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

## *Working Paper Series*

Working Paper 12-WP-SEMRU-06

Investigating Scale Heterogeneity in Latent Class Models

Marco Boeri
Edel Doherty
Danny Campbell
Alberto Longo

**SEMRU *Working Paper Series***

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

# Investigating Scale Heterogeneity in Latent Class Models****


## Abstract

This paper develops and compares two alternative approaches to accommodate scale heterogeneity (also referred to as heteroskedasticity) in latent class models. Our modelling approach compares two different representations of heteroskedasticity, respectively associating the heterogeneity in scale factor with respondent's characteristics (i.e. observed scale heterogeneity) or deriving it probabilistically (i.e. unobserved scale heterogeneity). The results reveal a number of benefits associated with this type of approach, particularly when heterosckedasticity can be linked to observed characteristics of the respondent. Our data comes from a discrete choice experiment eliciting recreational users preferences for farmland walking trails in Ireland.


.

******: This is a working paper, do not cite without authors' permission**

## 1 Introduction

For many years the assumption of homogeneity in preferences has dominated the early literature on non-market valuation of recreational goods with a few exceptions (e.g., Morey, 1981; Morey et al., 1993). In his seminal paper, Train (1998) emphasized that the explicit recognition of taste heterogeneity is important in the estimation of recreational choice to avoid biased welfare results. As a consequence, choice models that can capture individual level taste heterogeneity are now commonplace in estimation of choice behaviour.

The wider literature on modelling choice behaviour has noted a further important type of heterogeneity, namely scale heterogeneity, also referred to as heteroskedasticity (e.g., Louviere et al., 1999; Louviere and Eagle, 2006). This refers to unobserved heterogeneity in variance associated with the random component of the utility. In recent years effort has been made to develop models capable of accommodating both scale and individual taste heterogeneity within the same model. The desire to accommodate both types of heterogeneity simultaneously recognises that, as noted by Thiene and Scarpa (2010), '*addressing only preference or scale heterogeneity negates the fact that true choice behaviour is likely to be in some middle ground with some variation attributable to scale and some to taste*'.

The identification of individual scale parameter is, however, problematic as this is equivalent to perfect positive correlation across all random parameters of the indirect utility function. Hess and Rose (2012) demonstrate how the attempts in the literature to disentangle scale heterogeneity from heterogeneity in individual coefficients in discrete choice models are misguided. In their paper, they show how the various model specifications (eg. using log-normal distributions or WTP-space models) can simply be seen as different parameterisations using more flexible distributions, rather than identifying individual scale heterogeneity.

Instead of comparing models assuming continuous distributions for the random parameters (as in Hess and Rose, 2012), this paper focuses on discrete distributions. In particular, we explore alternative specifications of the LC model to also incorporate heteroskedasticity in estimation of site-choice models.

While Magidson and Vermunt (2007) developed a scaled-adjusted LC model that accommodates between class preference heterogeneity and within and between class unobserved scale heterogeneity (i.e. specified probabilistically), we add to the literature by developing two further heteroskedastic Latent class (HLC) specifications. The first HLC model accommodates between class preference heterogeneity and within and between class observed scale heterogeneity[1]. We also estimate a more general version of the scale adjusted LC model that does not suffer from the constraint that the scale classes are of the same proportions across all taste classes. For comparison purposes we estimate a standard LC model.

Our case-study explores preferences for farmland walking trails in the Republic of Ireland. In Ireland, the majority of land is owned privately as farmland and property rights are defined in such a manner that recreational users do not have a de-facto legal right of entry and landowners can prohibit walkers from entering their land. As a result, Ireland does not have a network of well defined countryside walking opportunities and many of the recreational walking opportunities in the Irish countryside are limited to public roads. Additionally, national parks in Ireland and other public lands for recreation are relatively limited. Previous research suggests that substantial supply-side potential exists to develop a network of countryside walking trails. The present study seeks to establish demand side preferences for these trails.

[1]The Heteroskedasticity is specified with a discrete distribution based on known groups–namely rural and urban respondents–within the data

## 2 Background Literature

The literature on accommodating scale heterogeneity in random utility models can be characterised by authors who account for scale heterogeneity between known groups within the data and by others who model scale heterogeneity without the need to identify *a priori* groups. The seminal paper in this area was developed by Swait and Louviere (1993) who outlined and tested a modelling approach to represent scale heterogeneity in a MNL model with a discrete variable based on known groups. Since then, other advances in the area have been initiated through a number of studies. For example, Swait and Adamowicz (2001) specify the scale parameter as the ability to choose which is a function of choice task complexity and respondent effort. In the recreational literature, DeShazo and Fermo (2002) parameterise scale as a function of a number of measures of choice sets complexity to examine the variability across individuals induced by different experimental treatments in a choice experiment (CE). Another early example is given by Scarpa et al. (2003) who estimates scale heterogeneity as a function of traders' experience related to valuing cattle breeds.[2] Finally, Breffle and Morey (2000) proposed a modelling approach assuming homogeneity in tastes and allowing for scale heterogeneity through a continuous mixing distribution without the need to specify discrete groups.

Despite the development of these approaches some researchers have questioned the accuracy of just accommodating either only taste or scale heterogeneity since it is possible that what the researcher is interpreting as preference heterogeneity could in fact be scale heterogeneity and *vice versa* (e.g., Louviere et al., 1999; Louviere and Eagle, 2006). Recognising that human behaviour is likely to be characterised by a combination of both preference and scale heterogeneity has given rise to models that attempt to

[2] Other early attempts to incorporate scale heterogeneity include Hu et al. (2006)— reference point effects in demand analysis, Cameron and Englin (1997)— experience in contingent valuation of environmental goods, Brownstone et al. (2000)— revealed and stated preferences data in transport and Hanley et al. (2005)— price vector effects for CEs.

accommodate both simultaneously. The models include the generalised multinominal logit model (G-MNL) (Fiebeg and Wasi, 2010), the scale-adjusted LC model (Magidson and Vermunt, 2007; Campbell et al., 2011) and the WTP-space model (Train and Weeks, 2005).

More recently, with particular emphasis on G-MNL and WTP-space models, Hess and Rose (2012) argue that these efforts to separately identify random scale heterogeneity have been misguided. In particular, they base their argument on the fact that, econometrically, a linear in parameters specification of the logit model perfectly confounds scale with taste sensitivity. They note that models estimated in this manner simply allow for more flexible distributions, thus uncovering from the data particular correlation structures within the heterogeneity that is being modelled whilst maintaining the scale/taste sensitivity confound.

This paper focuses on alternative methods to accommodate scale heterogeneity within a LC framework. In this context, Magidson and Vermunt (2007) proposed a scale-adjusted LC model, allowing for the simultaneous accommodation of preference and unobserved scale heterogeneity. In their paper, unobserved scale heterogeneity is probabilistically described by a discrete mixing distribution (not based on *a priori* groups) within each latent class. We further explore this approach by estimating and comparing it with two HLC models, firstly developed within this paper. In particular, we develop a HLC model that enables us to specify scale differences with a discrete variable based on known groups within the data. This approach is more closely aligned to Swait and Louviere (1993) and it enables more straightforward identification of scale influences within our data. Using this approach we can determine on a per class basis whether a particular group exhibits higher or lower variance. For our second contribution to this literature we generalise the scale-adjusted LC model by relaxing the constraint that the scale classes need to be in the same proportion across all taste classes, which is a restrictive feature of the scale-adjusted LC model.

