Comparing treatments to reduce Hypothetical Bias in Choice Experiments regarding Organic Food

Adelina Gschwandtner* and **Michael Burton

***University of Kent, UK; ** University of Western Australia, AU**

Hypothetical bias is one of the strongest criticism brought to stated preference methods. We evaluate and compare use of Cheap Talk and Honesty Priming as methods to mitigate such bias. Our study analyses the demand for organic food products in the UK and the results reveal a core of consumers with positive WTP for organic. However, when correcting for hypothetical bias, consumers appear to be willing to pay even more for other attributes. Most importantly, the results show that implementing mechanisms to correct for hypothetical bias are efficient to reduce WTP, with Cheap Talk having a higher overall significance than Honesty Priming.

Keywords: Choice Experiments, Willingness to Pay, Hypothetical Bias Treatments, Cheap Talk, Honesty Priming, Budget Constraint Reminder, Organic Food, Latent Class Model, Attribute Non-Attendance, Scale Heterogeneity

JEL Codes: C83, C90, D12, L81, Q18, Q21, Q51

Acknowledgments: The authors would like to acknowledge the support of Iain Fraser in the design of the CE, and they extend their thanks to the participants at the school seminar of the School of Agriculture and Environment of Western Australia. They would also like to express gratitude to the participants at the school seminar of the School of Economics at the University of Kent and the participants at the presentation of the paper at the 91st Conference of the Agricultural Economic Society in Dublin 2017 as well as the anonymous referees for their valuable comments.

*Corresponding author E-mail: a.gschwandtner@kent.ac.uk

1. Introduction

The present study analyses different methods designed to correct for hypothetical bias in choice experiments in order to analyse the demand for organic food products in the UK. As hypothetical bias is by far the strongest criticism regarding stated preference techniques (Cummings et al., 1986; Mitchell and Carson, 1989; Murphy et al., 2005; Carson and Groves, 2007), the present study analyses different methods designed to address it. Hypothetical bias appears to be especially severe in studies about organic food, for which consumers experience a 'warm glow' from overstating their true preferences (Lusk and Norwood, 2009; Schuldt and Schwarz, 2010; Norwood and Lusk, 2011). In these type of studies consumers appear to find it socially desirable to state that they wish to pay a premium for environmentally friendly production as well as higher animal welfare, which are attributes of organic food products (Lagerkvist et al., 2006; Lagerkvist and Hess 2011; Zander and Hamm, 2010).

In order to obtain more accurate measures for the WTP for organic attributes, the present study examines the effect of ex-ante scripts to reduce hypothetical bias in choice experiments. The two methods used are Cheap Talk and Honesty Priming. Cheap Talk involves making consumers aware of the fact that people tend to overstate in general their true WTP when related to goods such as organic products. Studies have shown that, if consumers are informed about this overstatement, the effect will be reduced or completely eliminated (Farrell and Rabin, 1996; Cummings and Taylor, 1999; Aadland and Caplan, 2003; Brown, Ajzen, and Hrubes, 2003; Carlsson et al., 2005; Landry and List, 2007; Mozumder and Berrens, 2007; Champ, Moore and Bishop, 2009; Mahieu, 2010; Jacquemet et al., 2011; Tonsor and Shupp, 2011, Penn and Hu, 2018), even though evidence regarding this tendency is mixed.[1] The second method employed is Honesty Priming. Under this method, consumers are asked to complete 10 statements using missing words. These missing words could be chosen from 2 options, a correct (true) one (such as "The earth is **round"**) and an incorrect one (such as "The earth is **square**"). Through this, literature has shown that consumers can be induced to answer truthfully in the following choice tasks (Maxwell et al., 1999; Chartland et al., 2008; De-Magistris et al., 2013). The method is borrowed from social psychology and is rooted in the conceptual priming literature. Conceptual priming is the activation of a cognitive representation in one context to unconsciously influence an unrelated context (Bargh and Chartrand, 2000).

[1] Loomis (2014) for example found that in 3 out of 7 studies that used Cheap Talk the hypothetical bias was eliminated, in 3 it was reduced and in one study it had no effect (Loomis, 2014; Table 1 page 38).

The main reason for choosing these two methods is the fact that they have been shown to be successful in some studies despite their simplicity. Another method which is relatively simple to implement that was not used in this study is Honesty Oath. Under Honesty Oath participants make a promise that they are always going to provide truthful answers and evidence supports the ability of this method to reduce hypothetical bias (Jacquemet et al. 2009, 2010, 2013). Kemper et al. (2019) use Query Theory to explain the mechanism behind the effectiveness of Honesty Oath. They find that Honesty Oath can change the content and order of a query and that Query Theory is a useful tool in order to analyse the thought process in valuation studies. Moreover, the study shows that Honesty Oath can be potentially another useful tool in reducing hypothetical bias in discrete choice experiments. Similarly to our results, the study shows that the Honesty Oath manages to reduce WTP values in choice experiments. Other methods such as Inferred Valuation or Bayesian Truth Serum might have been also successful in reducing hypothetical bias but also more difficult and expensive to implement (Lusk and Norwood, 2009, Norwood and Lusk, 2011, Weaver and Prelec, 2013).

In the implementation presented below, three different combinations of Cheap Talk and Honesty Priming are applied, two in which each method is applied separately and one in which the methods are applied together. In our setting, the Cheap Talk script included also a budget constraint reminder. As is often done in the elicitation of stated preferences, consumers were reminded that if they spend more money on a product, they have less money left for other goods. However, for simplicity, we will refer to this combination of Cheap Talk with a budget constraint reminder, just as Cheap Talk.

These combinations are further called hypothetical bias *treatments*. This setup allows us to formulate testable hypotheses that will offer insights about the subject in question. One such hypothesis is that hypothetical bias treatments are leading to a larger (in absolute terms) marginal utility of money, and hence lower willingness to pay (WTP) estimates. Another hypothesis is that one could expect that the treatment, which includes both methods together, will demonstrate a higher effect than the treatments, which apply the two methods individually. We also wanted to observe which treatment works best and for which type of consumers. We did not have an ex-ante theory regarding this, but we sought to observe these elements empirically and compare the impact of the two methods ex- post. To our knowledge, this is an innovative approach that has not been applied before and could contribute to the methodological discourse.[2]

[2] To the best of our knowledge, the only recent papers (published after 2005) that estimate WTP for organic food products in the UK are Wier et al. (2008), Griffith and Nesheim (2008, 2010, 2013), Zander and Hamm (2010), Janssen and Hamm (2012), Gerrard et al. (2013), and Gschwandtner (2018), but none of them looks at specific organic attributes and have therefore a different focus. The only study that looks at the WTP for specific credence attributes in food and compares

Comparing the two methods we find that, while both seem to be successful to a degree in influencing the choices of some people, Cheap Talk appears to be the most successful one. Both methods lead to lower WTP, but Cheap Talk has a significant impact in a larger number of classes, and it has a higher overall significance than Honesty Priming.[3] It seems to have the greatest impact on the classes consisting of elderly people who do not appreciate organic products.[4] Additionally, we find that, while using both methods together does not always seems to be more effective than using them individually, the methods seem to have been more successful for the meat products that generally have a higher price and an animal-welfare component. These are important insights that show that attempting to correct for hypothetical bias appears to be necessary, especially in the case of organic products and especially for meat products.

The reason for analysing organic food products in the UK is not only because they are especially prone to hypothetical bias but also because, although organic food sales have consistently increased worldwide, in the UK they have stayed relatively constant over the period 2005-2016 (Figure 1). In 2017 they started to recover; however, the organically farmed area in the UK is declining, implying that organic food imports are increasing and that the UK may not be experiencing both the economic and the environmental benefits of organic production.[5] In the outset of Brexit, the potential exit of the UK from the EU Common Agriculture Policy (CAP), and the redesign of the UK agricultural policy, the present results about the willingness to pay for organic might be of interest to several stakeholders.

Figure 1 should be included around here

The present study can contribute for example by informing UK producers which attributes of organic products are most valued by their consumers and therefore how they can modify their production or marketing. The Organic Trade Board has recently received 10.4 million € by the EU to run a three-year campaign to promote organic food in the UK.[6] At the same time, this analysis could inform retailers on which attributes of organic products they should focus in their advertising campaigns. Finally, the present analysis can be used for welfare analysis in order to estimate the economic value derived from changes in various attributes of organic products and from developing new products with these attributes. These estimates could be

---

between them (organic being embedded in the category of Environmental Condition) is, to our knowledge, an international trade analysis from New Zealand (Guenther et al., 2015). However, this study does not attempt to correct for hypothetical bias.

[3] As can be seen from Tables III and IV. HB2=Cheap Talk, HB3=Honesty Priming, HB1=all treatments together.

[4] Class 4 for both products.

[5] http://www.ruralbusinessresearch.co.uk

[6] http://ofgorganic.org/organic-trade-board-secures-eu-funding/

of interest in different consumer market context such as: agri-food with credence attributes, agri-food with different levels of quality and agri-food with small size.

The results show that there is a core group of consumers who appreciate and are willing to pay for organic products. However, consumers also appreciate and are sometimes prepared to pay even more for other attributes, such as higher quality products, low chemical usage, and environmentally friendly production. For example, in the case of chicken, the average WTP for environmentally friendly production is £0.64/400gr, and the WTP for higher quality is £1.30/400gr, while the average WTP for the organic label is negative.[7]

Section 2 describes the design of the choice experiment. Section 3 describes the survey instrument and the data, section 4 the methodology and section 5 the results. Section 6 concludes.

2. Choice Experiment Design

The concept of WTP is rooted in consumer choice theory and is a measure of utility which assigns a level of satisfaction with the consumption of a good or a combination of goods. The optimal choice of a consumer is derived from its preferences and the budget constraints it faces. From the change in the optimal choices with the prices, a demand curve can be derived. Demand curve approaches in order to elicit the WTP and hence the valuation of an (environmental) good have been traditionally divided in stated and revealed preference methods based on the data collection methods. Stated methods use survey techniques and revealed methods use actual or simulated data (Garrod and Willis 1999). Empirical studies that compare the results obtained with the two methods find a stated to revealed ratio usually (but not exclusively) above one (Murphy et al., 2005; Gschwandtner, 2018; Penn and Hu, 2018). The gap between stated and revealed preferences is often referred to as hypothetical bias and it is the main aim of the present study to address it.

There are several theories and models that explain consumer choice but the literature seems to be in agreement that the five stage model developed by Engel, Kollat and Blackwell (1978) and extended by John Dewey (1910), is a good representation of consumer purchasing behaviour. According to this model the five stages through which the consumer reaches a buying decision are: problem recognition, search, alternative evaluation purchase, choice, and outcomes. These five stages are the most widely accepted, as evidenced in a majority of consumer behaviour textbooks (see, for example, Assael, 1998; Hawkins et al., 2003;

[7] Weighted averages calculated over latent classes (Tables III and IV).

Blackwell et al., 2005). At each stage additionally factors such as context and culture, knowledge, lifestyle and personality, motives, values and individual attitudes, emotions, beliefs and habits play a specific role in the decision process (Darley et al. 2010). Arguably, stated preference elicitation methods cover only some of these stages and not all and this might be the reason why a potential bias is observed.

Another theory, in line with the five stages model by Engel, Kollat and Blackwell (1978) that explains consumer buying behaviour is the Theory of Planned Behaviour (TPB). This theory states that attitude toward behaviour, subjective norms, and perceived behavioural control, work together to influence behavioural intention and finally actual behaviour (Ajzen, 1991). TPB has been often used in order to understand and predict consumer's motivation to buy organic food (see Scalco et al., 2017 for a meta-analysis related to organic consumption based on TPB). According to TPB, consumers go through a process of evaluation, then making plans to purchase and only then go through the process of purchase itself. Revealed preference is obviously at the final point of this process where consumers actually act, but in previous 'planning' stages there is scope for bias. Hence, stated behaviour is likely most prone to the bias. However, this bias can be significantly reduced if the elicitation method is incentive-compatible, consequential and includes ex-ante treatments against hypothetical bias such as the ones used in the present study and/or ex-post mechanisms such as for example certainty follow-up (Miller at al., 2011; Penn and Hu, 2018).

