University of Kent

School of Economics Discussion Papers

Combining Stated and Revealed Preferences for valuing Organic Chicken Meat

Adelina Gschwandtner Jose Eduardo Ribeiro

Cesar Revoredo-Giha Michael Burton

December 2021

KDPE 2113

Combining Stated and Revealed Preferences for valuing Organic Chicken Meat

By Adelina Gschwandtner, Jose Eduardo Ribeiro, Cesar Revoredo-Giha and Michael Burton[*]

Abstract

The present paper uses a jont stated preference (SP) and revealed preference (RP) model in order to estimate the willingness to pay (WTP) for the organic attribute among other key environmental attributes in chicken meat. The stated preference model is based on the respondent's choice from hypothetical choice sets in a choice experiment. The revealed preference model is using a comprehensive data-set of scanned supermarket shopping's to model the choice for chicken meat in a similar manner. The attributes in the stated preference model are based on the ranges of the actual levels of attributes found in supermarket and are presented to respondents using a fractional factorial design. The joint SP-RP approach takes advantage of the benefits of both approaches and addresses econometric issues and biases from both. The results show that the two models appear to reflect similar underlying preferences and can be meaningfully combined. Furthermore, the results show that when combining the RP and SP information, the consumers appear to be willing to pay a larger amount for the organic attribute in chicken meat than when the SP and RP approach's are applied separately. The paper contributes to the literature by being the first to estimate the WTP for organic chicken using a joint estimation approach.

Keywords: Choice Experiments, Revealed Preferences, Joint Estimation
JEL Classification: C25, Q18, Q51

[*]The first two authors are from the School of Economics at the University of Kent Canterbury/UK. The third author is from SRUC/Edinburgh/UK. The fourth author is from the University of Western Australia. The first author is the corresponding author: a.gschwandtner@kent.ac.uk

1 Introduction

The present study analyzes the WTP for the organic and other environmental attributes in chicken meat using a joint stated and revealed preference setting. Estimating the marginal utility provided by the attributes of chicken meat offers essential information to stakeholders and policymakers to identify and quantify the welfare impact of potential intervention points. These can go from narrowing potential information gap about consumer preferences towards environmental attributes (e.g. health, animal welfare and environmental friendliness), and use of appropriate labelling, to identification of the true benefits of production and consumption of conventional and organic food products. This is thus an opportunity to better educate consumers, or to inform policy makers about the impact on welfare from alternative fiscal policies, e.g. subsidies to local, environmentally friendly, low chemical usage production processes, or taxation of socially undesirable attributes.

Environmental valuation methods have been traditionally categorized as stated (direct) and indirect (revealed). Indirect methods such as the hedonic pricing method (HPM) or the travel cost method (TCM), use actual choices made by consumers in real market transaction to develop models of choice. They are best suited to estimate the use value[1] that consumers derive from a good or service or their attributes. They are called indirect because they typically use the costs that consumers are incurring for the consumption of the good or service as a mean to reveal/infer their preferences. Stated preference methods, such as the contingent valuation method (CVM) or choice experiments (CE) ask directly the consumers about their preferences which are then typically translated into their willingness to pay (WTP). Here the consumers do not actually make any behavioural changes, they only state that they would behave in a specific manner. They are best suited to estimate non-use values[2] for which a market does not actually exist. Because the organic attribute contains both use-values such as perceived better taste and better health and non-use values such as perceived environmental friendliness and higher animal welfare, it is important to use both methods when estimating the WTP for it.

Both methods have their advantages and disadvantages. The biggest disadvantage of stated preference techniques is the hypothetical bias.[3] Because respondents are in a hypothetical situation they might over- or understate their preferences depending on the desired answer. When consumers are confronted with a socially desirable good, they might over-state their WTP. This is sometimes referred to as 'Social Desirability Bias', 'Warm Glow' or simply 'Hypothetical Gap'. When consumers are confronted with a good that they will or might have to pay for in future, they might act strategically and under-stated their preferences. This is sometimes referred to as 'Strategic Bias'. When consumers are asked to recall what they have paid in the past, they might not recall correctly which is referred to as 'Recall Bias'. All these biases derived from the hypothetical nature of the experiment can be summarized under the term 'Hypothetical Bias'. Even though there are several methods to deal with it and choice experiments are designed to reduce it by giving consumers only specific choices to chose from, hypothetical bias remains until today the strongest criticism brought to stated preferences techniques (Cummings, 1986; Cummings and Taylor, 1999; Hausman, 2012; J. Whitehead, T. Haab, and J.-c. Huang, 2012; Mitchell and Carson, 2013; Penn and Hu, 2018; Gschwandtner and Burton, 2020).

However, SP currently provide the only viable alternative for measuring non-use values[4] and they are the most common method used to elicit values in cases in which the environmental quality

[1]Use value are benefits which are derived from the direct consumption of a good. Examples from environmental valuation are the direct benefits from fishing or logging.

[2]Non-use values are unrelated to the use of a good and arise often from the pure existence of the resource.

[3]Other types of biases such as 'anchoring bias', 'sampling bias', 'selection bias', 'response bias', 'non-response bias', 'survivor bias', bias', 'acquiescence bias', 'question order bias', 'answer order bias' won't be discussed here but are also common in SP (List and Gallet, 2001).

