



Controlling for the Effects of Information in a Public Goods Discrete Choice Model

Mikołaj Czajkowski · Nick Hanley · Jacob LaRiviere

Accepted: 7 November 2014 / Published online: 27 November 2014
© The Author(s) 2014. This article is published with open access at Springerlink.com

Abstract This paper develops a reduced form method of controlling for differences in information sets of subjects in public good discrete choice models, using stated preference data. The main contribution of our method comes from accounting for the effect of information provided during a survey on the mean and the variance of individual-specific scale parameters. In this way we incorporate both scale heterogeneity as well as observed and unobserved preference heterogeneity to investigate differences across and within information treatments. Our approach will also be useful to researchers who want to combine stated preference data sets while controlling for scale differences. We illustrate our approach using the data from a discrete choice experiment study of a biodiversity conservation program and find that the mean of individual-specific scale parameters and its variance in the sample is sensitive to the information set provided to the respondents.

Keywords Information effects · Uncertainty · Discrete Choice modelling · Combined datasets

JEL Classification Q51 · C51 · D03 · D83 · D61

M. Czajkowski (✉)
Department of Economic Sciences, University of Warsaw, Warsaw, Poland
e-mail: miq@wne.uw.edu.pl

N. Hanley
Department of Geography and Sustainable Development, University of St Andrews, St Andrews,
Scotland, UK
e-mail: ndh3@st-andrews.ac.uk

J. LaRiviere
Department of Economics & Baker Center for Public Policy, University of Tennessee,
Knoxville, TN, USA
e-mail: jariv11@utk.edu

1 Introduction

Stated preference methods such as Contingent Valuation (CV) using Discrete Choice Experiments (DCE) involve presenting detailed information to a randomly selected sample of the population and subsequently eliciting their preferences and willingness to pay. Rather early on in the application of CV to the valuation of environmental goods, researchers found that providing respondents with new or differently-phrased information on these goods could change measures of willingness to pay (Thaler 1980). This should not come as a surprise, as the same phenomena apply to goods and services traded in markets—what people know about the characteristics of a good, and of substitutes and complements, can co-determine their values (Milgrom 1981; Munro and Hanley 2002).

In this paper, we present an approach for controlling for the effects of different information sets provided to respondents in a very popular class of generalized multinomial logit (G-MNL) models. Specifically, we allow information to affect both preferences and the mean and variance of individual-specific scale parameters in a random utility model (RUM). In this way we incorporate both observed and unobserved preference and scale heterogeneity to investigate differences across and within information treatments. As a result, our technique better controls for different information effects than existing methods in the literature.

Random Utility forms the basis for many applications of both stated and revealed preference approaches. In a random utility model, the formulation of the utility function leads to an empirical model in which the observed choices of an individual are used to link choice alternatives with utility levels. An agent's utility depends on a deterministic component V and a random component ε (McFadden 1974). The deterministic component, V , is comprised of estimated preference parameters which map attributes and individual characteristics to utility. The introduction of the error term ε is due to the researcher's inability to observe all attributes of choice and all significant characteristics of respondents which influence her choices (McFadden 1976).¹ Pragmatically, this makes it possible to explain why apparently equivalent individuals (equal in all attributes which can be observed) may make different choices.

Variation of the random component of utility (σ_ε) relative to the magnitude of the deterministic component is often called the scale parameter. As the scale parameter increases, the size of the deterministic portion of utility increases relative to the idiosyncratic portion of utility. Thus as the scale parameter increases, respondents' choices appear less random from the econometrician's perspective.² As a result, it influences the confidence intervals of WTP estimates.

Consistent with the literature, our study is predicated upon the notion that altering the information set presented to a subject in a stated preference study could affect the predictive power of the econometrician's RUM (Carson and Czajkowski 2014). For example, if subjects have a more complete information set, it could lead to the value of certain attributes of a public good being estimated more precisely. In this paper, we propose a methodology for making the preference and scale parameters a function of the information set a respondent holds and, as a result, provide a method for accounting for the effects of information in random utility-based

¹ Although the econometrician does not observe the random component of utility, the consumer does. Put another way, in the classical formulation of RUMs with full information, there is no randomness in the consumer's choice from the consumer's perspective. Note too that this term could alternatively represent variation in a decision maker's choice from a set of all possible decision rules (Manski 1977).

² Breffle and Rowe (2002) point out that choice complexity can co-determine the degree of randomness in people's choices with regard to environmental goods, and should thus be included in econometric estimation where possible.

stated preference methods. Using a simple theoretical model, we highlight how this approach is consistent with modeling agents as Bayesian updaters who incorporate newly presented information provided during a study into their stated preferences. Our approach also provides a convenient way for combining datasets from two different but related CV studies for joint estimation thereby increasing estimator efficiency.³ Seen in this light, our approach is based on an extension of the generalized multinomial logit (G-MNL) model (Fiebig et al. 2010). Most generally, the approach allows for a flexible treatment of both observed and unobserved preference and scale heterogeneity.

We illustrate our approach by applying it over preferences for biodiversity conservation. Biodiversity conservation is a well-suited public good for our study for two reasons. First, the non-use benefits of enhanced biodiversity conservation in the case examined here are non-rival and non-excludable in consumption. Second, there is no well-functioning market for biodiversity meaning that stated preference methods are common in evaluating management options for this type of good. Our results confirm that the estimated contribution of the deterministic portion of respondents' utility relative to the stochastic element both varies across individuals and is sensitive to the information set they are given. In contrast, changes in the information provided did not influence the preference (taste) parameters of respondents' utility functions. We find, then, that from the econometrician's perspective, the information set provided to respondents may affect the precision with which the deterministic portion of utility is estimated, rather than the level of coefficient estimates. This finding is consistent with subjects refining their preferences for biodiversity conservation over the econometrician's observable characteristics as they are presented with different information sets.

The remainder of the paper is structured as follows. Section 2 begins with an overview of information effect studies in stated preferences. We then conduct a brief theoretical modeling exercise which shows variance in preferences for a good could vary with informative signals about the good if consumers are Bayesian updaters. Section 3 offers precise discussion of how scale and preference heterogeneity has been modelled in discrete choice studies. We then set out a new approach to represent differences in unobserved preference and scale heterogeneity in combined datasets, namely differences in mean preference and scale coefficients, as well as differences in their variances. The design and implementation of a choice experiment with two information treatments is then described. Results from applying this framework to our study follow, and we conclude with some observations on implications for future work.

2 Information Effects in Stated Preferences

The effect of information about environmental goods on willingness to pay was one of the early concerns amongst stated preference researchers, and reflects a long-standing interest in information, complexity and human behavior in decision science (Payne 1976).⁴ Munro and Hanley (2002) consider eight contingent valuation studies finding statistically significant effects of different information sets presented to subjects on mean WTP. Munro and Hanley

³ It has long been known that failing to account for scale differences when combining datasets can lead to biased estimates (Morikawa 1989; Ben-Akiva and Morikawa 1990). Several methods have been proposed to deal with this problem (e.g., Swait and Louviere 1993, see Sect. 3.2 for details).

⁴ Even for goods purchased in a market, more information may increase the precision of beliefs if agents are strictly neoclassical in their preferences (Nelson 1970). For example, in most models of experience goods—goods for which agents have uncertainty over their preferences—the expected variance of utility conditional on purchase for goods in a market declines as consumers have more information and/or experience (Bergemann and Välimäki 1996, 2006).

(2002) also formally show in an expected utility model how the mean and variance of WTP might be affected by more “positive” information about an environmental good.⁵ Other studies attempt to test whether certain neoclassical and behavioral economic theories are the causes behind these changes. Specifically, the effect of costless signals and cheap talk, bounded rationality, and Bayesian updating conditional on previous levels of familiarity with a good have been examined and been found to be important (MacMillan et al. 2006; Aadland et al. 2007; Hoehn et al. 2010; LaRiviere et al. 2014) or unimportant (Alberini et al. 2005; Czajkowski et al. forthcoming) for estimating mean WTP for non-market goods. The literature shows, then, that influencing the information set of agents in a CV study can affect estimated mean WTP.⁶

A less studied aspect of the CV literature is how information provided in a CV study could influence the econometrician’s ability to predict mean WTP levels (Czajkowski et al. forthcoming). In a CV study, the subject is often presented with information about the public good being valued before the econometrician elicits their WTP. It is reasonable that the provided information could interact with the subject’s previous information set and experiences to influence their stated preferences. For example, Bayesian updating due to a subject’s previous information set interacting with newly provided information in CV study has been shown to affect estimated preference parameters and potentially affect σ_ε relative to V (Christie and Gibbons 2011). Efficiently and consistently estimating σ_ε relative to V is vital because it allows the econometrician to correctly infer the level and confidence intervals of WTP.