# 3 Methodology

## 3.1 Modelling data from discrete choice experiments

Statistical analyses of the responses obtained from DCEs are grounded in the Random Utility Maximization (RUM) theory (Thurstone, 1927; Manski, 1977), which assumes that the respondent's choices are driven by the maximization of his utility function:

$$U_{nit} = \beta' x_{nit} + \epsilon_{nit}. \tag{1}$$

Where $n$ indicates the respondent, $i$ the chosen alternative, $t$ the choice occasion, $x$ is a vector of attributes, $\beta$ is a vector of parameters to be estimated, and $\epsilon$ is a random component, unobserved by the researcher, assumed to be *iid* Gumbel distributed[3]. Given this functional form, the probability for the individual $n$ of choosing alternative $i$ over any other alternative $j$ in the choice set is defined as:

$$\Pr(i_{nt}) = \frac{\exp(\lambda \cdot \beta' x_{nit})}{\sum_{j=1}^{J} \exp\left(\lambda \cdot \beta' x_{njt}\right)}, \tag{2}$$

As under the MNL assumptions it is not possible to estimate both $\beta$ and $\lambda$ the latter is conveniently fixed to 1. It is possible to explore the issue of Scale heterogeneity by mean of the heteroscedastic MNL (HMNL) model as specified in Swait and Adamowicz (2001)[4]. Instead of assuming $\lambda = 1$ for everybody in the sample, the HMNL model relies on the assumption that the scale parameter is heterogeneous and its heterogeneity can depend upon observed characteristics of either respondents (e.g. Scarpa et al., 2003) or choice situations (e.g. DeShazo and Fermo, 2002). In their paper, Swait and Adamowicz (2001) assume that tastes are constant and that only the scale parameter ($\lambda$) varies across the sample. In their conclusions, they propose an extension based on an

[3]For a derivation of the model see Ben-Akiva and Lerman (1985)

[4]For more comments and details on the derivation, refer to Swait (2006).

exploration of simultaneous representation of taste and scale heterogeneity. Given the flexibility of LC models, now it is possible to accomplish this task of accommodating heterogeneity across scale and preferences by estimating HLC models as described in what follows.

## 3.2 Taste and scale heterogeneity: Heteroskedastic Latent class models

As traditionally estimated, LC models are a semi-parametric variant of the MNL model based on a finite mixing distribution. This type of models is based on the assumption that individuals can be implicitly sorted, up to a probability, into a set of $C$ classes, each of which is characterised by unique class-specific utility parameters, $\beta_c$, for the attributes in the choice sets. Given membership to class $c$, the probability of respondent $n$'s sequence of choices $y_n$ over the $T_n$ choice occasions (i.e., $y_n = \langle i_{n1}, i_{n2}, \ldots, i_{nT_n} \rangle$), is:

$$\Pr(y_n | c, x_{nit}) = \prod_{t=1}^{T_n} \frac{\exp(\beta_c' x_{nit})}{\sum_{j=1}^{J} \exp\left(\beta_c' x_{njt}\right)}. \tag{3}$$

Considering now the probability of membership to a class $c$ defined as $\pi_c$ (where $0 \leq \pi_c \leq 1 \; \forall c$ and $\sum_{c=1}^{C} \pi_c = 1$),[5] the probability of a sequence of choices is:

$$\Pr(y_n | x_n) = \sum_{c=1}^{C} \pi_c \left( \prod_{t=1}^{T_n} \frac{\exp(\beta_c' x_{nit})}{\sum_{j=1}^{J} \exp\left(\beta_c' x_{njt}\right)} \right). \tag{4}$$

The modelling approach proposed in what follows builds on the idea of a scale-adjusted LC model (Campbell et al., 2011; Magidson and Vermunt, 2007) and the HMNL model (Swait, 2006).

[5] Membership probability can be based only on a constant (Scarpa and Thiene, 2005) or be informed by socioeconomics covariates (Boxall and Adamowicz, 2002). In our paper, we follow the former approach in order to facilitate a more direct comparison between models, and we leave to further research the specification of heteroskedastic LC models informed by socioeconomics covariates.

Two HLC models are derived: the first accounts for observed heteroskedasticity (in which we assume that the analyst can observe the characteristics on which the heterogeneity in the scale factor depends), while the second accommodates unobserved (probabilistic) heteroskedasticity.

The first model, that we call Observed Heteroskedastic Latent class (ObsHLC), can be described as a LC built on a HMNL model as developed by Swait and Adamowicz (2001) therefore by allowing $\lambda$ estimated for an observed group of respondents. While Swait and Adamowicz (2001) proposed this within a MNL modelling framework, we extend this to a LC specification allowing for within class scale heterogeneity. As a result the choice probability becomes:

$$\Pr(y_n|x_n) = \sum_{c=1}^{C} \pi_c \left( \prod_{t=1}^{T_n} \frac{\exp(\lambda_c \cdot \beta_c' x_{nit})}{\sum_{j=1}^{J} \exp\left(\lambda_c \cdot \beta_c' x_{njt}\right)} \right). \tag{5}$$

The second model, that we call Probabilistic (or Unobserved) Heteroskedastic Latent class (ProbHLC) builds on the model developed by(Magidson and Vermunt, 2007) and further applied by Campbell et al. (2011). The model developed in this paper allows the researcher to relax the constraint requiring the same probability for the scale classes across all taste classes (eg. the Scale adjusted LC model (Magidson and Vermunt, 2007) allows to estimate for each class c, s classes with different scale parameters $\lambda_{cs}$, but in all classes those s scale classes must have the same membership probability, in our ProbHLC the scale classes have different membership probability in each taste class. Assuming that within each class there are $S$ classes with different scale factor, each associated with a scale membership probability $\pi_{cs}$, the model can be represented as:

$$\Pr(y_n|x_n) = \sum_{c=1}^{C} \pi_c \left[ \sum_{s=1}^{S} \pi_{cs} \left( \prod_{t=1}^{T_n} \frac{\exp(\lambda_{cs} \cdot \beta_c' x_{nit})}{\sum_{j=1}^{J} \exp\left(\lambda_{cs} \cdot \beta_c' x_{njt}\right)}, \right) \right], \tag{6}$$

where $\sum_{s=1}^{S} \pi_{cs} = 1$ and $0 \leq \pi_{cs} \leq 1$ $\forall s$ in each taste class $c$ and $\pi_c \leq 1$ $\forall c$ and $\sum_{c=1}^{C} \pi_c = 1$. Note that if $\pi_{cs}$ is the same over classes $\pi_{cs} = \pi_s$ the model in Equation 6 is a Scale Adjusted latent class (ScaleAdjLC).

As we are interested in how the scale parameter differs in each group from a baseline group (for which the scale factor is fixed to one to avoid specification problems), we specify $\lambda = 1 + \eta$, ($\eta$ being the difference in scale from the baseline group or class) subject to the constraint $\eta > -1$. We then estimate for each group how its scale parameter differs from the baseline group.[6]

The models were estimated with Pythonbiogeme (see Bierlaire, 2003, 2009) using maximum log-likelihood estimation procedures. In order to deal with the problem of local maxima in LC models, we used the CFSQP algorithm (Lawrence et al., 1997) and we run the estimations between 100 and 200 times (depending of the model) using random starting values.[7]

### 3.3 Welfare analysis

Given that one of the main objectives of environmental recreational site choice studies is the assessment of users' welfare, we develop compensating variation (CV, also referred to as Consumer surplus) estimates to determine users' welfare associated with specific policy changes.

For models that incorporate random variation, in order to compute both willingness to pay (WTP) and CV estimates it is necessary to obtain the individual-specific posterior estimates. As shown by Scarpa and Thiene (2005) for LC models, individual-specific posterior class probabilities can be computed using Bayes' theorem. They also illustrate how the individual-specific posterior parameter estimates can be computed

[6]This simplification, which allows to reduce the computational efforts of estimating the scale factor $\mu = exp(\lambda\eta)$, is possible as the software used for estimations (Biogeme) is very efficient in handling constraints on estimated parameters.

[7]This was coded in 'PERL' and used in combination with Pythonbiogeme run under Ubuntu 10.04 LTS - the Lucid Lynx. See Boeri (2011) for a more in-depth discussion of the software.

using the predicted class membership probabilities as weights of the average of the parameters over classes. Following Scarpa and Thiene (2005) therefore, it is possible to retrieve the individual-specific WTP.

Additionally once the individual-specific posterior parameters are retrieved, the CV, conditional to class membership, is simply another weighted average as follows:

$$
\begin{aligned}
\widehat{\mathrm{CV}}_n &= \sum_{c=1}^{c=C} \hat{\pi}_{nc} \widehat{\mathrm{CV}}_{nc} \\
&= \sum_{c=1}^{c=C} \hat{\pi}_{nc} \left\{ -\left(\hat{\beta}^{€}_{nc}\right)^{-1} \left[ \ln\left(\sum \exp\left(\hat{\lambda}\hat{\beta}'_{nc} X_n^1\right)\right) - \ln\left(\sum \exp\left(\hat{\lambda}\hat{\beta}'_{nc} X_n^0\right)\right)\right]\right\} \quad (7)
\end{aligned}
$$

where $X_n^0$ reflects the attribute levels for respondent $n$ prior to the policy change and $X_n^1$ is the level of the attributes after the policy change.