Options to measure WTP using revealed preferences are: market data and experiments. Experiments could be further subdivided in laboratory experiments, field experiments, or auctions. Auctions have received more attention recently, sometimes extending the Becker, De Groot and Marschak (BDM) (1964) bid mechanism to include eye-, mouse or web-tracking in order to obtain more accurate results (Liaukonyte et al. 2015). With respect to stated preferences, WTP could be elicited with the contingent valuation method (CVM) or through hypothetical discrete choice experiments (DCE). According to literature, CVM appears to be more suitable to evaluate a good in its wholeness, while DCE usually is better to evaluate a set of attributes which describe the specific good. Hence, estimated WTP values could vary according to the method used (Moser et al. 2011). Moreover, literature has shown that tangible attributes such as 'price' are weighted relatively more than intangible attributes such as 'prestige' in real choice experiments as compared to pure stated preference experiments (Horsky et al. 2004).

Additionally, the way the questions are asked in the survey or the way the attributes are displayed on the choice card in a choice experiment can have an impact on consumer behaviour. For example, it has been shown that the order in which questions are asked can lead to different outcomes (Bradley and Daly, 1994; Carlsson and Martinsson, 2001; Carlsson et al., 2012; Nguyen et al., 2015). This might have impact on marketing decisions. Products with specific attributes placed ‘earlier in the queue’ might receive a greater attention from the consumer and might be marketed at a higher price. This is in accordance with Query Theory developed by Johnson, Haeubl and Keinan (2007) which suggests that the order in which queries are processed influences choice behaviour. Equally, results might differ when lexicographic decision-making rules are accounted for in the modelling of discrete choice responses (Campbell et al. 2006). More importantly, questions about the willingness to accept (WTA) usually lead to much higher values than questions about the willingness to pay (WTP) (see Horowitz and McConnell 2002 for a review). At the same time, whether the questions are asked in a face-to-face interview or online might lead to a different result with online WTPs usually being lower due to reduced ‘warm glow effect’ (see for example Duffy et al. 2005 and citations herein). The present setup was designed to minimize potential overstatement by using an online choice experiment which randomizes the order of the question about the WTP. The ex-ante treatments and the ex-post mechanisms applied in this study are designed to minimize any potential biases arising and constitute its main methodological contribution.

Some more recent examples for studies that use revealed preference methods to elicit the WTP for organic products are: Griffith and Nesheim (2010), Briggeman and Lusk (2010), Griffith and Nesheim (2013), Bazoche et al. (2013), Waldrop et al. (2017). Current references of studies using stated preference methods to elicit the WTP for organic products are: Zander and Hamm (2010), Hu et al. (2011), Janssen and Hamm (2012), Rousseau and Vranken (2013), Caputo et al. (2013), Gerrard et al. (2013), Meas et al (2014), Feldmann and Hamm (2015), Zander et al. (2015), Kemper et al. (2019). Studies that use both stated and revealed preferences to elicit WTP for organic products published in more recent years are rarer: Lusk and Norwood (2009), Brooks and Lusk (2010), Norwood and Lusk (2011), Gschwandtner (2018). However, the list is not exclusive.

Choice experiments usually ask respondents to choose from a hypothetical choice set or to rank or judge attributes. The approach used here is developing choice sets in which the choices are described by bundles of attribute values usually but not exclusively associated with organic products. A set of attributes is chosen that reflects the characteristics of two products: chicken breast and carrots. The analysis focuses on these two product categories, because we would like to understand the differences between the most bought organic products (vegetables) and the least bought organic products (meat).[8] The list of the 7 attributes and the levels chosen for the analysis are presented in Table I.

Table I should be included around here

The attribute list in Table I also shows the way the levels of the attributes were communicated to the consumer on the choice cards. Since the Soil Association Organic Logo is more frequently seen in the UK than the EU Organic Logo, both logos were used to describe an organic product. The prices chosen are conditioned upon prices for the two products that can be found in UK shops and local markets.[9] The highest price is considered to be the price at which most consumers would stop buying the respective product ("choke price"). Note that the attribute levels that varied in the choice sets are discrete, even though in some cases they reflect underlying continuous variables, as it is often done in the literature (see for example Adamowicz et al., 1994). Discrete attribute levels are a consequence of the statistical design process used to create the choice sets.

For most of the attributes, only two levels have been chosen to make the design more feasible. Chemical usage can be average, or it can be low, thus reflecting low use of artificial pesticides and antibiotics for animals. Environmental friendliness refers to the use of environmentally friendly practices such as engaging in ecological processes, recycling, rotating crops, fitting the cycles, and maintaining the ecological balances in nature. Two levels exist, average and high, with the high level symbolized by the 'Eco-Friendly' logo. High animal welfare (which is an attribute available only for chicken) is graphically labelled with the 'Freedom Food' logo, which appears in the UK. The Freedom Food standards are designed to ensure the highest animal welfare and that all farm animals have a good life and are treated with compassion and respect. For example, chickens can be free range or kept indoors with plenty of space and natural light to move around and flap their wings. Objects like straw bales are made available for the chickens to peck at, and natural light helps to keep them active and healthy. The product can come at two qualities, average and high. The 'Best Before' date refers to the expiry date of the product. If the product expires in less than one week, it carries the label 'Soon (< 1 week)' and is depicted graphically with a sign saying 'Hurry up'. If the product expires after one week or longer, it carries the label '1 week or longer' and is depicted

[8] Even though organic dairy products are also often bought, the organic designation is mostly associated with produce.
[9] They were derived from the pilot study or retrieved from the homepages of the shops where the products were bought.

graphically to the consumers with a yellow box containing the information, ‘You can use this product after one week or longer’. Even though there are conventional products that expire before one week, in general they are assumed to last longer than organic products. We would like to understand how important this attribute is to the consumer.

Although clearly the attributes must not be correlated with each other, it is easy to observe how an organic product may be more expensive produced in more environmentally friendly conditions, and involve higher animal welfare than a conventional product. Therefore, it would be difficult if not impossible to use them together in a revealed preference model without encountering multicollinearity. The statistical design in the choice experiment ensures that the impact of each attribute can be estimated independently.

The set of attributes and their levels are setting the space to be spanned in the choice experiment. The present choice experiment is unlabelled with respondents selecting between product1 (Option A) and product2 (Option B), or no choice (Option C). If each attribute is treated as discrete, there are 2^5 * 3*2 (192) possible alternatives for chicken and 2^4 * 3*2 (96) alternatives for carrots. Obviously, it is impossible to confront consumers with all of these alternatives; therefore, a subset was chosen using an efficient design. The problem of choice set construction can be viewed as sampling from the universe of possible pairs of products. The most important condition to be fulfilled is the IIA property. The respondent was faced with three non-labelled alternatives, the profile of two of which being drawn from the design, while the third option was ‘neither of these’, implying that they would not purchase the commodity that week.

In the present paper, we present only the results estimating the main effects model, and we do not use interaction effects.[10] Moreover, we do not use a complete factorial design, but we choose just a fraction of it, such that it enables the estimation of the parameters to keep the number of standard errors as low as possible. The design strategy produces optimally efficient estimates of the parameters based on the notion of D-optimality. A D-optimal efficient design minimizes the D-Error, which employs the determinant of the asymptotic variance covariance matrix of a single respondent. Constructing the design requires priors, which were obtained from a pilot study run with 60 individuals previously to the actual CE. The final design consists of 32 choice sets per product using the main effects design strategy.presents examples of choice card/tasks to illustrate how the design was implemented into the survey. As can be

[10] Even though interaction effects could in principle exist and it may be interesting to analyse them, it would complicate the design even further.

seen from these examples, ‘chemical usage’ and ‘organic label’ were never both present in the same design. It would have made little sense to have the two attributes in one choice set as they are alternative ways of representing a similar thing. ‘Organic’ implies already per definition low chemical usage. For this reason the existing design was duplicated, in one there was the ‘Organic Label’ and in the second this attribute was replaced with ‘Chemical Usage’. The motivation for doing the replication was to compare consequences of using either a ‘Chemical Usage’ or ‘organic label’ as attributes and to see if it makes a difference. It appears that it does as the model where the parameters of both attributes are restricted to be equal is different to the model where they are estimated separately. Consumers do seem to perceive them as different attributes. The hypothesis that the two models are different could not be rejected in a likelihood ratio test.[11]

It is unrealistic to expect individuals to respond to all 2X32 choice sets in an interview setting. Consequently, we divided the experiment into four sets of eight choices for each product by using an additional four-level column as a factor in the design. Grouping the choice tasks in such a manner ensures that each block of choice sets is approximately equivalent. Therefore, the respondents had to perform 16 randomly chosen choice tasks in the survey, which is a large number of choices but is typically used in the literature (see Adamowicz et al. 1994, Balcombe et al., 2016, Burton et al., 2016). Each respondent received a set of instructions for completing the survey and the choice task together with background information about organic and a detailed description of the attributes. Three different hypothetical bias treatments were employed. A set of socio-economic characteristics was elicited together with the choice tasks in the survey and will be described in greater detail in the Data section below.

3. Data / Survey Instrument

The data were collected in April 2016 via an online survey performed across UK by a professional market research company after we discussed the survey instrument with a focus group.[12] Originally, 60 observations were collected for the pilot from which priors were derived. After running the pilot, about 520 observations were collected for the main survey, from which

[11] In order to test if it makes a difference to use ‘organic label’ or ‘chemical usage’, the latent class models shown in the paper were replicated using instead of ‘organic label’ and ‘chemical usage’ the sum of both. This effectively restricts the parameters on each to be the same. The hypothesis that the models using ‘organic label’ and ‘chemical usage’ separately are different could not be rejected (alpha=0.05, df=5). Results can be obtained from the author upon request.

[12] The focus group consisted of organic buyers and non-buyers from the University of Kent and from Canterbury Christchurch University. The company providing the data is ‘Qualtrics’ (www.qualtrics.com).

505 were maintained as valid and used in the further analysis. This number of observations should be representative for the UK population according to Thompson (1987).[13]

The questionnaire consisted of three main parts. The first (A) inquired about the actual purchases of the individuals (revealed preferences=RP), the second (B) about purchases in the experimental set up (Choice Experiment for elicitation of stated preferences=SP), and the last part (C) contained questions about socio-characteristics of the individuals.

The first part contained background information and informed the consumer that their answers will help design supermarket pricing policy and may have consequences for the future; this implied consequentiality of the project to the consumer. The person that does the shopping in the household was asked to complete the questionnaire and to answer as truthfully as possible. Consumers were reminded that, even if this is a hypothetical situation, it is important that they try to answer as if in a real shopping situation. The hypothetical scenario involved a situation in which the government would be interested in encouraging the production and consumption of organic products and therefore would like to find out how much consumers pay for organic products if they buy any, and how much they would be willing to pay for organic products even if they don't buy any now.

In the first part of the questionnaire, consumers were asked if they bought chicken breast and/or carrots in the last month. If they did, then consumers were asked about the quantity bought (in kg), the shop where they purchased the product(s), whether the products were organic or not, whether they were the shop's own brand or not, and whether they would expire in less than one week or not. If the consumers did not buy any chicken breast or carrots in the last month, they were excluded from the sample.[14]

The second part of the survey, which concerned the stated preferences, contained a comprehensive description of the attributes, instructions on how to answer the choice tasks, and the choice tasks themselves. As hypothetical bias is the strongest criticism brought to stated preferences techniques, the choice experiment contained 3 different treatments against hypothetical bias. The first treatment contains both Cheap Talk and Honesty Priming, the second treatment contained just Cheap Talk, and the third treatment just Honesty Priming.