[4]Such as existence, option, bequest, vicarious and altruistic values.

change involves a large number of attribute changes (Adamowicz, J. Louviere, and Williams, 1994). Moreover, they are the only one able to estimate the WTP for goods, products and combinations of attributes that my not exist in reality yet (J. C. Whitehead and Lew, 2020). Therefore, they are especially suited for new policies or products beyond the limited range of historical experience (J. C. Whitehead, Pattanayak, et al., 2008). Stated preference methods are also in general flexible, fully in the control of the researcher and can be easy to obtain. Most importantly, RP methods suffer from collinearity between attributes and it is often impossible to isolate a specific factor that affects choice while using them. However, it is precisely this isolation of a specific attribute that is required when performing welfare analysis. For example an organic food product might include a higher quality, environmentally friendliness, low chemical usage in production and higher animal welfare, but the researcher might be interested in separating all these effect which might be impossible with a RP approach. The stated preference application we present here does exactly that and can be considered a type of contingent behavior analysis where we ask individuals to respond to a set of questions which require them to choose one alternative from three options. In each question, the individual is faced with choices of chicken meat products containing various attributes such as the ones mentioned above, including organic plus a non-buy alternative. The choice made by the individual indicates a preference for the attribute of one alternative over the other making the process consistent with random utility theory.

RP techniques are more accurate and do not suffer from hypothetical bias as they are based on actual choices in a market situation. Therefore they produce more valid estimates of WTP. They are generally regarded as rich in information of the chosen alternative and usually contain a wealth of data. However, they are more difficult to obtain and the data collection process is more problematic/expensive. One of their biggest disadvantages however, is that they are mostly truncated at 1 and that it is difficult to built in the non-buy alternative. As described by D. A. Hensher (2010) this is maybe their biggest disadvantage that they are problematic on the attributes describing the non-chosen alternative as typically only the purchased alternative is observed. There are also other disadvantages of RP such as for example that they depend on market equilibrium assumption that might not hold in reality. HPM assumes perfect competition in the housing market for example, where the method has started and still has its most applications, which is often not given due to mobility costs and asymmetric information. HPM is also riddled with the issues of heteroskedsticity and multicollinearity between attributes which lead to insignificant coefficients and make it difficult to estimate the changes in the variables of interest. Moreover, as pointed out in his seminal work by Rosen (1974), HPM is faced with endogeneity derived from the simultaneity of supply and demand equilibrium or marginal price and quantity of attributes, something that is rarely addressed in the literature (Wooldridge, 1996; Combris, Lecocq, and Visser, 1997; Bishop and Timmins, 2011; Gopalakrishnan et al., 2011; Ribeiro, Gschwandtner, and Revoredo-Giha, 2021). In this paper we use revealed preferences from a large sample of scanned food shopping's, hence actual market transactions happened in supermarkets in the UK. In the revealed preference case the individuals choose one alternative from all food alternatives available to them. The model explains the choice of food as a function of attributes which include the organic attribute. Since both the SP and the RP models reflect the same process of making food choices based on attributes it is possible to combine them in a joint analysis. The RP scanner data comes from Kantar UK Panel (sampled from just below 27,000 households) in 2016 and the SP comes from a choice experiment carried out in the same year. The scanned data is used to construct the set of alternatives individuals are faced when choosing their purchases in a consistent way with the choice experiment. Both samples are representative for the UK consumer population. Based on the utility maximisation assumption and the random utility treatment, the rule for both SP and RP choice for each individual follows the same probability rule in their discrete decisions. Results show that individual preferences for key attributes across the two datasets are not significantly different when individuals characteristics are controlled for.

Combining the two methods makes it possible to take advantage of the strengths and reduce the weaknesses of both methods. RP data grounds the hypothetical choices of SP data with real choices addressing the strongest concern related to the SP method. RP can be used to detect and mitigate hypothetical bias but also to validate SP results and test for convergent validity (Gschwandtner, 2018). Convergent validity exists if two methods that measure WTP yield measures that are not statistically different (J. C. Whitehead, Pattanayak, et al., 2008). Hence combining the two methods can validate both types of data. The strength of the RP methods are the weaknesses of the SP methods and the other way round. For example SP can help with the multicollinearity issue in RP and with filtering out the impact of a specific attribute. Von Haefen and Phaneuf (2008) have shown how combining SP and RP data can break the multicollinearity and endogeneity problems present in RP data. In the present study multicollinearity is minimised through the wider range of attributes and greater freedom through their randomisation in the choice experiment. The joining of the SP and RP also avoids the simultaneity caused when price is set as a dependent variable, i.e. simultaneously defined by supply and demand equations. Recognizing the complementarity of the two methods has led to the so called 'joint estimation paradigm' (Cameron, 1992; Adamowicz, J. Louviere, and Williams, 1994). At the same time the combination of the two methods is sometimes addressed as 'the data enrichment paradigm' as it adds the two types of data together leading to more observations with more robust results. Combining SP with RP allows us to exploit the strengths of each method whilst minimizing their relative weaknesses and helps to understand how participation and market size will change when new products are introduced or when environmental policy will change (J. J. Louviere, D. A. Hensher, and J. D. Swait, 2000; J. C. Whitehead, Pattanayak, et al., 2008).

To mitigate the strong assumption of independence of irrelevant alternatives, the estimations include interaction terms between attributes and consumer characteristics as source of heterogeneity within the population. For the joint estimation, this study use a heteroscedastic conditional logit model (CLHet), which addresses heterogeneity across the SP and RP samples. At the same time the heteroscedastic model accommodates scale differences between the SP and RP data in the joint estimation. The present study contributes to the literature by being one of the few joint estimations applications in environmental economics and to our knowledge, the first one with respect to organic food. By this it not only helps to address the hypothetical bias problem present in stated preferences related to socially desirable goods such as organic, but it also helps mitigating issues related to multicollinearity and endogeneity present in revealed preferences. Most importantly it manages to estimate in a robust manner the WTP for environmental attributes related to organic and their combination that might not exist yet in reality. The joint estimation approach is especially important in relationship to organic products which contain not only use-values but also non-use values. By this it helps to inform organic producers, supermarkets and other retailers and policy makers about the WTP of UK consumers for 'environmental attributes' such as organic in meat products at a crossroads point where the UK is redesigning its agricultural policy after leaving the EU. Moreover, it analyzes in a consistent manner the WTP for food product attributes that address health, environmental and animal welfare issues at a time when sustainability and environmental protection are at the top of the policy agenda worldwide (Assembly, 2015; Ban, 2016).