## 4 Survey design and data description

The paper uses data from a DCE used to elicit public preferences for the development of farmland walking trails in Ireland. Given the strong policy focus, the study design involved gathering opinions from a wide-range of stakeholders interested in addressing public access concerns within Ireland as well as conducting focus groups with members of the general public.

After extensive discussions with key stakeholders and insights from focus group participants, it was decided to use labels to reflect the diversity of farmland in Ireland, and, hence, the potential for diverse types of farmland walking trails. The labels reflected the main types of potential farmland walks that could be implemented at a national level: Hill, Bog, Field and River walks. In the final version of the questionnaire, five attributes were decided upon to describe the walking trails. These attributes were chosen on the basis of their choice relevancy to members of the general public as well as their suitability and relevance for farmland recreation. Questions were also included in the pilot and main questionnaires to explore respondents' acceptance of the choice scenarios presented to them.

The first attribute, 'Length', indicated the length of time needed to complete the walk from start to finish (all walks were described as looped (circular) so that people using the walks did not have to walk back along the same route). This attribute was presented with three levels with the shortest length between 1–2 hours, the medium length between 2–3 hours and the longest length between 3–4 hours. The levels of the Length attribute were presented using interval levels to reflect the fact that not everyone walks at the same pace. These levels were informed by discussions at focus groups as well as information on the current recreation walking activity of the Irish population. The second attribute, 'Car Park', was a dummy variable denoting the presence of car parking facilities at the walking trail. The third attribute, 'Fence', was a dummy variable used to indicate if the trail was fenced-off from livestock. This attribute only applied to the field and river walk alternatives, since these are the most likely types of walks where livestock would be encountered. The fourth attribute, 'Path and Signage', was a dummy variable to distinguish if the trail was paved and signposted. These three attributes represented the infrastructural features that were deemed important and realistic for farmland walking trails based on findings from the qualitative part of the study. The final attribute, 'Distance', denoted the one-way distance (in kilometres) that the walk is located from the respondent's home. The attribute was presented with six levels (5, 10, 20, 40, 80 and 160 kilometres) reflecting realistic distances that would be travelled in Ireland for a recreational day trip. This attribute was later converted to a 'Travel Cost' per trip using estimates of the cost of travelling by car from the Irish Automobile Association. Findings from focus group discussions indicated that this represented a conceptually realistic and acceptable payment mechanism.

Each respondent was asked to complete a panel of twelve choice tasks, which were constructed using a a Bayesian efficient design (Ferrini and Scarpa, 2007). For each task, respondents were asked to choose between a combination of the experimentally designed alternatives and a "stay at home" option. When making their choices, respondents were asked to consider only the information presented in the choice task and to treat each task separately. Respondents were further reminded that distant trails would be more costly in terms of their time and money.

[Figure 1 about here.]

The survey was administered to a sample of Irish residents in 2009 using face-to face interviews. A quota controlled sampling procedure was followed to ensure that the survey was nationally representative for the population aged 18 years and above. The quotas used were based on known population distribution figures for age, gender and region of residence taken from the Irish National Census of Population, 2006.

### 4.1 Urban-Rural exploration of scale heterogeneity

In the HMNL and ObsHLC models we accommodate scale heterogeneity based on *a priori* expectations of scale heterogeneity between known groups in the dataset. This requires that we have expectations regarding likely scale differences in the dataset between groups. In this case study, we exploit potential differences in scale between respondents who reside in either rural or urban locations.[8] Our focus on urban-rural differences reflects a priori expectations that these respondents may have different levels of familiarity and experience of countryside recreation in general and farmland recreation specifically. As a result we may expect that rural and urban residents may have different scale parameters. In addition, it is further possible that there may be differences within each subgroup (i.e. between rural residents or urban residents themselves)

[8]For the purpose of this case-study we define rural respondents as those who reside outside the main cities in Ireland and urban respondents as those who live in one of these cities. This classification reflects the ease with which respondents located outside the main cities can access farmland compared to their urban counterparts. The sample breakdown is 281 and 189 rural and urban respondents respectively.

in their scale parameters. [9] As we show below, we can use the ObsHLC model to account for these possibilities. While we restrict our analysis to exploring rural-urban differences our modelling approach can be applied to accommodate a host of factors that could cause scale differences, such as choice task complexity or respondent effort in stated preference studies (DeShazo and Fermo, 2002; Swait and Adamowicz, 2001), experience levels (Scarpa et al., 2003) or differences between revealed and stated preference data (Brownstone et al., 2000), for example.

## 5 Estimation results

In this section we present the results of models estimation. We firstly compare the models in terms of model fit and then we present the estimates from the best LC models specifications – all the LC models with 6 classes are reported in Table 2 and 3, however only results from ObsHLC and ScaleAdjLC are discussed for reasons of brevity.

All models are estimated from our balanced panel of 470 respondents, providing a total of 5,640 observations. The explanatory variables for choice probabilities are four dummy variable attributes: Length–over two hours, Car Park, Fence, Path and Signage (equal to one if the option contained, respectively, a walk longer than two hours, Car park facilities, fence and/ or path and signage, and zero otherwise), the 'Travel Cost' coefficient, and four alternative specific constant, one for each type of walk (hill, bog, field and river–the baseline being the "stay at home" option).

### 5.1 Model comparisons

In Table 1 we compare different LC models specifications. To decide the number of classes with different preferences, we use the information criteria discussed in Hurvich

[9] We should also expect differences in preferences between urban and rural respondents, that could be modelled introducing covariates in the class membership probabilities of our LC models, however this is not the focus of the present paper .

and Tsai (1989)[10].

[Table 1 about here.]

Based on the criteria repeated in Table 1 for each of the four types of LC models (namely LC, ObsHLC, ScaleAdjLC and ProbHLC), the best fit for the data is provided by the six classes specification of the three HLC, while the eight classes specification provides the best fit for the data within the context of the traditional LC model.[11] In addition by comparing the goodness of fit criteria across the four groups of LC models, we notice that the specifications with 6 classes for both ObsHLC and ScaleAdjLC outperform the LC and the ProbHLC models. Additionally we note that the increased number of parameters estimated in the ProbHLC (compared to the ScaleAdjLC ) are not justified by its increased flexibility and potential ability to explain the data more accurately. However, this is likely to be data specific and therefore, an open empirical question. For the ObsHLC and the ScaleAdjLC model specifications the BIC and crAIC suggest that both fit the data in a similar manner.

## 5.2 Model estimates

To reduce the number of tables we do not present the results from the MNL and HMNL models, however, our results from both models confirm that respondents have negative and significant coefficients associated with the travel cost attribute and for length of walks that are over two hours duration. Additionally respondents show positive and significant preferences towards the coefficients representing carpark, fence and path

[10] The information criteria statistic (C) is specified as $-2lnL + K\delta$ where $lnL$ is the log likelihood of the model at convergence, $K$ is the number of estimated parameters in the model, and $\delta$ is a penalty constant. There are a number of different types of information criteria that can be employed depending on the value taken by the penalty constant $\delta$. For $\delta = 3$ we have the Akaike Information Criteria (AIC); for $\delta = lnL(N)$ we obtain the Bayesian Information Criteria and finally, for $\delta = 2 + 2(K+1)(K+2)/(N-K-2)$ we have the corrected AIC (crAIC), which increases the penalty for the number of extra parameters estimated (Hynes et al., 2008).

[11] However, inspection of the model highlights many non-significant coefficients and membership probabilities. We further tested models specifications with 9 classes, but we found these models to be problematic because of identification problems and in some instances the models did not converge.

and signage attributes. The coefficients representing the walk alternatives are positive albeit the coefficient for bog walk is not significant. Our HMNL model, which accommodates observed scale heteroegenity, reveals that respondents from rural areas are estimated to have lower variance when making their choices, resulting in a higher scale parameter, compared to urban respondents in this model.

Estimates from LC models are presented in Table 2 and 3. As the ObsHLC and ScaleAdjHLC models with 6 classes respectively represented the best models based on certain IC criteria, our discussion and analyses focuses on these specifications.

[Table 2 about here.]

[Table 3 about here.]