---

[13] Equation (1) on page 43 defines the sample size $\mathrm{n} = \max_m z^2(\frac{1}{m})(1-\frac{1}{m}) \big/ d^2$, where m=nr of categories, (choices) =3 in our case, d= allowed sampling error of 0,05, z= upper $(\alpha/2\mathrm{m}) \times 100\text{th percentile of the standard normal distribution can}$ be found in the tables for α=0.05 and Φ(z)= 0.99 being equal to 2.3. Therefore, $\mathrm{n} = \frac{2.3^2\left(\frac{1}{3}\right)\left(1-\frac{1}{3}\right)}{0.05^2} \approx 487$<505.

[14] Which unfortunately makes the stated and the revealed sample not perfectly compatible and, therefore, impossible to join. Nevertheless, the information from the revealed part can be insightfully used in the stated part.

Furthermore, these treatments were randomly distributed to the consumers in four blocks as follows:

Block 1 (HB1) contained the first treatment: Cheap Talk and Honesty Priming.

Block 2 (HB2) contained the second treatment: Cheap Talk only.

Block 3 (HB3) contained the third treatment: Honesty Priming only.

Block 4 (HB4) contained no hypothetical bias treatment and was used as a reference group.

By comparing the WTP between the four blocks, it can be determined whether the hypothetical bias treatments have influenced responses and hence have reduced hypothetical bias. Questions regarding the ranking of the attributes according to their importance to the consumers and the attribute non-attendance concluded the stated preference part.

In the last part of the questionnaire, a wealth of socio-economic characteristics, including scales with reasons for and against organic purchasing, were elicited. The set of questions was carefully constructed after consulting the recent literature with respect to consumption behaviour regarding organic food (Yiridoe et al., 2005, Hemmerling et al., 2015).
From the multitude of variables that we could statistically describe, we chose some important socio-economic characteristics related to consumption of environmental/organic goods to assess the representativeness of the sample, which are described below.

Table II should be included around here.

Mean **Age** in our sample is around 50, with a median of 52 and a modal value of 59. This is above the UK predicted mean (median) age by the Office of National Statistics (ONS) of 40, and therefore, presumably our sample is not representative in terms of age. However, this is not necessarily relevant, since the impact of age on organic purchase is inconclusive (Aertsens et al., 2009). Moreover, as will be explained later, we control for this bias.[15]

Income is presented to the consumer in intervals of £ 1000 from below £500 (category 1) net per month until over 4500 £ (category 10). According to the ONS statistics, the average salary in the UK in 2015/2016 was £27.600.[16] This corresponds to a net monthly disposable income of £1610, which is close to our sample average of £1524.95.[17] Since income is arguably the

[15] Additionally, the higher average might be driven by the fact that our sample does not include people below 18 years old as opposed to the estimate of the ONS, which is correct, as people under 18 are probably in general not responsible for the household shopping.
[16] https://www.incometaxcalculator.org.uk/average_salary.php
[17] 1610=27000*0.7/12, assuming an average tax rate of 30%.

most important characteristic determining the consumption of environmental and ethical goods such as organic products, this is reassuring.[18]

Education is divided into eight categories corresponding to education levels, starting with less than high school up to PhD. It can be observed that the average value is between 3 and 4, which the categories are corresponding to 'Some College' and a'2-year College Degree'. This corresponds to 13.5 years of education, which is above the UNO statistics reported for the UK of 12.3.[19] This implies that we have a sample of over-educated people, as often is the case in online surveys.[20] Even if it is in general expected that higher educated people buy more organic products and care more about attributes like animal welfare and environmental friendliness, the results with respect to education are in general inconclusive in the literature (Aertsens et al., 2009). Moreover, as will be explained later, we calculate our variables as deviations from the true population average education level, and therefore, the bias if existent should be minimized.

Children is a variable that has often been found to be associated with consumption of organic or a healthy diet in general. Some articles find that the impact of children is positive, as parents want to provide their children with healthy nutrition (Thompson and Kidwell 1998, McEachern and Willock 2004), while others find that they impact negative mainly due to income effects (Loureiro and Hine 2002, Tiffin and Arnoult 2010).

The ONS reports the number of families with dependent children according to the number of children (0, 1, 2, 3, or more). We have calculated the percentages and, based on this, the expected number of children in each category in our sample. Then we have employed a Chi^2 Test to compare the expected with the observed number of children for each category, and it appears that the number of children living in the household in our sample is representative of the UK population which is reassuring.[21]

Gender plays a significant role in food consumption in general and especially with respect to the consumption of environmental and organic goods (Byrne et al. 1991, McEachern and McClean 2002, Lea and Worsley 2005, Arbindra et al. 2005, Radman 2005, Stobbelaar et al. 2007). Therefore, the sample was chosen to consist of 60% women and 40% men in order to reflect the fact that women not only are more frequently responsible for food shopping in the

[18] T-test statistics= -1.77 < 2.6 for 100 degrees of freedom at the 99% confidence level, and therefore, we can accept the hypothesis that our sample mean is not different from the population mean.
[19] http://hdr.undp.org/en/content/mean-years-schooling-adults-years
[20] T-test for comparison of the means = 16.96 > 2.6 (critical value for 100 degrees of freedom at 99% confidence level).
[21] The Chi^2 value is 8.10< <70.6, which is the critical value for more than 100 observations (we have 505); therefore, with a 99% confidence level, we can argue that the sample is not significantly different from the UK population in terms of children.

household but also buy more environmental goods. Further variables used in the study are described in the Appendix in order to conserve space.

4. Methodology

The response to the choice between the three constructed choice alternatives (product1, product2, no option) can be modelled in a random utility framework. The overall utility can be expressed as the sum of a systematic component, which is expressed as a function of the attributes presented, and a random component:

$$U_{in} = v_{in} + e_{in} \qquad (1)$$

where U_{in} is the utility of individual n from choosing alternative i, v_{in} is the systematic utility component, and e_{in} is the random component. Alternative i is chosen over alternative j if $U_{in} > U_{jn}$. *The* probability of person n choosing alternative i is given by:

$$\pi_n(i) = \Pr(v_{in} + e_{in} \geq v_{jn} + e_{ij} \,;\forall\, j \in C_n) \qquad (2)$$

where C_n is the choice set for individual n. If we consider V_{in} to be a conditional indirect utility function that has a linear form, we can write it as follows:

$$V_{in} = \beta_1 + \beta_2 x_{in2} + \beta_3 x_{in3} + \cdots + \beta_k x_{ink} + \alpha(Y - P_i) \qquad (3)$$

where x_{ink} are the attributes of the alternatives described above, Y is income, and P_i is the price of alternative i. Assuming that the error terms are Gumbel distributed with a scale parameter μ, the probability of choosing alternative i is then given by:

$$\pi_n(i) = \frac{exp^{\,\mu V_{in}}}{\sum_{j \in C_n} exp^{\mu V_{jn}}} \qquad (4)$$

The scale factor μ is usually assumed to be equal to 1. However, in the present case, $\mu=2$ yielded a better fit in terms of BIC. Note that it is important to differentiate between preference and error heterogeneity, and failure to adjust for scale heterogeneity might lead to biased results (Louviere and Eagle, 2006).

In the present study, a Latent Class Model – Attribute Non-Attendance (LCM - ANA) has been employed. The Latent Class Model is a semi-parametric extension of the Multinomial Logit Model, which facilitates the investigation of heterogeneity on a class (segment) level, and it relaxes the assumptions regarding the parameter distribution across individuals (Greene and Hensher, 2003). This approach endogenously groups individuals into classes of homogenous preferences (Scarpa and Thiene, 2005, Hammitt and Herrera-Araujo, 2017) and estimates their probability of membership according to their designated class depending on their socio-economic characteristics (Kikulwe et al., 2011).

Hensher et al. (2005) state that respondents may not always use all attributes when making their decision in choosing an alternative; some may, intentionally or not, be ignored. According to Mariel et al. (2013), respondents do not use all attributes when making their decision, and if this information is not taken into account, the estimate of their willingness to pay could be influenced. In the present study, a condition for the non-attendance of a particular attribute setting its parameter to zero was applied if the coefficient was not significantly different from zero in the latent class model. Campbell et al. (2008) support that including this information provides a better-fitted model and yields more accurate results.

One of the main aims of the present study is to quantify individuals' willingness to pay (WTP) for each attribute within the choice set. The WTP is calculated as the ratio of each attribute's coefficient over the monetary value coefficient (Loureiro and Umberger, 2007; Kerr and Sharp, 2009; Greene, 2012) and is interpreted as a change in value associated with an increase of the attribute by one unit. This measure can then be used in order to estimate the levels of welfare associated with various products and their attribute combinations in order to decide which one is most valued by the consumer. It should be noted that if the denominator is zero the WTP is undefined.

5. Empirical Results

5.1. Latent Class Results

The description of the results will be done first for chicken and then for carrots. It will start by describing the results for each latent class of consumers taking into account their attribute non-attendance. It then proceeds with the description of the WTP results setting the emphasis on hypothetical bias treatments and their impact.

In determining the optimal number of classes the analysis started with the simple one-class model and increased the number of classes iteratively to the point at which the BIC and CAIC values started to increase.

As mentioned above it is important to differentiate between preference and scale heterogeneity as the differences in parameters may result from spurious segments that differ only in scale and not in their preferences or willingness to pay. We control for this by estimating a scale-adjusted latent class (SALC) model. We find that a model with 5 preference classes and 2 scale segments gives the best fit for each type of product. For chicken the segment with the lower variance consists of 84% and for carrots of 37% of the sample. [22] [23]

In order to analyse the impact of the hypothetical bias treatments, we estimated the models using price interaction dummies for the three treatments, leaving the fourth group, which had no treatment as a comparison base. The results are presented at the bottom of tables III and IV before the s-class and the covariates results. The following covariates were used in order to explain class membership: actual organic purchases[24], age, income, and pro-organic attitude.[25]

Table A2 in the Supplementary Material presents the answers to the debriefing questions with respect to attribute non-attendance. According to consumers, their most ignored attribute was Chemical Usage (ca 40% of people reported having ignored this attribute) followed by Environmentally Friendly (ca 36% ignored it), Best Before and Organic Label (ca 30% each). A surprisingly high percentage of people (around 27%) claimed to have ignored the price attribute. This is in accordance with the results for chicken, for which the class that ignored price consisted of about 27% of the respondents (Class 2). The least ignored attributes seem to be Animal Welfare and Quality. Quality may not be surprising, since it seems to be the attribute that is valued by most classes of people and seems to be the attribute with the highest WTP. Most notably however, only around 15% of the sample reported *not* having ignored any attributes. Therefore, estimating a model without accounting for attribute non-attendance (ANA) might have led to wrong results.

The general form of the class *k* utility functions for the 8 attributes (L=Organic Label, E=Environmentally Friendly, Q=Quality, B=Best Before, C=Chemical Usage, A=Animal Welfare, SQ=No Choice Option or Status Quo, P=price, P1=Price*HB1dummy, P2= Price*HB2dummy, P3= Price*HB3dummy) is:

[22] Information criteria for latent class models for chicken are presented in Table A1 in the Supplementary Material.
[23] The scale factor for chicken is exp(-2.26.) and for carrots it is exp (-1.99) i.e. ~0.1 for both products as can be seen from Tables III and IV.
[24] A dummy called BuyOrg indicating if the person bought organic in the last 2 weeks.
[25] Derived from a scale consisting of 10 questions related to organic products. More detail in the Appendix.

$$U_k = \beta_{kL} * L_k + \beta_{kE} * E_k + \beta_{kQ} * Q_k + \beta_{kB} * B_k + \beta_{kC} * C_k + \beta_{kA} * A_k + \beta_{kSQ} * SQ_k + \beta_{kP2} * P2_k + \beta_{kP3} * P3_k + \beta_{kP1} * P1_k + \beta_{kP} * P_k$$

Given the number of attributes in the model, the number of possible combinations of ANA is large. Therefore, we adopt the pragmatic strategy of identifying the insignificant attributes in the five class models estimated without restriction, and restricting these to zero, allowing for the price coefficients to be free in order to estimate WTP, if necessary.[26] The resulting model structure can be found in the Appendix in the Supplementary Material.