The motivation behind our approach is to account for how providing different information sets to subjects can affect the estimated scale parameter (e.g., the ratio of σ_ε relative to V) in a random utility model. It is relatively straightforward to understand how information could influence the estimated scale parameter (and hence, potentially also the predictability of subjects’ WTP estimates) for the econometrician. We now present a simple theoretical model showing how information can influence the variance of observed WTP from the econometrician’s perspective if subjects are Bayesian updaters.⁷

Consider a model in which there are two possible states of the world $x \in \{A, B\}$. An agent trying to infer the true state of the world has a prior that the state of the world is A : $pr(x = A) = \rho$ such that $pr(x = B) = 1 - \rho$. In our case, ρ might be the probability an agent believes their home is at risk from flooding. In this case, A may be associated with an increased value the agent receives from additional flood protection (e.g., in state A the agent’s home is at high risk of flooding). As a result, state A is positively related to the level of utility from an additional public works project that defends against flooding whereas state B is negatively related to the utility the agent gains from the public works project. A different example would have that state A corresponds to the re-introduction of a species with no

⁵ Positive information is defined by the authors as information which increases the subjective probability that an environmental good has more desirable attributes and fewer un-desirable attributes. They also show that an individual’s WTP is increasing in positive information about the characteristics of a good, implying that mean WTP is also increasing in positive information.

⁶ Further, the literature has considered the impacts of changing the manner in which information is conveyed to respondents in the context of the complexity/familiarity of the resource in question (Hoehn et al. 2010).

⁷ See Czajkowski et al. (forthcoming) for an alternative specification using a normal distribution in the context of how experience with a new good can affect preferences for it via Bayesian updating. Note that Bayesian updating is only one possible updating process. The point of this section is to show how one widely used updating model relates additional information to WTP in the simplest possible framework. In the subsequent empirical specification, there is no scope for testing this model relative to others since marginal pieces of information were not provided. Rather, in this study, entirely different information sets were provided to subjects. For a more complete evaluation of the Bayesian model relative to others, see Czajkowski et al. (2014b).

adverse effects on other flora and fauna, whereas in state B the species re-introduction causes negative impacts on existing animals and plants. Without loss of generality, then, V_A can be thought of as the agent's value for the public project in state A and V_B can be thought of as an agent's value for the project in state B.

Define the information a subject receives in a CV study at time t as s_t . s_t informs the subject as to state of the world such that $s_t \in \{a, b\}$ and $pr(s_t = a|A) = pr(s_t = b|B) = \theta \in (0.5, 1)$. Note that signals are informative but not perfectly informative given the support of θ : with probability $1 - \theta$ the subject will receive a signal of b even if the true state of the world is A. This set up is meant to describe a situation in which a subject doesn't know the true state of the world (e.g., does know their valuation for a public good with certainty at the time of a survey) and so must infer if from imprecise signals (e.g., a signal s_1 , in the information provided during the survey).

The expected value and variance of the public good conditional on a subject's prior are

$$E(V|\rho) = \rho V_A + (1 - \rho) V_B$$

$$\text{var}(V|\rho) = V_A V_B \rho (1 - \rho)$$

The variance of utility is single peaked with a maximum at $\rho = 0.5$ and equal to zero for $\rho \in \{0, 1\}$. Consider the properties of utility conditional on receiving the CV study's information signal, s_t , given the above model given a prior, ρ , that the true state of the world is A. As long as the information causes the consumer's posterior, $\rho_{t+1}|s$, to be updated toward either zero or one, the variance in a representative consumer's expected valuation would decrease.⁸ As a result, the more informative signals an agent receives, the lower the variance from consumption as ρ is updated with new information.⁹ Importantly, the agents in this model do not have random preferences conditional on a state of the world. Rather, new information can affect the agent's belief about the true state of the world and therefore affect their WTP for a good.

Now consider that there are two classes of signals, one class noisier than the other. This situation mirrors what would occur in a survey in which one information packet is written by an interest group and another by an objective surveyor. If that case, the noisier signal (e.g., the information from the interest group) would lead to different updating than the cleaner signal (e.g., the information from the objective surveyor). As a result the observed variance in WTP for a good would vary by signal type.

The above theoretical exercise shows how additional information can affect a representative agent's variance in WTP for a good via Bayesian updating. In a discrete choice model, variance in WTP is summarized by the relative magnitude of the unobserved (to the econometrician) portion of utility (e.g., the random component (ε)). The magnitude of that unobserved component is dictated by the scale parameter (σ_ε). As a result, the simple model above shows that different information sets could have different effects on the estimated scale parameter. Further, the information set a subject enters a survey with is unobserved to the econometri-

⁸ Here we assume that the representative consumer has unbiased priors. That is, the average high type consumer has a prior shading them toward actually being a high type: $\rho > 0.5$. If all consumers had biased priors than on average variance would increase then decrease.

⁹ For example, in the model presented here, the posterior probability an agent is a high type conditional on observing a signal $s_t = a$ is $\rho_{t+1} = \rho\theta / (\rho\theta + (1 - \rho)(1 - \theta))$. Czajkowski et al. (forthcoming) show explicitly how both mean and variance in WTP could be affected asymmetrically by additional experience using a normal conjugate priors.

cian.¹⁰ Importantly, the subject's unobserved information can interact with the information provided during the survey process leading to changes in the variance of WTP conditional on covariates. From the econometrician's perspective, this also manifests as new information affecting the magnitude of the estimated deterministic portion of utility relative to the idiosyncratic portion, or the scale parameter. As a result, if different information sets were presented to various subjects in a CV study, it is reasonable to expect their scale parameters could be influenced differently.

It is not uncommon for agents in CV studies to be presented with different information sets in order to either satisfy various stakeholders or be part of an economic field experiment (Carson et al. 1994; MacMillan et al. 2006). In that situation, it is reasonable for the econometrician to allow different information sets to have heterogeneous effects on subjects' scale parameters. Indeed, the above exercise motivates an econometric approach for allowing different information sets provided to subjects in a CV study to asymmetrically affect the scale parameter in RUMs.

In the next section, we show how allowing an information fixed effect to enter the estimated scale parameter is a relatively straightforward extension of discrete choice models. Allowing heterogeneous treatment effects of information on the scale parameter permits heterogeneous predictability (e.g., 'perceived randomness') of agent's choices by the econometrician. This is the main contribution of our econometric model.

3 Modelling Discrete-Choice Sata

Before introducing the how information fixed effects can enter the scale parameter in an econometric model, we first briefly introduce the standard Random Parameters Logit (RPL) model (McFadden and Train 2000; Hensher and Greene 2003). The RPL model allows for an economic agent's preferences to vary with their observable characteristics. In the RPL model respondent i 's utility associated with selecting alternative j out of a set of J available alternatives at time occasion t can be represented as:

$$U_{ijt} = \sigma \beta'_i \mathbf{x}_{ijt} + \varepsilon_{ijt}, \quad (1)$$

where \mathbf{x}_{ijt} is a vector of respondent- and alternative-specific choice attributes, and β_i represents a vector of individual-specific taste parameters associated with marginal utilities of the choice attributes, such that they follow a multivariate distribution $\beta_i \sim f(\mathbf{b}, \Sigma)$, with means \mathbf{b} and variance-covariance matrix Σ .¹¹ Finally, since the stochastic component of the utility function ε is typically assumed to follow the extreme value distribution (with a known mean and variance), the parameter σ can be thought as introducing the required amount of perceived randomness into respondents' choices by scaling the deterministic part of their utility function—the higher the scale, the more deterministic (predictable) the choices from the econometrician's perspective.¹²

A method allowing to simultaneously model the preference and scale heterogeneity is the G-MNL model (Fiebig et al. 2010). In this model, the utility function takes the form:

¹⁰ Presumably some types of information will be correlated with observable characteristics of subjects. It is reasonable to expect, though, that some information will still be unobservable to the econometrician even conditioning on covariates.

¹¹ One variant of this is the Latent Class (LC) model in which the random parameters follow a discrete, rather than continuous, distributions (e.g., Shen 2009).

¹² Note that the utility function is ordinal by nature; scaling does not alter respondents' underlying preferences.

$$U_{ijt} = [\sigma_i \mathbf{b} + \gamma \mathbf{v}_i + (1 - \gamma) \sigma_i \mathbf{v}_i]' \mathbf{x}_{ijt} + \omega_{ijt}. \quad (2)$$

Similarly to the RPL model, the coefficients in the utility function are individual-specific (\mathbf{b} represents the population means of the parameters, while \mathbf{v} are individual-specific deviations from these means). Unlike in the RPL, however, the scale coefficient is also individual-specific, with $\sigma_i \sim LN(1, \tau)$ or $\sigma_i = \exp(\bar{\sigma} + \tau v_i)$ with $v_i \sim N(0, 1)$. Since it is still necessary to normalize scale, we want $E\sigma_i = \exp(\bar{\sigma} + \tau^2/2)$. This may be achieved by assuming $\bar{\sigma} = -\tau^2/2$. This way the scale is no longer constant across respondents; instead it is assumed to follow a lognormal distribution, with the new parameter τ reflecting the level of scale heterogeneity in the sample.

