Comparing Serial, and Choice Task Stated and Inferred Attribute Non-Attendance Methods in Food Choice Experiments

Vincenzina Caputo, Ellen J. Van Loo, Riccardo Scarpa, Rodolfo M. Nayga, Jr. and Wim Verbeke[1]

(Original submitted April 2016, revision received October 2016, accepted May 2017.)

Abstract

A number of choice experiment (CE) studies have shown that survey respondents employ heuristics such as attribute non-attendance (ANA) while evaluating food products. This paper addresses a set of related methodological questions using empirical consumer data from a CE on poultry meat with sustainability labels. First, it assesses whether there are differences in terms of marginal willingness to pay estimates between the two most common ways of collecting stated ANA (serial and choice task level). Second, it validates the self-reported ANA behaviour across both approaches. Third, it explores the concordance of stated methods with that of the inferred method. Results show that WTP estimates from serial-level data differ from those from choice task-level data. Also, self-reported measures on choice task ANA are found to be more congruent with model estimates than those for serial ANA, as well as with inferred ANA.

Keywords: *Attribute non-attendance; serial stated attribute non-attendance; choice task stated attribute non-attendance; inferred attribute non-attendance; choice experiments; sustainable food labels.*

JEL classifications: *C33, C35.*

---

[1] Vincenzina Caputo is at the Department of Agricultural, Food and Resource Economics, Michigan State University, MI, USA. E-mail: vcaputo@msu.edu for correspondence. Ellen J. Van Loo and Wim Verbeke are at the Department of Agricultural Economics, Ghent University. Riccardo Scarpa is in the Business School, University of Durham, UK. Rodolfo M. Nayga, Jr. is at the Department of Agricultural Economics and Agribusiness, University of Arkansas, and also affiliated with the Norwegian Institute of Bioeconomy Research and NBER. The authors would like to thank the Editor, David Harvey, and two anonymous journal referees for their helpful and comprehensive comments and suggestions. This work was partially supported by the AgBioResearch of Michigan State University. This work was partially supported by the USDA National Institute of Food and Agriculture, [Hatch] project [1013332].

1. Introduction

Modelling food choice behaviour in a random utility framework requires an adequate understanding of which food attributes are actively evaluated by each respondent and which ones are not. Such understanding is not only essential to develop appropriate individual utility functions to be used in estimation, but it is also crucial for improving CE survey designs and determining the reliability and validity of welfare estimates. These important considerations are, however, often neglected in food choice studies, especially those involving stated preference surveys using choice experiments (CEs). For example, some CE respondents may ignore some of the food attributes used to describe the product profiles while evaluating the set of alternatives in a choice task. In the CE literature, this issue is commonly called 'attribute non-attendance' (ANA) behaviour. To progress research in this area, we examine (1) the estimation effects of alternative ways of modelling stated ANA behaviour, and (2) the concordance in results between both serial and choice task stated ANA and those obtained through the inferred ANA method.

Stated ANA methods rely on asking respondents follow-up questions on whether specific attributes were ignored when evaluating alternatives in a choice task. Self-reported statements can be asked at the end of the entire choice task sequence (i.e. serial stated ANA) (Hensher *et al*., 2005; Rose *et al*., 2005; Campbell, 2007; Hensher and Rose, 2009; Scarpa *et al*., 2009; Cameron and DeShazo, 2010; Balcombe *et al.* 2011; Alemu *et al.*, 2013; Thiene *et al.*, 2012; Kragt, 2013; Colombo and Glenk, 2014; Glenk *et al.*, 2015) or after each individual choice task (i.e. choice task stated ANA) (Puckett and Hensher, 2008, 2009; Scarpa *et al*., 2010). Alternatively, the inferred ANA method infers ANA behaviour through the estimation of analytical models and is most often based on the latent class framework (Hess and Rose, 2007; Scarpa *et al.*, 2009; Hensher and Greene, 2010; Campbell *et al*., 2011; Hensher *et al.*, 2012; Caputo *et al*., 2013) and more rarely on the variable selection method. The most popular latent class model is the equality-constrained latent class model (ECLC). In the ECLC model, classes do not refer to differences in preference intensities as in the standard latent class models. Instead, they differ on the basis of the particular pattern of attributes with no impact on utility. The coefficients for the attributes with a recognised impact on utility (non-zero) may be assumed to be either the same or different across classes (Scarpa *et al.*, 2009; Campbell *et al.*, 2011; Caputo *et al.*, 2013). Other inferred methods include the combined latent class mixed logit (Hess *et al*., 2013), the random parameter mixed panel logit models (Hess and Hensher, 2010) and the ECLC with scale and preference heterogeneity model (Thiene *et al.* 2015).

The choice modelling literature illustrates how ignoring ANA behaviour in CEs has repercussions for market share predictions and welfare measure estimates (Hensher *et al.*, 2005; Lancsar and Louviere, 2006; Hensher, 2006, 2008; Scarpa *et al.*, 2009, 2010; Carlsson *et al.*, 2010; Hensher and Greene, 2010; Campbell *et al.*, 2011; Hole, 2011). However, there is, as yet, no consensus on the best way to account for ANA behaviour. For instance, should stated ANA information be collected at the end of a sequence of choice tasks or after each individual choice task? Studies on choice task ANA (Puckett and Hensher, 2008, 2009; Scarpa *et al.*, 2010) show that ANA behaviour often varies along the series of choice tasks presented to respondents, pointing to the inadequacy of an assumed uniform ANA behaviour, as implied in the serial choice task approach. In this regard, an important methodological question is whether and to what extent collecting ANA information after each choice task (i.e. choice task ANA) influences subsequent choice behaviour as opposed to asking ANA information after the respondent has gone through the whole series of choice tasks in the CE (i.e. serial ANA). While the previously mentioned studies have examined either serial or choice task ANA, to date, only Scarpa *et al.* (2010) have compared these two stated ANA approaches in a public good context. However, in their study they did not actually collect serial ANA information. Instead, they collected choice task ANA and then reconstructed serial ANA based on the reported choice task ANAs. Thus, their serial ANA data might have been affected by the ANA questions asked during the CE at the choice task level. Their findings suggest that accounting for choice task ANA significantly improves model fit and yields marginal WTP estimates that seem to be more accurate for the public goods in question (i.e. natural park features in their study).

Moreover, there have been concerns in the literature about possible measurement errors in stated ANA. These concerns refer to (a) whether the self-reported ANA behaviours collected at either the serial or choice task level are consistent with the true ANA behaviour, and (b) to what extent they can be affected by recall problems and approximations (Scarpa *et al.*, 2010). Measurement errors can exist when, for example, respondents who indicated to have ignored a given attribute have actually not fully ignored it, but most likely have just given to it a lower importance (Campbell and Lorimer, 2009; Carlsson *et al.*, 2010). Obviously, measurement error in ANA behaviour can affect the reliability and/or validity of the stated ANA methods. So, should researchers rely on self-reported ANA information when modelling consumer choice behaviour? If it is assumed that self-reported ANA information is accurate, then the attributes reported as ignored are selectively removed from the individual utility in the data estimation process (Campbell and Lorimer, 2009; Alemu *et al.*, 2013; Scarpa *et al.*, 2013). We refer to this modelling approach as the 'Conventional ANA' model. However, incorrectly constraining an attribute self-reported as ignored to have a zero impact on the utility function could lead to a mis-specified choice model (Hole *et al.*, 2013). Scarpa *et al.* (2013) noted that one way to validate self-reported stated ANA statements is to specify an indirect utility function that separately estimates two coefficients for each of the attributes, depending on whether the respondent identified the attribute as having played a role in the evaluation of alternatives or not. We refer to this second modelling approach as the 'validation method'. Studies employing this validation method have demonstrated discrepancies between what survey respondents self-reported and what this approach suggests they actually did (Campbell and Lorimer, 2009; Hess and Hensher, 2010). Other studies concluded that a separate treatment on the basis of such self-reported ANA did not improve model fit (Balcombe *et al*., 2011). This method has, so far, only been applied to validate self-reported serial ANA statements but not yet to choice task ANA statements (Campbell and Lorimer, 2009; Alemu *et al.*, 2013; Scarpa *et al.*, 2013).

Finally, measurement errors have also been mentioned as a possible reason for the lack of concordance (no one-to-one correspondence) in the CE outcomes when using stated ANA and inferred ANA. A number of researchers (e.g. Hess and Hensher, 2010; Kragt, 2013; Scarpa *et al.*, 2013) have compared results between serial stated ANA with inferred methods. Their findings generally suggest that (i) there is little concordance between serial stated and inferred ANA, and that (ii) inferred ANA models provide better model fit (e.g. Hess and Hensher, 2010; Kragt, 2013; Scarpa *et al*., 2013) than models based on serial stated ANA. These findings suggest that inferring ANA econometrically could be a valuable alternative, also considering the possible measurement error discussed above. However, it remains difficult to know which method could better represent the 'true' ANA behaviour (Collins, 2012). While the stated ANA approach is vulnerable to measurement errors, the inferred method has the drawback of requiring the researcher to make decisions on how to take ANA into account in the models (e.g. number of latent choice-behavioural classes, structure of preferences, etc.). Thus, the relative merits of using the inferred method could largely depend upon subjective choices made by researchers given the data at hand.