In the ObsHLC model –presented in Table 2– a negative coefficient is associated with the longer length walk attribute except for classes four and five (where the coefficient is not significant for these two classes). This suggests that Irish residents have a strong preference for walks that are generally of a shorter duration, which most likely reflects the broader walking patterns of Irish residents. For instance, Curtis (2002) in a survey of the recreational behaviour of the Irish population, has shown that most Irish people who participate in recreational walking, mostly engage in relatively shorter walks. The travel cost coefficient is negative as expected and is significant in all but one class (Class 5).

In the case of class one, the 20 percent of the sample who are assigned to this class have positive and significant preferences for the carpark, fence and path and signage attributes as well as the walk alternatives. This class may characterise people who like walking, but generally prefer quite a structured walk with facilities. Respondents in class two dislike longer length walks and walks that are located further from their homes but generally do not care for facilities. Respondents assigned to this class also show positive and significant preferences for the walk alternatives. These results suggest that respondents in the first two classes have preferences for all the walk alternatives (to various degrees) but generally differ in which attributes they care for. For classes 3-6, there is some variation in terms of which attributes and walk alternatives that respondents show a preference for. Respondents in class 3 have significant preferences for river walks compared to the stay at home option. The other walk alternatives are not significant nor are the coefficients representing the Fence and Path and Signage attributes. Most of the coefficients representing the attributes in class 4 are not significant at the five per cent level, while all the coefficients representing the walk alternatives are significant at the 10 percent level and the travel cost coefficient is marginally below the five percent significance threshold. Respondents in class 5 have a significant coefficient for the path and signage attribute and for the river and field walk alternatives only. For class 6, the coefficients representing the longer length attribute, the travel cost coefficient and the hill walk are significant. In terms of the probability of class membership the class sizes are relatively evenly distributed, with no one particularly large class or very small class.

The ObsHLC model can be viewed as a natural extension to the HMNL. An advantage of this model is that it is possible to determine on a per class basis whether rural or urban respondents exhibit higher or lower variance. In addition, we can explore the added dimension of preference heterogeneity within this framework compared to the HMNL model. Using this approach, we find that rural respondents in class 3 have a higher scale parameter implying less uncertainty (lower variance) in their preferences. This class exhibit a scale parameter for rural respondents similar to the outcome from the HMNL model. Class 2 and 4 do not exhibit significant scale heterogeneity between rural and urban respondents assigned to this class. Whereas classes 1, 5 and 6 are associated with higher variance (lower scale parameter) for rural respondents assigned to these classes. Potentially, the proportion of rural respondents assigned in one of these classes may be more aware of present difficulties associated with accessing farmland for recreation in Ireland. As a result, this could be reflected by more uncertainty (and higher variance) regarding their preferences for farmland walking trails.

Results from the ScaleAdjLC model are presented in Table 3. While we do not go into a detailed overview of the model, a number of features are noteworthy. First, while the model produces similar class probabilities compared to the ObsHLC model, there are some differences in terms of the estimated coefficients. For instance, there is a class (class 2) in this model where the coefficients associated with all the walk types are all negative relative to the stay at home option, albeit only the coefficient associated with the bog walk alternative is significant. In addition, similar to the previous model, the coefficients for the attributes are significant in some classes and not in others. This is a relatively common finding in LC models, as it reflects the heterogeneity in preferences between respondents in different classes. The interpretation of scale heterogeneity is arguably less meaningful in the ScaleAdjLC model than in the ObsHLC model. This is because we cannot state what factors may influence this heterogeneity within and between classes. We are only able to determine what proportion of respondents are estimated to have higher or lower variance in each class. In class 1 and 3, for instance, we find that approximately 40 per cent of respondents have a lower scale parameter compared to the remaining respondents in these classes. In Class 4 and 6 we do not find any significant differences in scale heterogeneity.

## 5.3 Post-estimations and policy implications

### *5.3.1 Welfare analysis*

In Table 4 WTP for each attribute of the different types of walk and CV for four policy scenarios based on individual-specific posterior parameters are shown for the models. We report a range of measures to illustrate the heterogeneity associated with WTP in this data. We present median WTP estimates along with the mean estimates as our mean estimates appear to be influenced by outliers. Based on the ObsHLC model, respondents have a median WTP of €11 to avoid walking trails that are over two hours duration. This translates into a willingness to travel further distances to access shorter walking trails. Respondents are willing to pay most for the path and signange attribute and least for using trails that are fenced-off from livestock. Similar estimates are retrieved from the models for the WTP associated with the walk alternatives. In terms of model comparison, there are noticeable differences between the ProbHLC model and the other LC specifications in terms of the retrieved WTP estimates. For instance in the case of the length and the path and signage attributes the median WTP estimates are approximately 40 and 80 per cent lower respectively in the ProbHLC model compared to the ObsHLC model.

[Table 4 about here.]

Examining the change in quality or quality of an environmental good is an important policy consideration. As a result we estimate the CV associated with four scenarios. The hypothetical scenarios that we consider are:

- Scenario 1: No car park is provided close to any of the farmland walk;
- Scenario 2: No path and signage is provided for any of the farmland walk;
- Scenario 3: There is always a car park provided close to all the farmland walk;
- Scenario 4: There is always a path and signage available at all the farmland walk.

Our scenarios are useful in the context of farmland trails because it provides policy-makers with information on the welfare impacts associated with the potential design of trails. For instance, policy-makers can trade off the welfare impacts associated with providing a path and signage at the walks with the cost of providing these facilities.

The results in Table 4 reveal that the scenarios associated with not having a car park or a path and signage have a negative impact on respondents' welfare, while the latter two cases, related to having a car park and a path and signage leads to positive overall welfare effects. A noteworthy finding is that the magnitudes of the welfare effects of the first two scenarios are larger than the welfare effects of the latter two scenarios. This suggests some evidence of loss aversion behaviour whereby respondents' place a greater welfare impact from not having these facilities compared to the welfare effect of having these facilities.

#### 5.3.2 Direct Elasticities

As a final investigation we derive direct choice elasticities for the walk alternatives. Given a one percent change in the level of an attribute, the choice elasticities provide the percentage change in the probability of choosing the type of walk characterised by that specific attribute.

[Table 5 about here.]

Direct elasticity values are computed for each attribute (longer length , car park, fence and path and signage) for each type of walk (hill, bog, field and river). Results based on individual-specific posterior parameters retrieved from the models are reported in Table 5. The elasticity of demand for the longer length attribute is relatively more elastic across the walk alternatives. This suggests that respondents' choice of a particular walk changes relatively more as a result of a change in the level of the length attribute. Additionally, the elasticities for this attribute are quite different across the walk alternatives. For instance demand for the bog walk alternative would be most responsive to changes in the length attribute, while demand for the river walk alternative would be least responsive. For the other attributes demand is relatively inelastic. For example, demand for the walk alternatives would not change substantially by changes in the level of car park facilities. Since these attributes were presented as dummies, this suggests that demand for a walk type does not change much based on whether a car park is provided or not, even though welfare estimates are impacted by the provision of these facilities. We also note that there are some differences in the retrieved elasticities particularly for the longer length and car park attributes across the models.

## 6 Discussion and conclusions

This paper examined and compared alternative ways of incorporating scale heterogeneity in a LC modelling framework using data on site-choice for farmland walking trails in Ireland. We contrasted a modelling approach that only incorporates preference heterogeneity, namely the standard LC model, with representations which incorporate scale heterogeneity, namely the ObsHLC, the ScaleAdjLC and ProbHLC models. We specifically examined and compared the impacts of assuming a discrete representation of scale based on known groups in the data (ObsHLC model) as well as to the case were the scale heterogeneity is unobserved and probabilistically assigned within classes (ScaleAdjLC and ProbHLC models). This is the first study to undertake such a comparison. It is also the first study to incorporate observed scale heterogeneity in a LC model and to develop a more flexible scale-adjusted LC model – the ProbHlc model. Our focus on a LC specification stems from the fact that the model has a number of advantages in the recreational site choice context, through allowing us to identify discrete groups with particular recreational preferences, which in the case of recreational goods is highly useful (Hynes et al., 2008).

The focus of the study fits within the wider discrete choice literature that attempts to accommodate both scale and taste heterogeneity in discrete choice models, assuming discrete distribution of both scale and taste heterogeneity. In terms of the overall comparison of models we note that the ObsHLC and the ScaleAdjLC represented the best fit for this data based on a number of information criterion. While we acknowledge that the ProbHLC model represents a more general model, as it does not require the same proportion of respondents in each scale class and therefore represents a substantial advance to both the traditional LC model and the scale adjucted LC model developed by Magidson and Vermunt (2007), we do not find any evidence, in the case of this study, the extra parameters were not warranted and did not improve the fit of the model. We do believe, however, that is likely to be data specific and an open empirical question.