The coefficient $\gamma \in [0, 1]$ controls how the variance of residual taste heterogeneity varies with scale. If $\gamma = 0$ the individual coefficients become $\beta_i = \sigma_i (\mathbf{b} + \mathbf{v}_i)$, while if $\gamma = 1$ they are $\beta_i = \sigma_i \mathbf{b} + \mathbf{v}_i$. These are the two extreme cases of scaling (or not scaling) residual taste heterogeneity in the G-MNL model (type I and type II respectively), however, all intermittent solutions are possible.

Finally, it should be noted that the preference (taste) and scale parameters are not separately identifiable, as they are always observed as a in multiplicative form. [Hess and Rose \(2012\)](#) demonstrate that in the case when (1) all parameters are modelled as random and (2) all parameters are allowed to be correlated, introducing the random scale coefficient is equivalent to allowing for a more flexible distribution for the taste-scale mixture. In many cases, however, introducing a random scale coefficient is useful because it allows one to account for all (random or non-random) parameters for a particular individual becoming larger or smaller, relative to the utility function error term (whose variance is normalized to one). In this way, a single parameter allows us to observe how the deterministic part of respondent's utility function varies relatively to the random component, from the perspective of the analyst. This approach provides a convenient way of observing and interpreting the level of heterogeneous predictability ('perceived randomness') of agent's choices by the econometrician. In our case, as described in the next section, we not only make the scale coefficient random, but also introduce information-set-specific covariates into its mean and variance, thus proposing a useful, reduced form method of empirical investigation of the effects of information and updating in a public goods discrete choice model.¹³

3.1 Methods for Accounting for Information-Related Effects on Preference and Scale Heterogeneity

The information sets respondents hold may in some cases be observed. A typical way of controlling for the effects of information on preference parameters is making the means (and possibly variances) of the taste parameters a function of information-related covariates (\mathbf{z}), so that $\beta_i \sim f(\mathbf{b} + \phi' \mathbf{z}_i, \Sigma + \psi' \mathbf{z}_i)$. By comparing the means or the variances of random preference (taste) parameters in one treatment with the parameters in another, the modeler is able to identify the relative changes in the observed preferences resulting from different information treatments. This way, it is possible to investigate the effects of different information sets on respondents' preferences (tastes) while allowing for unobserved preference heterogeneity.

¹³ In addition, our data required imposing particular constraints on the correlation structure of the random parameters, as it would not make sense to allow the parameters of the attributes between two different studies (datasets) to be correlated. Because of these correlation constraints, the model without random scale coefficient would not fully take the scale heterogeneity into account, as indicated by the necessary conditions put forth by [Hess and Rose \(2012\)](#).

The main contribution of our paper, however, comes from accounting for not only the effects of information on preferences, but also the effects on scale. Since in the case of the G-MNL model individual scale is a random variable, there are two possible effects which can be taken into account—the systematic differences in the mean scale, and the systematic differences in its variance. We propose to account these effects by making the mean of the random scale parameter and its variance functions of information-related covariates:

$$\sigma_i \sim LN(1 + \phi'z_i, \tau + \xi'z_i). \quad (3)$$

Exploring issues related to the observed scale differences by including observation-specific covariates of (non-random) scale has been done before. There are several models which allow for this, including the covariance heterogeneity model (DeShazo and Fermo 2002), the error components model (Hensher et al. 2008; Savage and Waldman 2008), modeling Gumbel variance directly by using socio-economic characteristics (Scarpa et al. 2003), the heteroskedastic extreme value model (Salisbury and Feinberg 2010), multiplicative errors model (Fosgerau and Bierlaire 2009) and, perhaps most notably, the heteroskedastic MNL model (e.g. Hensher et al. 1998; Dellaert et al. 1999; Swait and Adamowicz 2001; Caussade and Ortúzar 2005). None of these approaches allow for unobserved preference and scale heterogeneity, however.

The possibility of including covariates of scale, while allowing for unobserved scale heterogeneity, was first mentioned by Fiebig et al. (2010), although the authors did not pursue this approach. We apply their concept and extend it by also including information-related covariates in the scale variance (τ). The rationale of our approach is as follows. Just as two samples can differ with respect to mean WTP and its variance, they can differ with respect to (1) how random or how deterministic the respondents appear (on average) from the econometrician's perspective and (2) how differentiated each sample of respondents is, in terms of whether the respondents have similar scale parameters. The former effect is captured by the mean of the individual scale parameters, the latter by the variance. Our approach, therefore, allows for a greater flexibility in accounting for scale differences between groups of observations.

In our case, the information sets were dataset-specific. Therefore, we use dataset-specific covariates of mean scale and its variance to control for the possible effects of information, so that¹⁴:

$$\sigma_i = \exp(\bar{\sigma} + \exp(\lambda'z_i) \tau v_i + \theta'z_i). \quad (4)$$

In this formulation, even though the absolute levels of scale or its variance are not identified, positive coefficients θ indicate observations with a higher scale, i.e. less uncertainty surrounding choices, in comparison with the reference treatment. Positive coefficients of λ , on the other hand, represent observations which higher scale heterogeneity, e.g., a group of respondents who are more diversified in terms of how predictable their choices are, in relation to the reference group.¹⁵

¹⁴ The model is identified as long as there is at least one common taste parameter across treatments; there are otherwise no restrictions on the observed variables which can simultaneously be used as covariates of individual taste parameters, individual scale and its variance. The mean scale is relative—it needs to be arbitrarily chosen (normalized) for at least one group of observations, and the scale for the other groups can be interpreted in relation to the reference level.

¹⁵ As a result, the identifiable parameters in the above eqnarray are τ -indicating the level of scale heterogeneity, and λ and θ , which represent shifts in the mean and the variance of the random scale coefficient in comparison with the reference group (treatment). v_i is a standard normal variate, while $\bar{\sigma}$ is the mean scale parameter in the reference group, necessary for normalizing the overall scale level in this group to 1.

The resulting extension of the G-MNL model is flexible enough to capture observed and unobserved preference heterogeneity, as well as observed and unobserved scale heterogeneity. In the case of our empirical application, it allows us to control for different information levels of respondents. We note, however, that it can easily be applied to model other phenomena.

3.2 Accounting for Scale Differences When Combining Datasets

Our approach has one other practical application: accounting for scale differences when two or more datasets are combined. It has long been recognized that if observations from two datasets are to be combined, accounting for scale differences is necessary (Swait and Louviere 1993). This is because utility function parameters are confounded with scale and so failing to take this into account (i.e. assuming the scale parameter in two datasets is the same) may lead to misleading conclusions being drawn. Scale may vary across data sets due to e.g., differences in sampling or in the information provided to respondents. Only after the scale differences have been accounted for it is possible to formally test the equality of utility function parameters and their variances if unobserved preference heterogeneity is allowed for (Hensher et al. 1998).¹⁶

Several methods to control for scale differences between datasets have been suggested. Ben-Akiva and Morikawa (1990) proposed a procedure to efficiently estimate the scale differences between two data sources. Their procedure simultaneously maximizes a joint likelihood function for observations from two or more datasets. A relative scale factor can be estimated for each type of data (except one which is arbitrarily chosen as the base level; Morikawa 1989). Bradley et al. (1992, 1994) incorporated the one-step estimation approach of Morikawa and Ben-Akiva into the nested logit setting. They call this approach the logit-based scaling approach (Hensher and Bradley 1993; Bradley and Daly 1994). Perhaps the most commonly used method of controlling for scale differences when combining datasets was proposed by Swait and Louviere (1993). It allows data from two sources to be combined by exploring through a grid search a range of plausible relative scale factors for which the parameters of the models estimated for two samples are statistically equal. This “tedious, but straightforward” procedure results in unique maximum for the log likelihood function, at least for the linear-in-parameters MNL model for which the log likelihood is concave and remains probably the most commonly used way to test for scale differences between two datasets, at least in environmental economics applications (e.g. von Haefen and Phaneuf 2008; Christie and Azevedo 2009; Olsen 2009; Brouwer et al. 2010).

