Operationalization of ecosystem services for choice experiments: the effect of relevance in the valuation of agri-environmental policies

Annika Tienhaara[1], Heini Ahtiainen[1] & Eija Pouta[1]

[1] Natural Resources Institute Finland (LUKE)

Contact:
annika.tienhaara@luke.fi
heini.ahtiainen@luke.fi
eija.pouta@luke.fi

Contribution presented at the XV EAAE Congress, “Towards Sustainable Agri-food Systems: Balancing Between Markets and Society”

August 29th – September 1st, 2017

Parma, Italy

Operationalization of ecosystem services for choice experiments: the effect of relevance in the valuation of agri-environmental policies

Abstract

Valuing ecosystem services with stated preference methods requires operationalizing these services for the purposes of the survey. This study presents a process for selecting agricultural ecosystem services for attributes of a choice experiment and analyzes how the relevance of these attributes affects respondents' preferences, allowing for individual preference and scale heterogeneity. The results show that a non-significant cost attribute is associated with low relevance of the attributes, and that respondents who consider the attributes relevant have less uncertainty in their answers. The findings emphasize the importance of attribute selection when the object of the valuation is complex, such as certain ecosystem services.

Keywords: agriculture, ecosystem services, choice experiments, environmental valuation, relevance of attributes

1. Introduction

Ecosystem services (ES), providing a link between the ecosystem and human well-being, are increasingly included in environmental policy assessments. The ecosystem service framework and classification have developed rapidly since the Millennium Ecosystem Assessment (MA, 2005). The ES typology has been applied in various fields and contexts, and also developed further in parallel classifications, such as The Economics of Ecosystems and Biodiversity (TEEB 2010) and Common International Classification of Ecosystem Services (CICES) (Haines-Young & Potschin 2012). ES are typically classified into provisioning, regulating and cultural services. Some classifications also have a category for supporting or habitat services. For valuation purposes, ES are often divided into intermediate and final services depending on their connection to human welfare (Fisher, Turner & Morling 2009).

The fundamental aim of the ecosystem service concept is to guarantee that all of the contributions that ecosystems provide for people are taken into account in decision making. One option to assess these contributions is to identify and value changes in ecosystem services resulting from environmental policies. Those ES, for which markets and prices do not exist, can be valued using economic valuation methods, such as discrete choice experiments (CE), which elicit citizens' willingness to pay for changes in the environment using surveys.

In the CE, respondents choose between two or more discrete alternatives that are described with several attributes. As the attribute levels vary and a price variable is included as one of the attributes, respondents' willingness to pay (WTP) for alternative or attribute level is indirectly revealed based on the choices they make (e.g. Hanley et al. 2001). As individuals who respond to the valuation survey have limited capacity to process information, only few attributes can be included. Consequently, also some ES that are affected by the policy might be left out.

To a certain degree, selecting relevant ES for the CE is relatively easy. Typically, there is no need to use stated preference valuation techniques for ES that have markets and prices, e.g. some provisioning services, such as food and lumber. Intermediate services that contribute to the final services can also be left out from the potential attributes, because their value is embedded in the final services and including both intermediate and final services could lead to double counting (Fisher et al. 2009).

Still, a demanding challenge for the analyst is to choose the relevant ES for the CE study. Stated preference literature suggests some guidelines for the attribute selection using various types of qualitative processes, such as focus groups and stakeholder meetings (Bateman et al. 2002, Blamey et al. 2002; Coast et al. 2012; Abiiro et al. 2014), but there are only a few studies that have developed or tested these processes empirically (Coast and Horrocks 2007; Armatas, Wenn & Watson 2014; Jeanloz et al. 2016). Armatas et al . (2014) identified the attributes for water-based ecosystem services by applying rank orderings and the Q-method. Jeanloz et al. (2016) provided and tested a five-stage qualitative process for attribute selection in relation to ecosystem services. This process included literature review and consultations, discussion protocol, scoring and ranking characteristics, analysis, and final attribute selection.

These methods of attribute selection guide researcher to select attributes that, on average, have higher importance for people than other attributes. Nevertheless, stated preferences valuation methods rely on the analyst's ability to identify, select, define and articulate the goods (Armatas et al. 2014). Even though the analyst would follow the guidelines and have previous experience and expertise of attribute selection, there is a chance that some survey respondents consider that relevant ES have been excluded.

The relevance of the attributes presented in the CE can determine the attribute processing strategies adopted by a respondent (Hensher 2006; Hensher, Rose & Greene 2012). Respondents who consider the valuation task to be totally or partly irrelevant due to low importance of the attributes may indicate protest behavior, or discontinuous, zero and low preferences (e.g. Hensher, Rose & Greene 2005; Campbell, Hensher & Scarpa 2011; Alemu et al. 2013). Attribute non-attendance (ANA) occurs when one or more attributes are ignored by the respondent when making choices, thus implying zero preferences (Campbell et al. 2011). It is generally agreed that there are always respondents who will ignore the attributes to some extent (Alemu et al. 2013). Hence, it is important to explore the underlying reasons of why respondents ignore the attributes (Carlsson, Kataria & Lampi 2010; Scarpa, Thiene & Hensher 2010). The low relevance of attributes which stems from difficulties in choosing ecosystem services to the CE is one possible reason for attribute non-attendance.

In this study, we explore the relevance of the ES included in the choice experiment, and how the heterogeneity in relevance affects respondents' choice behavior. The data come from a CE that includes the ES from the Finnish agri-environmental policy as the attributes. Many previous choice experiments of agri-environmental programs have related to the characteristics of agricultural or rural landscapes or multifunctional agriculture. The ES typology has been the base of some CE studies concerning agricultural ecosystems (Bernués et al. 2015; Takatsuka et al. 2009 Rodríguez-Ortega, Bernués & Alfnes 2016). Takatsuka et al. (2009) selected four ES for valuation: greenhouse gas emissions, water quality, soil quality and scenic views of cropping farms. In defining attributes, Bernués et al. (2015) started with the multifunctionality framework and translated the functions identified by stakeholders into four different ES for economic valuation. Rodríguez-Ortega et al. (2016) selected the most important ES provided by Mediterranean high nature value farmland: agricultural landscape maintenance, biodiversity conservation, forest wildfire prevention and availability of local quality products. The importance was assessed from the biophysical point of view, but also evaluated based on the socio-cultural perspectives using focus groups.