The paper proceeds as follows. In the next section we are going to give an overview of the joint estimation literature in environmental valuation to the present day. Then we are going to present the econometric models used in the study. Afterwards we are going to describe our two interesting data-sets: the stated and the revealed. In the fourth section we are going to present the joint estimation results and compare them with the stated and revealed individual ones. In this section we are also going to present willingness to pay estimates that can be used for policy evaluations. The fifth section concludes the paper with a short discussion.

2 Joint estimation literature

This section presents an overview over joint SP and RP literature in environmental valuation. While there have been a number of RP and SP joint estimation applications in other literature's, notably in marketing and transportation where joint estimation is rooted (Ben-Akiva and Morikawa, 1990; D. A. Hensher and Bradley, 1993; Ben-Akiva, Bradley, et al., 1994; J. Swait, J. J. Louviere, and Williams, 1994; D. Hensher, J. Louviere, and J. Swait, 1998; Feit, Beltramo, and Feinberg, 2010; Ellickson, Lovett, and Ranjan, 2019) and there have been quite a few RP and SP separate estimations to test for validity (Lew and Larson, 2011; Lew and Larson, 2012; Larson and Lew, 2013; Griffith and Nesheim, 2013; Lew and Larson, 2014; Gschwandtner, 2018) there are yet relatively few applications of joint estimation in environmental valuation.

Cameron (1992) is probably the first application of joint estimation in environmental economics. She combined responses from a sated Contingent Valuation (CV) survey with the ones from a revealed Travel Cost (TC) survey in order to calculate welfare estimates from recreational fishing, a domain where many applications of joint estimation would follow. Typically, the value of recreational fishing has been estimated using TC. However, travel cost information is usually limited to preferences of current users.[5] Contingent Valuation responses can extend to preferences of non-users and can help estimated measures of non-use demand which could be especially important in the case of recreational fishing and in environmental issues in general. The study shows how the two types of information can be combined in a joint model to produce a more comprehensive picture of preferences than what would be available from either information source used separately. It finds that consumers would need to be compensated approx. $3500 for a total loss of access to the fishing site. Five years later Kling (1997) performed simulation experiments using a variation of the model suggested by Cameron (1992) to confirm the gains in increased precision and reduced bias from combining contingent valuation (SP) and travel cost (RP) data especially in small samples.

Adamowicz, J. Louviere, and Williams (1994) joined SP and RP data for water-base recreational resources that included not only fishing but also swimming, the presence of a beach and other facilities for running and standing water bodies such as streams and lakes. The study is the first in environmental valuation to join data from a choice experiment with data from revealed preferences choice sets and to estimate the relative scale parameter between the two types of data. It shows that while the SP and RP models when estimated independently appear to reflect different preferences they are in fact similar and can be meaningfully joined. Even thought the main focus of the study is on the econometric estimation, it may be worth mentioning that one of the parameters that contributes strongest to the choice probability is having a fully serviced campsite and that a 10% increase in the fishing catch rate would result in an average increase in the benefit per trip of $1.74 in the RP model, $0.11 in the SP model and $0.43 in the joint model. Three years later Adamowicz, J. Swait, et al. (1997) replicate the finding of RP-SP parameter equality, once variance heterogeneity is accounted for, and show that joint RP-SP models are superior to RP models alone. This later study also collects 'perception data' with respect to recreational moose hunting sites attributes but concludes that in the specific case the costs of collecting the data outweigh the benefits of slightly improving the fit of the models. Again, even though the focus is on the econometric specification it may be worth mentioning that in this case the joint estimation models result in welfare measures from increases in moose population which are 2-6 times larger than in the SP model alone.

An especially interesting application of combining revealed and stated preferences is the one by Von Haefen and Phaneuf (2008) because it not only shows how joining the two types of data can help to overcome difficulties associated with small choice sets and multicollinearity but it also shows how

[5] Zonal travel cost method can make some inferences about non-users from the same zone but they assume a representative consumer which in reality might not exist.

it can break endogeneity. The authors revisit the SP and RP moose hunting data sets in order to show that by fusing them, they can circumvent the limitations associated with RP two-step estimators that require large choice sets, variation in the observed attributes, and instruments for endogenous attributes. However they fail to replicate the tests for parameter consistency across the RP and SP data found in previous studies. One interesting result from this study related to the increased moose population scenarios is that these are qualitatively different whether alternative specific constants (ASC) are included or not. The inclusion of ASCs seems to result in much larger point estimates ($61.02 versus $2.99), but the larger estimate also has a substantially larger standard error (20.6 versus 2.37). Five years later, Phaneuf and co-authors combine revealed hedonic pricing data from real estate transactions with stated preferences from a choice experiment via generalized method of moments (GMM) estimation (Phaneuf, Taylor, and Braden, 2013). They apply the joint estimation to value remediation of a lake contaminated with toxic chemicals known to cause cancer and neurological defects in humans in Buffalo, New York, and find evidence in support of estimates arising from their approach. The results show that the value of the properties increase as they are "moving" further from the offending site and that welfare estimates when using the joint SP/RP models 'are more than twice as large as the comparable estimate from the SP model'.