We note that the methods presented above do not allow for unobserved scale heterogeneity, despite a growing body of literature suggesting that modelling of unobservable scale differences may be a significant component in accounting for overall heterogeneity (e.g. Louviere et al. 2002). Once unobserved scale heterogeneity is allowed, groups of observations (e.g. datasets) can differ not only in terms of preferences (e.g. means of random parameters), preference heterogeneity (e.g. variances of random parameters), and mean scale, but also

¹⁶ Many studies compare implicit prices (i.e. ratios of parameters) derived from separately estimated models, as this way dataset-specific scale cancels out. However, this approach leads to the possibility of confounding differences in preferences for the attributes with the marginal utility of income (cost). In addition it does not allow the researcher to easily test for dispersion of random parameters, if unobserved preference heterogeneity is allowed. Finally, we note that in the case of unobserved preference heterogeneity, many researchers fail to properly conduct the test for equality of mean WTPs, as they do not take the all the information about the empirical distribution of WTPs into account, usually focusing on their means and associated standard errors only, rather than also take the standard deviations (and their standard errors) of the distributions of WTPs into account.

with respect to the scale heterogeneity (i.e. scale variance). The approach we propose allows one to simultaneously take all these differences into account. In addition, it does not require that all taste parameters are assumed equal across the sample. We therefore argue it is more flexible and, at the same time, more convenient to use when data from two or more sources is combined.

4 Case Study and Survey Design

This section applies the empirical model introduced in Sect. 3 to a specific issue in biodiversity conservation. The management of Red Grouse (*Lagopus lagopus scotticus*) in the UK uplands provides an interesting case study. Management of moorlands for Red Grouse shooting since the mid-nineteenth century has led to declines in many species of predators (Newton 1998), since the aim of grouse management is to maximize numbers of birds available for shooting in the autumn. One particular conflict which has arisen in this context concerns the management of Hen Harriers (*Circus cyaneus*) on sporting estates. Hen harriers are listed as endangered due to population declines in the last 200 years (Baillie et al. 2009). Economic costs to grouse moor owners arise because harriers prey on grouse (Thirgood et al. 2000), and arguments between the conservation lobby and the sporting estate community have become polarized over time (Redpath et al. 2004; Thirgood and Redpath 2008). Evidence shows that (1) Hen Harrier densities can increase to the extent that they make management for grouse shooting economically unviable; (2) illegal killing has resulted in a suppression of harrier populations in both England and Scotland (Etheridge et al. 1997); and (3) that enforcement of current laws prohibiting lethal control has been ineffective (Redpath et al. 2010). Golden Eagles are often found in Hen Harrier habitat, and are also top predators which have been subject to illegal persecution, particularly in managed grouse moors (Watson et al. 1989; Whitfield et al. 2007).

To understand public preferences over the conservation of Hen Harriers on heather moorland, we designed a choice experiment (Hanley et al. 2010). The choice experiment design consisted of four attributes. These were:

- Changes in the population of Hen Harriers on heather moorlands in Scotland. The levels here were a 20% decline (used as the status quo), maintaining current populations, and a 20% increase in the current population.
- Changes in the population of Golden Eagles on heather moorlands in Scotland. The levels here were a 20% decline (used as the status quo), maintaining current populations, and a 20% increase in the current population.
- Management options. These included the current situation, moving Hen Harriers (“MOVE”), diversionary feeding (“FEED”) and tougher law enforcement (“LAW”). These levels were included as labelled choices. That is, in each choice card, 4 options were available. One represented the status quo, and then 3 choice columns showed variations in other attribute levels given a particular, labelled management strategy.
- Cost of the policy. We told respondents that “*the cost level indicated is the amount of extra tax which a household like yours might have to pay if the government went ahead with that option.*” The levels used were 0 (the status quo), 10, 20, 25, and 50 GBP.

Table 1 gives an example of a choice card; each questionnaire included 6 choice cards.¹⁷

¹⁷ Households were contacted by mail, and a 3-stage Dillman procedure followed in terms of reminder letters and new copies of the survey instrument. Respondents were told that all potential changes to management

Table 1 Example choice card

	DO NOTHING Maintain current management	LAW Stricter law enforcement	FEED Feeding stations away from grouse	MOVE Move eggs and chicks to new sites
HEN HARRIER	20 % population decline	Maintain current population	Maintain current population	Maintain current population
GOLDEN EAGLE	20 % population decline	20 % population increase	Maintain current population	20 % population decline
COST	£0	£50	£50	£10
YOUR CHOICE (please tick one only)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

The choice experiment was designed to minimize the determinant of the AVC matrix of the parameters (*D-error*) given the priors on the parameters of a representative respondent's utility function using a Bayesian efficient design (Scarpa and Rose 2008). The parameters of this distribution were derived from a preliminary model estimated on data available from a pilot study. Pilot surveys were undertaken using in-person surveys of a random sample of Edinburgh households.¹⁸

There were two different versions of the survey, which differed only in the information provided to respondents. The first survey (study 1), reported in Hanley et al. (2010), used an information pack developed solely by the research team, based on existing research findings. The second survey (study 2) used an information pack which was re-written by a group of stakeholders engaged in moorland ownership, management and grouse shooting. In each case, the information pack covered the following items:

- a description of what we meant by “the uplands” in the UK, and how some uplands areas are managed as grouse moors;
- the contribution that grouse shooting makes to the Scottish economy;
- the contribution of grouse management to maintaining heather moorlands, rather than allowing moorlands to be converted to rough grassland or plantation forestry;
- a description of the Hen Harrier, including conservation status and threats from illegal persecution;
- a description of Golden Eagles, their conservation status and current threats to the species;
- three alternatives for moorland management aimed at Hen Harriers.

Given that the biodiversity issues involved in moorland conservation are likely to be unfamiliar to many respondents, we might anticipate that differences in information provided will impact their choices. The most significant differences between how these items were

would be costly, and that these costs would need to be paid by taxpayers. No explicit statement was made
Footnote 17 continued

about how choices would affect future policy decisions, only that “the aim of the research is to inform future government policy” (Taylor et al. 2010). Respondents who chose the status quo, zero cost option in each choice card were asked why this was, in order to separate out protest bidders from people who did not value Hen Harrier or Golden Eagle conservation in moorlands. The survey also included a series of socio-economic and behavioral questions, for example including household income, and whether the respondent was a hunter or had ever been hunting. A copy of the questionnaires can be obtained from the authors on request.

¹⁸ The design for the pilot study was also to optimize for D-efficiency using informative Bayesian priors. These priors were based on existing studies that we found to be relevant to this exercise and on informal information that we gained during the earlier development phase for this survey instrument.

Table 2 A comparison of sample characteristics

Characteristic	Population ^a	Sample 1	Sample 2
Female (%)	52	52	56
Age (years)	42	52.1	50.7
Household income after tax (GBP/year)	32, 128	33,424	29,544

^a Statistics for based on UK's Office for National Statistics reports for 2010–2011, <http://www.ons.gov.uk/>

relayed to respondents in the two treatments were that (1) moorland management is depicted as more beneficial to a variety of species in survey 2 (2) information was provided about how organizations involved in shooting and conservation are trying to find a solution to this conservation conflict in survey 2 (3) moorlands were described as “an important part of our cultural heritage” and “internationally important species of animals and plants” in survey 2, but not in survey 1; (4) a higher figure was provided for the number of jobs generated by moorland management for red grouse shooting in survey 2¹⁹; (5) Hen Harriers were depicted as less threatened in survey 2 than in survey 1; and (6) more information was provided on Golden Eagles in survey 1, such as how they mate for life, how population numbers have recovered during the twentieth century, and how illegal persecution is still carried out.

5 Results

In the first survey, we obtained 233 responses from 1,000 mail outs, a 23 % response rate. In the second survey, we obtained 347 responses from 1,700 mail outs, a 20 % response rate. Table 2 presents the comparison of sample characteristics. Overall, the samples were very similar to each other with respect sex, age and income. Both samples suffered from some over-representation of older respondents. We believe this to be an artefact of the mail format of our survey. The ratio of male and female respondents and the mean household income was very close to the national average.

The observations from the two studies were combined and modelled using the approach outlined in Sect. 3. We applied three different estimators to provide an illustration for our approach and allow for a case-study comparison with existing methods; the results are reported in Table 3. The first approach is an MNL model, followed by the RPL model, and the conventional G-MNL model. The last approach we report allows for not only the mean scale to differ between studies, but also scale parameters variance (λ). In all cases we followed the standard practice in joint estimation on data combined from different sources (Ben-Akiva et al. 1994) by allowing for the scale coefficient to vary between two studies. This is represented by the scale correction factor θ which is included for the observations from study 2 (study 1 is used as a reference). This dataset-specific coefficient reveals significant differences between the two studies in terms of how predictable respondents' choices were (effective scale was higher in the second study).