The current literature on ANA in choice modelling is mostly in the field of transportation (Hensher *et al.,* 2005; Hensher, 2006, 2008; Hensher and Greene, 2010), environmental valuation (Campbell *et al.,* 2008; Scarpa *et al.,* 2009, 2010; Carlsson *et al.,* 2010; Campbell *et al.,* 2011; Kragt, 2013), and health economics (Hole, 2011; Mc-Intoch and Ryan, 2002; Lancsar and Louviere, 2006; Hole *et al.,* 2013). Only three studies have examined ANA in food choice modelling (Bello and Abdulai, 2016; Caputo *et al.,* 2013; Scarpa *et al.,* 2013). Bello and Abdulai (2016) measured the impact of consumer non-attendance behaviour on *ex-ante* hypothetical bias mitigation methods using only the serial ANA approach. Scarpa *et al*. (2013) studied inferred and stated ANA but did not collect choice task ANA responses. Caputo *et al*. (2013) only inferred ANA using latent class models but did not analyse stated ANA.

Amongst the many methodological issues that have so far not been answered, we focus on three that we deem important for the modelling of ANA in food choice. First, we investigate whether there is any systematic difference in terms of CE outcomes (e.g. WTPs and model performance) across the two forms of stated ANA to test the robustness of previous findings (e.g. Scarpa *et al.*, 2010). We do so by implementing two experiments: the *Serial* experiment, in which the ANA questions are asked at the end of the entire sequence of CE questions, and the *Choice Task* experiment, in which the ANA questions are asked at the end of each CE question. Hence, in contrast to Scarpa *et al*. (2010), we directly collect ANA information at both the serial and choice task levels by exposing our sample of respondents to two independent treatments. Second, following Scarpa *et al.* (2013), we validate the self-reported serial and choice task ANA statements using the stated ANA model approach in which two coefficients for each attribute are estimated: one for the self-reported attended attributes and one for the self-reported ignored attributes. This allows us to identify whether there is any discrepancy between what survey respondents say they did when reporting ANA in our CE surveys across the entire series of choice tasks and in each separate choice task, and what they actually did do. Finally, we infer ANA using a latent class framework and then examine differences in results across the various methods to account for ANA (inferred, serial stated and choice task stated).

The rest of the article is structured as follows. The next section reports the experimental procedures used in the *Serial* and the *Choice Task* experiments, followed by a section that describes the empirical analysis. The results are then reported, followed by the conclusions.

2. Choice Experiment Design and the Experiments

We constructed a CE study on a chicken breast product in Belgium, which was described using a combination of five attributes: (i) organic label, (ii) animal welfare label, (iii) free-range claim, (iv) carbon footprint label, and (v) price. For the organic logo, three levels were considered: the EU organic logo, the Belgian private Biogarantie logo, and no organic logo. The levels for the free-range claim included those currently regulated in the European Union (EU) (EC, 2008): free-range, traditional free-range and, free-range total freedom. The levels of the price attribute were chosen based on the actual prices of chicken breast gathered during a store check in food stores in Belgium in February 2012, shortly before the survey was conducted. The levels used for carbon footprint were based on reported values in the literature for producing a chicken breast (Foster *et al*., 2006; Just Bare, 2010) and adopting a 20% and a 30% carbon footprint reduction as alternative levels. The definitions of the attributes and attribute levels are shown in Table 1.

(INSERT TABLE 1)

Based on these attributes, a D-optimal CE design was developed following the approach by Street and Burgess (2007).[2] We first generated an orthogonal factorial design for the first alternative, reducing the original 288 ($3^2 \times 4^2 \times 2$) combinations to just 16. Then, using the generators described by Street and Burgess (2007) a practical set of 16 pairs was obtained, with a D-efficiency of 95.7%. Finally, the 16 choice sets were divided into two blocks and the participants were randomly assigned to one of the two blocks. To increase the similarity with a real shopping experience, a no-buy alternative was added to each choice set. Following Scarpa and Rose (2008), the design was evaluated *ex post* in terms of its potential *D*-error. We calculated an efficient design based on the estimates obtained from the multinomial logit (MNL) model estimated from both the serial and choice task datasets. We found our design to require 103 and 89 design replicates in the *Serial* and *Choice Task* experiments, respectively, given that the two blocks were obtained with 206 and 178 participants. Since our sample size consisted of 344 and 257 subjects in the *Serial* and *Choice Task* experiments respectively, it far exceeded this requirement.[3] Hence, our designs seem to have performed adequately *ex post*, with the larger sample size compensating for the lack of efficiency in terms of *D*-error.

In the CE survey, each participant was presented with eight choice tasks. Each choice task included two experimentally-designed product profiles and a no-buy option (see example in Figure 1). A cheap talk script was included to mitigate the potential for hypothetical bias (Silva *et al*., 2011), and was presented to the participants before they were asked to engage in the choice tasks. The identification of what attributes were ignored was obtained from supplementary ANA questions asked of participants and recorded in two different ways. Participants were randomly assigned to one of two experiments. In the *Serial* experiment (serial ANA), the ANA questions were asked of

[2] We acknowledge that there are several alternative approaches to designing a CE and refer readers to Johnson *et al.* (2013) who give an overview of the most common experimental design approaches used in discrete choice studies.

[3] Design statistics are available upon request.

participants at the end of the sequence of choice tasks, while in the *Choice Task* experiment (choice task ANA) participants were asked what attributes they ignored after each choice task.[4]

Insert Figure 1 here

3. Empirical Analysis

The price attribute was treated as a continuous variable in all models, while the food quality labels were treated as dummy-coded attributes. We used dummy coding rather than effect coding because this allowed us to meaningfully restrict the parameters of self-reported ignored attributes to zero. As pointed out by Caputo *et al.* (2013), putting a zero restriction on an effect-coded variable (–1,1) would not be equivalent to a zero weight in the utility function, but rather to a weight which is intermediate between absence and presence of the attribute, which makes it collinear with the alternative-specific constant (ASC).

3.1. Modelling Serial and Choice Task Stated ANA using a RPL-EC Model

The serial and choice task CE datasets were used to estimate a Panel Logit model with Random Parameters and Error Component (RPL-EC) (Scarpa *et al.*, 2005, 2007; Hess and Rose, 2008). Accordingly, the utility function that individual *i* obtains from choice alternative *j* in choice situation *t* is as follows:

$$U_{ijt} = ASC + \alpha PRICE_{ijt} + \boldsymbol{\beta}_i{}'\mathrm{x}_{ijt} + 1_j(\eta_{it}) + \varepsilon_{ijt} \qquad (1)$$

Commented [RC1]: Typesetters, move equation numbers to right margin

Commented [CV2R1]: Done! Thank you!

where ASC is an alternative-specific constant representing the no-buy choice alternative; α is the marginal utility of price; $PRICE_{ijt}$ is the price of alternative *j* for person *i* at choice situation *t*; $\boldsymbol{\beta}_i$ is a vector of utility parameters for participant *i*; $\mathbf{x}_{ijt}$ is a *k*-dimensional vector of observed non-monetary food attributes and their levels related to alternative *j*, individual *i* and choice task *t* in the sequence. These are represented by the sustainability labels illustrated in Table 1: organic (*OrgEU* and *OrgBE*), animal welfare (*AW*), free-range (*FR*, *FRtrad* and *FRtot*), and reduced level of CO_2 emitted (*CO20* and *CO30*). $1_j(\cdot)$ is an indicator function that takes the value of 1 for both experimentally designed food profiles, and 0 otherwise; η_{it} is a zero-mean normally distributed respondent-specific idiosyncratic error component shared by the two hypothetical alternatives (i.e.

[4] For example, 'Did you ignore information about the organic label?'

those alternatives that portray a purchase decision), and is absent in the utility of the no-buy alternative (Scarpa *et al.*, 2007); ε_{ijt} is an i.i.d. extreme value error term. The coefficients of the sustainability labels are assumed to be independent and normally distributed, while the price coefficient is assumed to be fixed. This assumption may appear to be somewhat restrictive, but it provides us with the obvious advantage that the ratios of sustainability label and price coefficients (marginal WTPs) are normally distributed. In addition, it allows us to incorporate the individual self-reported ANA information for the price attribute when estimating choice models.

The RPL-EC specification is employed because it simultaneously accounts for heterogeneity in consumer preferences, by allowing the coefficients of the different claims to vary randomly over individuals and to deviate from the population mean, and for correlation across utilities, by identifying the additional variance of the utility of the experimentally designed alternatives, different from the no-buy option. The latter is of particular importance since the no-buy option is included in the choice tasks of our CE design. The no-buy option is actually experienced by participants while the experimentally designed alternatives are hypothetical. Hence, the utilities of the hypothetical options are likely to be more correlated between each other than with the no-buy option. In addition, hypothetical options require each respondent to conjure up a given profile of food attributes at each choice task and as a consequence, they tend to display a higher utility variance than the utilities of the no-buy option.