Abildtrup et al (2015) use a joint estimation approach to e analyse the determinants of the recreational value of forests in Lorraine France. Traditionally, this has been analyzed using travel costs (Willis and Garrod, 1991; King and Fraser, 2013) however, in the case of Lorraine, the easy access to a complex of forest poses a special challenge to valuation. Firstly, there are many alternative forests to choose between for a potential visitor and it is not easy to identify the specific forest to be evaluated since there are more than 5000 forest recreation units. Respondents do not necessarily know the name of the forest they visited, and most forests in Lorraine do not have individual names. To identify the forest visited, respondents were asked to identify the forest by clicking on it on an integrated and interactive map showing a satellite image of the Lorraine area in a web survey. Second, the easy access to forests in this region implies that a large share of the visitors either walk or bicycle to a forest, making it difficult to identify their travel costs. The authors apply an error-component mixed-logit model to simultaneously model the travel mode decision and the site selection decision and to combine revealed and stated preference data. The study uses both advantages of SP and RP data in order to estimate the effect on the WTP of changes in forest quality and access to it. One interesting result from this study appears to be that the WTP for an improvement of the forests such as making one hiking trail in each of the six forests appears to be higher for people walking (€0.28) than for people going by car (€0.11). The authors conjecture that this is so because the substitute forests with a hiking trail are more expensive to visit for people on foot than for people who have decided to go by car since these forests are further away from the town.

Probably the author that has written most about joint estimation is John C. Whitehead. Out of his very prolific work on the topic only few applications will be mentioned here. In an application to coastal erosion management in North Carolina, together with co-authors he estimates the benefits of increased beach width using a joint SP and RP approach (J. C. Whitehead, Dumas, et al., 2008). Two years later together with co-authors he compares the joint model with two other models that employ multiple site data: a count data demand system model and the Kuhn–Tucker demand system model using the same data-set. The study finds that trip change estimates from two of the three models are similar and convergent valid, though the willingness to pay estimates differ in magnitude (J. C. Whitehead, Phaneuf, et al., 2010). Respondents appear to be willing to take one extra trip per season as a result of an increase of the beach width by 30.5 meters (100 foot). In a more recent study he analyzes together with co-authors the WTP for coastal erosion management and finds that shoreline retreat is the management method that finds the largest support when comparing to beach nourishment and shoreline armoring (Landry, Shonkwiler, and J. C. Whitehead, 2020). This might seem surprising given that in contrast to active management of beach resources, coastal retreat is a passive management approach that entails moving structures and infrastructure to adapt to an

Table 1: Combined RP and SP studies

Combined Data	Stacked	Assumes IID
Comparison	NO	YES
Pooled	YES	YES
Panel	YES	NO
Mixed	NO	NO

evolving coastline. In contrast shoreline armoring is the practice of building physical structures such as seawalls and breakwaters to protect shorelines from coastal erosion. Beach replenishment involves periodically replacing eroded beach and dune sand. Each of these coastal management techniques have however, their advantages and disadvantages and the authors conclude that maybe the results reflect rather the interests of coastal private property owners than the preferences and concerns of recreational users and non-users.

Two of the most recent applications are in an area where the environmental joint estimation literature started 30 years ago: recreational fishing. J. C. Whitehead and Lew (2020) develop econometric models to estimate jointly revealed preference (RP) and stated preference (SP) models of recreational fishing behavior and preferences using survey data from the 2007 Alaska Saltwater Sport fishing Economic Survey. The RP data are from site choice survey questions, and the SP data are from a discrete choice experiment. The authors state that the RP data are more likely to estimate the cost element well, but they are unlikely to reflect the true benefits from angling. The SP data is likely to describe these benefits well but they will very likely under-estimate the cost component due to the hypothetical nature of the survey. Therefore, combining the two methods exploits the strengths of both methods. The study compares a number of models that have been used in the joint estimation literature (mixed logit, scaled and generalized multinomial logit, nested logit trick) and finds that they produce similar results to a generalized multinomial logit model that accounts for scale differences in RP and SP data. The results seem to suggest that the WTP for king salmon is higher than for halibut and that of halibut is higher than for silver salmon. Hindsley et al. (2021) use a similar setting but additionally control for attribute non-attendance (ANA). ANA arises when survey respondents ignore choice experiment attributes such as for example costs. The study uses ANA to identify respondents who may be ignoring the SP cost variable and finds that the SP cost coefficient accounting for ANA is 164% larger in absolute value than the SP coefficient from the model that does not account for ANA. The model accounting for ANA shows much more consistency between SP and RP and is statistically preferred. The results show that the WTP differ significantly between the models using ANA or not with the highest WTPs for fish that is kept (and not released after catch).[6]

As shown in Table 1, in addition to comparison studies aimed to test for convergent validity, there are different types of combined RP and SP studies. The main joint approaches are pooled, panel and mixed data studies. The different approaches are based on the assumptions about the error terms in both data, e.g. whether they are identically and independently distributed (IID). Pooled and panel data studies are classified as stacked data, i.e. RP and SP data have similar dependent and independent variables and the observations are added together. However, the joint data would violate the IID assumption as the error terms across respondents are correlated, i.e. there should be a correlation in the choices by the same individual within the respective data, although not across the SP and RP data as they come from different samples.

[6] As the present study uses two different treatments for hypothetical bias, we do not expect this to be a significant issue.

Comparison Studies: Comparison studies do not stack the RP and SP data, and assumes IID errors. In these studies, RP (travel cost or hedonic price) and SP (contingent valuation or choice experiment) methods are conducted in order to estimate willingness to pay independently. Ideally, comparison studies should require the RP and SP observations to come from the same sample respondents, but the SP and RP models are estimated separately. Results are then compared and serve as source of external validity, usually for the SP estimates. This is what is usually referred to as convergent validity. If results do not converge, the researcher would explore the bias driving both methods, and it should not be assumed that SP is not correct, as RP can have different sources of bias. SP can potentially suffer from hypothetical bias, but RP sometimes do not have enough variation, and/or might be influenced by endogeneity and multicollinearity. As explained joint estimations can mitigate these limitations as SP would contribute with enough variation, given that researchers can include attributes not available in the market, and RP would mitigate the hypothetical bias, as its observations come from a real market. Additional examples of comparison SP/RP studies include Wardman (1988): which concludes that SP studies are accurate estimates of individuals' preferences in travel behaviour; Laughland et al. (1996), which finds a low correlation between averting costs and contingent valuation of environmental improvements; and Scarpa et al. (2003), that concludes that CE, compared to HP, pass external validity in their estimation of values of cattle traits.