We examine the operationalization of the ES concept and classification in the choice experiment setting. We describe an example of a process of selecting the services, being simultaneously aware that the selected services are not the most relevant ones for each respondent. Thus, we identify different respondent groups based on the relevance of the attributes to them, and analyze how the relevance of the services affects choices and welfare estimates. We also produce benefit estimates for selected ES from agricultural environments.

2. Methods and data

Finnish agri-environment policy and ecosystem services

In Finland, the focus of the agri-environmental policy is on three main targets: water conservation, biodiversity and climate change mitigation. The current policy is designed to compensate the expenses from agri-environmental measures to farmers and it does not demand or ascertain the production of public goods or ES. It has been suggested that the agri-environmental policy in

Europe should be converted to more results or benefit-based direction in order to be effective and cost-efficient (for review Schwarz et al. 2008, Burton & Schwarz 2013, Russi et al. 2016). To develop this kind of a policy, it would be essential to know the value of various ES from the benefit-based policy for the beneficiaries, i.e. citizens.

Identifying ecosystem services for valuation

In this study, we applied the Common International Classification of Ecosystem Services (CICES) classification as a base for identifying agricultural ES for valuation (CICES 2016). The CICES classification was selected as it is a continuously developing European wide classification system that can be used also for the purposes of valuing ES (CICES 2016). To select the relevant ES from the CICES classification, we first selected 13 ES based on a literature review and expert judgement by agricultural economists and ecologists. The selected services were: 1) food, 2) agro-diversity, 3) bioenergy, 4) pollination, 5) habitats for animal nursery and reproduction 6) pest control, 7) soil productivity, 8) cultural heritage, 9) existence of species and ecosystems, 10) recreation environment, 11) landscape, 12) water quality and 13) climate change.

To select the attributes from these 13 ES for the choice experiment, we followed the steps from 1 to 8.

1. Analysis of the importance of ES for citizens based on previous survey data (N=800) (Pouta & Hauru 2015)
2. Evaluation of the importance of agri-environmental ecosystem services by stakeholders from the administration and NGOs (N=6)
3. Stakeholder discussion of the relevant ES based on step 2 (N=7)
4. Researchers' (N=9) summary of steps from 1 to 3 and analysis of market and non-market services, as well as final and intermediate services
5. Valuation experts' (N=10) evaluation of the questionnaire and the choice experiment
6. Attribute selection for the pilot study
7. Pilot study (N=202)
8. Researchers' decision on the attributes in the valuation task of final survey

After this process the selected services were landscape, existence of species and ecosystems, water quality due to agriculture, and climate change due to agriculture. The project group of environmental economists, ecologists and agri-environmental policy experts (N=12) developed these selected ecosystem services to measurable attributes and their levels. In defining the attribute levels, we looked for concrete indicators that could be affected by farming practices and thus targeted with the agri-environmental policy.

Data collection

The survey data were collected using an Internet survey during the spring of 2016**.** The sample was drawn from the Internet panel of a private survey company, Taloustutkimus, which comprises over 30 000 respondents who have been recruited to the panel using random sampling to represent the population (Taloustutkimus, 2017). After a pilot survey of 202 people, a random sample of 8391 respondents was selected, of which 2066 completed the survey. This corresponds to a response rate of 25%.

Grouping respondents

As only some of the thirteen ES produced on agricultural areas could be included in the CE, the respondents were likely to include both those for whom the selected ES were relevant and those that were more interested in excluded services. There could also be respondents with high interest to all thirteen ES or none of them. To separate these different groups of respondents, we applied a variant of multi-criteria analysis (MCA) for policy evaluation (e.g. Gezner et al. 2004; Blechinger & Shan 2011; Huang et al. 2011; Michailidou et al. 2016).

After introducing the basic facts about Finnish agriculture in the survey, the components for MCA were measured. The personal importance (e_i) of the thirteen ES produced by the Finnish agriculture and agricultural environment was assessed with a five point scale ranging from "very large importance" to "very small importance". This was followed by a description of the principles of the current agri-environmental programs, and asking the respondents' perceptions (p_i) of how Finnish agriculture has succeeded in producing the thirteen ES mentioned. The five point scale ranged from "extremely well" to "extremely poorly". If the current policy failed to produce an ES that respondent considered important, this ES was considered to be particularly relevant to the respondent in the new agricultural policy introduced in the CE.

Thus, we constructed a variable that indicates the relevance of ES to a respondent. The component of importance (e_i) and perception of the current state (p_i) of a particular ES were multiplied to form the relevance of that ES e_i*p_i . These products were summed for the group of ES that were included in the CE as attributes, as well as for the group of ES that were left out from the CE. These two sum variables $\sum e_i*p_i$ were used to form a fourfold table. As the general tendency of respondents was to consider services as important rather than unimportant, we used the upper quartiles from the ranges of the sum variables to separate the groups. The resulting groups were 1) neither included or excluded ES were relevant, 2) included ES were relevant, 3) excluded ES were relevant and 4) both included and excluded ES were relevant. These four groups were used to analyze how the relevance of the ES used in the choice experiment affected the respondents' preferences.

Choice experiment

In the next part of the survey, a new benefit-based agri-environmental policy was introduced to the respondents by telling that in the hypothetical new program farmers would get paid for producing environmental benefits. The effects of the program were described with four attributes: traditional rural biotopes and endangered species, typical agricultural landscape, climate effects and water quality effects. These attributes were described to the respondents with few sentences and information was given regarding the current state of the attribute as well as the different attribute levels. Table 1 presents the attributes together with their descriptions and levels.