Based on our analysis we find there are a number of benefits associated with specifying a discrete distribution for scale based on known groups within the data, as in the ObsHLC model. While the information criterion did not suggest that this outperformed the ScaleAdjLC. However beyond just comparing model fit, we believe that the key advantage of this model is its ability to identify the influence of class variability. Therefore, from a policy context this model can be used to inform on potential differences that exist between (groups of) individuals estimated to have homogeneous preferences (within class), and between individuals estimated to have heterogeneous preferences (between classes), in the variance associated with their choice behaviour. Additionally, incorporating observed scale heterogeneity into the estimation also aligns more closely with the concept of scale heterogeneity as was originally conceived by Swait and Louviere (1993). Further additions to the literature can be incorporated within this approach in particular through the inclusion of covariates to potentially drive class membership probabilities in these models, as is commonplace with the traditional LC model.

From a policy context it is evident that substantial heterogeneity exists both in the preferences that individuals hold for farmland walking trails. In particular we found significant taste heterogeneity associated with the attributes and alternative walking trails. In terms of our post-welfare analysis we determined that welfare is most negatively impacted by trails that are longer in length (over 2 hours) and is most positively impacted with the provision of a path and signage at the trails. Based on the median elasticity estimates, we also noted that demand for the walking trails is relatively unresponsive to whether a car park, fence or path and signage are provided at the walks and is relatively more elastic to whether the length of the walk is less than or greater than two hours.

These results lead us to conclude that implementation of a policy should take into consideration that variety of trails would be needed to accommodate heterogeneous preferences. Our findings further confirm that residents have preferences for car parking facilities close to the walking trails but these facilities would not be a requirement at every walking trail. The same conclusion holds for the path and signage attribute. In terms of the walk alternatives, relative to the stay at home option, residents show positive preferences for all the walk alternatives. Our welfare results confirm that river walks and hill walks are slightly preferred but the differences in preferences between the four walk alternatives are not substantial.

We also observed significant heterogeneity in the variance associated with choice behaviour. In the context of this study we observed differences in variance both between and within rural and urban respondents. This may reflect differences in familiarity, and their experiences related to countryside recreation including farmland recreation. While not specifically explored within this study, rural and urban residents may also differ in their individual tastes for the attributes and alternative walking trails.

Despite the substantial heterogeneity surrounding both preferences and variance our overarching finding is that significant demand side potential exists for the provision of farmland walking trails in Ireland. Given that the vast majority of land in Ireland is privately owned as farmland, this suggests that a wide provision of trails could be provided throughout the Irish countryside that would satisfy residents preferences. On the other hand, the recent improved profitability of farmland in Ireland as an agricultural resource may impede the supply-side realisation of the non-market benefits associated with farmland walking trails.

Overall, based on our analysis, the results indicate that significant demand exists among Irish residents for the provision of the farmland walking trails. farmland walk would have positive welfare benefit.

## References


Ben-Akiva, M. and Lerman, S. (1985). *Discrete choice analysis: theory and application to travel demand*, MIT Press, Cambridge, Massahusetts, U.S.A.

Bierlaire, M. (2003). BIOGEME: a free package for the estimation of discrete choice models, Proceedings of the $3^{rd}$ Swiss Transport Research Conference, Monte Verita, Ascona, Switzerland.

Bierlaire, M. (2009). *An introductory tutorial to BIOGEME Version 1.8*, Monte Verita, Ascona, Switzerland.

Boeri, M. (2011). *Advances in Stated Preference Methods: Discrete and continuous Mixing distributions in Logit models for representing variance and taste heterogeneity*, PhD thesis, Queen's University of Belfast, School of Biological Sciences - Gibson Institute for Land, Food and the Environment.

Boxall, P. C. and Adamowicz, W. L. (2002). Understanding heterogeneous preferences in random utility models: a latent class approach, *Environmental and Resource Economics* **23**(4): 421–446.

Breffle, W. S. and Morey, E. R. (2000). Investigating Preference Heterogeneity in a Repeated Discrete-Choice Recreation Demand Model of Atlantic Salmon Fishing, *Marine Resource Economics* **15**: 1–20.

Brownstone, D., Bunch, D. S. and Train, K. E. (2000). Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles, *Transportation Research Part B Methodological* **34**(5): 315–338.

Cameron, T. A. and Englin, J. (1997). Respondent experience and contingent valuation of environmental goods, *Journal of Environmental Economics and Management* **33**(3): 296–313.
**URL:** *http://ideas.repec.org/a/eee/jeeman/v33y1997i3p296-313.html*

Campbell, D., Hensher, D. A. and Scarpa, R. (2011). Non-attendance to attributes in environmental choice analysis: a latent class specification,. Forthcoming in Journal of Environmental Planning and Management. Accepted in September 2010..

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

DeShazo, J. R. and Fermo, G. (2002). Designing choice sets for stated preference Methods: the effects of complexity on choice consistency, *Journal of Environmental Economics and Management* **44**: 123–143.

Ferrini, S. and Scarpa, R. (2007). Designs with a-priori information for nonmarket valuation with choice-experiments: a {Monte Carlo} study, *Journal of Environmental Economics and Management* **53**(3): 342–363.

Fiebeg, D.G., K. M. L. J. and Wasi, N. (2010). The generalized multinomial logit:

accounting for scale and coefficient heterogeneity, *Marketing Science* **29(3)**: 393–421.

Hanley, N., Adamowicz, W. and Wright, R. E. (2005). Price vector effects in choice experiments: an empirical test, *Resource and Energy Economics* **27**(3): 227–234.

Hess, S. and Rose, J. (2012). Can scale and coefficient heterogeneity be separated in random coefficients models?, *Transportation* **forthcoming**.

Hu, W., Adamowicz, W. and Veeman, M. M. (2006). Labeling Context and Reference Point Effects in Models of Food Attribute Demand, *American Journal of Agricultural Economics* **88**(4): 1034–1049.
**URL:** *http://www.blackwell-synergy.com/doi/abs/10.1111/j.1467-8276.2006.00914.x*

Hurvich, M. and Tsai, C. (1989). Regression and time series model selection in small samples, *Biometrika* pp. 297–307.

Hynes, S., Hanley, N. and Scarpa, R. (2008). Effects on welfare measures of alternative means of accounting for preference heterogeneity in recreational demand models, *American Journal of Agricultural Economics* .

Lawrence, C., Zhou, J. and Tits, A. (1997). User's Guide for CFSQP Version 2.5: A C Code for Solving (Large Scale) Constrained Nonlinear (Minimax) Optimization Problems, Generating Iterates Satisfying All Inequality Constraints, *Technical Report TR-94-16r1*, Institute for Systems Research, University of Maryland, College Park, MD 20742, 1997.

Louviere, J. and Eagle, T. (2006). Confound it! that pesky little scale constant messes up our convenient assumptions., *Sawtooth Software Conference*.

Louviere, J., Meyer, R., Bunch, D., Carson, R., Dellaert, B., Hanemann, W., D.A., H. and Irwin, J. (1999). Combining sources of preference data for modelling complex decision processes, *Marketing Letters* **13**: 177–193.

Magidson, J. and Vermunt, J. (2007). Removing the scale factor confound in multinomial choice models to obtain better estimates of preference, *Procedings of the Sawtooth Software Conference*.

Manski, C. (1977). The structure of random utility models, *Theory and Decision* **8**: 229–254.

Morey, E. (1981). The demand for site-specific recreational activities: A characteristics approach, *Journal of Environmental Economics and Management* **8**(4): 345–371.
**URL:** *http://ideas.repec.org/a/eee/jeeman/v8y1981i4p345-371.html*

Morey, E., Rowe, R. and Watson, M. (1993). A Repeated Nested-Logit Model of Atlantic Salmon Fishing, *American Journal of Agricultural Economics* **75**: 578–592.

Scarpa, R., Ruto, E., Kristjanson, P., Radeny, M. and Rege, A. D. J. (2003). Valuing indigenous cattle breeds in Kenya: an empirical comparison of stated and revealed preference value estimates, *Ecological Economics* **45**(3): 409–426.