Given the possible differences in Gumbel error scale across the serial and choice task data, the interpretation of the coefficient values across the estimated models is not recommended (Greene and Hensher, 2003; Scarpa and Del Giudice, 2004). Hence, we test whether monitoring stated ANA at the serial or choice task level leads to different choice outcomes and then focus on the marginal WTP estimates. We proceed in two steps. In the first step, using the data from each experiment (i.e. *Serial* and *Choice Task*), we estimate two RPL-EC models, where the coefficients for the self-reported ignored attributes are constrained to zero during estimation). The implicit assumption of this model is that an observed choice provides no information concerning the respondent's preferences for those attributes that are ignored (Alemu *et al*., 2013). Hence, the coefficient estimates are conditional on the subset of those respondents who stated they have considered the attributes in the *Serial* experiment (Campbell and Lorimer, 2009), and on the subset of choice tasks in which the respondents claimed to have considered the attributes in the *Choice Task* experiment. For this reason, these models are referred to as 'Conventional ANA' models.

In the second step, we use the estimated coefficients and variance covariance matrices from these two models to perform the parametric bootstrapping method proposed by Krinsky and Robb (1986).[5] It results in a distribution of 1,000 marginal WTPs for each attribute. These 1,000 values are used to perform the combinatorial test suggested by Poe *et al.* (2005)[6] and test the following hypotheses:

$$H_0\colon (WTP_{Conventional-ANA\ Serial,k} - WTP_{Conventional-ANA\ Choice\ Task,k}) = 0, \text{ and}$$

$$H_1\colon (WTP_{Conventional-ANA\ Serial,k} - WTP_{Conventional-ANA\ Choice\ Task,k}) \neq 0.$$

If H_0 is rejected, we concluded that serial and choice task ANA produce significantly different WTPs. This leads us into the second issue we wish to investigate, namely which of the two stated ANA approaches (e.g. serial vs. choice task) is most adequate in capturing the ANA behaviour.

3.2. Validating Stated ANA: Serial and Choice Task

If respondents provide truthful responses to the ANA questions, then their choice behaviour should be consistent with self-reported ANA (Scarpa *et al*., 2013). To evaluate which of the two stated ANA approaches (i.e. serial and choice task) best agrees with self-reported ANA statements, we followed Scarpa *et al.* (2013) and estimated a second RPL-EC model, named ‘Validation model’, in which two coefficients are estimated for each of the attributes, depending on whether the attribute was stated as being either considered or ignored. Following Scarpa *et al.* (2013), we used a vector of *k* attendance indicators for each respondent *i*, one for each non-price attribute *k*. We denote the generic element of such vector as 1_{ik}(A=1) if respondent *i* stated having attended

[5] In particular, 1,000 observations for each attribute are drawn from multivariate normal distributions with means given by the estimated coefficients and covariance given by the estimated covariance matrix of the coefficients from each of the econometric models estimated for the *Serial* and *Choice Task* experiment. The 1,000 draws for each coefficient are then used to calculate the marginal WTP at each draw as the negative ratio between the parameter estimated of the non-monetary attribute and the parameter estimates for the price. The lower and the upper limit of the confidence interval for each attribute are given by the 26th and 975th sorted estimates of WTP.

[6] The Poe-test is performed to compare all possible combinations of the 1,000 bootstrapped values of marginal WTPs obtained from the two econometric models across *Serial* and *Choice Task* experiments. Hence, 1,000,000 (1,000 x 1,000) differences are calculated for each hypothesis test of interest.

attribute *k*, and 1_{ik}(A=0) if respondent *i* stated having ignored it. By denoting the utility coefficients conditional on attendance with the superscript 1 and those conditional on non-attendance with the superscript 0, the indirect utility function can be expressed as follows:

~~(2)~~

$$V_{ijt} = ASC + 1_i(A=1)\alpha^1 PRICE_{ijt} + 1_{ik}(A=1)\boldsymbol{\beta}_i^{1\prime}\mathbf{x}_{ijt}$$
$$+1_i(A=0)\alpha^0 PRICE_{ijt} + 1_{ik}(A=0)\boldsymbol{\beta}_i^{0\prime}\mathbf{x}_{ijt} + 1_j(\eta_{it}) \quad (2)$$

Commented [RC3]: Typesetters, move equation numbers to right margin

Commented [CV4R3]: Done! Thank you

where $1_i(\cdot)$ is an indicator of ANA for the price attribute by individual *i*, with $1_i(A=1)$ if individual *i* stated having attended to the price attribute, and $1_i(A=0)$ otherwise; as before, $1_{ik}(\cdot)$ is a *k*-dimensional vector of indicators of ANA for individual *i* and non-price attribute *k* (*k*=1,2,…,4), with $1_{ik}(A=1)$ if individual *i* stated having attended to attribute *k*, and $1_{ik}(A=0)$ otherwise. Hence, the utility coefficients α^1 and α^0 refer to the marginal utilities of price for the respondents who attended and not-attended the price attribute, respectively. Similarly, the vector of utility coefficients $\boldsymbol{\beta}_i^1$ refers to the coefficients for attended sustainability labels, while the utility coefficients $\boldsymbol{\beta}_i^0$ refer to those for the self-reported ignored attributes. The coefficients of the sustainability labels were assumed to be normally distributed, while the price coefficient was assumed to be fixed. The rest of the variables in equation (2) are defined as in equation (1).

The significance of the coefficient estimates for the ignored attributes can be used as a validation method. If the estimates for the food attributes stated as being ignored are different from zero, then this would indicate that respondents did not fully ignore these attributes. If this condition is verified, then there is evidence of discrepancies between what survey respondents reported and what they actually did in their choice behaviour (Campbell and Lorimer, 2009; Scarpa *et al*., 2013). Hence, this model also allows us to corroborate whether or not the hypothesis of the standard method (i.e. the 'Conventional ANA' model), which restricts the coefficients to zero for the self-reported ignored attributes is appropriate in our data. We explored this in both the *Serial* and *Choice Task* experiments.

3.3. Exploring the Concordance between Stated and Inferred ANA using an ECLC Model

To explore the concordance between the results from stated and inferred approaches, we also infer the incidence of ANA behaviour in the sample using an Equality Constrained Latent Class (ECLC) model for panel data (Hess and Rose, 2007; Scarpa *et al*., 2009; Campbell *et al.*, 2011; Caputo *et al.*, 2013). The ECLC models are different from standard latent class models, which are intended to explore preference heterogeneity. This is because ECLC models are based on classes embedding different forms of attendance to attributes (Scarpa *et al*., 2013) rather than different preference intensities. Hence, in the ECLC model, a specific form of ANA can be based on the estimation of class-specific membership probabilities with adequate restrictions on the utility coefficients. Membership probabilities from the ECLC model estimates can then be used to explore the concordance of the ECLC model with the frequencies of the self-reported ANA information at both the serial and the choice task levels (Kragt, 2013; Scarpa *et al*., 2013).

Formally, in the ECLC model, the unconditional probability of the observed panel of T choices by respondent i is a weighted average over the C classes with weight π_c, and each ANA class has indirect utility which embeds the zero-constrained coefficients:

~~(3)~~

$$P_{i,T} = \sum_c \pi_c \prod_t \frac{\exp(V_{ijtc})}{\sum_j \exp(V_{igtc})}. \qquad (3)$$

Commented [RC5]: Typesetters, move equation numbers to right margin

Commented [CV6R5]: Done! Thank you

In this application we estimate[7] an ECLC model accounting for both the presence of separate classes of taste intensity (AA) and various patterns of serial attribute non-attendance (ANA) (Caputo *et al.,* 2013). This is motivated by the fact that groups may differ not only in terms of patterns of attendance, but also in terms of taste intensities as demonstrated by the popularity of conventional latent class models, which were originally motivated by preference variation. Given that taste heterogeneity can co-exist with attribute processing, more than one AA class ought to be considered (Hensher *et al*., 2013). Similarly, the introduction of multiple AA classes may also reduce the chance of the attribute processing classes (ANA classes) capturing both taste heterogeneity and attribute processing.

4. Data and Results

4.1. Sample Characteristics and Stated ANA Statements: Descriptive Statistics

Data were collected by a market research company through an online survey conducted in Belgium in March 2012 and targeting the person of the household most often in charge of food purchases. A total of 601 participants completed the CE surveys and they were randomly assigned to either the

[7] The ECLC model was estimated in Latent Gold.

Serial (*n*=344) or the *Choice Task* (*n*=257) experiments. Socio-demographic characteristics were similar across the two sub-samples (all chi-square *p*>0.05) (see Table S1 in the online Appendix).