It is important to note that the reason why Chemical Usage appears to be the most ignored attribute is not because consumers presumed that the 'organic label' already contains the 'chemical usage' attribute anyway. As can be seen from the examples of choice cards in Figure A1 in the Supplementary Material, 'chemical usage' and 'organic label' were never both present in the same design. It would have made little sense to have the two attributes in one choice set as they are alternative ways of representing a similar thing. Organic implies already per definition low chemical usage. For this reason the existing design was duplicated, in one there was the 'organic label' and in the second this attribute was replaced with 'chemical usage'. The motivation for doing the replication was to compare consequences of using either a 'chemical usage' or 'organic' as a label and to see if it makes a difference. And it appears that it does as the model where the parameters of both attributes are restricted to be equal is different to the model where they are estimated separately. Consumers do seem to perceive them as different attributes. The hypothesis that the two models are different could not be rejected in a likelihood ratio test.[27]

Table A3 in the Supplementary Material compares the statistics parameter for the latent class model with ANA and without for chicken. It can be observed that BIC, AIC, AIC3, CAIC, and SBIC are smaller (better) for the model accounting for attribute non-attendance than without, as would be expected when a set of insignificant parameters is removed. Therefore, the model that accounts for the fact that some classes don't value (ignore) some attributes is an improvement. The results of the model using the ANA restrictions for chicken are presented in Table III and will be discussed below. The parameters are presented for each of the 5 classes (coefficients, standard errors and z-values) together with overall Wald-values.

Table III should be included around here.

[26] Alternatively, the ignored attributes could be given a lower weight (see for ex Hess et al., 2013 or Chalak et al., 2016).

[27] The likelihood ratio for chicken was 12.29 and for carrots 15.4 which both are above the critical value of 11.07 for alpha=0.05 and df=5. Results can be obtained from the author upon request.

The only attributes valued by people in class 1 are environmentally friendly production and quality. People in this class are sensitive to price and rather young as the coefficient of age is negative and significant. This class consists of about 33% of the sample.

Class 2 is the class of people that appreciates the organic attribute without considering the price as being an important deterrent, as the price coefficient of this class is not significantly different from zero. Unfortunately, because of this reason we won't be able to estimate the WTP of this class[28], and therefore, the WTP for organic chicken will be biased downwards, despite a considerable amount of people appreciating it (27%). Moreover, Class 2 also appreciates other attributes related to organics, such as environmentally friendly production and animal welfare but also the attribute 'Quality', which is appreciated by each class. Perhaps not surprisingly, the probability to belong to this class is positively influenced by organic consumption ('BuyOrg') and by a pro-organic attitude ('Pro-organic').

Class 3 is exactly the opposite. This class appears to *actively dislike* the organic attribute and other environmental attributes as the coefficient for these attributes is negative and significant.[29] The only attributes that this class seems to appreciate are 'Quality' and 'Best Before' (equivalent to a long expiry date). People in this class are sensitive to price, as the price coefficient is negative and highly significant and do not seem to have a high income as the coefficient of income is also negative and significant. This class is smaller consisting of about 20% of the sample.

Class 4 is similar to Class 3 in the sense that the only attributes that it really appreciates are 'Quality' and 'Price'. The main difference between Classes 3 and 4 is that, while Class 3 actively dislikes the organic label and other environmental attributes, Class 4 is indifferent to them (the coefficients are zero and not negative). The probability to belong to Class 4 is

[28] As it is technically infinite.

[29] It may seem surprizing that people appear to *dislike* organic products. However, the literature on organic shows that consumers sometimes do not trust organic products and consider the organic logo as a marketing trick used to increase prices and that trust in organic products seem to have decreased over time in some countries (Vitterso and Tangeland, 2015; Yin et al., 2016). Several 'organic scandals' where products sold as organic have been identified as not conforming to organic standards have led to consumer distrust with respect to organic food products (Italy 2011, Germany 2002 & 2013, Nebraska/USA 2018, Netherlands 2019). Moreover, concerns are raised that organically farmed animals actually have lower animal welfare due to higher exposure to natural enemies and due to lower doses of antibiotics allowed in case of illness. This concern has been often raised for example with respect to mastitis in milk-cows or with disease spread in pigs and poultry (Hovi, 2000; Sundrum, 2001; Hovi and Sundrum, 2003; Von Borell and Sorensen, 2004). Finally, there is a valid argument that organic agriculture has lower yields than conventional agriculture, and critiques argue that it would therefore need more land to produce the same amount of food resulting in more deforestation and biodiversity loss undermining the principle of environmentally friendly production (Seufert et al. 2012; Meemken and Quaim 2018).

positively influenced by age and negatively influenced by a pro-organic attitude. Hence, rather older people without a positive environmental attitude appear to belong to it.

Class 5 is the smallest class, consisting of only 5% of the sample population. The only attribute that this class seems to appreciate is 'Quality'. However, this class includes individuals that choose most frequently the opt-out attribute ('SQ') and have a positive price coefficient, meaning that they derive positive marginal utility from paying a higher price. People who buy organic and have a high income seem to belong to this class, which might explain their attitude towards price.

What is most interesting to observe is that the addition of the hypothetical bias interaction effects is overall significant, and this suggests that the WTP has changed as a result of the hypothetical bias treatments. In particular, the interaction of price with the HB2 dummy leads to a larger price coefficient in absolute value in Classes 3 and 4, so the WTP is smaller for that treatment compared to the baseline (no treatment) in the respective classes.[30] This would suggest that using Cheap Talk has been successful in reducing hypothetical bias, presuming that an overstatement was in place.[31] Even though the Wald-values may suggest that all treatments have an effect in aggregate, this does not necessarily mean that the treatment must have a significant effect for each and every class.[32] In Classes 1, 2, and 5, no treatment seems to have had a significant impact, while in Class 3, all treatments seem to have been effective. Class 3 is also a class that is very sensitive to price and consists of people with rather lower income which might explain why they are more sensitive to hypothetical bias treatments. In Class 4, Cheap Talk and both treatments together had a significant impact.

Therefore, we can conclude that the reduction in hypothetical bias was effective for Classes 3 and 4, the classes of people that are sceptical or indifferent towards organic products that either have a low income, are elderly, or ranked low on the pro-organic scale. For these types of consumers, using Cheap Talk alone or in combination with Honesty Priming seems to be effective in reducing potential overstatement of their WTP. Without these treatments,

---

[30] The interaction term has the largest z-value from all 3 HB treatments and is negative for all classes. This means that the price coefficient will be reduced for all classes and will therefore increase in absolute value (from -2 to -3, for example). When dividing the attribute coefficient by it, this will result in a lower WTP.

[31] We presume that we move towards the truth, but of course, we don't know with certainty.

[32] As often with credence attributes, we observe an overstatement of the WTP derived from a warm glow for the respondent in giving the desired answer, we interpret a reduction in the WTP as a movement towards the truth and hence success of the HB treatment. However, we don't really know if the reduction goes in the right direction, as we do not know the true WTP. We observe a difference between the WTP without HB treatment and the one with the treatment, but we don't actually know if this is in the direction of the truth.

respondents in these classes might get a warm glow from overstating their true WTP. The treatment appears to have determined them to change their marginal utility of money and induced them to reveal their true WTP, which is close to zero or even negative. It might be interesting to observe that, in the case of chicken, using both Cheap Talk and Honesty Priming (HB1) had the largest Wald-value, which might suggest that using the two treatments together was more effective than using them individually, as hypothesized. However, Honesty Priming alone (HB3) did not have an effect on Class 4 therefore, the observed effect of using Cheap Talk and Honesty Priming together in this class might be driven primary by Cheap Talk. Overall, using hypothetical bias treatments seems to have made an impact in the case of chicken, the product with the higher price and hence using them in eliciting the WTP for this product seems to be important. As this is a meat product, consumers might be inclined to overstated their WTP for animal welfare if hypothetical bias treatments are not in place.

Tables A4 and A5 in the Supplementary Material present parameter values for the models with ANA and without for carrots. As expected BIC values are improved when attribute non-attendance is corrected for. The results of the LCM- ANA results for carrots are presented in Table IV and will be discussed below.

Table IV should be included around here.

Class 1 seems to actively dislike environmentally friendly attributes, such as organic production, environmentally friendliness, low chemical usage, or quality as their coefficients are negative and significant. The only attribute that it seems to appreciate is a long expiry date (attribute Best Before). This class consisting of 36% of the sample seems to be made of rather young people as the attendance to it is negatively influenced by age. However, this class has a positive price coefficient, which makes the interpretation of the WTP for this class difficult.

Class 2 seems to be indifferent to most attributes, as their coefficients are either zero or insignificant. It seems to consist of relatively poorer respondents - as the coefficient of income is negative and significant - which are highly sensitive to price (price coefficient negative and highly significant as well). This class is the largest class, consisting of 38% of the sample.

Class 3 is a class of people that seem to appreciate the organic label and other environmental attributes, such as environmentally friendly production, low chemical usage, and also a higher quality but does not seem to appreciate a long expiry date, possibly due to their preference for fresh products. This class of people has a negative and significant price coefficient and consists of about 11% of the analysed population. Attendance in this class is positively influenced by a pro-organic attitude and negatively influenced by age.

Class 4 is indifferent about the organic label and all other attributes except price. This class seems to be extremely sensitive with respect to price, and maybe this is explained by the fact that attendance in this class is positively correlated with age. It consists of about 9% of the population.

Class 5 seems to appreciate the organic label as Class 3 does, but the price coefficient, even though negative, is not significant. Therefore, it won't be possible to estimate its WTP. This class is the smallest however, consisting of only 6% of the sample. Together, the classes that appreciated the organic label in the case of carrots are Classes 3 and 5, and make up 17% of the sample.

Again, what is maybe more interesting to observe is that also in case of carrots, hypothetical bias treatments seem to have had an impact. This can be seen from the parameters of the interactions between HB treatments and price. In Class 3 all interactions - and in Class 4 the interactions with HB2 and HB1 are negative and significant. The overall effect of HB treatments, however, is much smaller than in the case of chicken, as can be seen from the Wald-values. Class 4 is similar to that of chicken. It does not seem to appreciate almost any attribute, is highly price sensitive, and consists mainly of elderly people that did not score highly on the pro-organic scale. This is also the class that chooses most frequently the opt-out option (SQ). Honesty Priming (HB3) might not have worked as well in this class, as it is more time consuming and hence the observed significant effect of both treatments together (HB1) in this class might stem primarily from Cheap Talk (HB2) as in the case of chicken. In Class 3, HB3 seems to have worked in the opposite direction. The impact is positive and significant, suggesting that the treatment has increased marginal utility of money and has made consumers less sensitive to costs. This might seem surprising but could be explained by the fact that the treatment helped consumers reveal their true preferences for environmental attributes and their willingness to pay a higher price for them as this is the class that appears to value all environmental attributes.