Scarpa, R. and Thiene, M. (2005). Destination choice models for rock climbing in the Northeastern Alps: a latent-class approach based on intensity of preference., *Land and Economics* **81**(3): 426–444.

Swait, J. (2006). Commentary on Econometric Modeling Strategies for Stated Preference Experiments By David Layton, *Environmental and Resource Economics* **34**(1): 87–90.

Swait, J. D. and Adamowicz, W. (2001). Choice environment, market complexity and consumer behaviour: a theoretical and empirical approach for incorporating decision complexity into models of consumer choice, *Organizational Behaviour and Human Decision Processes* **86**: 141–167.

Swait, J. D. and Louviere, J. J. (1993). The role of the scale parameter in the estimation and comparison of multinomial logit models, *Journal of Marketing Research* **30**: 305–314.

Thiene, M. and Scarpa, R. (2010). An empirical investigation of individual wtps within couples under scale and taste heterogeneity: the case of household water, *World Congress of Environmental and Resource Economists, 28 June - 2 July*.

Thurstone, L. (1927). A law of comparative judgment, *Psychological Review* **34**: 273–286.

Train, K. (1998). Recreation demand models with taste differences over people, *Land Economics* **74**(2): 230–239.

Train, K. and Weeks, M. (2005). Discrete choice models in preference space and willing-to-pay space, *in* R. Scarpa and A. Alberini (eds), *Applications of simulation methods in environmental and resource economics*, Springer, Dordrecht.

## List of Figures

| | Bog walk | Field walk | River walk | Choose none |
|---|---|---|---|---|
| Length | 3–4 hours | 1–2 hours | 1–2 hours | I would not choose any of the walks. I would stay at home. |
| Car Park | Yes | No | No | |
| Fence | — | No | Yes | |
| Path and Signage | Yes | No | Yes | |
| Distance | 40 km | 10 km | 80 km | |



Figure 1: Example Choice Card

# List of Tables

Table 1: Comparison between LC models using different criteria for Number of Classes*

| **Model** | **LogLik.** | **K** | $\bar{\rho}^2$ | $\chi^2$ | **BIC** | **AIC** | **3AIC** | **crAIC** |
|---|---|---|---|---|---|---|---|---|
| MNL | -6,882.982 | 9 | 0.101 | 1,561.084 | 13,843.703 | 13,783.964 | 13,792.964 | 13,784.316 |
| HMNL | -6,863.299 | 10 | 0.103 | 1,600.450 | 13,812.974 | 13,746.598 | 13,756.598 | 13,747.067 |
| 2LC | -5,931.986 | 19 | 0.223 | 3,463.076 | 12,028.087 | 11,901.972 | 11,920.972 | 11,904.812 |
| 3LC | -5,521.889 | 29 | 0.276 | 4,283.270 | 11,294.270 | 11,101.778 | 11,130.778 | 11,111.395 |
| 4LC | -5,410.646 | 39 | 0.289 | 4,505.756 | 11,158.160 | 10,899.292 | 10,938.292 | 10,922.139 |
| 5LC | -5,308.935 | 49 | 0.301 | 4,709.178 | 11,041.114 | 10,715.870 | 10,764.870 | 10,760.583 |
| 6LC | -5,221.007 | 59 | 0.311 | 4,885.034 | 10,951.635 | 10,560.014 | 10,619.014 | 10,637.426 |
| 7LC | -5,173.348 | 69 | 0.316 | 4,980.352 | 10,942.693 | 10,484.696 | 10,553.696 | 10,607.853 |
| 8LC | -5,123.368 | 79 | **0.321** | **5,080.312** | **10,929.110** | **10,404.736** | **10,483.736** | **10,588.913** |
| 2ObsHLC | -5,903.298 | 21 | 0.227 | 3,520.452 | 11,987.986 | 11,848.596 | 11,869.596 | 11,852.380 |
| 3ObsHLC | -5,503.292 | 32 | 0.278 | 4,320.464 | 11,282.988 | 11,070.584 | 11,102.584 | 11,083.393 |
| 4ObsHLC | -5,349.495 | 43 | 0.296 | 4,628.058 | 11,070.408 | 10,784.990 | 10,827.990 | 10,815.424 |
| 5ObsHLC | -5,243.398 | 54 | 0.309 | 4,840.252 | 10,953.229 | 10,594.796 | 10,648.796 | 10,654.366 |
| 6ObsHLC | -5,172.844 | 65 | 0.317 | 4,981.360 | **10,907.135** | 10,475.688 | 10,540.688 | **10,578.839** |
| 7ObsHLC | -5,146.949 | 76 | 0.318 | 5,033.150 | 10,950.359 | 10,445.898 | 10,521.898 | 10,610.032 |
| 8ObsHLC | -5,107.203 | 87 | **0.322** | **5,112.642** | 10,965.881 | **10,388.406** | **10,475.406** | 10,633.906 |
| 2ScaleAdjLC | -5,803.145 | 22 | 0.240 | 3,720.758 | 11,796.318 | 11,650.290 | 11,672.290 | 11,654.615 |
| 3ScaleAdjLC | -5,418.23 | 33 | 0.289 | 4,490.588 | 11,121.502 | 10,902.460 | 10,935.460 | 10,916.472 |
| 4ScaleAdjLC | -5,311.541 | 44 | 0.301 | 4,703.966 | 11,003.138 | 10,711.082 | 10,755.082 | 10,743.645 |
| 5ScaleAdjLC | -5,225.556 | 55 | 0.311 | 4,875.936 | 10,926.182 | 10,561.112 | 10,616.112 | 10,624.003 |
| 6ScaleAdjLC | -5,168.404 | 66 | 0.317 | 4,990.240 | **10,906.892** | 10,468.808 | 10,534.808 | **10,576.739** |
| 7ScaleAdjLC | -5,164.198 | 77 | 0.316 | 4,998.652 | 10,993.494 | 10,482.396 | 10,559.396 | 10,653.039 |
| 8ScaleAdjLC | -5,125.094 | 88 | **0.320** | **5,076.860** | 11,010.300 | **10,426.188** | **10,514.188** | 10,680.199 |
| 2ProbHLC | -5,798.735 | 23 | 0.240 | 3,729.578 | 11,796.136 | 11,643.470 | 11,666.470 | 11,648.385 |
| 3ProbHLC | -5,414.511 | 35 | 0.289 | 4,498.026 | 11,131.339 | 10,899.022 | 10,934.022 | 10,915.663 |
| 4ProbHLC | -5,307.494 | 47 | 0.301 | 4,712.060 | 11,020.957 | 10,708.988 | 10,755.988 | 10,748.532 |
| 5ProbHLC | -5,221.775 | 59 | 0.311 | 4,883.498 | 10,953.171 | 10,561.550 | 10,620.550 | 10,638.962 |
| 6ProbHLC | -5,166.516 | 71 | 0.317 | 4,994.016 | **10,946.304** | 10,475.032 | 10,546.032 | **10,609.099** |
| 7ProbHLC | -5,133.720 | 83 | 0.319 | 5,059.608 | 10,984.364 | 10,433.440 | 10,516.440 | 10,646.805 |
| 8ProbHLC | -5,105.390 | 95 | **0.321** | **5,116.268** | 11,031.356 | **10,400.780** | **10,495.780** | 10,719.972 |

* The full output from each version of LC model with 6 classes (LC, ObsHLC, ScaleAdjLC and ProbHlc) are reported in the next tables. The outputs from the other models not reported in the paper are available from the authors' upon request.