Only 18% of the respondents in the *Serial* experiment reported having attended to all five attributes (see Table S2 in the online Appendix). The remaining 82% of the respondents stated having ignored at least one attribute. The carbon footprint attribute was stated as ignored most frequently, by 71% of the respondents (see Table S3 in the online Appendix~~Table 7~~). Although the meaning of each attribute level was explained to the participants prior to the CE, the low awareness of carbon footprint labels (Gadema and Oglethorpe, 2011) and the absence of this label in the Belgian poultry meat market (Van Loo *et al*., 2014) might explain its low stated attendance. However, we speculate it could also be related to other reasons such as attribute levels not being relevant. The other sustainability labels (organic, animal welfare and free-range) were stated to have been ignored by 42% to 50% of the respondents ~~(~~(see Table S3 in the online Appendix ~~Table 7~~). Also, as expected, the price attribute had the highest attendance and was reported to have been ignored by 26% of the respondents[8] ~~(~~(see Table S2 in the online Appendix). ~~Table 7)~~.

In the *Choice Task* experiment, we recorded information about ANA for each of the eight choice tasks (see Tables S3 and ~~S3~~ S4 in the online Appendix ~~and Table 7~~). In this experiment, the carbon footprint attribute was reported as ignored in 44% of all choice tasks and thus was the attribute reported as the least attended to. Price had the lowest stated non-attendance, consistent with the *Serial* experiment. Approximately 32% (for price) to 70% (for carbon footprint) of the respondents did not follow the same attribute processing behaviour in all eight choice tasks as they did not indicate having ignored the attribute or having attended to the attribute in all eight choice tasks (TableS~~3~~4 and S~~4~~5). Instead, these respondents stated that they ignored the attributes in between one and seven choice tasks out of the eight choice tasks (online Table ~~S3~~S4). This suggests that collecting information on attribute processing behaviour at the choice task level may be more informative than collecting information at the serial level where respondents are assumed to follow the same strategy for the whole sequence of choice tasks. Similar findings were reported in outdoor

[8] Similar shares of respondents who ignored the price/cost attribute are reported by other choice studies. In the field of food choice experiments, Scarpa *et al*. (2013) reported that about 38% and 35% of their respondents either rarely or never attended to the price level when choosing chicken and beef, respectively. In the field of environmental choice experiments, Carlsson *et al.* (2010) found out that about 24% and 31% of their respondents ignored the cost attribute in their study on flourishing lakes streams and clean air, respectively. Kragt (2013) reported 22% of their respondents stated that they ignored cost.

Commented [RC7]: Author, so far only table 1 has been cited – tables should be cited in numerical order

Commented [CV8R7]: Thank you for this really good comment. We now added a new table in the Appendix S3, and changed the numbering order of all the tables in the Appendix. Thank you!

recreation studies (Scarpa *et al.*, 2010), suggesting the advantages of monitoring ANA at the choice task level.

4.2. Estimates from the Conventional-ANA RPL-EC Models for Serial and Choice Task Experiments

Table 2 reports the coefficient estimates from the 'Conventional ANA' model with correlated random taste and error[9] for the *Serial* and *Choice Task* experiments. We remind the reader that the 'Conventional ANA' model constrains the coefficients to zero for those attributes that each respondent reported ignoring. In both *Serial* and *Choice Task* experiments, the alternative specific constants for the no-buy option have statistically significant and negative estimates, indicating that respondents favour the proposed hypothetical purchase alternatives. Also, the price coefficient estimates are, as expected, negative and statistically significant. Finally, the estimates of the standard deviations of the error components and sustainability labels are statistically significant in both experiments, suggesting the presence of heteroskedasticity across utilities (Scarpa *et al*., 2007) and preference heterogeneity.

In the *Serial* experiment, the coefficient estimates of all the sustainability labels are positive and statistically significant. The highest utility increment occurs when information on the 'free-range total freedom' label is presented (*FRtot*), followed respectively by 'traditional free-range' (*FRtrad*), organic Belgium (*OrgBE*), 'free-range' (*FR*), EU organic (*OrgEU*), 30% CO_2 reduction (*CO30*), animal welfare (*AW*), and 20% CO_2 reduction (*CO20*) labels. All coefficient estimates in the *Choice Task* experiment are also positive and statistically significant. The attributes by and large maintain the same relative rankings as in the *Serial* experiment estimates. The main difference being a drop of *FRtrad* from second to fourth rank and a rise of *AW* from seventh to fifth on the basis of the point estimates.

Insert Table 2 here

Table 3 reports the marginal WTP estimates for *Serial* and *Choice Task* experiments, along with the 95% confidence intervals based on the Krinsky and Robb (1986) bootstrapping procedure with 1,000 draws, as well as the results of the Poe-Test. The relative importance ranking of WTPs for

[9] Correlation across random taste coefficients is estimated by means of a Cholesky matrix (Cholesky matrix estimates are available upon request).

the labels changes across the experiments. Specifically, in the *Serial* experiment, the relative importance ranking of the marginal WTPs is: *FRtot, FRtrad, OrgBE, FR, OrgEU, CO30, AW* and *CO20*, while in the *Choice Task* experiment the rank order is *OrgEU, FRtot, OrgBE, FRtrad, AW, FR, CO30* and *CO20*. Most notably, when comparing the marginal WTPs across the *Serial* and *Choice Task* experiments, our hypothesis of equality of marginal WTPs across the *Serial* and *Choice Task* experiments is rejected for two (*OrgEU, AW)* out of the eight sustainability labels at the 5% significance level and rejected for an additional three attributes at the 10% significance level (*OrgBE, FRtot* and *CO20*). While the marginal conditional WTP is higher in the *Choice Task* experiment than in the *Serial* experiment for *OrgEU, OrgBE, AW* and *FRtot,* the opposite is true for *CO20*.

Insert Table 3 here

4.3. Validity of ANA statements across serial and choice task

Table 4 reports the coefficient estimates from the 'Conventional ANA' model with correlated random taste and error for both the *Serial* and *Choice Task* experiments.[10] In this model, those standard deviations estimates found to be insignificant in the restriction tests were restricted to zero, indicating absence of heterogeneity and fixed coefficients. We remind the readers that the 'Validation ANA Model' implies the estimation of separate coefficients for attributes reported as attended to and those reported as non-attended to.

In the *Serial* experiment, coefficient estimates for attributes reported as being considered show the expected signs and are significant. The coefficient estimates for the attributes reported as being ignored are also statistically significant, except for the organic labels (*OrgEU* and *OrgBE*). This implies that in this experiment, respondents who stated having ignored the free-range (*FRtrad),* animal welfare *(AW),* and reduced carbon footprint (*CO20* and *CO30*) labels (rather than stating having completely ignored) were assigned lower utility values than those who stated that these labels were attended to. Evidence of choice behaviour inconsistent with self-reported non-attendance was also found in previous food choice studies (Alemu *et al*., 2013; Campbell and Lorimer, 2009). Similar to our case, others showed that rather than ignoring attributes completely,

[10] Random coefficients with insignificant standard deviation estimates were fixed, implying absence of heterogeneity.

Commented [RC9]: Author, throughout the paper I have changed 'stated to have …' to 'stated having …'

Commented [EVL10R9]: Ok thanks

respondents might be putting lower weights on attributes that they claimed to have ignored (Alemu *et al*., 2013; Campbell and Lorimer, 2009; Carlsson *et al*., 2010; Scarpa *et al*., 2013).

Most notably, in the *Choice Task* experiment, the coefficient estimates for attributes reported as being ignored are not statistically significant for six (*FRtrad, AW, FR, FRtot, CO20* and *CO30*) out of the nine total number of attributes. These results suggest that ANA self-reporting at the choice task level is generally consistent with the choice behaviour that was actually adopted.

Insert Table 4 here

4.4 Estimates of the ECLC model: Are self-reported ANA concordant with inferred ANA?

A way of assessing the concordance between stated (serial and choice task) and inferred methods is to test their concordance in terms of the inferred ANA frequency for each attribute (Scarpa *et al*., 2013). The classes in our ECLC model differ in their nature. Some are preference heterogeneity classes while others are behavioural ANA classes, which have different sub-sets of attribute coefficients set to zero in accordance with different forms of ANA. Given five attributes, a total of 32 ANA class combinations are possible in our study. After having tested several specifications,[11] we only focus on a specification with 3 preference classes and 16 classes: 3 preference classes with 7, 6 and 3 classes differing not only in terms of preference structure but also in terms of different ANA patterns. In particular,

(i) preference class one (c=1) includes one complete attendance (AA1), one complete ANA (random choice), and five lexicographic preferences for single attributes (only one out of k matters, all other k–1 coefficients are set to zero);

(ii) preference class two (c=2) incorporates one complete attendance class (AA2), and ANA for a single attribute k; and

[11] The exact combinations of taste-differing classes and sub-sets of non-attendance are defined using a data-driven process. All classes for which probability membership was found to be lower than a given threshold were discarded. In addition, consistent with the literature, the Akaike Information Criteria (AIC), the Bayesian Information Criteria (BIC) and the modified Akaike Information Criteria (3AIC) were also used to drive our model selection. The lower the information criterion value, the better is the model fit.