The variables used in the model include alternative specific constants associated with different protection programs (*LAW*, *FEED*, *MOVE*), dummy-coded levels of improvement

¹⁹ These figures are contested in the public domain.

of Hen Harriers (HH_1, HH_2) and Golden Eagles (GE_1, GE_2), and the continuously coded cost (FEE). All preference parameters were assumed to be normally distributed (including the dis-utility of higher costs). We allowed all parameters to be study-specific (superscripts on variable names indicate the two different samples, except for cost (FEE), which was

Table 3 The results of the G-MNL model

Variable	MNL model scale	RPL model scale		Conventional G-MNL		New G-MNL model	
	study-specific	study-specific		model mean scale		mean and variance of	
	Coeff.	Mean	SD	Mean	SD	Mean	SD
	(SE)	(SE)	(SE)	(SE)	(SE)	(SE)	(SE)
LAW^1	-0.1276 (0.1289)	10.9770*** (1.9179)	17.1546*** (4.6479)	11.1789*** (2.2689)	10.2835*** (2.0164)	11.0462*** (2.3578)	10.8648*** (2.1273)
FEE^1	-0.0709 (0.1322)	11.4099*** (1.9345)	18.7138*** (4.9698)	11.7376*** (2.3108)	11.2513*** (2.0375)	11.5625*** (2.3680)	11.8904*** (2.1758)
$MOVE^1$	-0.332*** (0.1283)	10.3271*** (1.9024)	18.1803*** (4.9576)	10.4530*** (2.3157)	11.2828*** (1.8792)	10.3751*** (2.4152)	11.6050*** (2.0596)
HH_1^1	0.8083*** (0.1724)	1.5984*** (0.4221)	2.9508*** (0.7368)	3.7657*** (1.0071)	9.0646*** (1.6074)	3.0370*** (0.7846)	7.6760*** (1.4002)
HH_2^1	0.8452*** (0.1113)	1.9111*** (0.3417)	2.9566*** (0.7307)	4.4183*** (0.9290)	8.3690*** (1.4843)	3.5371*** (0.7355)	7.2888*** (1.3270)
GE_1^1	1.2981*** (0.1803)	2.9785*** (0.4762)	2.6110*** (0.6276)	6.0296*** (1.2481)	8.9063*** (1.4828)	5.3024*** (1.0273)	8.1872*** (1.5464)
GE_2^1	1.6332*** (0.1546)	3.4973*** (0.4635)	3.3279*** (0.8223)	7.2050*** (1.3454)	9.0671*** (1.4148)	6.1771*** (1.1045)	8.1992*** (1.4667)
LAW^2	-0.873*** (0.2754)	14.3795*** (4.2061)	11.5390*** (1.6093)	13.0938*** (3.7593)	15.8373*** (4.3945)	12.0039*** (3.7172)	14.4188*** (4.3896)
FEE^2	-1.0445*** (0.3095)	13.9061*** (4.1405)	12.1692*** (1.6355)	13.0730*** (3.7950)	15.5460*** (4.2893)	11.9657*** (3.7546)	14.1252*** (4.2400)
$MOVE^2$	-1.1658*** (0.338)	13.8246*** (4.0741)	11.6640*** (1.7392)	12.8068*** (3.7421)	15.6967*** (4.3378)	11.7394*** (3.7047)	14.2955*** (4.2929)
HH_1^2	1.1560*** (0.2606)	2.3578*** (0.5292)	3.7571*** (0.5936)	4.6089*** (1.1420)	4.7717*** (1.3008)	4.0516*** (0.9756)	4.5890*** (1.1379)
HH_2^2	1.5222*** (0.3732)	2.6998*** (0.5631)	3.8069*** (0.4892)	5.1522*** (1.1692)	5.3047*** (1.4418)	4.5623*** (1.0186)	4.9217*** (1.2019)
GE_1^2	2.0353*** (0.4746)	2.8598*** (0.6129)	4.1368*** (0.6245)	5.9484*** (1.3865)	4.7413*** (1.3021)	5.0887*** (1.0913)	4.0150*** (0.9782)
GE_2^2	2.2310*** (0.5448)	3.5109*** (0.7072)	4.5347*** (0.6123)	7.0218*** (1.5595)	5.5689*** (1.4318)	6.1589*** (1.2774)	4.7439*** (1.2482)
FEE	-2.6305*** (0.4742)	-5.0407*** (1.0398)	10.6097*** (1.2052)	-10.5165*** (2.4772)	16.4002*** (2.4418)	-8.8614*** (1.9715)	14.1493*** (2.4345)

Table 3 continued

Variable	MNL model scale study-specific	RPL model scale study-specific		Conventional G-MNL model mean scale study-specific		New G-MNL model mean and variance of scale study-specific	
	Coeff. (SE)	Mean (SE)	SD (SE)	Mean (SE)	SD (SE)	Mean (SE)	SD (SE)
Elements of the lower triangular of Cholesky matrix are skipped							
<i>Structural parameters</i>							
τ -scale variance	0 (fixed)	0 (fixed)		8.5543*** (1.1455)		7.2059*** (1.1175)	
θ -scale correction factor	0.2482 (0.2894)	0.7816*** (0.2584)		0.7816*** (0.2584)		0.7137*** (0.2614)	
λ -scale variance correction factor	0 (fixed)	0 (fixed)		0 (fixed)		-0.2621*** (0.0496)	
<i>Model characteristics</i>							
Log-likelihood	-4,316.53	-2,775.0651		-2,746.6137		-2,736.2711	
McFadden's pseudo R ²	0.0975	0.4198		0.4257		0.4279	
AIC/ <i>n</i>	2.5116	1.6597		1.6438		1.6384	
BIC/ <i>n</i>	2.5401	1.8165		1.8024		1.7988	
<i>n</i> (observations)	3,450	3,450		3,450		3,450	
<i>k</i> (parameters)	16	88		89		90	

*, **, *** represent statistical significance at 0.1, 0.05, 0.01 level, respectively

constrained to be equal in both studies for identification purposes.²⁰ The model allows for correlations between all random parameters within each study.²¹ As a result, the estimated utility function was of the following structure:

$$U_{itj} = [\sigma_i \mathbf{b} + \gamma \mathbf{v}_i + (1 - \gamma) \sigma_i \mathbf{v}_i]' \mathbf{X}_{itj} + \varepsilon_{itj}$$

where:

$$\sigma_i = \exp(\bar{\sigma} + \exp(\lambda S_i) \tau \varepsilon_{0i} + \theta S_i)$$

$$\gamma = \frac{\exp(\gamma^*)}{1 + \exp(\gamma^*)}$$

$$\mathbf{X} = \begin{bmatrix} LAW^1, FEED^1, MOVE^1, HH_1^1, HH_2^1, GE_1^1, GE_2^1, \\ LAW^2, FEED^2, MOVE^2, HH_1^2, HH_2^2, GE_1^2, GE_2^2, FEE \end{bmatrix} \quad (5)$$

and *S* is a binary variable associated with one of the studies. The estimated parameters are the mean tastes matrix **b** along with the elements of Cholesky decomposition of their variance-

²⁰ This is equivalent to introducing study-specific covariates to the means and variances of the taste parameters, as described in Sect. 3.1.

²¹ We constrained the relevant elements of the Cholesky matrix to equal zero to enforce the zero correlation that should hold asymptotically (due to random assignment) between variables associated with different studies. For example, it would make no sense for HH_1^1 (partial improvement of Hen Harriers in study 1) to be correlated with HH_1^2 (analogous attribute for study 2), as these attributes never appeared together.

covariance matrix Σ , and individual-scale-specific parameters τ , λ , θ .²² The estimation was performed in Matlab 8.1 using 1,000 shuffled Halton draws (Sándor and Train 2004). Since the log-likelihood function in the case of G-MNL is not necessarily convex, we used multiple starting points to ensure convergence at the global maximum. Standard errors of coefficients associated with standard deviations of random parameters were simulated using 10^6 draws (Krinsky and Robb 1986).

We start by noting that, overall, all attribute parameters in Table 3 are highly significant and of the expected sign. The statistical significance of the coefficients associated with the standard deviations of normally distributed parameters indicates that there is substantial un-observed preference heterogeneity with respect to all model parameters. The alternative specific constants associated with each protection program (*LAW*, *FEED*, *MOVE*) were relatively high. Coefficients associated with improvements in Hen Harriers (*HH*) and Golden Eagle (*GE*) populations indicate that overall respondents were more concerned with the latter than the former. The differences in preference parameters and WTP for different improvement levels are not linear, and could be interpreted as an asymmetric loss aversion effect (avoiding a 20% loss is more important than 20% gain) or alternatively as being due to sharply declining marginal WTP in some cases. In addition, the respondents in the second sample revealed a higher degree of preference heterogeneity with respect to protecting these bird species, as indicated by higher estimates of standard deviations associated with these parameters in study 2 in comparison with study 1. Comparing the relative importance of different attribute levels is not straightforward, however, as we allowed for correlations between the attributes. Because the random parameter associated with an attribute could be positively or negatively correlated with the cost, the implicit prices associated with these attributes do not necessarily reflect coefficients for the means associated with each attribute. We investigate this further below when we present simulated implicit prices of the attribute levels for both studies.