Pooled Data Studies: The most attractive way to join RP and SP is to pool observations from RP and SP data and stack them in the same sample. Usually coefficients are restricted to be equal across the SP and RP datasets. In this case errors are assumed to be independent and identically distributed. This implies that pooled data does not take into consideration the correlation in behaviour by the same individual in their respective data source, although it assumes that the coefficients across the RP and SP data are equal. Empirical studies using pooled data usually use ordinary least square, Poisson binomial or Tobit models for continuous choice, and multinomial-logit (MNL) model for multiple discrete choice. All these models assume the errors are IID. Examples are: Adamowicz, J. Louviere, and Williams (1994) - use of SP data into travel cost estimations; while Layman, Boyce, and Criddle (1996) combined estimation of recreation values derived from different fishery management conditions using travel cost (TC) and contingent valuation method (CVM); and Eiswerth et al. (2000), which is also pooled data from TC and CVM, and applied joint estimation to the valuation of water recreation.

Panel Data Studies: Panel data also stack RP and SP data, similarly to pooled data, but assumes errors to be correlated. However, error terms in the SP and RP data come from different sources, and are understandably inconsistent, thus leading to heteroscedasticity (J. Whitehead, T. Haab, and J.-c. Huang, 2012). Error correlation can be induced using methods such as heteroscedastic, fixed effects, random affects and auto-regressive models, including random effect Tobit, and random effect Poisson binomial models (J. C. Whitehead, Pattanayak, et al., 2008). Examples of panel data joint studies are: D. Hensher, J. Louviere, and J. Swait (1998), which relaxing heteroscedasticity in a combined choice model, it makes the case for SP data in market behaviour analyses to enrich RP insights; Grijalva, Bohara, and Berrens (2003), that estimate preferences of outdoor recreation using a joint TC and CVM pooled data; and Vass et al. (2018), which applies heteroscedastic c-logit to value the interventions and services in a healthcare context, allowing for scale differences between samples of the public and patients, reporting also an improvement from a multinomial logit estimation.

Mixed Data Studies: Mixed data do not stack RP and SP data, and these are treated as different framework. However, these are jointly estimated (though not stacked), and allows for error correlation, thus relaxing the IID assumption. Applications of mixed data studies include: Cameron (1992), which combines TC and CVM data to value recreational fishing; J.-C. Huang, T. C. Haab, and J. C.

Whitehead (1997), which uses comparative static analysis of SP and RP estimations, stressing that the difference in preference structures prevents joint estimations, save the quality change is caused by the same change in behaviour; Loomis (1997) uses data from TC and CVM inputs in a random effect probit to estimate the value of amenities of instream flow; and Brownstone, Bunch, and Train (2000) applies scaled multinomial and mixed logit models to joint SP/RP estimation of preferences for automobiles, addressing heterogeneity of preferences with best improved fit from mixed logit models.

This study stacks the SP and RP data together, and applies a heteroscedastic conditional logit model with interaction terms between attributes individual characteristics, combined to relax assumptions both source of heterogeneity (within and between SP and RP samples). This is further explored in the following sections.

Even though this summary might convey the impression that many applications of joint estimation in environmental literature exist, actually these are not too numerous, especially when compared to other literatures and the range of the topic covered is not too broad with many papers using the same data sets (beach erosion, fishing or hunting). Given the advantages of joint estimation one would expect that it would be applied more often especially when non-use values are involved as it is often the case in environmental valuation. The present paper contributes by applying the method to a product that includes both use and non-use values and hence where it is especially suited. To our knowledge this is the first study applying a joint estimation approach to organic food.

3 Econometric Models

This section is going to present the models used in the separate SP and RP estimations and the one used in the joint estimation together with some models used in the literature for comparison.

3.1 Discrete Choice models (SP and RP)

Following McFadden et al. (1973) and Hoffman and Duncan (1988), when selecting the set of available attributes, the choice behaviour of an individual randomly drawn from the population aims to maximise their utility U, following a function which can be written as:

$$U_{ij} = V(X_i, Z_j) + \epsilon(X_i, Z_j) \tag{1}$$

Where: V_{ij} is the value (utility) of a given alternative j to an individual i, and it is nonstochastic and reflects the tastes of the individual with chacteristics X_i, who faces J alternatives described by vectors of attributes Z_j. ϵ is the stochastic idiosyncrasies of the individual's tastes, which is unknown and treated as random. Thus, because of unobserved factors[7], one can only predict the probability that an individual i will choose an outcome with characteristic j over k if $U(i,j) > U(i,k)$, for $j \neq k$. In general this can be shown as:

$$P_{ij} = Pr(V_{ij} > V_{ik}), for\ all\ k \neq j \tag{2}$$

Assuming that the individual will choose the alternative that will give them highest utility (or at least are indifferent between alternatives), following their "behaviour rule", denoted by h_x, and the

[7]For this same reason, OLS estimation should be biased and inconsistent (Horrace and Oaxaca, 2006).

set of alternatives $B \subseteq z$, the conditional probability $P(Z \mid X, B)$ that the individual will choose alternative Z equals:

$$P_j \equiv P(Z \mid x, B) = \pi[h_\epsilon H \mid h_\epsilon(X, B) = X_j]$$

$$P_j = P\{[\epsilon(X, Z_k) - \epsilon(X, Z_j)] < [V(X, Z_k) - V(X, Z_j)], for\ all\ k \neq j\} \quad (3)$$

Thus:

$$P_j = P\{[V(X, Z_j) + \epsilon(X, Z_j)] \geq [V(X, Z_k)] + \epsilon(X, Z_k)] \quad (4)$$

The probability of selecting alternative j can be shown as:

$$P_{ij} = \frac{exp(V_{i,j})}{\sum_{m=1}^{j} exp(V_{i,m})},\ m \in J_i \quad (5)$$

Following the above theoretical background, modelling choices depends on the assumptions made about the unobservable $\epsilon_{X,Z}$.