Overall, also for carrots, HB2 (Cheap Talk) seems to have had the strongest impact, as it is reflected in the overall significance given by the Wald-value. However, in general, hypothetical bias treatments seem to have been less effective for carrots than for chicken, perhaps because the price of carrots is much lower than that of chicken, and there was not much space for overstatement and hence not much bias for which to correct in first place. This might also be the reason why using Cheap Talk and Honesty Priming together (HB1) was *not* more successful than using them individually, as the Wald-value would suggest. Hypothetical bias treatments had an impact also in the case of carrots but it seems to be much smaller than in the case of chicken.

5.2. WTP Results

In what follows, the WTP will be presented and discussed regarding all of the attributes for chicken for each class and for HB1.[33] The interpretation of the results, however, will be similar for all treatments, as the % change in WTP implied by the shift in the price coefficient is the same for each treatment in a specific class. Table V summarizes WTP results for chicken. The last column gives the average weighted WTP for each attribute.[34]

Table V should be included around here.

Class 1 had a positive and significant coefficient only for Environmentally Friendly production and 'Quality', and hence, only the WTP for these two attributes could be calculated, with the WTP for 'Quality' being almost twice as large as for 'Environmentally Friendly' (£3.67 compared to £1.94).[35] [36] Class 3 had a positive WTP for 'Quality' and 'Best Before' with the WTP for 'Quality' being 160% larger than for Best Before (£0.27 compared to £0.17). Class 4 had a positive WTP only for 'Quality' and equal to £1.01. Overall, the weighted average WTP for the attribute 'Quality' was the highest (£1.3) followed by the one of 'Environmentally Friendly' (£0.64) and the one of 'Best Before' (£0.04).[37]

However, it might be more interesting to discuss the % changes in WTP implied by the shift in the price coefficients as a result of hypothetical bias treatments reported in the bottom three rows of the table. It can be observed that, in Class 1 for all attributes, the WTP obtained with HB1 is only 0.87% of the one obtained using HB2 and it is only 0.89% of the one obtained with HB3. This implies that the WTP obtained with HB1, which uses both Cheap Talk and

---

[33] The basis for these calculations are not the attribute and price coefficients in Table III but ones in which individual prices for each treatment have been used, and therefore, the marginal utility from each treatment could be estimated (to be obtained upon request). The two models are behaviourally equivalent.

[34] The formula for calculating the weighted WTP is given by: *Average Weighted* $WTP_{ai} = \sum_{k=1}^{S}(WTP_{ak} * Pr_{ik})$ where WTP_{ak} is the class k WTP for the attribute a, and Pr_{ik} is the individuals i probability to be in that specific class. The sum is over all classes for each individual.

[35] Note that these are just the WTP figures obtained using HB1.

[36] This is also visible from the last column in Table V in which we multiply the WTP with the class size in order to obtain the weighted average WTP for all classes.

[37] It is important to remember that the WTP of Classes 2 and 5 cannot be calculated, because the price coefficient is either not significant or positive.

Honesty Priming together, is actually lower than the ones obtained using just Cheap Talk (HB2) or just Honesty Priming (HB3) alone. This supports the hypothesis that using both treatments together might be more effective than using them individually in this class. The most interesting result is however, the fact that HB1 seems to render a WTP that is only 70% of the one obtained when no treatment is in place (HB4). This suggests that using treatments against the hypothetical bias was necessary and has had an impact in this class.

Class 3 is the class of people that actively disliked the organic label and other environmentally friendly attributes, having a positive coefficient only for ‘Quality’ and ‘Best Before’ with a much larger WTP for ‘Quality’. All other attributes had either a zero or negative coefficient, and hence the WTP are accordingly. In this class however, using both Cheap Talk and Honesty Priming together (HB1) does not seem to lead to a lower WTP than when using the methods individually. The fractions HB1/HB2 and HB1/HB3 are both close to 1, suggesting that all treatments had a similar impact. This is reflected also in the fact that the coefficients of the three treatments are numerically very similar in size. This notwithstanding, using hypothetical bias treatments has the strongest impact in this class, which is not only reflected in the fact that all three interaction terms with HB treatments are highly significant but also in the fact that HB1 renders much lower WTP than if no treatment had been in place. The ration HB1/HB4 is only 0.46, meaning that the WTP obtained using both Cheap Talk and Honesty Priming together is only 46% of what it would have been if no treatment would have been applied. Again, this suggests that HB treatments have been necessary and had a successful impact also in this class. The strongest case for the effectiveness of hypothetical bias treatments can be made for this class in which all treatments seem to have caused the WTP to fall.

Class 4 is the class of people who did not care about any attributes but Quality and who chose often the opt-out option (‘SQ’). In this class, only HB1 and HB2 had a significant impact, and the effect of both appears to be very similar, as HB1/B2 is almost equal to 1. HB1 yields WTP values which are only 70% of the ones obtained using HB3, and hence, WTP values are lower for HB1 than HB3, which is what was originally hypothesized. Using both Cheap Talk and Honesty Priming together (HB1) is expected to be more effective than using just Honesty Priming (HB3). Using hypothetical bias treatments seems to have been effective in this class too, as WTP values using HB1 are only 70% of what they would have been without treatment.

Even though the weighted WTP values are only for one specific treatment (HB1), we can nevertheless say that the largest WTP in the case of chicken was for the attribute ‘Quality’, followed by ‘Environmentally Friendly’. Therefore, consumers seem to be interested in attributes related to organic production in the case of chicken, even if this is not reflected in

their WTP for the organic attribute specifically. It may be worth reminding again that there is a large class of people consisting of about 30% of the sample (Class 2) that appreciates the organic label and would probably have a high WTP for it. However, as the price coefficient of this class is not significantly different from zero, we cannot calculate its value.
Table VI presents the corresponding WTP for carrots.

Table VI should be included around here.

When analysing the WTP for the attributes of carrots, it must be mentioned again that only the price coefficients for Classes 2, 3, and 4 were negative and significant, and therefore, the WTP could be calculated only for these classes. It should also be reiterated that the HB price coefficients were mainly significant only for Class 4, and therefore, the HB treatments seem to have worked best in reducing HB for this class. However, this class was not relevant when estimating WTP, as the WTP was zero for each attribute. Therefore, it is also excluded from Table VI. For the estimation of the WTP, only Class 3 and sometimes Class 2 were relevant. Therefore, we do not expect that the hypothetical bias treatments have managed to reduce the WTP as much as in the case of chicken, and it may seem that what we observe is simply noise. This might explain some unexpected results, such as the fact that, in Class 2, HB1 renders a WTP almost twice as high as HB4, when in fact the WTP should be lower if treatments would have worked. The fact that in Class 3 HB1 renders a WTP that is 1.65 times lower than HB3 might be explained by the fact that HB3 was the only significant treatment in this class and by the fact that the treatment has actually shifted the respondents to not caring about funds and having a positive impact on their WTP.

The other relative effects between HB treatments are close to 1 and suggest no large differences between the treatments. The general picture that emerges is that HB treatments seem not to have been as successful for carrots as they have been for chicken. This might have been driven by the fact that carrots are missing the animal welfare component, and hence there is less scope for social desirability bias than in the case of chicken. Moreover, the price of carrots is relatively low, leaving little space for overstatement and hence potential hypothetical bias.

Bearing this in mind, it seems that, in the case of carrots, the weighted WTP for 'Organic Label' is highest (0.65 £/kg), followed by the one of 'Quality' (0.27 £/kg). The average WTP for low chemical usage is 0.21 £/kg, and that for environmentally friendly production is 0.15 £/kg. In general, the WTP for environmental attributes like low chemical usage, environmentally friendly production, and the organic label (which encompasses both) seems to be larger when choosing carrots than when choosing chicken. Perhaps since chicken is already a more expensive meat product, the WTP for additional attributes is lower, or maybe our respondents consider that these environmental attributes are more important in the case of carrots. This is supported by the fact that there seems to exist a core of environmentalists (Class 3 and Class 5=17% of the respondents) in the latent class model who appreciate all of these attributes.[38]

6. Discussion and Conclusions

The findings of this paper should be of interest to economists and practitioners for several reasons. Firstly, there has been almost no recent formal economic analysis of the WTP of British consumers for organic attributes, and the present study not only estimates WTP for the organic label but also determines related attributes together with the estimated hypothetical bias. The analysis has been done for two different products representing two food categories, meat and produce, on a large sample of British consumers.

Secondly, as a method of valuing non-market goods, the choice experiment approach has proven to be a superior variant of stated preferences techniques, as it forces respondents to focus on a respondent task to make trade-offs between attributes and offers consistency with the random utility model. Moreover, biases created through the order of the questions are avoided with the help of a random design. In addition, a WTP can be estimated for each specific individual attribute. We distinguish between 6 attributes for carrots ('Organic Label', 'Environmentally Friendly', 'Quality', 'Best Before', 'Chemical Usage', and 'Price') and 7 for chicken breast ('Animal Welfare' in addition to the attributes listed for carrots) and report the WTP for each of these attributes.

Thirdly, and possibly most importantly, the present study both uses ex-ante and ex-post methods in order to address the problem of hypothetical bias and calibration in choice experiments. Correcting for hypothetical bias in general is important, as hypothetical bias is by far the strongest criticism brought to stated preferences techniques (Cummings et al., 1986; Mitchell and Carson, 1989; List and Gallet, 2001; Little and Berrens, 2004; Murphy et al., 2005; Carson and Groves, 2007; Little, Broadbent, and Berrens, 2012), but it appears to be especially important in the case of organic products, as people experience a warm glow from

[38] Conditional logit results are reported in the Appendix in the Supplementary Material (Table A6 for chicken and Table A7 for carrots) for comparison. The results show that the overall sample does seem to appreciate the organic label for both products. Moreover, the price coefficient is on average negative and significant and hence an average WTP for organic could be estimated. However, the analysis is not as refined as the latent class model where it can be determined precisely which class of people likes and which one dislikes the organic label and what determines the attendance to these classes.

giving a socially desirable answer related to the credence attributes of organic production (Lusk and Norwood, 2009; Schuldt and Schwarz, 2010; Norwood and Lusk, 2011).

For this reason the present study employs two established ex-ante methods for reducing hypothetical bias – Cheap Talk and Honesty Priming – in three different combinations/treatments and shows that, while both methods have been effective in reducing hypothetical bias to a specific degree, Cheap Talk appears to be more successful than Honesty Priming.[39] In our setting the Cheap Talk script also included a budget constraint reminder. As often done in the elicitation of stated preferences, consumers were reminded that, if they spend more on a product, they have less money left for other goods. However, for simplicity, we refer to this combination of Cheap Talk with budget constraint reminder simply as Cheap Talk. The results also show that Cheap Talk appears to work best in classes of people who do not appreciate organic products and might tend to overstate their WTP if treatments would not be in place. Using both treatments together is in general but not always more effective than using them individually, and they seem to work best for meat, which is the product that has on average a higher price and includes an animal welfare attribute.

Moreover, the present analysis divides ex-post the sample in latent classes, accounts for attribute non-attendance and a further refined division of the sample in scale classes according to their variance heterogeneity. As stressed in the literature failure to differentiate between preference and scale heterogeneity might render biased results (Louviere and Eagle, 2006). These are important methodological improvements that could advance the state of practice with respect to hypothetical bias in discrete choice experiments. To our knowledge, no previous study on organic food has performed this type of analysis, and therefore, the results are expected to be more accurate and reliable and hence better suited for policy recommendations.

Policy recommendations could relate for example to the impact of the present results on different consumer market contexts such as: agri-food with credence attributes, agri-food with different levels of quality or agri-food with small size. Welfare analysis could be employed to estimate the economic value derived from changes in the attributes analysed in the present study. Profitability estimates from developing new products with these attributes could be made after taking into account potential higher costs of production and lower yields (Seufert et al. 2012). It appears that while quality is an attribute worth developing, other organic attributes might not automatically lead to higher profitability especially for smaller sized farms.