(iii) preference class three (c=3) includes one complete attendance (AA3), and ANA for a pair of attributes k sharing the desire to fulfill a similar sustainability dimension (rather than a common metric), such as organic and carbon footprint labels and animal welfare and free-range labels. Comparable attributes such as organic and carbon footprint labels, as well as animal welfare and free-range labels, could satisfy a related dimension (e.g. environmental sustainability or animal friendliness). So, there is a possibility that respondents eager to simplify complexity of choice enacted decision heuristics resulting in some kind of aggregation into the same overarching dimension (see Hensher *et al*., 2013). For instance, both free-range and animal welfare labels are ethical claims related to farming systems (e.g. animal housing, stocking density, outdoor access); and the popularity of free-range farming is mainly related to animal welfare issues given that animals raised under free-range conditions are not confined in intensive production systems (Van Loo *et al*., 2014). Thus, hypothesising that these labels fulfil a similar dimension when considered in alternative evaluation, and are either jointly ignored or jointly attended to, makes sense.

Commented [RC11]: Should this be a new para?

Commented [EVL12R11]: No this belongs to iii)

Estimates of our ECLC model are reported in two separate tables: Table 5 displays the estimates of the class-specific membership probabilities, while Table 6 reports the coefficient estimates from the ECLC model. The largest preference class has an estimated membership probability of 54.1%, shared across seven ANA classes, as described in Table 5, which also reports the second and third preference classes shared respectively by six and three ANA classes.

Insert Table 5 here

Insert Table 6 here

Estimates of the attribute coefficients across the three preference classes are shown in Table 6. For preference class 1, all coefficients are significantly different from zero and the pattern of signs is consistent with our expectations, except for the negative effect for the *OrgEU* label. Specifically, the archetype member of preference class 1 tends to prefer the *CO30* label, followed by *FRtot, FRtrad*, *CO20*, *FR*, *AW* and *OrgBE*. In preference class 2, all estimated coefficients are significant and with expected signs. In this class, members generally prefer free-range labels, such as *FRtot*, *FRtrad* and *FR* followed by the *OrgBE*, *OrgEU*, *CO2*, *CO30* and *AW*. Finally, estimates in preference class 3 report a negative effect for price and the no-buy alternative-specific constant. Consumer choices from this preference class are only significantly affected by the *OrgEU* label, *FRtrad* and *AW* labels.

Table 7 displays the frequencies of self-reported ANA for the *Serial* and *Choice Task* experiments and contrasts these with the inferred frequencies of ANA from the ECLC model. Results suggest that frequencies at the choice task level are similar to frequencies based on the inferred method (ECLC), with relative differences between 3% and 8% for organic, free-range and carbon footprint labels. We find larger relative differences only for the attributes animal welfare label and price: for the former, they were 39% self-reported versus less than 27% in the ECLC models, resulting in a relative difference of 46.6%. For price, they were 20.4% for the self-reported ANA versus 30.5% in the ECLC models, resulting in a relative difference of –33.0%.

Larger differences are found when comparing the frequencies obtained from the serial ANA with those from the ECLC model (relative differences ranging from 11% to 65%). Thus, concordance between inferred and stated ANA is better for the choice task ANA than for the serial ANA. This has also been reported by Kragt (2013), who also found little concordance between ECLC and serial ANA.

Insert Table 7 here

4.5. Comparing model fits

Table 8 reports the Akaike Information Criteria (AIC), the Bayesian Information Criteria (BIC), and the modified Akaike Information Criteria (3AIC) that are used to compare data fit across models. We also report the model fit of the Full Attendance RPL-EC model, in which ANA was not accounted for, i.e. this model assumes that respondents attended all attributes (see model estimates in Table ~~S5~~ S6 in the online Appendix). When looking at the performance of Stated ANA models, we can conclude that addressing ANA using self-reported choice task ANA is better than addressing ANA using self-reported serial task ANA in terms of model fit. Importantly for the *Choice Task* experiment, the 'Conventional ANA' model has a better performance than the validation model. On the other hand, for the *Serial* experiment, the 'ANA-Validation' model is the best model. The difference in best model performances across the *Serial* and *Choice Task* experiment is also confirmed by the coefficient estimates in the 'ANA-Validation' model (Table 4), which are mostly statistically significant for the non-attenders in the *Serial* experiment, and statistically insignificant in the *Choice Task* experiment. This implies that in the *Serial* experiment, respondents did not truly ignore the attributes that they self-reported as ignored; thus the 'Conventional ANA' model may lead to biased results.

Insert Table 8 here

5. Conclusion and Future Research

Past studies have shown that addressing ANA behaviour using both stated and inferred methods affects both market share prediction and welfare estimates (Scarpa *et al*., 2014). However, only a few studies have explored the merits of stated *vis-à-vis* inferred ANA. Our study is the first that examines this issue in food choice data, notably in terms of assessing (i) how WTP estimates are affected by the incorporation of serial vs. choice task stated ANA by using both conventional and validation modelling approaches; and (ii) whether WTP estimates from both stated ANA methods (i.e. serial and choice task) are concordant with those from the inferred method.

Our results generally suggest that firstly, the ‘Conventional ANA’ model applied to serial versus choice task self-reported ANA leads to differences in marginal WTP estimates. The marginal WTPs for the sustainability labels in the *Choice Task* experiment are generally higher than those in the *Serial* experiment (5 out of 8 cases). By contrast, Scarpa *et al*. (2010) found higher WTPs when accounting for serial ANA than for choice task ANA. However, their study differed from ours in two important ways: 1) in their study, self-reported serial ANA information was not recorded in the survey. Rather, it was derived from information reported at the choice task level; 2) in our study, ANA behaviour is modelled using *RPL-EC* models, rather than multinomial logit models.

Secondly, self-reported ANA at the choice task level suggests that few respondents follow the same attribute processing strategies throughout the entire sequence of choice tasks. As such, collecting ANA information at the choice task level is more desirable than at the serial level. This makes sense because ANA information gathered at the serial level may fail to capture the changes in attribute processing behaviour along the CE sequence. This is further confirmed by the estimates obtained from the validation model (ANA-Validation) which show that, when allowed to take a value different from zero, most of the coefficients of the self-reported ignored attributes in the *Serial* experiment are indeed significantly different from zero. Instead, only a few coefficients are found to be significantly different from zero in the *Choice Task* experiment. This suggests that there is higher concordance between what respondents reported they ignored and what their choices show. This result corroborates the finding reported by Scarpa *et al*. (2010) that the intra-respondent variation of attribute attendance at the single choice task level is of substantial importance to the welfare estimates and model fit when addressing ANA behaviour. This higher concordance of choice task data extends to inferred ANA probabilities obtained with the ECLC model.

Finally, in terms of the empirical fit to the data, the choice task stated ANA models appear to outperform models based on both serial stated ANA and inferred ANA. Taken together, these findings provide guidance on how one should collect ANA information and what model specifications should be used when incorporating ANA behaviour in CE models. In particular, our overall results suggest that although the collection of choice task ANA data requires more effort from the respondents as compared to the serial ANA, the advantages of accounting for choice task ANA might outweigh its additional cost and effort.

The literature and our findings lead to a number of interesting areas for future research on the issue of ANA. For example, stated ANA studies usually assume that respondents ignore a specific attribute, irrespective of its levels. However, it is possible that respondents only ignore subsets of food attribute levels (Erdem *et al*., 2015). Future research on attribute processing strategies with respect to stated ANA might evaluate ANA based on the attribute levels. Moreover, people may follow certain attribute processing strategies based on the attribute level present in the choice task. Hensher *et al*. (2012), for instance, suggest research on the use of respondent-specific attribute ranges as certain attributes might only be relevant if a respondent-specific threshold level is reached. Future research might also investigate how ANA is linked to the complexity of the task (e.g. number of attributes, number of attribute levels, number of choice sets, ranges of attributes) (Hensher, 2006; Carlsson *et al*., 2010; Collins and Hensher, 2015), the importance and relevance of single food attributes, and the specific relevance of attribute levels (Hensher *et al*., 2012). Finally, attendance to attribute can be corroborated with neurological data derived from eye-tracking investigations measuring times of eye-fixation (Balcombe *et al*., 2015), other measures of cognitive effort or by recording access to attribute information during computerised surveys (Kaye-Blake *et al*., 2009; Kravchenko, 2016). While this area of research is still in its infancy and can be challenging, it is a promising area for future research.

References

Alemu, M. H., Mørkbak, M. R., Olsen, S. B. and Jensen, C. L. ‘Attending to the reasons for attribute non-attendance in choice experiments’, *Environmental and Resource Economics*, Vol. 54(3), (2013) pp. 333–359.

Balcombe, K., Burton, M. and Rigby, D. ‘Skew and attribute non-attendance within the Bayesian mixed logit model’, *Journal of Environmental Economics and Management,* Vol. 62(3), (2011) pp. 446–461.

Balcombe, K., Fraser, I. and McSorley, E. ‘Visual attention and attribute attendance in multi-attribute choice experiments’, *Journal of Applied Econometrics,* Vol. 30(3), (2015) pp. 447–467.

Bello, M. and Abdulai, A. ‘Impact of ex-ante hypothetical bias mitigation methods on attribute non-attendance in choice experiments’, *American Journal of Agricultural Economics,* Vol. 98(5), (2016) pp. 1486–1506.