Next, the survey explained that the new agri-environmental program would be financed with taxes and that depending on the extent of the program, the cost to taxpayers would vary, but all taxpayers would participate in financing the programme. It was told that the current program also causes expenses to citizens, based on the expert judgement approximately 40 euros per individual per year. Consequentiality was enforced by stating that the information from the choice tasks would help decision-makers to revise the agri-environmental program.

After introducing the attributes, the respondents were presented with six choice tasks. Each choice task comprised of three alternatives: the status quo alternative, described as maintaining the current program, and two alternatives with higher levels of ES compared to the current state. Each alternative was described with four ES attributes, their levels and the cost attribute. The status quo alternative was identical across choice tasks. An example of a choice task is shown in Table 2.

To allocate the attribute levels to the choice tasks in the CE, an efficient experimental design was constructed. Efficient designs are used to generate parameter estimates with standard errors that are as low as possible, and thus to get the maximum information from each choice situation (see e.g. Rose and Bliemer 2009). The generation of efficient design requires the specification of priors for the parameter estimates. We employed zero priors in the design of the pilot survey. In the final study, we employed a Bayesian D-efficient design using Ngene (v. 1.0.2), taking 500 Halton draws for the prior parameter distributions and using the parameter estimates obtained from the pilot study. Bayesian designs take into account the uncertainty related to the parameter priors. We used a Bayesian prior for the attribute plants in cultivation and fixed priors for all other attributes.

Overall, we generated a design with 36 choice tasks, blocking them in 6 subsets, which resulted in six choice situations for each respondent. Four different versions of the design were created using four different cost scales. Cost scales varied from 5-300€, 5-500€, 40-500€ and 40-300€. Apart from the varying cost scale, the design was identical between the four different versions. The final design had a D-error of 0.08829.

3. Statistical models

We used the latent class and the scale-adjusted latent class model to examine respondents' choice of agri-environmental policies, allowing for heterogeneity in both preferences and scale. As we were especially interested in how the relevance of the selected attributes in the CE is related to respondents' choices, the relevance groups were included as covariates in all models.

The choice experiments, based on random utility theory, have traditionally been modeled with the conditional logit model (CL) (Mcfadden 1974). However, CL model assumes similar preference structure across all respondents. This is a somewhat unrealistic assumption and defining heterogeneous respondent segments has been an interest in many studies. The latent class model is one approach that allows for heterogeneity in preferences (Boxall and Adamovicz 2002). The latent class model simultaneously divides individuals into latent segments and estimates a choice model in these classes. In each latent class, preferences are assumed to be homogeneous, but between the segments, they are assumed to vary. The estimation is first carried out for one class, then two classes, three classes and so on. In each step, the explanatory power of the model is assessed to select the optimal number of classes. For this purpose, several information criteria, including Bayesian (BIC), Akaike's (AIC) and corrected-AIC (CAIC) information criteria, can be used. The latent class model also enables the estimation of willingness to pay for ES with various attribute combinations for different respondent segments.

We also explored whether attribute selection could cause attribute non-attendance (A-NA), i.e. respondents ignoring attributes irrelevant to them. Hensher et al. (2005) were the first to include attribute non-attendance in their modelling. Much of the past research has asked respondents directly whether or not they paid attention to the attributes in a choice experiment. However, the reliability of the responses to such questions has raised concerns (Hensher et al. 2012). There has been a rise in the number of research that identify attribute non-attendance by analyzing observed response pattern through model inference instead of supplementary questions (eg. Lagarde 2013, Hensher et al. 2012, Scarpa et al. 2009). Attribute non-attendance (A-NA) models are often constructed in a way that an attribute that is not attended to is constrained to zero and attributes that are attended to, take the same value across all classes. However, our aim was to identify preference heterogeneity among classes so this constraint was relaxed and values were allowed to differ between classes. Our attribute non-attendance model allows some respondents to belong to latent class with zero utility weights, while non-zero attributes are assumed for other classes. Respondents in the total non-attendance class ignore all attributes and give their responses based on chance (Scarpa et al. 2009). In this class, all coefficients were fixed to zero. We did not model all possible attribute non-attendance combinations, since our aim was only to reveal those respondents who respond completely by chance and to see whether the relevance of the attributes in CE affects the non-attendance.

There has been increasing discussion about the scale parameter in choice modelling (Fiebig et al. 2009, Louviere et al. 2002, Greene & Hensher 2010). The scale parameter reflects respondent's certainty and it is inversely related to response variance. Scale-Adjusted Latent Class (SALC) models allow for differences in scale, in addition to heterogeneity in preferences, hence resulting in segments that differ purely in preference. SALC models were introduced by Magdison and Vermunt (2007). In standard discrete choice models, there can be confounding between scale factors and utilities, and as a result, two respondent segments might appear to have different utilities due to a scale factor despite having exactly the same preferences. Comparison between class estimates is therefore difficult since it is not clear whether the differences in parameters are a result of true difference in preferences or just difference in error variability (Louviere & Eagle 2006). SALC models contain two latent variables X and S. X indicates segments differing in their preferences and S denotes the differing scale parameter. S can be either categorical or continuous. We used two scale classes in our modelling. The scale parameter for the first subgroup is always fixed to 1 for identification purposes and the values for other scale parameters are assumed to be non-negative (Magdison & Vermunt 2007). SALC model estimates the log-scale factor to guarantee non-negative values and therefore estimates for scale factors are obtained by exponentiating the scale factor parameters. SALC models have been used in various studies ranging from pesticide use (Glenk et al. 2012) to educational research (Burke et al 2015) and water quality (Thiene, Scarpa & Louviere 2015).