Table 2: LC and ObsHLC estimates, 5,640 observations.

| | Class 1 | | Class 2 | | Class 3 | | Class 4 | | Class 5 | | Class 6 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | est. | \|*t*-rat.\| | est. | \|*t*-rat.\| | est. | \|*t*-rat.\| | est. | \|*t*-rat.\| | est. | \|*t*-rat.\| | est. | \|*t*-rat.\| |
| **LC** | | | | | | | | | | | | |
| $\beta_{\text{Length (over 2 hours)}}$ | -0.395 | 4.68 | -0.337 | 2.8 | -3.1 | 12.19 | -1.05 | 6 | -1.56 | 8.19 | -0.453 | 2.49 |
| $\beta_{\text{Car Park}}$ | 0.107 | 1.35 | 0.659 | 5.67 | 0.86 | 4.32 | 0.671 | 4.21 | 0.0348 | 0.24 | 0.1 | 0.65 |
| $\beta_{\text{Fence}}$ | 0.0661 | 0.67 | 0.232 | 1.55 | 0.194 | 0.86 | 0.285 | 1.63 | 0.388 | 2.04 | 0.102 | 0.69 |
| $\beta_{\text{Path and Signage}}$ | 0.21 | 2.35 | 1.13 | 7.85 | 0.0569 | 0.28 | -0.188 | 1.07 | 0.614 | 3.6 | 0.259 | 1.58 |
| $\beta_{\text{cost}}$ | -0.0107 | 5.29 | -0.151 | 14.03 | -0.221 | 6.42 | -0.308 | 12.96 | -0.0264 | 4.93 | -0.00227 | 0.71 |
| $\beta_{\text{hill}}$ | 2.82 | 11.04 | 2.62 | 9.38 | -0.28 | 0.96 | 2.57 | 9.62 | 0.271 | 1.16 | 0.233 | 0.53 |
| $\beta_{\text{bog}}$ | 1.83 | 6.77 | 2.39 | 8.5 | -0.0266 | 0.08 | 2.11 | 7.9 | -0.576 | 2.37 | 0.158 | 0.35 |
| $\beta_{\text{field}}$ | 2.35 | 8.13 | 2.52 | 8.84 | 0.294 | 0.93 | 2.58 | 8.82 | 0.0548 | 0.23 | 1.81 | 4.91 |
| $\beta_{\text{river}}$ | 2.71 | 10.49 | 2.59 | 8.84 | 0.784 | 2.47 | 3.52 | 11.1 | 0.177 | 0.72 | 3.21 | 9.39 |
| $\pi_{\text{class}}$ | 0.2253 | | 0.1895 | | 0.1724 | | 0.1662 | | 0.1241 | | 0.1225 | |
| $\mathcal{L}(\hat{\beta}) = -5,221.007$ | | | | | | | | | | | | |
| **obsHLC** | | | | | | | | | | | | |
| $\beta_{\text{Length (over 2 hours)}}$ | -0.514 | 3.54 | -0.449 | 4.33 | -1.97 | 5.5 | -0.0167 | 1.53 | -0.37 | 1.88 | -2.74 | 6.83 |
| $\beta_{\text{Car Park}}$ | 0.823 | 5.66 | 0.0223 | 0.24 | 0.615 | 3.47 | 0.0107 | 1.43 | 0.123 | 0.78 | 0.142 | 0.61 |
| $\beta_{\text{Fence}}$ | 0.341 | 2 | 0.0445 | 0.39 | 0.193 | 1.3 | 0.00354 | 0.74 | 0.0981 | 0.65 | 0.727 | 2.6 |
| $\beta_{\text{Path and Signage}}$ | 1.35 | 6.19 | 0.101 | 0.9 | 0.12 | 0.9 | -0.00599 | 1.08 | 0.457 | 2.7 | 0.44 | 1.85 |
| $\beta_{\text{cost}}$ | -0.186 | 9.78 | -0.0164 | 5.86 | -0.183 | 7.83 | -0.00728 | 1.94 | -0.00266 | 0.77 | -0.0481 | 5.7 |
| $\beta_{\text{hill}}$ | 2.87 | 9 | 4.58 | 7.72 | -0.182 | 0.85 | 0.0539 | 1.62 | 0.107 | 0.22 | 0.913 | 2.71 |
| $\beta_{\text{bog}}$ | 2.65 | 8.21 | 3.13 | 5.95 | 0.0678 | 0.29 | 0.0472 | 1.62 | 0.193 | 0.38 | -0.557 | 1.33 |
| $\beta_{\text{field}}$ | 2.67 | 8.47 | 3.87 | 6.95 | 0.25 | 1.17 | 0.0543 | 1.62 | 2.3 | 6.54 | 0.598 | 1.69 |
| $\beta_{\text{river}}$ | 2.83 | 8.47 | 4.51 | 7.74 | 0.581 | 2.8 | 0.0778 | 1.98 | 3.5 | 9.75 | 0.592 | 1.58 |
| $\eta_{\text{urban}}$ | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| $\eta_{\text{rural}}$ | -0.203 | 1.97 | -0.0464 | 0.36 | 0.63 | 2.26 | 45.6 | 1.59 | -0.743 | 11.14 | -0.263 | 2.07 |
| $\pi_{\text{class}}$ | 0.1982 | | 0.1840 | | 0.1797 | | 0.1660 | | 0.1587 | | 0.1134 | |
| $\mathcal{L}(\hat{\beta}) = -5,172.844$ | | | | | | | | | | | | |

Note: Observed heteroscedasticity is retrieved based on respondents from Rural and Urban (Baseline) areas.

Table 3: ProbHLC and ScaleAdjLC estimates, 5,640 observations.

| | Class 1 | | Class 2 | | Class 3 | | Class 4 | | Class 5 | | Class 6 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | est. | \|*t*-rat.\| | est. | \|*t*-rat.\| | est. | \|*t*-rat.\| | est. | \|*t*-rat.\| | est. | \|*t*-rat.\| | est. | \|*t*-rat.\| |
| **ScaleAdjLC** | est. | \|*t*-rat.\| | est. | \|*t*-rat.\| | est. | \|*t*-rat.\| | est. | \|*t*-rat.\| | est. | \|*t*-rat.\| | est. | \|*t*-rat.\| |
| $\beta_{\text{Length (over 2 hours)}}$ | -0.43 | 3.11 | -1.48 | 5.43 | -0.561 | 3.19 | -0.424 | 2.57 | -0.877 | 5.07 | -2.42 | 4.65 |
| $\beta_{\text{Car Park}}$ | 0.72 | 5.19 | 0.365 | 4.44 | 0.231 | 1.51 | 0.0029 | 0.03 | 0.591 | 3.89 | -0.0333 | 0.11 |
| $\beta_{\text{Fence}}$ | 0.233 | 1.32 | 0.217 | 2.49 | 0.159 | 1.06 | 0.045 | 0.38 | 0.256 | 1.67 | -0.529 | 1.37 |
| $\beta_{\text{Path and Signage}}$ | 1.36 | 6.99 | 0.0704 | 1.05 | 0.434 | 2.55 | 0.0886 | 0.72 | -0.0697 | 0.47 | 0.0547 | 0.17 |
| $\beta_{\text{cost}}$ | -0.178 | 12.33 | -0.0272 | 5.06 | -0.00234 | 0.72 | -0.0148 | 2.82 | -0.251 | 6.88 | -0.667 | 3.68 |
| $\beta_{\text{hill}}$ | 3.19 | 9.69 | -0.146 | 1.64 | 0.247 | 0.48 | 4.44 | 3.16 | 2.2 | 6.33 | 1.29 | 1.73 |
| $\beta_{\text{bog}}$ | 2.91 | 8.92 | -0.722 | 3.35 | 0.614 | 1.54 | 2.86 | 2.93 | 1.82 | 5.6 | 2.93 | 3.7 |
| $\beta_{\text{field}}$ | 3.14 | 8.88 | -0.168 | 1.6 | 2.18 | 6.12 | 3.67 | 3.1 | 2.19 | 6.12 | 2.85 | 3.23 |
| $\beta_{\text{river}}$ | 3.13 | 9.02 | -0.0763 | 0.78 | 3.45 | 9.5 | 4.29 | 3.16 | 3 | 6.44 | 3.39 | 3.9 |
| $\eta_1$ | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| $\eta_2$ | -0.647 | 7.13 | 3.84 | 3.99 | -0.752 | -12.23 | 0.000 | 0 | 0.647 | 1.70 | 0.000 | 0 |
| $\pi_{s_1}$ | 0.5965 | | | | | | | | | | | |
| $\pi_{s_2}$ | 0.4035 | | | | | | | | | | | |
| $\pi_{\text{class}}$ | 0.2143 | | 0.1907 | | 0.1821 | | 0.1671 | | 0.1538 | | 0.0921 | |
| $\mathcal{L}(\hat{\beta}) = -5,168.404$ | | | | | | | | | | | | |
| **ProbHLC** | | | | | | | | | | | | |
| $\beta_{\text{Length (over 2 hours)}}$ | -4.49 | -7.9 | -0.438 | -3.33 | -0.0104 | -6.14 | -0.233 | -2.16 | -0.94 | -3.08 | 1.68 | 0.72 |
| $\beta_{\text{Car Park}}$ | 1.1 | 4.42 | 0.726 | 5.21 | 0.00635 | 4.04 | 0.0554 | 0.8 | 0.436 | 1.53 | -2.08 | -1.49 |
| $\beta_{\text{Fence}}$ | 0.66 | 2.48 | 0.248 | 1.45 | 0.00266 | 1.44 | 0.036 | 0.58 | 0.121 | 0.55 | 2.82 | 1.07 |
| $\beta_{\text{Path and Signage}}$ | 0.245 | 1.14 | 1.35 | 6.94 | -0.00157 | -0.84 | 0.182 | 2.16 | -0.0899 | -0.33 | 4.64 | 1.57 |
| $\beta_{\text{cost}}$ | -0.161 | -6.67 | -0.173 | -12.99 | -0.00321 | -12.41 | -0.00045 | -0.34 | -0.0447 | -4.6 | -0.0402 | -1.21 |
| $\beta_{\text{hill}}$ | -0.39 | -1.38 | 2.87 | 8.7 | 0.0249 | 8.24 | 0.182 | 0.88 | 23.5 | 2.11 | 21.1 | 1.67 |
| $\beta_{\text{bog}}$ | -1.08 | -3.02 | 2.57 | 8.1 | 0.0222 | 7.93 | -0.0404 | -0.19 | 22.7 | 2.04 | -405 | 0 |
| $\beta_{\text{field}}$ | -0.317 | -0.94 | 2.79 | 7.61 | 0.026 | 8.49 | 0.833 | 3.05 | 23 | 2.06 | 4.9 | 0.72 |
| $\beta_{\text{river}}$ | 0.261 | 0.82 | 2.79 | 8.26 | 0.0358 | 10.59 | 1.36 | 3.26 | 23.9 | 2.15 | 12.4 | 1.39 |
| $\eta_1$ | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| $\eta_2$ | -0.708 | 21.58 | -0.712 | 10.46 | inf | - | 1.5 | 2.3 | -0.817 | 10.09 | -0.922 | 19.57 |
| $\pi_{s_1}$ | 0.5749 | | 0.7858 | | 0.0618 | | 0.3624 | | 0.5173 | | 0.1751 | |
| $\pi_{s_2}$ | 0.4251 | | 0.2142 | | 0.9382 | | 0.6376 | | 0.4827 | | 0.8249 | |
| $\pi_{\text{class}}$ | 0.2639 | | 0.2055 | | 0.1666 | | 0.1649 | | 0.1396 | | 0.0595 | |
| $\mathcal{L}(\hat{\beta}) = -5,166.516$ | | | | | | | | | | | | |