3.1.1 Multinomial logit model:

McFadden et al. (1973) introduced multinomial logit (MNL) in which individual choices of consumption are based on their observable characteristics, i.e. one nominal dependent variable given one or more dependent variable, such as income, and thus focuses on analysis of individuals, and their characteristics as explanatory variables, thus[8]:

$$V_{ij} = f_1(X_i) \quad (6)$$

The model is based on the assumption that individuals' unobserved tastes and preferences are homogeneous, i.e. the probability of choosing an alternative with characteristic j over choosing characteristic k does not depend on the attributes of the other alternatives. Hence the coefficient of the product attributes are assumed to be the same across all respondents. Thus MNL focuses on the respondents, and their characteristics are the explanatory variables. The choice probability in the MNL model is:

$$P_{ij} = \frac{exp(X_i\beta_j)}{\sum_{m=1}^{j} exp(X_i\beta_m)},\ m \in J_i \quad (7)$$

In the MNL model, therefore, choice probabilities can only be affected by the different impact individuals have on the different alternatives, i.e. the coefficient β_j for each explanatory variable shows the effect individual characteristics on the probability of choosing each alternative.

[8] As opposed to simple logit, mulltinomial logit can have more than two discrete outcomes.

3.1.2 Conditional logit model:

McFadden et al. (1973) also introduced the conditional logit (C-logit), which, in contrast to MNL, focuses on the set of alternatives available to individuals, thus conditional to the number of cases. Therefore, the explanatory variables are the characteristics of the other alternatives, and all observed factors or explanatory variables which can be represented in Z_{ij} (Cushing et al., 2007).[9] This model has been more applied than MNL in models for discrete choice studies, as it relaxes the homogeneity assumption.

The choice probability in the conditional logit is shown as:

$$P_{ij} = \frac{exp(Z_{ij}\alpha)}{\sum_{m=1}^{j} exp(Z_{im}\alpha)},\ m \in J_i \tag{8}$$

Where α shows the effect of attributes on the probability of choosing each alternative.

The differences between MNL and C-logit are more evident when re-arranging equations 7 and 8 (dividing them by their respective numerator), thus MNL can be re-written as:

$$P_{ij} = 1/\sum_{m=1}^{j} exp[X_i(\beta_m - \beta_j)],\ m \in J_i \tag{9}$$

And the C-logit model as:

$$P_{ij} = 1/\sum_{m=1}^{j} exp[(Z_{im} - Z_{ij})\alpha],\ m \in J_i \tag{10}$$

In equation 9, the probability depends on the difference between alternatives, while in equation 10 it depends on the difference between the characteristics of the alternatives, i.e. the highest utility V_{ij} (given by the choice that the individual values the most) depends on the attributes Z_j, following an unspecified function form such as:

$$V_{ij} = f_2(Z_{ij}) \tag{11}$$

Therefore, the differences between MNL and C-logit reflect how the researcher aims to model individual behaviour based on the choice hypotheses. As Hoffman and Duncan (1988) argue, utility is primarily regarded as a function of individual's level of consumption, which is equivalent to the exogenous level of income and the price set they face, thus enforcing equation 11. Moreover, while MNL models can offer relevant insight about individuals who made the choices, they are not properly suited to test important hypotheses of why choices are made. Therefore, in the present study, C-logit will be used for the separate analyses of the SP and RP datasets.

[9] Instead of having one line per individual like in the classical logit model, there will be one row for each category of the variable of interest, per individual.

3.2 Joint estimation model

Pooling RP and SP data is tempting, but can be challenging. Carson et al. (1996) show a positive correlation between contingent valuation (most popular SP method) and RP methods, indicating that RP and SP methods are based on common preferences, thus joint estimations would be valid. Phaneuf, Taylor, and Braden (2013) also show that although RP and SP methods differ in their estimation approaches, i.e. in SP methods the magnitude estimates are estimated from the utility function while RP uses the price function, the baseline marginal WTP $\frac{\delta P}{\delta q}$ derived from SP approaches is equal to the marginal implicit price of attributes in RP. Also, both RP and SP are based on the random utility theory.

More closely related to this study, Brooks and Lusk (2010) estimated a pooled log-likelihood function joining scanned milk purchase data and responses from a choice experiment in order to estimate WTP for selected milk attributes, including organic, cloned animal, and rBST[10]. Most joint studies use MNL or Nested Logit to estimate the joint parameters. However, considering the nature of this study, especially the estimation of individual preferences from the CE, heteroskedastic conditional logit (HC-logi) appears to be more suitable. The model relaxes the IID assumption, allows for scale differences and for the use of interaction terms which further allows for variance within the SP and RP samples.

3.2.1 Heteroscedastic Conditional Logit and Scale Parameter

The present study uses some techniques to relax IIA and heterogeneity effects with conditional logit with alternative specific variables and interaction terms between alternative and case variables. In addition, the opt-out option suggested by J. C. Whitehead, Pattanayak, et al. (2008), named status quo in our choice experiment and then adapted in the revealed data, helps to relax IID assumption.[11] However, joining the stated and revealed data usually gives also rise to issues with stated-dependence and scale differences between the two datasets. For this study, the former does not pose a problem, given that the subjects from the two datasets are not the same. Traditionally, in joint estimations the researcher surveys individuals to collect their choices from a set of hypothetical alternatives, then collects from the same individuals their actual choices, for example from current or past trips to a site in travel cost methods, or to recall their shopping behaviour, thus giving room for state-dependence or compliance bias. Therefore, scale is the main issue this study has to address.