4. Results

The results show that policy programs were popular among the respondents as the status quo option was selected on average in 7% of the choice sets. However, there were significant differences (p-value<0.001) between all relevance groups, except between the groups for whom none of the ES were relevant and mostly excluded ES were relevant. The respondents who considered neither included nor excluded ES to be relevant (13% of the respondents) chose the status quo in 11% of the choice sets. In turn, those respondents to whom all ES were relevant (67% of the respondents) were the most willing to support new policy programs, as they chose the status quo only in 6% of their choices. Similarly, those who felt mostly ES that were excluded from the CE to be relevant (8% of the respondents) selected the status quo more often than those respondents who considered mostly ES included in CE to be relevant (12% of the respondents). Portions of choosing the status quo for these two groups were 10% and 7%, respectively.

Preference heterogeneity

The conditional logit was estimated as the reference model. The results are reported in the first column of Table 3. We then began our specification search for preference heterogeneity and latent class models from 1 up to 8 classes were estimated. Estimation was carried out using Latent Gold Choice 5 and effects coding. The effects of adding a scale parameter and total attribute non-attendance class were then tested for each class count. Policy relevance was used as a covariate in all of the models. The first model specification examining the heterogeneity of preferences resulted in 5 classes. Table 3 presents the results of the latent class model for the choice of the policy program showing that there was significant heterogeneity between the groups for all attributes.

First class was the largest comprising one third of the respondents. These respondents frequently chose the policy programs over the current program and almost all of the program attributes were significant. The cost of the program was significant, but the coefficient was really small indicating a low sensitivity to cost. The relevance covariate shows that Class 1 considered all ES to be relevant. These respondents could be described as "environmentalists" based on their support for improvements in environmental quality and low sensitivity to cost attribute, which indicates high willingness to pay for environmental improvements. Class 2 was large as well, with approximately 30% of the respondents. Similarly to Class 1, these respondents considered all ES to be relevant. Additionally, ES that were included in the CE were particularly relevant for these respondents. This is why this group was named "scenario focused". All attributes were significant and of the expected sign, except for plants in cultivation. Third class (17% of the respondents) did not consider all ES as policy relevant. ASCs reveal that this class still chose the policy programs over the current program, but noticeably fewer attributes were significant for this class. These respondents had the highest marginal utility of money and can therefore be called "cost sensitives". Class 4 (10%) consisted of respondents who did not consider any of the ES as relevant and chose more often the status quo, i.e. the current program than other classes. Hence, the fourth group was named "not interested". This group had only couple of significant attributes and the highest level of water quality had an unexpected negative sign. This indicates that the lower level of water quality improvement was sufficient for the class. The cost attribute was significant for this group. The smallest class was Class 5 with only 8% of the respondents. The attributes included in the CE were not relevant to these respondents. Instead, the excluded attributes were perceived as relevant. Accordingly, Class 5 is called “outsiders”. The lacking relevance of attributes is evident since only climate effects had significant coefficients. In addition, Class 5 was the only class for which the cost attribute was not significant. ACSs show that these respondents were likely to choose the first of the two programs. This may imply that the low relevance of attributes in CE caused them to use simplifying strategies and select the option in the middle.

Attribute non-attendance

To examine the relationship between the relevance of attributes in CE and attribute non-attendance, we assumed one class to answer completely by chance, i.e. all attribute coefficients in this total non-attendance class were set to zero. The model with total non-attendance class also resulted in five classes. The results from this model were very similar to the basic latent class model. The size of total non-attendance class was 6.76% of the respondents. This class was positively related to the relevance of ES that were excluded from the CE (0.57, significance level 0.01). Furthermore, ES included in the CE were not relevant to these respondents (-0.60, significance level 0.1).). This suggests that the respondents, who did not consider the attributes in the CE relevant, were likely to answer purely by chance.

Scale heterogeneity

In order to further improve the basic latent class model, we added a scale parameter in the model specification. Table 4 reports the results of scale adjusted latent class model. The relevance of the CE attributes was used as a covariate for both preferences and scale. Model with 5 classes was chosen based on BIC and CAIC.

All of the classes in SALC model resembled greatly the classes in the basic latent class model, with the exception of Class 2 being largest class in SALC model. Class 1 had significant coefficients mostly for the SQ level and for the higher improvement level. Lower level of improvement was significant only for traditional biotopes and endangered species. The cost attribute had very low coefficient in this class. These results suggest that respondents in class 1 support greater policy effort even at higher cost. Class 2 comprised 33.6% of the respondents. This class considered all ES to be relevant and supported new policy programs. Class 2 also had more significant attributes than the other classes. All attributes were highly significant (except for plants in cultivation) and of the expected sign. Class 3 (15.7% of the respondents) did not perceive any of the ES relevant. Still, these respondents chose new policy programs more frequently than the current program. Few of the program attributes were significant. In addition, the respondents in class 3 were clearly the most sensitive concerning the cost of the program. The fourth class, with 11.2% of the respondents, considered none of the ES to be relevant. Even so, there is a slight indication that ecosystem services that were excluded from the CE might be relevant to these respondents. The non-relevance of the attributes leads to favoring the SQ, as these respondents noticeably chose SQ over the new programs. This group had nearly no significant attributes, besides the cost. For Class 5, comprising 7.8% of the respondents, ES that were excluded from the CE were relevant. This is reflected in the rather confusing behavior of choosing SQ and the first alternative over the second alternative. The cost attribute was not significant.

The SALC model showed a significant difference in the scale parameters between the scale classes (p-value 0.000). The scale factor for the scale class 1 was fixed to 1 and the scale factor estimated for the scale class 2 was 0.288, suggesting that scale class 2 had a greater error variance and hence was more uncertain relative to scale class 1. The relevance of the attributes in the CE was used as a covariate also for the scale. As can be seen in the bottom section of Table 5, relevance significantly affected the scale, but only for the respondent group that perceived attributes in CE to be relevant. This group had negative coefficient for scale class 2, indicating that the respondents for whom the attributes were relevant were more likely to belong to the more certain scale class 1.