Table 4: Welfare analysis based on individual-specific posterior parameters

| | Conditional WTP for 6LC | | Conditional WTP for 6obsHLC | | Conditional WTP for 6ScaleAdjLC | | Conditional WTP for 6ProbHLC | |
|---|---|---|---|---|---|---|---|---|
| | Median | Mean | Median | Mean | Median | Mean | Median | Mean |
| long walk | -14.7 | -37.2 | -11.2 | -31.6 | -11.5 | -48.1 | -6.4 | -12.5 |
| car park | 4.2 | 8.6 | 3.4 | 8.1 | 4.0 | 19.0 | 4.9 | 5.3 |
| fence | 1.6 | 7.9 | 1.9 | 7.4 | 1.5 | 12.2 | 2.0 | 3.1 |
| path and signage | 7.5 | 19.2 | 7.2 | 24.5 | 6.4 | 31.5 | 1.5 | 4.3 |
| hill | 17.1 | 74.6 | 15.6 | 65.9 | 16.8 | 24.6 | 16.6 | 147.6 |
| bog | 15.3 | 46.9 | 14.3 | 48.9 | 13.8 | 44.4 | 6.8 | 14.9 |
| field | 16.6 | 131.0 | 14.4 | 154.1 | 17.4 | 151.4 | 16.2 | 139.9 |
| river | 17.1 | 199.3 | 15.2 | 216.0 | 17.6 | 236.3 | 16.2 | 148.6 |
| | Comp. Variation for 6LC | | Comp. Variation for 6obsHLC | | Comp. Variation for 6ScaleAdjLC | | Comp. Variation for 6ProbHLC | |
| | Median | Mean | Median | Mean | Median | Mean | Median | Mean |
| Scenario 1 | -1.9 | -4.2 | -1.8 | -11.9 | -2.1 | -9.8 | -2.2 | -2.5 |
| Scenario 2 | -5.6 | -12.5 | -4.5 | -24.3 | -3.9 | -21.9 | -0.4 | -3.3 |
| Scenario 3 | 1.4 | 3.5 | 1.5 | 8.7 | 1.7 | 7.7 | 1.7 | 1.7 |
| Scenario 4 | 1.5 | 5.1 | 1.5 | 8.2 | 0.8 | 8.4 | 0.2 | 0.6 |

Table 5: Elasticities based on conditional parameters

| Elasticities based on 6LC | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **LONG WALK (over 2 hours)** | | **CAR PARK** | | **FENCE** | | **PATH AND SIGNAGE** | |
| | Median | Mean | Median | Mean | Median | Mean | Median | Mean |
| hill | -0.2705 | -0.6333 | 0.0029 | 0.1808 | - | - | 0.0571 | 0.1664 |
| bog | -0.3462 | -0.7001 | 0.0000 | 0.1801 | - | - | 0.0874 | 0.1877 |
| field | -0.2931 | -0.6092 | 0.0031 | 0.1740 | 0.0000 | 0.0692 | 0.0598 | 0.1598 |
| river | -0.1625 | -0.5466 | 0.0000 | 0.1505 | 0.0000 | 0.0656 | 0.0459 | 0.1471 |
| Elasticities based on 6obsHLC | | | | | | | | |
| | **LONG WALK (over 2 hours)** | | **CAR PARK** | | **FENCE** | | **PATH AND SIGNAGE** | |
| | Median | Mean | Median | Mean | Median | Mean | Median | Mean |
| hill | -0.1863 | -0.5122 | 0.0014 | 0.1281 | - | - | 0.0714 | 0.2087 |
| bog | -0.3398 | -0.5720 | 0.0000 | 0.1309 | - | - | 0.0943 | 0.2304 |
| field | -0.1932 | -0.5000 | 0.0099 | 0.1233 | 0.0000 | 0.0704 | 0.0859 | 0.2001 |
| river | -0.0762 | -0.4567 | 0.0000 | 0.1152 | 0.0000 | 0.0698 | 0.0577 | 0.1770 |
| Elasticities based on 6ScaleAdjLC | | | | | | | | |
| | **LONG WALK (over 2 hours)** | | **CAR PARK** | | **FENCE** | | **PATH AND SIGNAGE** | |
| | Median | Mean | Median | Mean | Median | Mean | Median | Mean |
| hill | -0.4281 | -0.6084 | 0.0000 | 0.1630 | - | - | 0.0635 | 0.2658 |
| bog | -0.4381 | -0.6621 | 0.0000 | 0.1679 | - | - | 0.0670 | 0.2864 |
| field | -0.3381 | -0.5689 | 0.2386 | 0.1510 | 0.0000 | 0.0374 | 0.0656 | 0.2416 |
| river | -0.1655 | -0.4969 | 0.0000 | 0.1338 | 0.0000 | 0.0302 | 0.0550 | 0.2072 |
| Elasticities based on 6ProbHLC | | | | | | | | |
| | **LONG WALK (over 2 hours)** | | **CAR PARK** | | **FENCE** | | **PATH AND SIGNAGE** | |
| | Median | Mean | Median | Mean | Median | Mean | Median | Mean |
| hill | -0.1619 | -0.8156 | 0.0041 | 0.2497 | - | - | 0.0000 | 0.2420 |
| bog | -0.3340 | -0.8946 | 0.0000 | 0.2431 | - | - | 0.0000 | 0.2988 |
| field | -0.1489 | -0.7922 | 0.0050 | 0.2302 | 0.0000 | 0.1114 | 0.0000 | 0.2711 |
| river | -0.0744 | -0.7264 | 0.0000 | 0.2020 | 0.0000 | 0.1093 | 0.0000 | 0.2538 |

Author Contact Details: Marco Boeri, Queen's University Belfast. Email:m.boeri01@qub.ac.uk

Socio-Economic Marine Research Unit