Cameron, T. A. and DeShazo, J. R. ‘Differential attention to attributes in utility-theoretic choice models’, *Journal of Choice Modeling*, Vol. 3(3), (2010) pp. 73–115.

Campbell, D. ‘Willingness to pay for rural landscape improvements: Combining mixed logit and random-effects models’, *Journal of Agricultural Economics*, Vol. 8(3), (2007) pp. 467–483.

Campbell, D., Hutchinson, W. G. and Scarpa, R. ‘Incorporating discontinuous preferences into the analysis of discrete choice experiments’, *Environmental & Resource Economics,* Vol. 41(3), (2008) pp. 401–417.

Campbell, D. and Lorimer, V. S. ‘Accommodating attribute processing strategies in stated choice analysis: Do respondents do what they say they do?’ in European Association of Environmental and Resources Economics, Annual Conference, Amsterdam, (2009)~~, June 2009~~.

Campbell, D., Hensher, D. A. and Scarpa, R. ‘Non-attendance to attributes in environmental choice analysis: a latent class specification’, *Journal of Environmental Planning and Management,* Vol. 1, (2011) pp. 1–16.

Caputo, V., Nayga, R. M. Jr. and Scarpa, R. ‘Food miles or carbon emissions? Exploring labeling preference for food transport footprint with a stated choice study’, *Australian Journal of Agricultural Economics*, Vol. 57, (2013) pp. 1–18.

Carlsson, F., Kataria, M. and Lampi, E. ‘Dealing with ignored attributes in choice experiments on valuation of Sweden's environmental quality objectives’, *Environmental & Resource Economics*, Vol. 47(1), (2010) pp. 65–89.

Collins, A.T. ‘Attribute nonattendance in discrete choice models: Measurement of bias, and a model for the inference of both nonattendance and taste heterogeneity’, PhD dissertation (Institute of Transport and Logistics studies, Business School, The University of Sydney, 2012).

Collins, A. T. and Hensher, D. A. 'The influence of varying information load on inferred attribute non-attendance', in S. Rasouli and H. Timmermans (eds.), *Bounded Rational Choice Behaviour: Applications in Transport* (Bingley, UK: Emerald Group Publishing Limited, 2015) pp. 73–94.

Colombo, S. and Glenk, K. 'Social preferences for agricultural policy instruments: Joint consideration of non-attendance to attributes and to alternatives in modelling discrete choice data', *Journal of Environmental Planning and Management,* Vol. 57(2), (2014) pp. 215–232.

European Commission (EC). European Commission Regulation (EC) No 543/2008 of 16 June. Laying down detailed rules for the application of Council Regulation (EC) No 1234/2007 as regards the marketing standards for poultry meat. *Official Journal* ~~829~~ *of the European Union*, L157, (2008), pp. ~~830~~ 46–87

Erdem, S., Campbell, D. and Hole, A. R. 'Accounting for attribute-level non-attendance in a health choice experiment: Does it matter?' *Health Economics,* Vol. 24(7), (2015) pp. 773–789.

Foster, C., Green, K., Bleda, M., Dewick, P., Evans, B., Flynn, A. and Mylan, J. 'Environmental impacts of food production and consumption: A report to the Department for Environment, Food and Rural Affairs', Manchester Business School (Defra, London, 2006). Available online at: http://www.ifr.ac.uk/waste/Reports/DEFRA Environmental%20Impacts%20of%20Food%20Production%20%20Consumption.pdf (Last accessed: 5 January 2013).

Glenk, K., Martin-Ortega, J., Pulido-Velazquez, M. and Potts, J. 'Inferring attribute non-attendance from discrete choice experiments: Implications for benefit transfer', *Environmental and Resource Economics,* Vol. 60(4), (2015) pp. 497–520.

Greene, W. H. and Hensher, D. A. 'A latent class model for discrete choice analysis: Contrasts with mixed logit', *Transportation Research Part B: Methodological,* Vol. 37(8), (2003) pp. 681–698.

Hensher, D. A. 'Revealing differences in willingness to pay due to the dimensionality of stated choice designs: an initial assessment', *Environmental and Resource Economics*, Vol. 34(1), (2006) pp. 7–44.

Hensher, D. A. 'Joint estimation of process and outcome in choice experiments and implications for willingness to pay', *Journal of Transport Economics and Policy,* Vol. 42(2), (2008) pp. 297-322.

Hensher, D. A. and Rose, J. M. 'Simplifying choice through attribute preservation or non-attendance: Implications for willingness to pay', *Transportation Research Part E,* Vol. 45(4), (2009) pp. 583–590.

Hensher, D. A., Collins, A. T. and Greene, W. H. ‘Accounting for attribute non-attendance and common-metric aggregation in a probabilistic decision process mixed multinomial logit model: a warning on potential confounding’, *Transportation*, Vol. 40(5), (2013) pp. 1003–1020.

Hensher, D. A. and Greene, W. H. ‘Non-attendance and dual processing of common-metric attributes in choice analysis: a latent class specification’, *Empirical Economics,* Vol. 39(2), (2010) pp. 413–426.

Hensher, D. A., Rose, J. and Greene, W. H. ‘The implications on willingness to pay of respondents ignoring specific attributes’, *Transportation*, Vol. 32(3), (2005) pp. 203–222.

Hensher, D. A., Rose, J. M. and Greene, W. H. ‘Inferring attribute non-attendance from stated choice data: Implications for willingness to pay estimates and a warning for stated choice experiment design’, *Transportation,* Vol. 39(2), (2012) pp. 235–245.

Hess, S. and Rose, J.M. ‘A latent class approach to modelling heterogeneous information processing strategies in SP studies’ (Oslo Workshop on valuation methods in transport planning, Oslo, 2007).

Hess, S. and Hensher, D. A. ‘Using conditioning on observed choices to retrieve individual-specific attribute processing strategies’, *Transportation Research Part B: Methodological,* Vol. 44(6), (2010) pp. 781–790.

Hess, S. and Rose, J. M. ‘Should reference alternatives in pivot design SC surveys be treated differently?’ *Environmental and Resource Economics,* Vol. 42(3), (2008) pp. 297–317.

Hess, S., Stathopoulos, A., Campbell, D., O’Neill, V. and Caussade, S. ‘It’s not that I don’t care, I just don’t care very much: Confounding between attribute non-attendance and taste heterogeneity’, *Transportation*, Vol. 40(3), (2013) pp. 583–607.

Hole, A. R. ‘A discrete choice model with endogenous attribute attendance’, *Economics Letters*, Vol. 110(3), (2011) pp. 203–205.

Hole, R., Kolstad, J. R. and Gyrd-Hansen, D. ‘Inferred vs. stated attribute non-attendance in choice experiments: a study of doctors’ prescription behaviour’, *Journal of Economic Behavior & Organisation,* Vol. 96, (2013) pp. 21–31.

Johnson, F. R., Lancsar, E., Marshall, D., Kilambi, V., Mühlbacher, A., Regier, D. A., Bresnahan, B. W., Kanninen, B. and Bridges, J. F. P. ‘Constructing experimental designs for discrete-choice experiments: Report of the ISPOR conjoint analysis experimental design good research practices task force’, *Value in Health,* Vol. 16(1), (2013) pp. 3–13.

Just Bare. ‘The carbon footprint of Just BARE boneless skinless chicken breast’, Life cycle assessment summary report, 2010. Available online at:

http://www.gnpbusiness.com/files/Just_BARE_SummaryLCAReport_FNL6-30-11.pdf (Last accessed: 5 January 2013).

Kaye-Blake, W.H., Abell, W.L. and Zellman, E. ''Respondents' ignoring of attribute information in a choice modelling survey', *The Australian Journal of Agricultural and Resource Economics*, Vol. 53(4), (2009) pp. 547–564.

Kragt, M.E. 'Stated and inferred attribute attendance models: a comparison with environmental choice experiments', *Journal of Agricultural Economics*, Vol. 64(3), (2013) pp. 719–736.

Kravchenko, A. 'The value of irrigation water in New Zealand', Ph.D. Thesis (University of Waikato, Hamilton New Zealand, 2016).

Krinsky, I. and Robb, A.L. 'On approximating the statistical properties of elasticities', *The Review of Economics and Statistics*, Vol. 64(4), (1986) pp. 715–719.

Lancsar, E. and Louviere, J. 'Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences?' *Health Economics*, Vol. 15(8), (2006) pp. 797–811.

McIntosh, E. and Ryan, M. 'Using discrete choice experiments to derive welfare estimates for the provision of elective surgery: Implications of discontinuous preferences', *Journal of Economic Psychology*, Vol. 23(3), (2002) pp. 367–382.

Poe, G.L., Giraud, K.L. and Loomis, J.B. 'Computational methods for measuring the difference of empirical distributions', *American Journal of Agricultural Economics*, Vol. 87, (2005) pp. 353–365.