Willingness to pay

As the knowledge on the value of ES is essential in designing new agri-environmental policies, we also estimated willingness to pay for different attribute levels. As the models used effects coding, it was possible to calculate the monetary values for all attribute levels. Table 8 shows willingness to pay estimates that were calculated both based on the conditional logit and the basic latent class model without the scale parameter. Non-significant measures are not reported and they can be interpreted as zero.

Reported measures are annual willingness to pay per individual for ten-year period between 2017 and 2026. As the models were estimated using effects coding, the willingness to pay for moving from one attribute level to another is obtained by calculating the difference between them. For example, for class 2, willingness to pay for decreasing agricultural greenhouse-gas emissions 30% from the current level is 31€. Confidence intervals were calculated with the delta method.

5. Discussion and conclusions

This study examined how the concept of ecosystem services can be operationalized for the choice experiment method in order to value agri-environmental policies. Our main focus was on studying how the relevance of the selected ecosystem services, specified as attributes in the CE, affected respondents´ choices. We identified four respondent groups based on the importance of the ecosystem services provided by the agricultural environment and perceptions on the success of the agricultural policies in providing these ecosystem services, i.e. the relevance of the ecosystem services to the respondents.

The results showed that the low relevance of the ecosystem service attributes was associated with a non-significant coefficient of the cost attribute, indicating that the cost did not affect the choices of those who considered the ecosystem service attributes irrelevant. The non-significance of the cost can lead to fat tail problems and unrealistically high willingness to pay estimates if it is not detected. This highlights the importance of modelling preference heterogeneity, as assuming homogenous preferences while some respondents ignore the cost can distort the results. Low relevance of the ecosystem service attributes also created conditions for attribute non-attendance. Those respondents, who considered only the excluded attributes relevant, were more likely to make their choices completely by chance.

In contrast, many of the attributes had a significant effect on utility for the respondent groups that perceived the attributes in the CE to be relevant. These respondents divided into two large classes that differed with respect to their sensitivity to the cost. The "environmentalists" selected ES improvements even with higher cost levels, whereas "scenario-focused" had lower willingness to pay estimates while still favoring the improvements.

Introducing the scale parameter to the estimation revealed that there was scale heterogeneity alongside preference heterogeneity across respondents. The respondents' choices appeared less random from the perspective of the analyst, when the CE included only attributes that were relevant to the respondent. Including relevant attributes in the CE is thus important to obtain more reliable results.

Nevertheless, this leads to an unresolved question in the CE design. Is it better to have a wide range of attributes that are likely to include also relevant ones or a more limited number of attributes to reduce the cognitive burden of the CE task? Our results support the importance of a careful selection process for the attributes, especially when the object of the valuation is complex, such as certain ecosystem services. Our selection process can be regarded as rather successful as attributes included in the CE were relevant to 79% of the respondents. Even so, the attribute describing a typical agricultural landscape in terms of the number of plants in cultivation was problematic, as it was clearly less significant compared to other attributes. The agricultural landscape is proven to be an important ecosystem service (Pouta & Hauru 2015), but converting it to simple indicators was challenging, especially for the cultivated plants. The levels for cultivated plants were close to another, ranging from 3 to 5 plants, possibly causing the non-significance. However this range was an expert opinion on realistic scenarios. This illustrates the challenges in choosing the right indicators to describe the ES in stated preference valuation surveys.

Further studies are needed on the use of ecosystem services in stated preference studies and determining the relevance measures for the attributes in the choice experiment. In our study, using the importance of agricultural ecosystem services and perceptions on the successfulness of policies to provide these services for constructing the relevance measures ended up in rather small differences in measures of relevance across respondents. Ranking scale could be a better fit for this purpose.

Our results suggest that the relevance of the attributes selected for the CE clearly affects respondents' choices. This implies that if a comprehensive CE or the careful selection of attributes in CE is not possible and, as a result, relevant attributes are excluded, anomalies and random choices are likely to occur. However, this can partly be addressed by allowing both preference and scale heterogeneity in the modelling.

6. References

Abiiro, G., Leppert, G., Mbera, G., Robyn, P., De Allegri, M., 2014. Developing attributes and attribute-levels for a discrete choice experiment on micro health insurance in rural Malawi. *BMC Health Services Research* 14 (1): 235.

Ajzen, I., and Fishbein, M. (1980). *Understanding Attitudes and Predicting Social Behavior*, Prentice-Hall, Englewood Cliffs, N.J.

Alemu, H., Mørkbak, M., Olsen, S. & Jensen, C. 2013. Attending to the reasons for attribute non-attendance in choice experiments. *Environmental and Resource Economics* 54: 333-359.

Armatas, C., Venn, T. & Watson, A. 2014. Applying Q-methodology to select and define attributes for non-market valuation: A case study from Nortwest Wyoming, United States. *Ecological Economics* 107: 477-456.

Bateman, I., Carson, R., Day, B., Hanemann, M., Hanley, N., Hett, T., Jones-Lee, M., Loomes, G., Mourato, S., Özdemiroglu, E., Pearce, D., Sugden, R. & Swanson, J. 2002. *Economic valuation with stated preference techniques: a manual.* Edward Elgar, Cheltenham, UK.

Bernués, A., Rodríguez-Ortega, T., Alfnes, F., Clemetsen, M. & Eik, L. 2015. Quantifying the multifunctionality of fjord and mountain agriculture by means of sociocultural and economic valuation of ecosystem services. *Land Use Policy* 48: 170-178.

Blamey, R.K., Bennett, J.W., Louviere, J.J., Morrison, M.D. & Rolfe, J.C.,2002. Attribute Causality in Environmental Choice Modelling. *Environmental and Resource Economics* 23: 167–186.