Scale parameter can differ significantly between RP and SP data as the parameters of tastes and preferences of individuals influencing the choice for a specific attribute should be influenced by the experimental design in the SP. In the CE there is an implicit encouragement for respondents to consider the utility effect from their choices more cautiously. As illustration, the error term ϵ in equation 1 is associated with the alternative to be selected in the rational choice of individual i. From equation 3, adapted from Vass et al. (2018), one can analyse the probability unobserved heterogeneity between alternatives $\epsilon_{i,k} - \epsilon_{i,j}$ is less than the observed $V_{i,k} - V_{i,j}$. In the Gumbel distribution, variance is defined as $\frac{\pi^2}{6\mu}$. Thus, the normalised scale parameter μ, is inversely related to the variance of the error term, i.e. $\sigma^2_{\epsilon_{i,j}} = \frac{\pi^2}{6\mu}$. Multiplying through the choice probabilities in equation 5 by μ gives:

[10] recombinant bovine growth somatotropin (rBGT) is a bovine hormone, which is a modification of bovine natural somatotropin, which has risen concerns amongst the population of possible links with cancer.

[11] Please note that applying mixed logit, which completely relaxes the assumption of IIR and IID, given the randomised order of alternatives present in the design of the choice experiment used in this study, would imply that a significant number of observations from the SP data would be lost.

$$P_{ij} = \frac{exp(\mu V_{i,j})}{\sum_{m=1}^{j} exp(\mu V_{i,m})},\ m \in J_i \tag{12}$$

Therefore, the scale parameters and preferences cannot be separately identified, thus each observable variable has an effect from the scale parameter, reported as $\mu\beta$ (scaled preference parameters). As μ increases, variance decreases and estimates β are shown larger. However, these estimated coefficients have no interpretation, considering that utility has no scale, and may differ between samples or survey models (Vass et al., 2018). Therefore, $\mu\beta$ prevents us from establishing the source of differences in preferences as the level of utility from choosing one alternative or variance of the error term.

The HC-logit allows μ_i to be a function of individuals characteristics X, i.e. λ_i is parameterised as $exp(X_{i\gamma})$, where γ is a vector of parameters which incorporates the effect if individuals on the scale parameter (Hole et al., 2006). Therefore, the choice probability can be expressed as an alternative (heteroscedastic) conditional logit model, which allows for unequal variance across individuals, rewritten as:

$$P_{ij} = \frac{exp(\mu_i \beta Z_{ij})}{\sum_{j=1}^{J} exp(\mu_i \beta Z_{ij})} \tag{13}$$

In the HC-logit, the $exp(Z_{i\gamma})$ treats μ_i as positive for all individuals, and collapses to C-logit when $\gamma = 0$. In addition, following DeShazo and Fermo (2002), $\mu\beta$ is estimated by the log-likelihood function:

$$LL = \sum_{i=1}^{I} \sum_{j=1}^{J} y_{ij} lnPij \tag{14}$$

Where $y_{ij} = 1$ if alternative j is chosen by individual i $y_{ij} = 0$ if not.

4 Data

This section presents the SP and RP datasets used in this chapter. The data provided by the choice experiment is formed by collected information from UK 505 households, including socio-economic characteristics (household size, income, education, etc.), behaviour towards the environment and organic food, health status, lifestyle, and happiness, altogether offering 12,120 observations. Each alternative to be chosen in each round, of 3 alternatives, counts as one observation. Therefore, subjects provided the survey with an average of 8 rounds of choice cards ($505 * 8 * 3 = 12,120$). These were collected online by a professional survey company in 2016, after a pilot project conducted with with 60 subjects. The RP is from the Kantar Worldpanel and constitutes scanned purchases from 26,658 households in the same year, from which 9,948 households contributed to 58,170 used observations. Although the SP and the RP datasets have potentially different respondents, they are both representative samples for British consumers.

4.1 SP Data and Consumer Behaviour

The choice experiment used in this study was conducted in April 2016 by a professional market research company, and involved a maximum of 30 minutes online interviews with respondents across the UK. The interview was divided into four sections: data about recalled purchases, instructions, Choice Experiment, and Socio-Demographics, including attitudes towards organic food, environment, health and happiness.

In the first section of the survey, respondents were given general information about the experiment, and were asked about their purchasing behaviour. More specifically, they were asked about their past shopping experience, including product information such as quantity, shop, price, labels (e.g. organic), quality, and expiry date. In the second section respondents were asked some warm up questions. This section, combined with the fourth section, incorporated hypothetical bias treatments. The first treatment was a 'Cheap Talk' script where the participants were made aware that people in general tend to act differently when they face hypothetical decisions and were urged to act as in a real shopping situation. The 'Cheap Talk' script included also a 'Budget Constraint Reminder' that reminded respondents that when they spend more on one item they have less left for other goods. The second treatment was 'Honesty Priming' where by answering a series of positive statements with obviously true or false words, respondents were primed into answering truthfully in the following choice tasks. These two measures aimed to reduced potential hypothetical bias, as demonstrated by Farrell and Rabin (1996), Cummings and Taylor (1999), Carlsson, Frykblom, and Lagerkvist (2005), Jacquemet et al. (2011), Tonsor and Shupp (2011), Gschwandtner and Burton (2020) and others. In the experiment, these treatments were randomly assigned to respondents, thus forming four treatment combinations: 'Cheap Talk' + 'Honesty Priming', 'Cheap Talk' only, 'Honesty Priming' only, and 'No Treatment' for reference. A more detailed description about the hypothetical bias treatments used in the CE and their impact can be found in Gschwandtner and Burton (2020).