Puckett, S. M. and Hensher, D.A. 'The role of attribute processing strategies in estimating the preferences of road freight stakeholders', *Transportation Research Part E: Logistics and Transportation Review*, Vol. 44(3), (2008) pp. 379–395.

Puckett, S. M. and Hensher, D. A. 'Revealing the extent of process heterogeneity in choice analysis: an empirical assessment', *Transportation Research Part A*, Vol. 43(2), (2009) pp. 117–126.

Rose, J.M., Hensher, D.A. and Greene, W.H. 'Recovering costs through price and service differentiation: Accounting for exogenous information on attribute processing strategies in airline choice', *Journal of Air Transport Management* 11(6), (2005) pp. 400–407.

Scarpa, R. and Del Giudice, T. 'Market segmentation via mixed logit: Extra-virgin olive oil in urban Italy', *Journal of Agricultural & Food Industrial Organisation*, Vol. 2(1), (2004) pp. 1–18.

Scarpa, R., Ferrini, S. and Willis, K. 'Performance of error component models for status-quo effects in choice experiments applications of simulation methods in environmental and resource economics', in R. Scarpa and A. Alberini (eds.), *Applications of Simulation Methods in Environmental and Resource Economics* (The Netherlands: Springer-Verlag, 2005) pp. 247–273.

Scarpa, R., Gilbride, T. J., Campbell, D. and Hensher, D. A. 'Modelling attribute non-attendance in choice experiments for rural landscape valuation', *European Review of Agricultural Economics*, Vol. 36(2), (2009) pp. 151–174.

Scarpa, R. and Rose, J. M. 'Design efficiency for non-market valuation with choice modelling: How to measure it, what to report and why', *Australian Journal of Agricultural and Resource Economics* , Vol. 52, (2008) pp. 253–282.

Scarpa, R., Thiene, M. and Hensher, D. A. 'Monitoring choice task attribute attendance in nonmarket valuation of multiple park management services: Does it matter?' *Land Economics*, Vol. 86(4), (2010) pp. 817–839.

Scarpa, R., Willis, K. and Acutt, M. 'Valuing externalities from water supply: Status-quo, choice complexity and individual random effects in panel kernel logit analysis of choice experiments', *Journal of Environmental Planning and Management*, Vol. 50(4), (2007) pp. 449–466.

Scarpa, R., Zanoli, R., Bruschi, V. and Naspetti, S. 'Inferred and stated attribute non-attendance in food choice experiments', *American Journal of Agricultural Economics,* Vol. 95(1), (2013) pp. 165–180.

Silva, A., Nayga, R. M. Jr, Campbell, B. L. and Park, J. L. 'Revisiting cheap talk with new evidence from a field experiment', *Journal of Agricultural and Resource Economics,* Vol. 36(2), (2011) pp. 280–291.

Street, D. and Burgess, L. *The Construction of Optimal Stated Choice Experiments* (Hoboken, NJ: John Wiley & Sons Inc, 2007).

Thiene, M. and Meyerhoff, J. 'Scale and taste heterogeneity for forest biodiversity: Models of serial nonparticipation and their effects', *Journal of Forest Economics,* Vol. 18(4), (2012) pp. 355–369.

Thiene M., Scarpa, R. and Louviere, J. 'Addressing preference heterogeneity, multiple scales and attribute attendance with a correlated finite mixing model of tap water choice', *Environmental and Resource Economics,* Vol. 62(3), (2015) pp. 637–656.

Train, K.E. *Discrete Choice Methods with Simulation* (Cambridge, UK: Cambridge University Press, 2003).

Van Loo, E. J., Caputo, V., Nayga, R. M. Jr. and Verbeke, W. 'Consumers' valuation of sustainability labels on meat', *Food Policy*, Vol. 49, (2014) pp. 137–150.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Table S1: Demographics across experiment

Table S2: Number and attributes ignored by the respondents in the Serial experiment

Table S3: Attributes ignored by the respondents in the Serial and Choice Task experiments

Table ~~S3~~S4: Attributes ignored by the respondents across choice tasks in the Choice Task experiment

Table ~~S4~~S5: Number of attributes ignored across choice tasks in the Choice Task experiment

Table ~~S5~~S6: Parameters estimates and unconditional WTPs from the full attendance model

Commented [RC13]: Author this table not cited in the text please add citation to text

Commented [EVL14R13]: In text citation was added

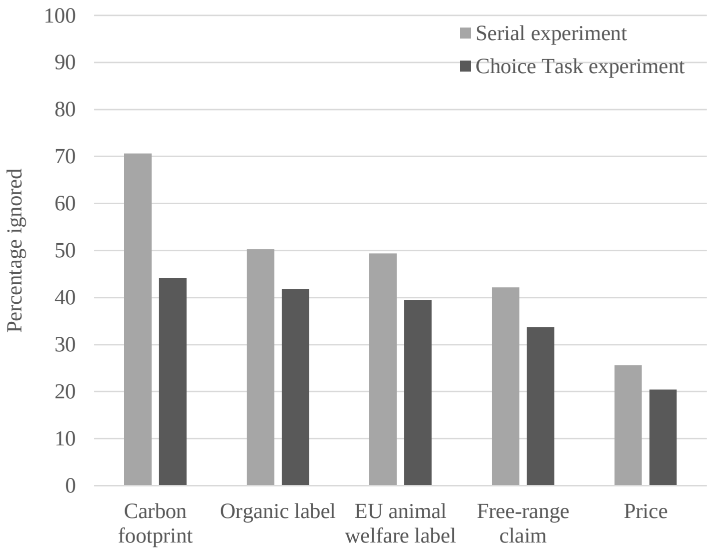
Commented [CV15R13]: As per my previous comment, I added a new table and changed the numbering order of the others reported in the Appendix. I also made a few edits/additions/changes. Thank you.

Figures

Figure 1. Example of choice set question

	Alternative A	Alternative B	Alternative C
Organic logo	EU Organic logo	No logo	
Animal welfare label	EU Animal welfare label	No label	
Free-range claim	Traditional free-range	Free-range—total freedom	Neither alternative A nor B is chosen
Reduced carbon footprint label	No label	5.6 kg CO_2 compared to 7 kg CO_2	
Price	€20/kg	€25/kg	
I prefer	O	O	O

Figure 2. Stated ANA across attributes and experiments (% choice tasks in which an attribute was ignored)



Notes: The frequencies of choice tasks in which the respective attribute was stated as ignored differs significantly between *Serial* and *Choice Task* experiment for each of the attributes *t* (all five Chi-square test have p <0.05). In the *Serial* experiment, data from 2,752 choice tasks (344 respondents) were used, while for the *Choice Task* experiment 2,056 choice tasks were used (257 respondents).

Tables

Table 1

Attributes and levels for the choice experiment

Attributes	**Levels considered**
Organic label	- No organic label - Biogarantie label (*OrgBE*) - EU Organic label (*OrgEU*)
Animal welfare protection label	- No animal welfare label present - European animal welfare label (*AW*)
Types of free-range farming claim	- No free-range claim - Free-range (*FR*) - Traditional free-range (*FRtrad*) - Free-range-total freedom (*FRtot*)
Reduced carbon footprint label (CO_2 emitted)	- No carbon footprint label - 20% reduction: 5.6 kg CO_2e compared to 7 kg CO_2 (*CO2*) - 30% reduction: 4.9 kg CO_2e compared to 7 kg CO_2 (*CO3*)
Price	- €10/kg - €15/kg - €20/kg - €25/kg

Table 2
Conventional ANA model (i.e. ignored parameters set to zero) across Serial and Choice Task experiments

	Serial experiment (N[1]=2,752)				**Choice Task experiment (N=2,056)**			
	β	|z-values|	σ	|z-values|	β	|z-values|	σ	|z-values|
No-buy	–8.28***	10.23	–	–	–6.19***	7.45		
Sd. of ERC			7.71***	18.16			8.76***	12.90
Price	–0.28***	18.15	–	–	–0.32***	13.20	–	–
OrgEU	1.37***	6.04	1.95***	7.19	3.12***	7.93	2.18***	5.29
OrgBE	1.62***	8.97	1.37***	6.40	2.55***	6.38	2.46***	8.01
AW	1.06***	7.99	0.83***	4.90	1.93***	8.51	1.56***	5.36
FR	1.41***	7.39	0.92***	3.22	1.77***	5.49	1.90***	4.57
FRtrad	1.65***	8.13	0.91***	2.91	2.07***	5.77	2.08***	4.53
FRtot	1.97***	8.82	1.62***	5.74	2.83***	7.32	2.16***	5.13
CO20	0.82***	3.35	1.33***	2.91	0.37	1.58	1.37***	4.56
CO30	1.30***	4.24	2.07***	3.61	1.19***	3.53	2.22***	5.39

Notes: [1] Number of observations (choices); ***, **, * indicate significance at 1%, 5%, 10% levels, respectively.