Comparing the different approaches, modelling unobserved heterogeneity of scale (the change from the RPL to the conventional G-MNL model) allows for an improvement in model fit, as indicated by the highly significant increase in the value of the simulated LL function, as measured by the LR test (p value <0.0001). The scale variance coefficient τ is significantly different from zero, which indicates the presence of significant unobserved scale heterogeneity in the sample—there were significant differences between respondents in terms of how deterministic or how random their choices appeared from the modeler's perspective. The third and the most flexible extended G-MNL estimator additionally allows for the variance of individual scale parameters to differ between the two studies. By introducing an additional component λ in the scale variance τ of the observations from study 2 we were able to allow for not only the mean scale, but also its variance to differ between the two datasets. This also proved to be a statistically significant improvement (p value resulting from the LR test <0.0001). The significant negative coefficient on λ indicates that the respondents of study 2 displayed lower scale variance, than the respondents of study 1, despite having a higher mean scale.

Our results show that allowing for a more flexible estimator, in terms of allowing for random scale and its variance to differ between the two datasets, allows for an improvement in model fit by the AIC, BIC and pseudo R^2 metrics. This allows the econometrician to avoid potential bias resulting from the problem of misspecification, to which nonlinear models are particularly vulnerable (Greene 2011). In addition, it allows for a useful comparison of the

²² An additional structural parameter γ of the G-MNL model is not reported, since in the estimation it approached 0. Due to the numerical problems it caused in the estimation we constrained it to 0, effectively using the G-MNL model type II in which residual taste heterogeneity is scaled in the same way as the means of the distributions (Fiebig et al. 2010).

potential of our model to predict the choices of the respondents in study 1 and 2. In our case, the respondents' choices in the second study appeared less random from the econometrician's perspective. At the same time, respondents in the second study exhibited less variation in the econometrician's ability to predict their preferences (i.e., lower scale variance across subjects). Given that sampling was random, we attribute these differences in scale to changes in the information provided to respondents.²³

5.1 Implicit Prices

Finally, we turn to the analysis of respondents' welfare measures associated with the changes in each attribute level. The WTP for particular attribute levels are calculated as marginal rates of substitution between respective utility function components (a public good attribute for the monetary attribute) and so they allow an additional insight into respondents' preferences.

Since we impose that all parameters in our model (including cost) are normally distributed, the resulting ratio distribution of WTP has infinite moments (i.e. does not have well defined mean and standard deviation (Fieller 1932; Meijer and Rouwendal 2006; Carson and Czajkowski 2013). We therefore employed the following procedure to simulate the median of each distribution of WTP:

1. We took $n = 10^5$ draws from the multivariate normal distribution described by a vector of estimated parameters and the asymptotic variance-covariance matrix;
2. For each of n draws we decomposed the resulting vector of parameters to a vector of means and standard deviations of random parameters associated with 15 choice attributes;
3. For each of the n draws we took $m = 10^5$ draws from multivariate normal distribution described by means and variances of the parameters drawn in step 2;
4. For each of the $n \cdot m$ draws we calculated implicit prices of the attribute levels, and calculated their medians, standard deviations and 95 % quantile ranges.

The results are presented in Table 4. From a policy perspective, the interesting questions relate to the willingness to pay estimates for each management alternative (*LAW*, *FEED*, *MOVE*) and for changes in the two raptor populations, as determined by (1) the modelling strategy and (2) the information provided to respondents. Regarding the former, observing sometimes substantial differences in WTP depending on which modelling approach was used is not uncommon (e.g., Czajkowski et al. 2014a). In our case, the largest differences come from accounting for unobserved preference heterogeneity (i.e. the change from the MNL to any of the models with random taste parameters), particularly for the alternative specific constants (ASC) as represented by *LAW*, *FEED* or *MOVE*. The MNL model seems to suggest no status quo effect in some cases and a positive status quo effect in others, while the other specifications are consistent with an anti-status quo effect—all of the ways of implementing the new plan are preferred to doing nothing but there is no differentiation on average in terms of preferring one approach. These differences can be explained with the presence of substantial preference heterogeneity with respect to different protection programs, as can be seen from the ratios of means of the ASCs to their variances in the RPL or GMNL model (Table 3). As a result, the point estimates of implicit prices of *LAW*, *FEED* and *MOVE* in the MNL model are not significantly different from 0 or are much different from the medians of

²³ It is tempting to speculate as to the reasons for these differences. For example, the stakeholders' information package used in the second study framed its information slightly more than the research team's information set in survey 1. While this could highlight the role of framing in driving behavior, discussion of that issue is beyond the scope of the current study.

Table 4 Simulated median implicit prices of the attribute levels (GB Pounds per household per year)

	MNL model scale study-specific		RPL model scale study-specific		Conventional G-MNL model mean scale study-specific		New G-MNL model mean and variance of scale study-specific	
	Implicit price (SE)	95 % CI	Implicit price (SE)	95 % CI	Implicit price (SE)	95 % CI	Implicit price (SE)	95 % CI
<i>LAW</i> ¹	-4.85 (4.59)	[-13.84, 4.14]	44.12 (15.85)	[19.74, 81.55]	49.45 (11.61)	[30.03, 75.49]	51.65 (13.54)	[29.61, 82.62]
<i>FEED</i> ¹	-2.69 (4.82)	[-12.14, 6.75]	42.94 (15.62)	[19.18, 79.95]	49.56 (11.73)	[29.94, 75.83]	51.64 (13.53)	[29.49, 82.38]
<i>MOVE</i> ¹	-12.62 (4.34)	[-21.13, -4.11]	42.24 (15.52)	[18.62, 78.92]	48.33 (11.51)	[29.09, 74.06]	50.36 (13.34)	[28.64, 80.86]
<i>HH</i> ₁ ¹	30.73 (3.18)	[24.50, 36.95]	10.24 (5.14)	[2.76, 22.69]	18.17 (5.72)	[8.60, 30.90]	17.78 (6.08)	[7.92, 31.66]
<i>HH</i> ₂ ¹	32.13 (3.90)	[24.49, 39.77]	12.51 (5.57)	[4.18, 25.72]	20.18 (5.90)	[10.25, 33.25]	20.11 (6.42)	[9.68, 34.76]
<i>GE</i> ₁ ¹	49.35 (4.58)	[40.37, 58.33]	14.38 (6.29)	[4.74, 28.99]	24.98 (6.87)	[13.15, 39.94]	24.27 (7.17)	[12.18, 40.22]
<i>GE</i> ₂ ¹	62.09 (6.67)	[49.01, 75.16]	17.10 (6.83)	[6.62, 33.18]	29.17(7.50)	[16.27, 45.63]	29.29 (8.07)	[15.66, 47.17]
<i>LAW</i> ²	-33.19 (8.51)	[-49.87, -16.50]	69.75 (19.02)	[38.28, 112.67]	45.78 (9.94)	[28.80, 67.69]	49.17 (11.83)	[29.68, 75.87]
<i>FEED</i> ²	-39.71 (9.23)	[-57.80, -21.61]	73.37 (19.69)	[40.62, 117.73]	47.46 (10.12)	[30.14, 69.88]	51.01 (12.09)	[31.18, 78.45]
<i>MOVE</i> ²	-44.32 (9.94)	[-63.80, -24.83]	65.06 (18.20)	[35.19, 106.35]	41.57 (9.46)	[25.60, 62.50]	45.13 (11.36)	[26.41, 71.00]
<i>HH</i> ₁ ²	43.95 (5.86)	[32.46, 55.44]	7.17 (4.00)	[1.22, 16.78]	12.94 (4.49)	[5.48, 23.02]	12.13 (4.59)	[4.71, 22.67]
<i>HH</i> ₂ ²	57.87 (9.43)	[39.39, 76.34]	9.42 (4.14)	[3.25, 19.32]	15.76 (4.70)	[7.92, 26.37]	14.34 (4.80)	[6.64, 25.45]
<i>GE</i> ₁ ²	77.37 (11.19)	[55.44, 99.30]	14.73 (5.77)	[5.99, 28.36]	22.01 (5.91)	[12.17, 35.24]	22.17 (6.39)	[11.82, 36.78]
<i>GE</i> ₂ ²	84.81 (13.68)	[58.00, 111.62]	18.45 (6.34)	[8.60, 33.29]	26.95 (6.65)	[15.85, 41.90]	26.29 (7.12)	[14.75, 42.65]

the distributions of WTP allowing for unobserved preference heterogeneity.²⁴ The change in WTP when moving from the RPL to the G-MNL model is much less evident, with a modest (although not statistically significant) increase in medians and decrease in confidence intervals of the simulated WTPs. This is the result of allowing for random scale, which increases the model fit and reduces the standard errors associated with model parameters. Accounting for observed (information set specific) scale variance does not seem to be causing significant changes in WTPs, or their confidence intervals, indicating that at least in the case of our dataset, the bias resulting from failing to account for information's effect on scale variance differences may be small.