[39] Wald-values larger for HB2 for both products in Tables III and IV.

Small-size agriculture is usually less efficient and large-scale producers have been quick to expand production to niche markets once they have become established (Crespi and Marette, 2009; Saitone et al., 2015; Saitone and Sexton, 2017). Hence, for small-sized farms, depending on the product, it might be more important to focus stronger specific attributes such as quality.

The results show that the WTP for organic attributes is on average larger for carrots than for chicken. Attributes that were also important in the choice of carrots were low chemical usage and environmentally friendly production. For both products, the attribute Quality seems to be important, since this attribute has the highest average WTP for chicken and the second highest for carrots (1.3 £/400gr for chicken and 0.27 £/kg for carrots). The results stay the same even after excluding insignificant parameters and correcting for hypothetical bias and corroborate those of Guenther et al. (2015), according to which in four different countries, including the UK, Quality was found to be highly valued among several credence food attributes, including the organic label.[40] Griffith and Nesheim (2008, 2010) also find that quality is the most valued attribute of organic products by UK consumers.[41] Therefore, even if, in the case of both products, there is a core group of people consisting of about 20-30% of the sample who appreciate the Organic label per se, there are also other attributes that consumers seem to appreciate, such as environmentally friendly production, low chemical usage, and most importantly high quality. Even if products exist that have these attributes independently of organic products, as they are also part of them, emphasizing them both in production and in retail could increase the sales of organic products and help the organic sector in the UK to grow stronger. This is especially interesting in the light of a potential Brexit and the redesign of the UK agricultural policy.

Most importantly, the results in the present study show that implementing mechanisms to correct for hypothetical bias has a significant impact and that in some cases willingness to pay values are down to 46% of what they would have been if no treatment would have been in place. This suggests the usefulness and even necessity to correct for hypothetical bias in stated preference studies.

[40] Table 3-2 on page 9 (with `Organic' being embedded in 'Environmental Condition').
[41] Tables 11 and 12 on pages 50 and 51.

References

Aadland D., and A. J. Caplan (2003). Willingness to Pay for Curbside Recycling with Detection and Mitigation of Hypothetical Bias. *American Journal of Agricultural Economics 85*: 492–502.

Adamowicz W., Louviere J., and M. Williams (1994). Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities. *Journal of Environmental Economics and Management 26*: 271-292.

Aertsens J., Verbeke W., Mondelares K., and G. van Huylenbroeck (2009). Personal determinants of organic food consumption: a review. *British Food Journal 111(10)*: 1140-1167.

Arbindra P. R, Moon W., and S. Balasubramanian (2005). Agro-biotechnology and organic food purchase in the United Kingdom. *British Food Journal 107(2)*: 84-97.

Assael H. (1998). Consumer behavior and marketing action.Cincinnati:South-Western College Publishing.

Ajzen, I. (1991). The theory of planned behaviour. *Organizational Behavior and Human Decision Processes* 50 (2): 179–211. doi:10.1016/0749-5978(91)90020-T

Balcombe K., Fraser I., Lowe B., and D. Souza Monteiro (2016). Information Customization and Food Choice. *American Journal of Agricultural Economics 98(1)*: 54-73.

Bargh J. A., and T. L. Chartrand (2000). A Practical Guide to Priming and Automaticity Research, In *Handbook of Research Methods in Social Psychology*, ed. H. Reis and C. Judd, 253-285. New York: Cambridge University Press.

Bazoche, P., Combris, P., Giraud-Héraud, E., Seabra Pinto, A., Bunte, F., and Tsakiridou, E. (2013). Willingness to pay for pesticide reduction in the EU: nothing but organic? *European Review of Agricultural Economics 41(1)*: 87-109.

Becker G.M., DeGroot M.H., and Marschak J. (1964). Measuring utility by a single-response sequential method. *Behav Sci. 9*(3): 226–32. doi:10.1002/bs.3830090304

Blackwell, R. D., Miniard, P. W., & Engle, J. F. (2005). Consumer behavior. Cincinnati: South-Western College Publishing.

Bradley, M., and Daly, A. (1994). Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data. *Transportation 21*(2): 167-184.

Briggeman, B. C., and Lusk, J. L. (2010). Preferences for fairness and equity in the food system. *European Review of Agricultural Economics 38(1)*: 1-29.

Brooks, K., & Lusk, J. L. (2010). Stated and Revealed Preferences for Organic and Cloned Milk: Combining Choice Experiment and Social Scanner Data. *American Journal of Agricultural Economics 92 (4)*: 1229–1241. doi:10.1093/ajae/aaq054.

Brown T. C., Ajzen I., and D. Hrubes (2003). Further Tests of Entreaties to Avoid Hypothetical Bias in Referendum Contingent Valuation. *Journal of Environmental Economics and Management 46*: 353–361.

Burton M., Rogers A., and C. Richert (2016). Community acceptance of biodiversity offsets: evidence from a choice experiment. *Australian Journal of Agricultural and Resource Economics 59*: 1-20.

Byrne P. J., Toensmeyer U. C., German C. L., and H. R. Muller (1991). Analysis of consumer attitudes toward organic produce and purchase likelihood. *Journal of Food Distribution Research* 22: 49-62.

Campbell, D., Hutchinson, W. G., & Scarpa, R. (2006). Lexicographic preferences in discrete choice experiments: Consequences on individual-specific willingness to pay estimates.

Campbell D., Hutchinson W. G., and R. Scarpa (2008). Incorporating discontinuous preferences into the analysis of discrete choice experiments. *Environmental and Resource Economics 41(3)*: 401-417.

Caputo, V., Nayga Jr, R. M., & Scarpa, R. (2013). Food miles or carbon emissions? Exploring labelling preference for food transport footprint with a stated choice study. *Australian Journal of Agricultural and Resource Economics 57(4)*: 465-482.

Carson R. T., and T. Groves (2007). Incentive and informational properties of preference questions. *Environmental and Resource Economics 37(1)*: 181-210.

Carlsson F., Frykblom P., and C. J. Lagerkvist (2005). Using cheap talk as a test of validity in choice experiments. *Economics Letters 89(2)*: 147-152.

Carlsson, F., and Martinsson, P. (2001). Do hypothetical and actual marginal willingness to pay differ in choice experiments?: Application to the valuation of the environment. *Journal of Environmental Economics and Management 41(2)*: 179-192.

Carlsson, F., Mørkbak, M. R., and Olsen, S. B. (2012). The first time is the hardest: A test of ordering effects in choice experiments. *Journal of Choice Modelling 5(2)*: 19-37.

Chalak A., Abiad M., and K. Balcombe (2016). Joint use of attribute importance rankings and non-attendance data in choice experiments. *European Review of Agricultural Economics*, *43(5)*: 737-760.

Champ P. A., Moore R., and R. C. Bishop (2009). A Comparison of Approaches to Mitigate Hypothetical Bias. *Agricultural and Resource Economics Review 38*: 166–180.

Chartland, T.L., Huber J., Shiv B. and Tanner, R. J. (2008). Nonconscious goals and consumer choice. *Journal of Consumer Research 35*: 189-201.

Crespi, J.M. and Marette, S. (2009). The procompetitive effect of demand-enhancing check-off programs. *American journal of agricultural economics*, *91*(2): 389-401.

Cummings R. G., Brookshire D. S., and W. D. Schulze (Eds.) (1986). Valuing Environmental Goods; A State of the Arts Assessment of the Contingent Valuation Method. Rowman and Allanheld, Totowa, NJ.

Cummings R. G., and L. O. Taylor (1999). Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method. *American Economic Review 89*: 649-665.

Darley, W.K., Blankson, C. and Luethge, D.J. (2010). Toward an integrated framework for online consumer behavior and decision making process: A review. *Psychology & marketing*, *27*(2): 94-116.

De-Magistris T., Garcia A., and R. M. Nayga, Jr. (2013). On the Use of Honesty Priming Tasks to Mitigate Hypothetical Bias in Choice Experiments. *American Journal of Agricultural Economics 95(5)*: 1136-1154.

Dewey, John (1910). How we think. Boston: D. C. Heath and Company. OCLC 194219.

Duffy, B., Smith, K., Terhanian, G., and Bremer, J. (2005). Comparing data from online and face-to-face surveys. *International Journal of Market Research 47(6)*: 615-639.

Engel, James F., Blackwell, Roger D., and Kollat, David T. Consumer Behavior. Hinsdale, Ill.: Dryden Press, 1978.

Farrell J., and M. Rabin (1996). Cheap Talk. *Journal of Economic Perspectives 10(3)*: 103-118.

Feldmann, C., and Hamm, U. (2015). Consumers' perceptions and preferences for local food: A review. *Food Quality and Preference 40*: 152-164.

Gerrard C., Janssen M., Smith L., Hamm U., and S. Padel (2013). UK consumer reactions to organic certification logos. *British Food Journal 115(5)*: 727-742.

Garrod, G. and Willis, K.G., 1999. Economic valuation of the environment. *Books*.

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

Greene W. H. (2012). Econometric Analysis: International Edition (7th ed). Pearson Education Limited.

Griffith R. and L. Nesheim (2008). Household Willingness to Pay for Organic Products, CEPR Discussion Paper No. DP6905. Available at SSRN: https://ssrn.com/abstract=1240215

Griffith R., and L. Nesheim (2010). Estimating household's willingness to pay, Cemmap Working Paper CWP24/10. Available at Cemmap: https://ssrn.com/abstract=1240215

Griffith R. and L. Nesheim (2013). Hedonic methods for baskets of goods. *Economics Letters 120*: 284-287.doi:https://doi.org/10.1016/j.econlet.2013.04.040

Gschwandtner, A. (2018). The Organic Food Premium: A Local Assessment in the UK. *International Journal of the Economics of Business 25(2)*: 313-338.

Guenther M. C., Saunders P. Rutherford D. P., and T. Driver (2015). *Maximising Export Returns (MER): Consumer attitudes towards attributes of food and beverages in export markets relevant to New Zealand*. AERU Research Report No. 336, Lincoln University, Canterbury, New Zealand.

Hammitt J. K., and D. Herrera-Araujo (2017). Peeling back the onion: Using latent class analysis to uncover heterogenous responses to stated preference surveys. *Journal of Environmental Economics and Management 75*: 41-53.

Hawkins D. I., Best R. J., and Coney K. A. (2003). Consumer behaviour: Building marketing strategy. Boston: McGraw-Hill/Irwin.

Hemmerling S., Hamm U., and A. Spiller (2015). Consumption behaviour regarding organic food from a marketing perspective - a literature review. *Organic Agriculture 5(4)*: 277-313. http://dx.doi.org/10.1007/s13165-015-0109-3

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

Hess S., Stathopoulos A., Campbell D., O'Neil V., and S. Caussade (2013). It's not that I don't care, I just don't care very much: confounding between attribute non-attendance and taste heterogeneity. *Transportation 40*: 583-607.

Horowitz, J. K., and McConnell, K. E. (2002). A review of WTA/WTP studies. *Journal of Environmental Economics and Management 44(3)*: 426-447.

Horsky, D., Nelson, P., & Posavac, S. S. (2004). Stating preference for the ethereal but choosing the concrete: how the tangibility of attributes affects attribute weighting in value elicitation and choice. *Journal of Consumer Psychology 14(1-2):* 132-140.

Hovi, M., Roderick, S. (2000). Mastitis and mastitis control strategies in organic milk. *Cattle Pract*. *8*: 259 – 264.

Hovi, M., Sundrum, A., and Thamsborg, S.M., (2003). Animal health and welfare in organic livestock production in Europe: current state and future challenges. *Livest. Prod. Sci*. *80*: 41 – 53.

Hu, W., Batte, M. T., Woods, T., and Ernst, S. (2011). Consumer preferences for local production and other value-added label claims for a processed food product. *European Review of Agricultural Economics 39(3)*: 489-510.