Table 3
WTP estimates from the conventional ANA model (i.e. ignored parameters set to zero) across Serial and Choice Task experiments

	Serial experiment	**Choice Task experiment**	
	Mean (st. dev.) [Confidence intervals]	Mean (st. dev.) [Confidence intervals]	P-Value[2]
OrgEU	4.93[1]*** (0.79) [3.3916–6.4920]	9.63*** (1.03) [7.5905–11.6613]	0.0001
OrgBE	5.81*** (0.62) [4.6184–7.0596]	7.82*** (1.07) [5.7857–9.9732]	0.0519
AW	3.83*** (0.48) [2.9062–4.7675]	5.95*** (0.65) [4.7178–7.2563]	0.0036
FR	5.10*** (0.70) [3.7598–6.4.19]	5.44*** (0.88) [3.6503–7.1392]	0.3767
FRtrad	5.96*** (0.72) [4.5155–7.3502]	6.39*** (1.00) [4.4823–8.4717]	0.3631
FRtot	7.10*** (0.76) [5.6031–8.6403]	8.72*** (1.02) [6.7822–107990]	0.0974
CO20	3.00*** (0.89) [1.3183–4.7583]	1.16* (0.72) [–0.2285–2.4932]	0.0518
CO30	4.68 (1.14) [2.5001–6.8137]	3.66*** (1.04) [1.6668–5.6768]	0.2538

Note: ***, **, * indicate significance at 1%, 5%, 10% levels, respectively.

[1] Numbers are means of 1,000 bootstrapped WTP estimates from the 'Conventional ANA' models across the *Serial* and *Choice Task* experiments calculated using the Krinsky–Robb bootstrapping method.
[2] *p*-values testing whether the marginal WTP distribution for each attribute from the *Serial* experiment equals to the marginal WTP distribution for the corresponding attribute from the *Choice Task* experiment. The *p*-values are based on the non-parametric combinatorial method proposed by Poe *et al.* (2005) to 1,000 bootstrapped WTP estimates from the 'Conventional ANA' models across the *Serial* and *Choice Task* experiments calculated using the Krinsky–Robb bootstrapping method.

Table 4

Estimates from the Validation ANA model (i.e. estimated two coefficients for each attribute) across Serial and Choice Task experiments

	Serial experiment (N^{a}=2,752)				**Choice Task experiment (N=2,056)**			
	β	|z-values|	σ	|z-values|	β	|z-score|	σ	|z-values|
No-buy	–7.01***	11.73			–6.94***	8.77		
Sd. of ERC			5.88***	7.68			6.27***	9.99
Considered								
Price	–0.26***	18.90	-	-	-0.30***	13.75	-	-
OrgEU	1.09***	7.25	-	-	2.43***	7.19	1.89***	4.50
OrgBE	1.35***	10.30	0.42**	2.17	1.98***	6.99	2.03***	5.94
AW	0.97***	8.47	0.68**	2.27	1.73***	8.67	1.18***	2.84
FR	1.25***	7.12	1.04***	3.35	1.56***	6.10	0.95**	2.23
FRtrad	1.47***	7.84	0.85**	2.24	1.76***	6.07	1.37***	3.13
FRtot	1.89***	9.44	1.32***	3.40	2.54***	7.42	1.71***	3.78
CO20	0.70***	3.69	1.08***	3.92	0.46**	2.21	1.22***	3.65
CO30	1.20***	4.60	1.77***	3.54	1.13***	3.91	1.60***	2.77
Ignored								
Price	–0.05***	4.16	-	-	–0.08***	3.78	-	-
OrgEU	–0.07	0.58	-	-	–0.04*	1.91	0.54*	1.78
OrgBE	–0.09	0.88	-	-	–0.41***	2.63	-	-
AW	0.21**	2.50	-	-	-0.08	0.60	-	-
FR	0.32**	2.29	-	-	0.09	0.39	-	-
FRtrad	0.35**	2.15	-	-	0.04	0.16	-	-
FRtot	0.49***	3.28	-	-	0.18	0.80	-	-
CO20	0.18**	2.21	-	-	0.10	0.61	-	-

Formatted Table

Commented [CV16]: I restructured the table to avoid inconsistencies of the formatting of the numbers!

Formatted: Centered

Formatted: Centered

Formatted: Centered

Formatted: Centered

Formatted: Centered

Formatted: Centered

Formatted: Centered

Formatted: Centered

CO30	0.42***	3.82	-	-	-0.04	0.23	-	-
Ignored								
Price	-0.05***	4.16	-	-	-0.08***	3.78	-	-
OrgEU	-0.07	0.58	-	-	-0.04*	1.91	0.54*	1.78
OrgBE	-0.09	0.88	-	-	-0.41***	2.63	-	-
AW	0.21**	2.50	-	-	-0.08	0.60	-	-
FR	0.32**	2.29	-	-	0.09	0.39	-	-
FRtrad	0.35**	2.15	-	-	0.04	0.16	-	-
FRtot	0.49***	3.28	-	-	0.18	0.80	-	-
CO20	0.18**	2.21	-	-	0.10	0.61	-	-
CO30	0.42***	3.82	-	-	-0.04	0.23	-	-

Note: ***, **, * indicate significance at 1%, 5%, 10% levels, respectively; [a] Number of observations (choices).

Formatted Table
Formatted: Centered
Formatted: Centered
Formatted: Centered
Formatted: Centered

Table 5
Class memberships from the ECLC model (N=344)

Preference classes (Probabilities %)	**Description ANA behaviour**	**Probabilities (%)**
Preference class 1 (54.14%)	AA1 (complete attendance)	23.73
	AA-PRICE (Only price attended)	14.93
	AA-ORG (Only organic labels attended)	0.07
	AA-AW (Only animal welfare labels attended)	3.69
	AA-FREE (Only free-range labels attended)	4.16
	AA-CO (Only carbon footprint labels attended)	2.35
	ANA (complete ANA)	5.21
Preference class 2 (23.63%)	AA2 (full attendance)	2.34
	ANA-PRICE (only price ignored)	15
	ANA-ORG (only organic labels ignored)	0.07
	ANA-AW (only animal welfare labels ignored)	0.1
	ANA-FREE (only free-range labels ignored)	6.04
	ANA-CO (only carbon footprint labels ignored)	0.08
Preference class 3 (22.23%)	AA3 (complete attendance)	7.34
	ANA-AW+FREE (animal welfare and free-range labels ignored)	0.12
	ANA-ORG+CO (organic and carbon footprint labels ignored)	14.77

Table 6

Estimates from the ECLC model (N[1]= 2,752)

	Preference Class 1 (AA1)		Preference Class 2 (AA2)		Preference Class 3 (AA3)	
Parameters	β	|z-values|	β	|z-values|	β	|z-values|
PRICE	–0.94***	6.61	–0.20***	5.31	–0.22***	9.10
OrgEU	–1.07**	2.27	1.27***	4.74	5.56**	2.46
OrgBE	0.66*	1.86	1.65***	6.07	7.19	1.62
AW	1.45***	5.54	1.16***	5.88	0.55***	3.05
FR	2.16***	4.13	2.22***	5.59	0.36	1.34
FRtrad	2.80***	4.06	2.60***	6.49	0.90***	3.33
FRtot	3.37***	4.80	3.06***	6.92	0.26	0.91
CO20	2.36***	3.13	1.22***	4.49	1.47	0.65
CO30	5.64***	3.88	1.18***	4.24	4.95	1.11
NOBUY	–22.02***	7.60	3.35***	5.88	–3.44***	7.33

Notes: [1] Number of observations (choices); ***, **, * indicate significance at 1%, 5%, 10% levels, respectively.

Table 7

Frequencies of self-reported ANA (serial and choice task) versus inferred ANA latent classes (ELCL)

	Serial experiment (S)	Choice Task experiment (T)	Inferred ANA (ECLC)	(S-ECLC)/ ECLC	(T-ECLC)/ ECLC
	% Respondents	*% Choice tasks*	*% Respondents*	*% Relative difference*	*% Relative difference*
Organic labels	50.29	41.83	45.18	11.31	–7.42
EU animal welfare label	49.42	39.49	26.94	83.44	46.60
Free-range claim	42.15	33.71	32.41	30.06	4.00
Carbon footprint	70.64	44.21	42.91	64.62	3.03
Price	25.58	20.43	30.48	–16.07	–32.98
Complete Attendance	17.73	34.58	33.41	–46.92	3.51
Complete ANA	7.56	5.16	5.21	45.87	7.49
N	344[1]	2,056[2]	344	344	344

Notes: [1] Number of respondents; [2] Number of total choices (e.g. 8 per respondent).

Table 8
Summary statistics of model fit

		Serial experiment		Choice Task experiment		Inferred ANA
	Complete Attendance	Standard ANA	ANA Validation	Standard ANA	ANA Validation	ECLC
N	2,752	2,752	2,752	2,056	2,056	2,752
LL	–1,780.14	–1,711.37	–1,714.22	–1,123.11	–1,116.55	–1,665.77
BIC/N	1.452	1.402	1.404	1.296	1.361	1.340
AIC/N	1.334	1.284	1.286	1.146	1.158	1.243
AIC3/	1.354	1.304	1.306	1.173	1.194	1.260
N. Par.	55	55	55	55	74	45

Note: N. Par refers to number of parameters.