Boxall, C.P. and Adamowicz, L.W. 2002. Understanding Heterogeneous Preferences in Random Utility Models: A Latent Class Approach. *Environmental and ResourceEconomics* 23: 421–446.

Burke, P. Aubusson, P., Schuck, S. Buchanan, J. Prescott, A. 2015. How do early career teachers value different types of support? A scale-adjusted latent class choice model. *Teaching and Teacher Education* 47: 241-253.

Burton, R. and Schwarz, G. 2013. Result-oriented agri-environmental schemes in Europe and their potential for promoting behavioural change. *Land Use Policy* 30: 628–641.

Campbell , D.; Hensher, D.A. & Scarpa, R. 2011. Non-attendance to attributes in environmental choice analysis: a latent class specification. *Journal of Environmental Planning and Management* 54 (8): 1061-1076.

Carlsson, F., Kataria, M. & Lampi, E. 2010. Dealing with ignored attributes in choice experiments on valuation of Sweden´s environmental quality objects. *Environmental and Resource Economics* 47: 65-89.

Coast, J., Horrocks, S., 2007. Developing attributes and levels for discrete choice experiments using qualitative methods. *Journal of Health Services Research & Policy 12* (1): 25–30.

Coast, J., Hareth, A., Sutton, E.J., Horrocks, S.A., Vosper, A.J., Swancutt, D.R., Flynn, T.N., 2012. Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. *Health Econ*omics 21: 730–741.

Fiebig, D. G., Keane, M. P., Louviere, J., & Wasi, N. 2010. The generalized multinomial logit model: accounting for scale and coefficient heterogeneity. *Marketing Science* 29(3): 393-421.

Fisher, B., Turner, K. & Morling, P. 2009. Defining and classifying ecosystem services for decision making. *Ecological Economics* 68: 643-653.

Glenk, K., Hall, C., Liebe, U. & Meyerhoff, J. 2012. Preferences of Scotch malt whisky consumers for changes in pesticide use and origin of barley. *Food Policy* 37: 719-731.

Greene, W. & Hensher, D. 2010. Does scale heterogeneity across individuals matter? An empirical assessment of alternative logit models. *Transportation* 37: 413-428.

Haines-Young, R. & Potschin, M. 2012. CICES Version 4: Response to consultation. Centre for Environmental Management, University of Nottingham.

Hanley, N., Mourato, S., and Wright, R. E. (2001). Choice modelling approaches: a superior alternative for environmental valuation? *Journal of Economic Surveys* 15: 435–462.

Hensher, D. 2006. How do respondents process stated choice experiments? Attribute consideration under varying information load. *Journal of Applied Econometrics* 21: 861-878.

Hensher, D., Rose, J. & Greene, W. 2005. The implications on willingness to pay of respondents ignoring specific attributes. *Transportation* 32: 203–222.

Hensher, D., Rose, J. & Greene, W. 2012. Inferring attribute non-attendance from stated choice data: an application for willingness to pay estimates and a warning for stated choice experiment design. *Transportation* 39: 235–245.

Jeanloz, S., Lizin, S., Beenaerts, N., Brouwer, R., Van Passel S. & Witters, N. 2016. Towards a more structured selection process of attributes and levels in choice experiments: A study in a Belgian protected area. *Ecosystem Services* 18: 45-57.

Lagarde, M. (2013). Investigating attribute non-attendance and its consequences in choice experiments with latent class models. *Health Economics* 22: 554–567.

Louviere, J. & Eagle, T. 2006. Confound it! That pesky little scale constant messes up our convenient assumptions. Proceedings, 2006 Sawtooth Software Conference, Sawtooth Software, 211-228.

Louviere, J., Street, D., Carson, R., Ainslie, A., Deshazo, J. R., Cameron, T., Hensher, D., Kohn, R. & Marley, T. (2002). Dissecting the random component of utility. *Marketing Letters* 13(3): 177-193.

Magdison, J. & Vermunt, J.K. 2007. Removing the scale factor confound in multinomial logit choice models to obtain better estimates of preference. Proceedings, 2007 Sawtooth Software Conference, Sawtooth Software, 139-154.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. In Zarembka, P.E. (ed.), *Frontiers in Econometrics*. New York: Academic Press, 105–142.

Pouta, E. & Hauru, K. Citizens' evaluation of the importance of ecosystem services. In Jäppinen, J-P & Heliölä J. (eds.) Towards A Sustainable and Genuinely Green Economy. The value and social significance of ecosystem services in Finland (TEEB for Finland) The Finnish Environment 1/2015.

Rodríguez-Ortega, T., Bernués, A. & Alfnes, F. 2016. Psychographic profile affects willingness to pay for ecosystem services provided by Mediterranean high nature value farmland. *Ecological Economics* 128: 232-245.

Rose, J. M., and Bliemer, M. C. J. (2009). Constructing efficient stated choice experimental designs. *Transport Reviews* 29: 597–617.

Russi, D., Margue, H., Oppermann, R., Keenleyside, C. 2016. Result-based agri-environment measures: Market-based instruments, incentives or rewards? The case of Baden-Württemberg. *Land Use Policy*: 54: 69-77.

Scarpa, R., Gilbride, T.J., Campbell, D.; Hensher, D.A. 2009. Modelling attribute non-attendance in choice experiments for rural landscape valuation. *European Review of Agricultural Economics* 36: 151-174.

Scarpa, R.; Thiene, M. & Hensher, D. 2010. Monitoring choice task attribute attendance in nonmarket valuation of multiple park management services: does it matter? *Land Cconomics* 86 (4): 817-839.

Schwarz, G., Moxey, A., McCracken, D., Huband, S. & Cummins S. 2008. An analysis of the potential effectiveness of a Payment-by-Results approach to the delivery of environmental public goods and services supplied by Agri-Environment Schemes. Report to the Land Use Policy Group, UK. Pareto Consulting and Scottish Agricultural College, Macaulay Institute p. 108

Takatsuka, Y., Cullen, R., Wilson, M. & Wratten, S. 2009. Using stated preference techniques to value four key ecosystem services on New Zealand arable land. *International Journal of Agricultural Sustainability* 7(4): 279-291.