The approach advanced in this paper can also determine whether providing an alternative information set shifts utility function taste parameters, scale parameters or both and how they can be reflected in respondents' WTP. We start by noting that although there were noticeable changes in the WTPs between information treatments, these differences were generally not statistically significant at the 95 % level. Overall, we found that the respondents did not attach importance to how increases in raptor populations were achieved. This finding holds across information treatments and across econometric methods.²⁵

The absolute values of WTP for each management alternative were substantially different comparing information pack one with information pack two in the MNL and the RPL model. However, this difference does not emerge from either of the G-MNL models. This could be an indication that, at least in the case of our dataset, allowing for unobserved scale heterogeneity (while unobserved taste heterogeneity is also accounted for) controls for a substantial part of the effect of different information sets. Allowing for observed scale heterogeneity provided only a small improvement in this respect.

With regard to changes in populations for either Hen Harriers or Golden Eagles, information pack two (study 2) resulted in smaller absolute WTP values for both species in all econometric treatments relative to information pack one, although these differences are not significant at the 95 % level. As expected, higher implicit prices are observed for a 20 % population increase in either species relative to stabilizing populations at current levels. Comparing WTP for a 20 % increase in either Hen Harriers or Golden Eagles across models for a given information set shows changes in absolute values (e.g., from 10.24 GBP for Hen Harriers in the RPL model, to 18.17 GBP in the conventional G-MNL model, to 17.78 GBP in our modified G-MNL model), but again, these differences are not significant. Finally, in our dataset we find that our modified G-MNL model leads to only modest changes- both increases and decreases- in the standard error of estimated WTP for attributes.²⁶

²⁴ It can also be noted that the estimated standard deviations of ASCs for the RPL model change between information treatment 1 and 2 while this is not the case for the G-MNL models—standard deviations estimated for these models fall between these for the two treatments in the RPL model. It is difficult to draw conclusions solely on this observation, however, since parameters were allowed to be correlated and hence increased standard deviations of parameter estimates do not necessarily translate directly into increased dispersion of respondents' WTP.

²⁵ We note that many of the confidence intervals have gotten smaller between the RPL and the conventional G-MNL model. The new GMNL model appears to shift the upper bound of the confidence interval upward (except for two parameters which have small downward shifts) which is sometimes (but not always) accompanied by a downward shift in the lower bound. That is possibly because the new GMNL model's implicit prices are on average a few percent higher than those resulting from the conventional GMNL model, which in turn are a few percent higher than ones from the RPL model. These differences are not statistically significant, however.

²⁶ It is entirely possible, though, that the G-MNL framework could lead to larger differences in estimated WTP levels and standard errors. One intriguing area of future research, beyond the scope of this paper, is to identify the situations in which the G-MNL leads to larger changes.

Overall, the differences in WTP resulting from information treatments may be difficult to observe because (1) there is large uncertainty and standard deviation associated with all of the estimated WTPs and (2) information provided in the questionnaire can be contested by some respondents. In general, however, our approach shows how one can investigate how changes in an information set can shift not only the location parameters, but also the scale parameter, and how these changes can get reflected in respondents' preference parameters (discussed before) and WTP. Future research to identify precisely when the effect of information would manifest in the location versus scale parameters is needed. Structural models such as ours may be best suited for field experiments embedded in surveys and/or simulation studies aimed at addressing this question.

6 Discussion and Conclusions

For many years, researchers have been interested in the effects of information provided to respondents on their stated values. In a random utility model, changes in information can affect both the relative magnitudes of the estimated deterministic and estimated random components of utility, something which has not so far been highlighted in stated preference work. This paper demonstrates an econometric approach which allows both preference and scale variance heterogeneity to be included, which we argue to be useful for considering how information provided in a CV study can influence the econometrician's ability to predict subjects' choices. The method allows for heterogeneity in preferences across and within information treatments, as well as variations in relative scale within and between treatments. This may be of particular interest to researchers who wish to investigate the effects of presenting different types or varying amounts of information to respondents, for example to reflect conflicting views or uncertainty over the non-market impacts of a project. Whilst we illustrate the method in the context of a public good, it would be applicable to the use of stated preference methods for measuring demand for private goods where different consumers have access to differing levels of information.

Results show that the estimated random element of choice varied across the two information treatments, both in terms of its mean magnitude and its variance. Respondents given a more complete and positively framed information set displayed higher mean relative scale and lower scale variance (e.g., more predictable choices from the econometrician's perspective). This is in addition to within-treatment unobservable variations in the random component of utility. We also find considerable evidence of preference variation in the deterministic component of choices, but small differences in willingness to pay between treatments. Our new G-MNL estimation offered a significant improvement in fit over either a random parameters logit or a standard G-MNL model. Moreover, there were significant differences in parameter estimates between our preferred G-MNL model and the other two models.

None of the above addresses some key questions for information provision in stated preferences. These include how well-informed preferences should be before policy-makers and regulators rely on them for making decisions using cost-benefit analysis, and how much information should be provided to respondents (MacMillan et al. 2006). Moreover, information is often contested, a very relevant example being disagreements within the stakeholder community over the effects of grouse moor management on biodiversity and its contribution to local economic activity, which partly motivated the design of this choice experiment (Thirgood and Redpath 2008). We do not have a quantitative measure of how much information respondents received or assimilated in each treatment, and so are unable to relate this to the precision of willingness to pay estimates (and thus their credibility with decision-makers).

Finally, we note that the paper suggests some fruitful avenues for follow-up work. These include an investigation of the effects of measures of prior familiarity with an environmental good on scale heterogeneity, and a treatment which measures awareness of the characteristics of a good before and after the provision of new information, then relates this to preference and scale heterogeneity. Individual's responses to complexity in information sets could be related to observables such as education and age. There are also interesting questions relating to the interplay between information provision, learning and consequentiality. Structural models, field experiments embedded in surveys and/or simulation may be best suited to address this important question.

Acknowledgments Data collection was funded as part of the EU Framework 6 project "Hunting for Sustainability". We thank two referees for comments on an earlier version of this paper. The first author gratefully acknowledges support of Polish Ministry of Science and Higher Education as well as Foundation for Polish Science.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Aadland D, Caplan A, Phillips O (2007) A Bayesian examination of information and uncertainty in contingent valuation. *J Risk Uncertain* 35(2):149–178
- Alberini A, Rosato P, Longo A, Zanatta V (2005) Information and willingness to pay in a contingent valuation study: the value of S. Erasmo in the Lagoon of Venice. *J Environ Plan Manage* 48(2):155–175
- Baillie SR, Marchant JH, Leech DI, Joys AC, Noble DG, Barimore C, Grantham MJ, Risely K and Robinson RA (2009) Breeding Birds in the Wider Countryside: their conservation status (1972–1996): a report of the BTO's integrated population monitoring. British Trust for Ornithology Research Report No. 516, Thetford
- Ben-Akiva M, Bradley M, Morikawa T, Benjamin J, Novak T, Oppewal H, Rao V (1994) Combining revealed and stated preferences data. *Market Lett* 5(4):335–349
- Ben-Akiva M, Morikawa T (1990) Estimation of travel demand models from multiple data sources. In: 11th international symposium on transportation and traffic theory, Yokoham
- Bergemann D, Välimäki J (1996) Learning and strategic pricing. *Econometrica* 64(5):1125–1149
- Bergemann D, Välimäki J (2006) Dynamic pricing of new experience goods. *J Polit Econ* 114(4):713–743
- Bradley M, Daly A (1992) Estimation of logit choice models using mixed stated preference and revealed preference information. In: 6th international conference on travel behaviour, Quebec
- Bradley M, Daly A (1994) Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data. *Transportation* 21(2):167–184
- Breffle WS, Rowe RD (2002) Comparing choice question formats for evaluating natural resource tradeoffs. *Land Econ* 78(2):298–314
- Brouwer R, Dekker T, Rolfe J, Windle J (2010) Choice certainty and consistency in repeated choice experiments. *Environ Resour Econ* 46(1):93–109
- Carson RT, Czajkowski M (2013) A new baseline model for estimating willingness to pay from discrete choice models. In: International choice modelling conference, Sydney
- Carson RT, Czajkowski M (2014) The discrete choice experiment approach to environmental contingent valuation. In: Hess S, Daly A (eds) *Handbook of choice modelling*. Edward Elgar, Northampton
- Carson RT, Wilks L, Imber D (1994) Valuing the preservation of Australia's Kakadu conservation zone. *Oxford Econ Paper* 46:727–749
- Caussade S, Ortúzar JDD, Rizzi LI, Hensher DA (2005) Assessing the influence of design dimensions on stated choice experiment estimates. *Transp Res Part B Methodol* 39(7):621–640
- Christie M, Azevedo CD (2009) Testing the consistency between standard contingent valuation, repeated contingent valuation and choice experiments. *J Agric Econ* 60(1):154–170
- Christie M, Gibbons J (2011) The effect of individual 'ability to choose' (scale heterogeneity) on the valuation of environmental goods. *Ecol Econ* 70(12):2250–2257

