organisation such as a kayak or surf club or an environmental organisation such as Birdwatch Ireland or Greenpeace	0.47	0.5
Travel Cost	Return travel cost from home to beach	15.28	17.43
Travel Cost Substitute Site	Return travel cost to the alternative site most frequently visited by respondent	13.77	15.32
Water Sport Participation	Dummy variable indicating whether trip to beach involved a water sport	0.15	0.36

Table 2. Negative Binomial Contingent Behaviour Models accounting for Truncation and Endogenous Stratification

	Negative Binomial Panel Count Model	Latent Class Negative Binomial Panel Count Model	
		Latent Class 1	*Latent Class 2*
Age	0.156*** (0.035)	0.104*** (0.031)	0.215*** (0.052)
Income	-0.003 (0.002)	-0.005** (0.002)	0.003 (0.002)
Incidental Visit to Beach	-1.202*** (0.145)	-1.481*** (0.197)	-0.680*** (0.221)
Member of Recreation or Environmental Organisation	0.404*** (0.094)	-0.052 (0.101)	0.180 (0.114)
Contingent Behaviour	0.481(0.388)	0.292* (0.172)	0.666*** (0.210)
Travel Cost	-0.033*** (0.009)	-0.047** (0.019)	-0.064*** (0.017)
Travel Cost Substitute Site	0.032*** (0.009)	0.067*** (0.022)	0.039** (0.016)
Water Sport Participation	0.553*** (0.145)	0.166 (0.155)	0.437** (0.182)
Constant	0.463 (0.420)	3.529*** (0.184)	0.534* (0.280)
Scale Parameter or Alpha in LC Model	1.345 (1.584)	0.051** (0.026)	0.722*** (0.176)
Class Probabilities		0.217*** (0.040)	0.783*** (0.040)
AIC	1735	1605	
BIC	1771	1679	
Hannan Quinn	1749	1635	
Log likelihood	-858	-781	

Standard errors are in parentheses. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level. The income variable has been rescaled by dividing by 1000.

Table 3. Consumer Surplus (CS) and Change in Trips Taken Estimates from Alternative Model Specifications (all figures are in per person).

Model Specification	Mean CS per Trip (€)	Change in number of trips taken a result of new walking trail	Change in CS as a result of new walking trail (€)
Pooled Poisson*	59.35 (43.09, 75.60)	5.63	334.23
Pooled Negative Binomial*	125.02 (-38.15, 288.20)	6.19	774.12
Basic Panel Poisson	35.88 (15.09, 56.66)	3.51	125.94
Basic Panel Negative Binomial	34.64 (1.31, 67.97)	4.87	168.56
Panel Models Accounting for Truncation and Endogenous Stratification			
Negative Binomial	30.54 (14.11, 46.96)	3.32	101.44
LC Negative Binomial: Class 1	21.43 (4.20, 38.65)	6.04	129.39
LC Negative Binomial: Class 2	15.67 (7.36, 23.98)	6.04	94.61
Weighted LC Negative Binomial**	16.93 (6.66, 27.21)	6.04	102.26

Ninety five percent confidence interval in brackets
* The model results of the Pooled Poisson and Pooled Negative Binomial models are not presented in this paper but are available from the authors upon request.
** This is the weighted consumer surplus per trip estimate estimated by considering the class probabilities in the NB Latent Class Model

SEMRU

Socio-Economic Marine Research Unit