Taloustutkimus. 2017. Internet panel. http://www.taloustutkimus.fi/in-english/products_services/internet_panel/. Accessed January 25. 2017.

TEEB. 2010. The Economics of Ecosystems and Biodiversity: Mainstreaming the Economics of Nature: A Synthesis of the Approach, Conclusions and Recommendations of TEEB. Earthscan, London and Washington.

Thiene, M., Scarpa, R. & Louviere, J.J. 2015. Addressing Preference Heterogeneity, Multiple Scales and Attribute Attendance with a Correlated Finite Mixing Model of Tap Water Choice. *Environmental and Resource Economics* 62: 637–656.

Tables

Table 1. Attributes of conservation programmes and their levels

Attribute	Description	Current state	Levels
Traditional rural biotopes and endangered species	Mowed or grazed seminatural grasslands (meadows, leas, pastures) can provide a habitat for several endangered species.	20 000 hectares, or about one per cent of agricultural land surface area. In Finland there are 2250 endangered species, about 700 live in seminatural grasslands.	• current area, 0 species protected • 30 % increase in area, 100 species protected • 60 % increase in area, 200 species protected
Typical agricultural landscape	Grazing animals and crops grown in open fields affect the diversity of the landscape.	Grazing animals rarely appear in landscape and few plant species are cultivated.	Grazing animals: • seldom seen • seen often during summer season • seen often during summer and the unfrozen season Plants in cultivation: • 3, 4, 5 species
Climate effects	Agricultural greenhouse gas emissions contribute to climate change. Greenhouse gas emissions can be reduced by various cultivation practices and capturing greenhouse gasses.	The agricultural sector produces 11% of Finland's greenhouse gas emissions.	0, 10, 30 % decrease in current emissions
Water quality effects Share of surface waters in good or excellent state	About half of the nutrient runoff to waters comes from fields. This is affected by the amount of fertilizers used, cultivation practices and annual weather conditions.	About 60% of the surface waters are in good or excellent condition.	60, 70, 80 %
Cost	Cost for taxpayers, €/year during 2017–2026.	40 €	5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 130, 160, 190, 300, 500 €

Table 2. Example of a choice set.

	Current program	**Alternative X**	**Alternative Y**
Traditional rural biotopes and endangered species	Present area, 0 species protected	Area is increased 60 %, 200 species protected	Area is increased 30 %, 100 species protected
Typical agricultural landscape • Grazing animals	Seldom seen	Seen often during summer season	Seen often during summer season
• Plants in cultivation	3 species	5 species	4 species
Climate effects Decrease from current emissions	0 %	0 %	30 %
Water quality effects Share of surface waters in good or excellent stage	60 %	80 %	60 %
Cost / taxpayer / year, during years 2017-2026	40 €	70 €	130 €
My choice	○	○	○

Table 3. Latent class model with policy relevance as covariate.

	CL	**Class 1**	**Class 2**	**Class 3**	**Class 4**	**Class 5**	**Overall**	
Pseudo R^2	0.0730	0.1126	0.3247	0.6310	0.0635	0.0182	0.4607	
Class Size	1	0.3383	0.3122	0.1717	0.0935	0.0843		
BIC	24065.0						19478.7	
Class names		***Environ-mentalists***	***Scenario focused***	***Cost sensitives***	***Not interested***	***Outsiders***	**Wald p-value**	**Wald (=) p-value**
Attributes				**Coefficients and significance levels**				
ASC_1 (SQ)	-0.22***	-2.14***	-1.06***	-0.77**	2.77***	0.26	0.000	0.000
ASC_2 (Program X)	0.25***	1.24***	0.64***	0.42**	-0.96***	0.30**		
ASC_3 (Program Y)	-0.03	0.90***	0.42***	0.35*	-1.80***	-0.56***		
<u>*Traditional rural biotopes and endangered species*</u>								
Present area, 0 species	0.25***	-0.76***	-0.39***	-0.12	-0.46	0.09	0.000	0.000
Area increased 30%, 100 species	0.17***	0.19***	0.28***	0.25**	0.16	-0.00		
Area increased 60%, 200 species	0.08***	0.57***	0.11*	-0.14	0.30	-0.08		
<u>*Grazing animals*</u>								

Seldom seen	-0.16***	-0.17***	-0.56***	-0.05	0.33	-0.09	0.000	0.000
Seen often during summer	0.03	0.05	0.16***	0.16	0.03	0.04		
Seen often during summer and melt season	0.14***	0.21***	0.40***	-0.11	0.36	0.05		
<u>*Plants in cultivation*</u>								
3 plants	-0.02	-0.02	-0.06	-0.06	-0.27	-0.02	0.018	0.021
4 plants	-0.00	-0.03	-0.06	0.27***	0.55**	-0.06		
5 plants	0.02	0.05	0.07	-0.21	-0.27	0.08		
<u>*Climate effects*</u>								
0% decrease	-0.13***	-0.41***	-0.28***	0.04	-0.28	-0.25**	0.000	0.001
10% decrease	0.00	0.08**	0.10**	-0.14	0.02	0.22**		
30% decrease	0.13***	0.33***	0.18***	0.10	0.27	0.02		
<u>*Water quality effects*</u>								
60%	-0.21***	-0.09*	-0.55***	-0.44***	0.14	-0.06	0.000	0.000
70%	0.05***	-0.06	0.18***	0.12	0.58***	-0.05		
80%	0.17***	0.15***	0.37***	0.32***	-0.73**	0.11		
<u>*Cost*</u>	-0.006***	-0.002***	-0.015***	-0.063***	-0.019***	-0.001	0.000	0.000
Covariates of preferences								
None of the ES relevant		-0.38***	-0.16	0.15	0.45***	-0.06	0.000	
Included ES relevant		0.19	0.25*	0.14	-0.12	-0.46*		
Excluded ES relevant		-0.30**	-0.33**	-0.08	0.16	0.55***		
All ES relevant		0.49***	0.25***	-0.21**	-0.49***	-0.04		

*z-test: *** 99% significance level; ** 95% significance level; * 90% significance level.*

Table 4. Scale adjusted latent class model (SALC) for policy program choice.