- Czajkowski M, Giergiczny M, Greene WH (2014a) Learning and fatigue effects revisited. Investigating the effects of accounting for unobservable preference and scale heterogeneity. *Land Econ* 90(2):323–350
- Czajkowski M, Hanley N, LaRiviere J, Simpson K (2014b) What is the causal effect of information and learning about a public good on willingness to pay? Evidence from a field experiment. University of Tennessee Working paper
- Czajkowski M, Hanley N, LaRiviere J (forthcoming) The effects of experience on preference uncertainty: theory and empirics for public and quasi-public environmental goods. *Am J Agric Econ*. doi:10.1093/ajae/aa087
- Dellaert BGC, Brazell JD, Louviere JJ (1999) The effect of attribute variation on consumer choice consistency. *Market Lett* 10(2):139–147
- DeShazo JR, Fermo G (2002) Designing choice sets for stated preference methods: the effects of complexity on choice consistency. *J Environ Econ Manag* 44(1):123–143
- Etheridge B, Summers RW, Green RE (1997) The effects of illegal killing and destruction of nests by humans on the population dynamics of the hen harrier *Circus cyaneus* in Scotland. *Appl Ecol* 34(4):1081–1105
- Fiebig DG, Keane MP, Louviere J, Wasi N (2010) The generalized multinomial logit model: accounting for scale and coefficient heterogeneity. *Market Sci* 29(3):393–421
- Fieller EC (1932) The distribution of the index in a normal bivariate population. *Biometrika* 24(3/4):428–440
- Fosgerau M, Bierlaire M (2009) Discrete choice models with multiplicative error terms. *Transp Res Part B Methodol* 43(5):494–505
- Greene WH (2011) *Econometric analysis*, 7th edn. Prentice Hall, Upper Saddle River
- Hanley N, Czajkowski M, Hanley-Nickolls R, Redpath S (2010) Economic values of species management options in human-wildlife conflicts: Hen Harriers in Scotland. *Ecol Econ* 70(1):107–113
- Hensher D, Bradley M (1993) Using stated response choice data to enrich revealed preference discrete choice models. *Market Lett* 4(2):139–151
- Hensher D, Greene W (2003) The mixed logit model: the state of practice. *Transportation* 30(2):133–176
- Hensher D, Louviere J, Swait J (1998) Combining sources of preference data. *J Econom* 89(1–2):197–221
- Hensher DA, Rose JM, Greene WH (2008) Combining RP and SP data: biases in using the nested logit ‘trick’—contrasts with flexible mixed logit incorporating panel and scale effects. *J Transp Geogr* 16(2):126–133
- Hess S, Rose JM (2012) Can scale and coefficient heterogeneity be separated in random coefficients models? *Transportation* 39(6):1225–1239
- Hoehn JP, Lupi F, Kaplowitz MD (2010) Stated choice experiments with complex ecosystem changes: the effect of information formats on estimated variances and choice parameters. *J Agric Resour Econ* 35(3):568–590
- Krinsky I, Robb AL (1986) On approximating the statistical properties of elasticities. *Rev Econ Stat* 68(4):715–719
- LaRiviere J, Czajkowski M, Hanley N, Aanesen M, Falk-Petersen J, Tinch D (2014) The value of familiarity: effects of knowledge and objective signals on willingness to pay for a public good. *J Environ Econ Manag* 68(2):376–389
- Louviere J, Street D, Carson R, Ainslie A, Deshazo JR, Cameron T, Hensher D, Kohn R, Marley T (2002) Dissecting the random component of utility. *Market Lett* 13(3):177–193
- MacMillan D, Hanley N, Lienhoop N (2006) Contingent valuation: environmental polling or preference engine? *Ecol Econ* 60(1):299–307
- Manski CF (1977) The structure of random utility models. *Theor Decis* 8(3):229–254
- McFadden D (1974) Conditional logit analysis of qualitative choice behaviour. In: Zarembka P (ed) *Frontiers in econometrics*. Academic Press, New York, pp 105–142
- McFadden D (1976) The revealed preferences of a government bureaucracy: empirical evidence. *Bell J Econ* 7(1):55–72
- McFadden D, Train K (2000) Mixed MNL models for discrete response. *J Appl Econom* 15(5):447–470
- Meijer E, Rouwendal J (2006) Measuring welfare effects in models with random coefficients. *J Appl Econom* 21(2):227–244
- Milgrom PR (1981) Good news and bad news: representation theorems and applications. *Bell J Econ* 12(2):380–391
- Morikawa T (1989) *Incorporating stated preference data in travel demand analysis*. MIT, Cambridge
- Munro A, Hanley ND (2002) Information, uncertainty, and contingent valuation. In: Bateman IJ, Willis KG (eds) *Valuing environmental preferences*. Oxford University Press, Oxford
- Nelson P (1970) Information and consumer behavior. *J Polit Econ* 78(2):311–329
- Newton I (1998) *Population limitation in birds*. Academic Press, London
- Olsen S (2009) Choosing between internet and mail survey modes for choice experiment surveys considering non-market goods. *Environ Resour Econ* 44(4):591–610

- Payne JW (1976) Task complexity and contingent processing in decision making: an information search and protocol analysis. *Organ Behav Human Perform* 16(2):366–387
- Redpath S, Amar A, Smith A, Thompson D, Thirgood S (2010) People and nature in conflict: can we reconcile hen harrier conservation and game management? In: Baxter JM, Galbraith CA (eds) *Species management: challenges and solutions for the 21st century*, The Stationery Office, pp 335–350
- Redpath SM, Arroyo BE, Leckie FM, Bacon P, Bayfield N, Gutiérrez RJ, Thirgood SJ (2004) Using decision modeling with Stakeholders to reduce human-wildlife conflict: a Raptor-Grouse case study. *Conserv Biol* 18(2):350–359
- Salisbury LC, Feinberg FM (2010) Alleviating the constant Stochastic variance assumption in decision research: theory, measurement, and experimental test. *Market Sci* 29(1):1–17
- Sándor Z, Train K (2004) Quasi-random simulation of discrete choice models. *Transp Res Part B Methodol* 38(4):313–327
- Savage SJ, Waldman DM (2008) Learning and fatigue during choice experiments: a comparison of online and mail survey modes. *J Appl Econom* 23(3):351–371
- Scarpa R, Rose JM (2008) Design efficiency for non-market valuation with choice modelling: how to measure it, what to report and why. *Aust J Agric Resour Econ* 52(3):253–282
- Scarpa R, Ruto ESK, Kristjanson P, Radeny M, Drucker AG, Rege JEO (2003) Valuing indigenous cattle breeds in Kenya: an empirical comparison of stated and revealed preference value estimates. *Ecol Econ* 45(3):409–426
- Shen J (2009) Latent class model or mixed logit model? A comparison by transport mode choice data. *Appl Econ* 41(22):2915–2924
- Swait J, Adamowicz W (2001) Choice environment, market complexity, and consumer behavior: a theoretical and empirical approach for incorporating decision complexity into models of consumer choice. *Organ Behav Hum Decis Process* 86(2):141–167
- Swait J, Louviere J (1993) The role of the scale parameter in the estimation and comparison of multinomial logit models. *J Mark Res* 30(3):305–314
- Taylor L, Morrison M, Boyle K (2010) Exchange rules and the incentive compatibility of choice experiments. *Environ Resour Econ* 47(2):197–220
- Thaler RH (1980) Toward a positive theory of consumer choice. *J Econ Behav Organ* 1:39–60
- Thirgood S, Redpath S (2008) Hen harriers and red grouse: science, politics and human-wildlife conflict. *J Appl Ecol* 45(5):1550–1554
- Thirgood SJ, Redpath SM, Rothery P, Aebischer NJ (2000) Raptor predation and population limitation in red grouse. *J Anim Ecol* 69(3):504–516
- von Haefen RH, Phaneuf DJ (2008) Identifying demand parameters in the presence of unobservables: a combined revealed and stated preference approach. *J Environ Econ Manag* 56(1):19–32
- Watson A, Payne S, Rae R (1989) Golden Eagles *Aquila chrysaetos*: land use and food in North-east Scotland. *Ibis* 131:336–348
- Whitfield DP, Fielding AH, McLeod DRA, Morton K, Stirling-Aird P, Eaton MA (2007) Factors constraining the distribution of Golden Eagles *Aquila chrysaetos* in Scotland: capsule between 1992 and 2003 persecution appeared to be the main influential factor. *Bird Study* 54(2):199–211