Jacquemet N., James A., Luchini S. and J.F. Shogren (2009). Preference elicitation under oath. Centre d'Economie de la Sorbonne (CES) Working Paper 09 (43).

Jacquemet N., James A., Luchini S. and J.F. Shogren (2010). Referendum under oath. Groupement de Recherche en Economie Quantitative d'Aix-Marseille (GREQAM) no. 2010-15.

Jacquemet N., James A., Luchini S., and J. F. Shogren (2011). Social Psychology and Environmental Economics: A New Look at Ex-Ante Corrections of Biases Preference Evaluation. *Environmental Resource Economics 48*: 413-433.

Jacquemet N., James A., Luchini S. and J.F. Shogren (2013). Preference elicitation under oath. *Journal of Environmental Economics and Management 65*: 110-132.

Janssen M., and U. Hamm (2012). Product labelling in the market for organic food: Consumer preferences and willingness-to-pay for different organic certification logos. *Food Quality and Preference 25*: 9-22.

Johnson, E. J., Häubl, G. and Keinan, A. (2007). Aspects of endowment: a query theory of value construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33(3): 461.

Kerr G. N., and B. M. H. Sharp (2009). Efficient design for willingness to pay in choice experiments: Evidence from the field, Paper presented at the New Zealand Agricultural and Resource Economics Society Conference.

Kemper N. P., Popp J. S. and R. M. Nayga (2019). A query theory account of a discrete choice experiment under oath. *European Review of Agricultural Economics,* 1-40, https://doi.org/10.1093/erae/jbz026.

Kikulwe E. M., Birol E., Wesseler J., and J. Falck-Zepeda (2011). A latent class approach to investigating demand for genetically modified banana in Uganda. *Agricultural Economics 42(5)*: 547-560.

Lagerkvist, C. J., Carlsson F., Viske D. (2006). Swedish Consumer Preferences for Animal Welfare and Biotech: A Choice Experiment.

Lagerkvist, C. J., and Hess, S. (2011). A meta-analysis of consumer willingness to pay for farm animal welfare. *European Review of Agricultural Economics 38*(1): 55-78.

Landry C. E., and J. A. List (2007). Using Ex Ante Approaches to Obtain Credible Signals for Value in Contingent Markets: Evidence from the Field. *American Journal of Agricultural Economics 89*: 420–429.

Lea E., and T. Worsley (2005). Australians' organic food beliefs, demographics and values, *British Food Journal 107(11)*: 855-69.

Liaukonyte, J., Streletskaya, N.A. and Kaiser, H.M. (2015). Noisy information signals and endogenous preferences for labeled attributes. *Journal of agricultural and resource economics* 1:179-202.

List, J.A., and C.A. Gallet (2001). What Experimental Protocol Influence Disparities between Actual and Hypothetical Stated Values? *Environmental and Resource Economics 20*: 241–54.

Little, J., and R. Berrens, (2004). Explaining Disparities between Actual and Hypothetical Stated Values: Further Investigation Using Meta-Analysis. *Economics Bulletin 3*: 1–13.

Little, J., C.D. Broadbent, and R.P. Berrens, (2012). Meta-Analysis of the Probability of Disparity between Actual and Hypothetical Valuation Responses: Extension and Preliminary New Results. *Western Economics Forum 11*: 1-12.

Loomis, J. B. (2014). Strategies for Overcoming Hypothetical Bias in Stated Preference Surveys. *Journal of Agricultural and Resource Economics 39(1)*: 34-46.

Loureiro M., and S. Hine (2002). A comparison of consumer willingness to pay for a local (Colorado-grown), organic, and GMO-free product. *Journal of Agricultural & Applied Economics 34 (3)*: 477-87.

Loureiro M., and W. J. Umberger (2007). A choice experiment model for beef: What US consumer responses tell us about relative preferences for food safety, country-of-origin labelling and traceability. *Food Policy 32*: 496-514.

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

Lusk J. L., and F. B. Norwood (2009). Bridging the gap between laboratory experiments and naturally occurring markets: An inferred valuation method. *Journal of Environmental Economics and Management 58*: 236-250. doi:10.1016/j.jeem.2008.12.003.

Mahieu, P. A. (2010). Does gender matter when using cheap talk in contingent valuation studies. *Economics Bulletin 30(4)*: 2955-2961.

Mariel P., Hoyos, D., and J. Meyerhoff (2013). Stated or inferred attribute non-attendance? A simulation approach. Economía Agraria y Recursos Naturales (*Agricultural and Resource Economics*) *13(1)*: 51-67.

Maxwell, S., Nye P., and N. Maxwell (1999). Less pain, same gain: The effects of priming fairness in price negotiations. *Psycology & Maketing 16*: 545-562.

McEachern, M. G., and P. McClean (2002). Organic purchasing motivations and attitudes: are they ethical? *International Journal of Consumer Studies 26(2)*: 85-92.

McEachern, M. G., and J. Willock (2004). Producers and consumers of organic meat: a focus on attitudes and motivation. *British Food Journal 106(7)*: 534-52.

Meas, T., Hu, W., Batte, M. T., Woods, T. A., and Ernst, S, (2014). Substitutes or complements? Consumer preference for local and organic food attributes. *American Journal of Agricultural Economics 97(4)*: 1044-1071.

Meemken, E.M. and Qaim, M. (2018). Organic agriculture, food security, and the environment. *Annual Review of Resource Economics*, *10*: 39-63.

Miller, K.M., Hofstetter, R., Krohmer, H. and Zhang, Z.J. (2011). How should consumers' willingness to pay be measured? An empirical comparison of state-of-the-art approaches. *Journal of Marketing Research*, *48*(1): 172-184.

Mitchell Robert, C. and Carson, R.T. (1989). Using surveys to value public goods: The contingent valuation method. *Washington: Resources for the Future*.

Moser, R., Raffaelli, R., and Thilmany, D. D. (2011). Consumer preferences for fruit and vegetables with credence-based attributes: a review. *International Food and Agribusiness Management Review 14*(1030-2016-82774): 121-142.

Mozumder, P., and Berrens, R. P. (2007). Investigating hypothetical bias: induced-value tests of the referendum voting mechanism with uncertainty. *Applied Economics Letters 14(10)*: 705-709.

Murphy J., Allen P. G., Stevens T. H., and D. Weatherhead (2005). A Meta-Analysis of Hypothetical Bias in Stated Preference Valuation. *Environmental and Resource Economics*, *30*: 313-325.

Nguyen, T. T., Haider, W., Solgaard, H. S., Ravn-Jonsen, L., and Roth, E. (2015). Consumer willingness to pay for quality attributes of fresh seafood: A labeled latent class model. *Food Quality and Preference 41*: 225-236.

Norwood, F. B., and Lusk. J. L., (2011). Social Desirability Bias in Real, Hypothetical and Inferred Valuation Experiments. *American Journal of Agricultural Economics 93(2)*: 528–534. doi:10.1093/ajae/aaq142.

Penn, J. M., & Hu, W. (2018). Understanding hypothetical bias: An enhanced meta-analysis. *American Journal of Agricultural Economics 100*(4): 1186-1206.

Radman M. (2005). Consumer consumption and perception of organic products in Croatia. *British Food Journal 107* (4-5): 263-73.

Rousseau, S., and Vranken, L. (2013). Green market expansion by reducing information asymmetries: Evidence for labeled organic food products. *Food Policy 40*: 31-43.

Saitone, T. L., Sexton, R. J. and Sumner, D. A. (2015). What happens when food marketers require restrictive farming practices? *American Journal of Agricultural Economics* 97(4): 1021–1043.

Saitone, T. L., and Sexton, R. J. (2017). Agri-food supply chain: evolution and performance with conflicting consumer and societal demands. *European Review of Agricultural Economics*, *44*(4): 634-657.

Scalco, A., Noventa, S., Sartori, R., and Ceschi, A. (2017). Predicting organic food consumption: A meta-analytic structural equation model based on the theory of planned behavior. *Appetite 112*: 235-248.

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

Schuldt, J. P., and N. Schwarz (2010). The 'organic' path to obesity? Organic claims influence calorie judgments and exercise recommendations. *Judgment and Decision Making* 5(3): 144-150.

Seufert, V., Ramankutty, N. and Foley, J. A. (2012). Comparing the yields of organic and conventional agriculture. *Nature* 485: 229–232.

Stobbelaar, D.J., Casimir, G., Borghuis, J., Marks, I., Meijer, L. and Zebeda, S. (2007). Adolescents' attitudes towards organic food: a survey of 15- to 16-year old school children. *International Journal of Consumer Studies 31*: 349-56.

Sundrum, A. (2001). Organic livestock farming: a critical review. *Livestock Production Science 67*(3): 207-215.

Thompson, S. K., (1987). Sample Size for Estimating Multinomial Proportions. *The American Statistician 41*(1): 42-46.

Thompson, G.D. and Kidwell, J. (1998). Explaining the choice of organic produce: cosmetic defects, prices, and consumer preferences. *American Journal of Agriculture Economics 80*(2): 277-87.

Tiffin R. and Arnoult M. (2010). The demand for a healthy diet: estimating the almost ideal demand system with infrequency of purchase. *European Review of Agricultural Economics 37*(4): 501-521.

Tonsor, G.T., and R.S. Shupp (2011). Cheap Talk Scripts Online Choice Experiment: Looking beyond the Mean. *American Journal of Agricultural Economics 93*(4): 1015-1031.

Vittersø, G., & Tangeland, T. (2015). The role of consumers in transitions towards sustainable food consumption. The case of organic food in Norway. *Journal of Cleaner Production 92*: 91-99.

Von Borell, E., and Sørensen, J. T. (2004). Organic livestock production in Europe: aims, rules and trends with special emphasis on animal health and welfare. *Livestock Production Science 90*(1): 3-9.

Weaver, R., and Prelec D. (2013). Creating truth-telling incentives with the Bayesian truth serum. *Journal of Marketing Research 50*(3): 289-302.

Waldrop, M. E., McCluskey, J. J., & Mittelhammer, R. C. (2017). Products with multiple certifications: insights from the US wine market. *European Review of Agricultural Economics 44*(4): 658-682.

Wier, M., O'Doherty Jensen K., Morch Andersen L. and K. Millock (2008). The character of demand in mature organic food markets: Great Britain and Denmark compared. *Food Policy*, *33*: 406-421.

Yin, S., Chen, M., Chen, Y., Xu, Y., Zou, Z., & Wang, Y. (2016). Consumer trust in organic milk of different brands: the role of Chinese organic label. *British Food Journal 118*(7): 1769-1782.

Yiridoe, E. K., S. Bonti-Ankomah, and R. C. Martin (2005). Comparison of consumer perceptions and preference toward organic versus conventionally produced foods: A review and update of the literature. *Renewable Agriculture and Food Systems 20*(4): 193-205. doi: 10.1079/RAF2005113

Zander, K. and U. Hamm (2010). Consumer Preferences for additional ethical attributes of organic food. *Food Quality and Preference 21*: 495-503.

Zander, K., Padel, S., and Zanoli, R. (2015). EU organic logo and its perception by consumers. *British Food Journal 117*(5): 1506-1526.