	Class 2	**Class 1**	**Class 3**	**Class 4**	**Class 5**	**Overall**	
Class Size Scale Class 1	0.211	0.232	0.116	0.077	0.054	0.690	
Class Size Scale Class 2	0.096	0.104	0.051	0.035	0.024	0.310	
Pseudo R^2						0.507	
BIC						19359.7	
CAIC						19445.7	
Class names	***Environ-mentalists***	***Scenario focused***	***Cost sensitives***	***Not interested***	***Outsiders***	**Wald p-value**	**Wald (=) p-value**
Attributes	**Coefficients and significance levels**						
SC_1 (SQ)	-11.28*	-1.32***	-1.37**	3.36***	4.07**	0.000	0.000
ASC_2 (Program X)	5.72*	0.74***	0.84**	-1.01**	1.59***		
ASC_3 (Program Y)	5.55*	0.58***	0.53*	-2.35***	-5.66***		
<u>*Traditional rural biotopes and endangered species*</u>							
Present area, 0 species	-1.01***	-0.49***	-0.08	-0.32	-1.67**	0.000	0.000
Area increased 30%, 100 species	0.28***	0.29***	0.34*	0.22	1.47*		
Area increased 60%, 200 species	0.73***	0.20***	-0.26	0.10	0.20		
<u>*Grazing animals*</u>							
Seldom seen	-0.27***	-0.59***	0.02	-0.46	-1.20**	0.000	0.040
Seen often during summer	0.01	0.14**	0.12	0.10	0.50		
Seen often during melt season	0.28***	0.45***	-0.14	0.36	0.70*		
<u>*Plants in cultivation*</u>							
3 plants	-0.04	-0.01	-0.12	-0.25	-1.14	0.13	0.13
4 plants	-0.05	-0.01	0.33**	0.46	-0.02		
5 plants	0.10*	0.02	-0.22	-0.21	1.16*		
<u>*Climate effects*</u>							
0% decrease	-0.55***	-0.32***	0.07	-0.39	-3.02***	0.000	0.001
10% decrease	0.07	0.16**	-0.21	0.20	0.36		
30% decrease	0.48***	0.16**	0.14	0.19	2.66***		
<u>*Water quality effects*</u>							
60%	-0.21***	-0.64***	-0.51**	-0.13	-0.08	0.000	0.000
70%	0.02	0.13**	0.09	0.61**	-0.71		
80%	0.19***	0.51***	0.41**	-0.47	1.55**		
<u>*Cost*</u>	-0.004***	-0.020***	-0.110***	-0.015***	0.000	0.000	0.000
Covariates of preferences							
None of the ES relevant		-0.12	0.29**	0.53***	-0.59**	0.000	
Included ES relevant		0.13	0.01	-0.25	-0.01		
Excluded ES relevant		-0.20	-0.05	0.26*	0.50***		
All ES relevant		0.19**	-0.25***	-0.55***	0.09		
Scale parameters							
Scale factor Class 1	1					0.000	
Scale factor Class 2	0.288***						
Covariates of scale							
None of the ES relevant	0.123					0.082	
Included ES relevant	-0.586**						
Excluded ES relevant	0.274						
All ES relevant	0.189						

*z-test: *** 99% significance level; ** 95% significance level; * 90% significance level.*

Table 5. Mean annual willingness to pay in 2016 € per individual for ten year duration (95% CI in parentheses).

	CL model	LC Class 1 Environmentalists	LC Class 2 Scenario focused	LC Class 3 Cost-sensitives	LC Class 4 Not interested	LC Class 5 Outsiders
Traditional rural biotopes and endangered species						
Present area, 0 species	-46 (-54 – -37)	-345 (-564 – -125)	-27 (-38 – -15)	-	-	-
Area increased 30%, 100 species	32 (24 – 39)	85 (32 – 137)	19 (11 – 27)	4 (1 – 8)	-	-
Area increased 60%, 200 species	14 (7 – 21)	260 (83 – 436)	8 (-1 – 16)	-	-	-
Grazing animals						
Seldom seen	-29 (-37 – -22)	-76 (-141 – -10)	-38 (-47 – -28)	-	-	-
Seen often during summer	-	-	11 (4 – 17)	-	-	-
Seen often during summer and melt season	25 (18 – 32)	98 (61 – 134)	27 (20 – 34)	-	-	-
Plants in cultivation						
3 plants	-	-	-	-	-	-
4 plants	-	-	-	4 (2 – 7)	29 (0 – 58)	-
5 plants	-	-	-	-	-	-
Climate effects						
0% decrease	-24 (-32 – -16)	-188 (-312 – -63)	-19 (-29 – -9)	1 (-3 – 5)	-	-
10% decrease	-	37 (-5 – 78)	7 (0 – 14)	-2 (-6 – 1)	-	-
30% decrease	24 (17 – 30)	151 (51 – 251)	12 (4 – 20)	2 (-2 – 6)	-	-
Water quality effects						
60%	-39 (-47 – -32)	-43 (-82 – -4)	-37 (-47 – -27)	-7 (-11 – -2)	-	-
70%	8 (2 – 15)	-	12 (5 – 19)	-	31 (5 – 57)	-
80%	31 (25 – 37)	69 (25 – 113)	25 (17 – 32)	5 (2 – 8)	-39 (-77 – 0)	-