Figures and Tables

Table I: Attributes

Attribute	Description	Image	coding
		EU UK	
Label	Organic label		1
	Conventional label	No label	0
Chemical Usage in Production (i.e. antibiotics for animals and artificial pesticides for carrots)	Average		0
	Low		1
Environmentally Friendly	Average	No label	0
	High		1
Animal friendly (for chicken only)	No Freedom Food	No label	0
	Freedom food		1
Quality	Average		0
	High		1
Best Before	Soon (<1 week)	Hurry up!	0
	1 week or longer	You can use this product 1 week or longer	1
Price (£) of Chicken Breast 400 Gramm (0.88lbs)	3.00, 3.50, 5.75, 6.64, 8.32		cardinal
Price (£) of Carrots 1kg (2.2lbs)	0.53, 0.75, 1.20, 1.33, 1.54		cardinal

Table II: Descriptive Statistics

Variable	***Mean***	***Median***	***Mode***	***Min***	***Max***	***St Dev***
Age	50.38	52	59	18	80	15.61
Income (£/month)	1524.95	1250	1250	1081.42	4750	1524.95
Education	3.78	4	3	1	8	1.59
Children*	1.73	2	1	1	6	0.88

*among households with children

Table III: Scale Adjusted Latent Class Models for *Chicken* with ANA and covariates, 5 classes, 2 scales

Attributes	Class1	z	Class2	z	Class3	z	Class4	z	Class5	z	Wald(=)
Organic Label	0	.	0.793***	2.930	-17.90***	-4.184	0	.	0	.	26.23
			(0.2708)		(4.2787)						
Environmentally Friendly	0.790***	6.054	0.475***	4.317	0	.	0	.	0	.	69.90
	(0.1305)		(0.1101)								
Quality	1.506***	6.671	0.518***	3.388	1.527***	3.146	1.126***	3.643	1.962*	1.857	14.02
	(0.2257)		(0.1529)		(0.4853)		(0.3091)		(1.0565)		
Best Before	0	.	0.185	1.731	0.949**	1.923.	0	.	1.379	1.553	6.94
			(0.1066)		(0.4937)				(0.8878)		
Chemical Usage	-0.416*	-1.884	0		-16.02***	-3.66	0	.	0	.	17.70
	(0.221)				(4.3764)						
Animal Welfare	0	.	0.717***	3.322	-15.98***	-3.426	0	.	0	.	22.38
			(0.2158)		(4.664)						
SQ	-3.56***	-9.377	-0.869**	-1.94	-73.7***	-4.000	-5.29***	-7.174	7.664***	3.549	69.56
	(0.3796)		(0.4473)		(18.4099)		(0.7369)		(2.1593)		
Price	-0.291***	-4.433	-0.000	-0.001	-2.54***	-2.890	-0.78***	-7.168	0.382**	1.939	51.19
	(0.0656)		(0.0728)		(0.8822)		(0.1091)		(0.1973)		
Price X HB2	-0.065	-0.899	-0.111	-1.342	-3.326***	-2.871	-0.35***	-3.670	-0.583	-1.053	66.85
	(0.0726)		(0.0827)		(1.1588)		(0.0944)		(0.5534)		
Price X HB3	-0.074	-1.038	0.062	0.974	-2.725***	-2.236	0.007	0.067	-0.050	-0.321	62.70
	(0.0714)		(0.0635)		(1.2186)		(0.1043)		(0.1568)		
Price X HB1	-0.120	-1.699	0.116	1.404	-3.024***	-2.604	-0.34***	-2.871	0.044	0.366	84.60
	(0.0706)		(0.0826)		(1.1612)		(0.1189)		(0.1194)		
Scale Model											
sClass1	0	.									
sClass2	-2.657***	-8.67									
	(0.3065)										

(continued)

Table III *(continued)*

Attributes	Class1	z	Class2	z	Class3	z	Class4	z	Class5	z	Wald(=)
Class Membership Model											
	(1.0156)	(1.9654)			(1.3882)		(3.7456)		(1.061)		
Age	-0.018**	-2.13	-0.030***	-3.002	0.002	0.274	0.030***	2.293	0.016	1.071	14.80
	(0.0084)		(0.0101)		(0.0083)		(0.0132)		(0.0147)		
Income	-0.000	-1.30	0.0001	0.400	-0.0003***	-2.669	0.000	0.907	0.0003**	1.964	10.39
	(0.0001)		(0.0001)		(0.0001)		(0.0002)		(0.0002)		
Pro-Organic	0.016	1.269	0.078***	4.144	-0.02	-1.641	-0.05***	-3.091	-0.021	-0.948	21.43
	(0.0128)		(0.0187)		(0.0122)		(0.0172)		(0.0022)		
Intercept	0.9331***	4.103	0.110	0.388	0.296	1.297	-0.169	-0.586	-1.170***	-3.735	61.77
	(0.2274)		(0.2844)		(0.2277)		(0.2882)		(0.3132)		
Class Size	0.33		0.27		0.23		0.12		0.05		

Notes: 449 cases and 3592 replications. Standard errors in parentheses.

Table IV: Scale Adjusted Latent Class Models for *Carrots* with ANA and Age, Income and Pro-Organic as covariates, 5 classes, 2 scales

Attributes	Class1	z	Class2	z	Class3	z	Class4	z	Class5	z	Wald (=)
Organic Label	-1.111***	-2.53	0	.	11.423***	2.821	0	.	2.578**	1.92	12.03
	(0.4398)				(4.0486)				(1.3436)		
Environ. Friendly	-1.158***	-2.50	0	.	2.677***	2.586	0	.	0	.	11.63
	(0.4623)				(1.0351)						
Quality	0	.	0	.	4.661***	3.730	0	.	0	.	13.91
					(1.2497)						
Best Before	1.312***	4.37	0.798	1.827	0	.	0	.	0	.	20.57
	(0.3005)		(0.4365)								
Chemical Usage	-1.134***	-2.54	0	.	3.621***	3.004	0	.	0	.	14.80
	(0.4471)				(1.2053)						
SQ	-22.58***	-3.62	-55.78***	-3.838	0	.	-52.23***	-3.51	4.160***	2.35	17.51
	(6.2409)		(14.5325)				(14.8805)		(1.7707)		
Price	1.286**	2.66	-27.7***	-3.627	-2.422***	-2.369	-38.922***	-3.47	-2.158	-1.16	19.64
	(0.4827)		(7.6265)		(1.0223)		(11.2023)		(1.8662)		
Price X HB2	-0.121	-0.23	1.831	0.534	0.375	0.525	-5.48***	-2.49	1.215	0.53	7.08
	(0.5325)		(3.4283)		(0.7141)		(2.1955)		(2.3106)		
Price X HB3	0.386	0.90	1.143	0.335	5.600***	2.153	-3.059	-0.99	1.335	0.66	5.11
	(0.4301)		(3.407)		(2.6017)		(3.0783)		(2.0318)		
Price X HB1	0.432	0.95	2.309	0.497	0.4973	0.582	-6.020***	-2.04	2.523	1.48	6.05
	(0.4543)		(3.8382)		(0.8553)		(2.9511)		(1.7104)		
Scale Model											
sClass1	0	.									
sClass2	-1.994***	-8.93									
	(0.2232)										

(continued)

Table IV *(continued)*

Attributes	Class1	z	Class2	z	Class3	z	Class4	z	Class5	z	Wald (=)
Class Membership Model											
BuyOrg	0.644 (0.9695)	0.67	0.196 (0.9786)	0.201	1.781 (1.0213)	1.743	-3.60 (3.7491)	-0.96	0.983 (1.0578)	0.93	11.30
Age	-0.027*** (0.0075)	-3.65	0.004 (0.007)	0.560	-0.030*** (0.0128)	-2.370	0.03*** (0.0121)	2.18	0.028* (0.0142)	1.94	23.89
Income	0.0001 (0.0001)	0.00	-0.0002* (0.0001)	-1.900	-0.0003 (0.0002)	-1.554	0.000 (0.0001)	0.82	0.0003 (0.0002)	1.73	8.86
Pro-organic	0.015 (0.0119)	1.25	-0.010 (0.095)	-0.084	0.084*** (0.0226)	3.709	-0.05*** (0.014)	-3.77	-0.036* (0.018)	-1.99	24.74
Intercept	1.220*** (0.1864)	6.54	1.118*** (0.1836)	6.089	-1.335*** (0.436)	-3.062	-0.230 (0.2729)	-1.10	-0.703*** (0.3106)	-2.27	66.20
Class Size	0.36		0.38		0.11		0.09		0.06		

Notes: 449 cases and 3592 replications. Standard errors in parentheses.

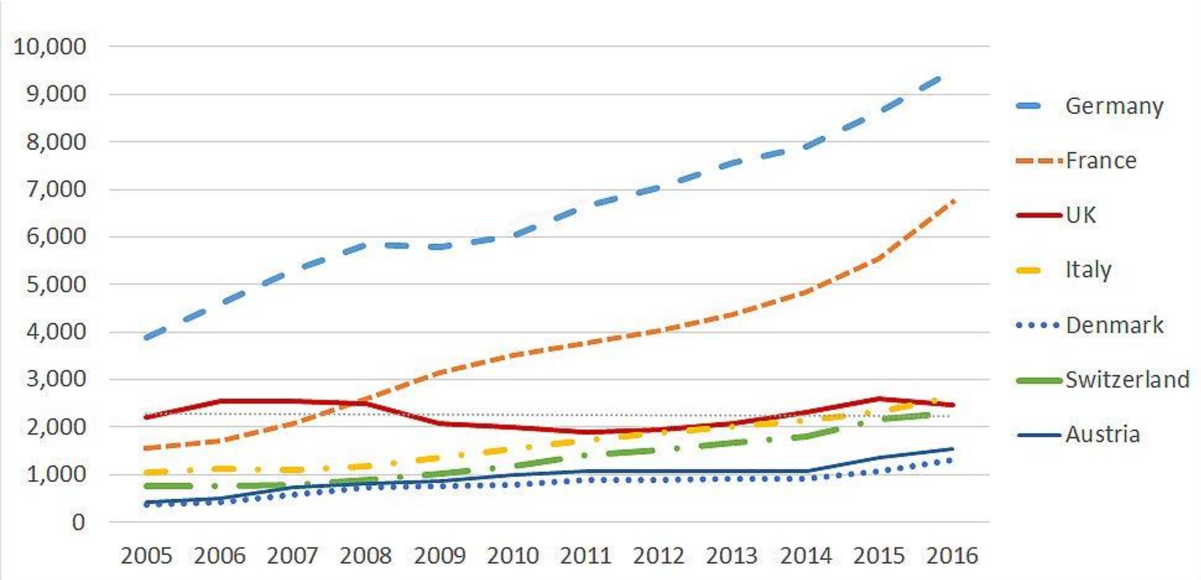
Table V: WTP (£/400gr) by attribute and class for hypothetical bias treatment HB1 for Chicken ANA-LCM

Attributes	***Class1***	***Class3***	***Class4***	***Weighted Average***
Class size (%)	0.33	0.23	0.12	
Organic Label	0.00	-3.21	0.00	-0.72
Environmentally Friendly	1.94	0.00	0.00	0.64
Quality	3.67	0.27	1.01	1.30
Best Before	0.00	0.17	0.00	0.04
Chemical Usage	-1.02	-2.88	0.00	-0.99
Animal Welfare	0.00	-2.88	0.00	-0.65
% Changes				
HB1/HB2	0.87	1.05	1.01	
HB1/HB3	0.89	0.94	0.70	
HB1/HB4	0.71	0.46	0.70	

Table VI: WTP by attribute, class and hypothetical bias treatment HB1 for Carrots in £/kg ANA-LCM

Attributes	***Class2***	***Class3***	***Weighted Average***
Class size (%)	0.38	0.11	
Organic Label	0.00	5.94	0.65
Environmentally Friendly	0.00	1.39	0.15
Quality	0.00	2.42	0.27
Best Before	0.03	0.00	0.01
Chemical Usage	0.00	1.88	0.21
% Changes			
HB1/HB2	1.02	1.06	
HB1/HB3	1.05	-1.65	
HB1/HB4	1.96	1.26	

Figure 1 Development of the Organic Market in Selected European Countries 2005-2016 (Sales in Million €)



Source: Organic Data Network