The third section was the experiment, i.e. the choice task itself. Respondents were presented with choice cards such as figure 2 (Appendix), in which each attribute is attached to a value, and respondents are posed with different combinations of product characteristics, from which they would choose one of two options (option A or option B), or none, defined here as status quo (SQ). Using this 'unlabeled' design, in our CE, following a fractional factorial design, there are 2^5 x 6^1 (192) possible alternatives for chicken breast, given that there are five dichotomous attributes, plus six levels for 'price'. The dichotomous attributes were 'Organic', 'Animal Welfare'[12], 'Environmentally Friendly', 'Quality' 'Best Before' (which states if the product is usable for one week or longer) and 'Low chemical usage'[13]. A description of the attributes and their levels can be found Figure 3 in the Appendix. The Appendix also includes Table 8 with descriptive statistics of the attributes and the hypothetical bias (HB) treatments. As it would be unfeasible to confront individuals with all possible alternatives within 30 minutes, a subsets of 16[14] randomly assigned choice cards were allocated for the interview.[15] A more detailed explanation of the choice experiment can be found in Gschwandtner and Burton (2020). In the final section, the questionnaire collected information about socio-economic characteristics, lifestyle, general behaviour towards organic, diet, frequency of exercises, and self-assessment of health and happiness. These responses were used for internal validity checks or detection of protest bids[16] and outliers. Moreover, these are used in the present study to analyse individuals characteristics and behaviour, and respective demand and WTP for organic and its associated attributes.

[12]Symbolized by the 'Freedom Food' label that stands for high animal welfare in the UK that.

[13]Which was used alternatively to 'Organic' as 'Organic implies low chemical usage.

[14]This is the typical number of choices used in the literature, e.g. Adamowicz, J. Louviere, and Williams (1994), Balcombe et al. (2016), and Burton, Rogers, and Richert (2017)

[15]The fractional factorial design was obtained by means of programming using the software called 'Ngene', which chooses a subset (fraction) of the full design such that it enables the estimation of the parameters with low D-errors.

[16]From individuals for example who may place a higher value than average, but refuse for ethical reasons.

The summary statistics of the variables used from the dataset provided by the choice experiment are described in Tables 2. As shown the experiment collected choices from 505 individuals, from which 60% were female, were on average 50 years old, 67% were married, with just below 1 child per household. On a scale from 1 (below high school) to 8 (professional degree), respondents on average placed their education level just below 4 (2-year college education, or 13.5 years in education). In terms of income, the average gross income of respondents was £2,454/month. Comparing with UK demographics from the Office of National Statistics (ONS), these figures do differ in terms of age, education and gender only. This is intuitive because these are the three variables that refer exclusively to the respondents, not their households, thus expected to be different from the average national figures, which include all individuals, not only the household member responsible for the grocery shopping. Most studies show that the main shoppers are mostly adult women, thus shifting upwards also age and years in education. Examples of studies that show that these are the typical characteristics of the average shopper are: Lea and Worsley (2005) (Australia), Arbindra and Wanki (2005) (USA) and Stobbelaar et al. (2007) (Netherlands). Therefore, one should explore the characteristics of household main shopper on product choice when determining the representatives of a sample and from this point of view the sample is representative.

Table 2: Summary statistics SP data: Individuals (N=505)

Variable	**Description**	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Female	gender dummy	0.602	0.49	0	1
Age	age of respondent	50.447	15.636	18	80
Married	dummy for marital status	0.667	0.472	0	1
Children	number of children in household	0.614	0.98	0	6
Vegetarian	dummy for vegetarian	0.044	0.204	0	1
Education	< high school to professional degree	3.782	1.589	1	8
High education	dummy for high education	0.392	0.489	0	1
Income	net income	2,454	1,918	250	10,500
High Income	income > UK average (£2,336*)	0.251	0.434	0	1
Professional	occupation dummy	0.206	0.405	0	1
Services	occupation dummy	0.051	0.221	0	1
Sales	occupation dummy	0.087	0.282	0	1
Farmer	occupation dummy	0.002	0.044	0	1
Construction	occupation dummy	0.016	0.125	0	1
Transports	occupation dummy	0.028	0.164	0	1
Government	occupation dummy	0.024	0.152	0	1
Retired	occupation dummy	0.269	0.444	0	1
Other	occupation dummy	0.19	0.393	0	1
Unemployed	occupation dummy	0.127	0.333	0	1
ProEnvir	green behaviour	54.473	8.609	16	70
ProOrganic	positive attitude	44.368	12.581	10	70
ConOrganic	do not buy organic	40.372	10.384	10	70
Happy 1	feeling happy	3.547	0.885	1	5
Happy 2	satisfied	3.606	0.951	1	5
Diet	dummy for respondent on diet	0.156	0.363	0	1
Healthy	healthy lifestyle, >50 (10-70)	0.921	0.270	0	1

*average monthly net income in 2016, source: ONS

The choice experiment that feeds the SP data also explored attitudes and behaviours of respondents towards organic products, the environment, their lifestyle, health and happiness. The main objective was to better understand how the consumer behavior is correlated with environmental attributes. This is very useful information for intervention purposes, e.g. education of individuals about the benefits of these attributes, and marketing and price strategies for suppliers. During the experiment, the survey asked subjects ten questions about the reasons for buying organic, about the environment, about their lifestyle and ten questions about their reasons for not buying organic. These were ranked from "strongly disagree" to "strongly agree" (1 to 7), and values were associated to the respondent, so that individuals with scores above the average were considered respectively pro-organic, pro-environment, with a healthy lifestyle, or con-organic respectively. For example the last entry in Table 2 is looking at respondents who scored on the 'healthy lifestyle scale' above 50 (out of maximum 70).[17] The table also shows that individuals in the sample were scoring between 3.5-3.6 from a maximum of 5 on two different happiness scales, so they seem to be consistently happy.[18] The questions asked related to these variables can be found n the Appendix. In addition, within these categories, each specific question denotes a behaviour that might be associated with the subject choice. For example, an individual that that agrees that they do not buy organic because it is expensive (a con-organic question) would be expected to not choose a card that includes organic and has a high price tag. A deviation from this would characterise a hypothetical bias.

The experiment also indicates that respondents behaviour associated with perceptions of organic, their approach towards the environment, and their lifestyle are significant when explaining the mean differences between the sub-groups for those who buy organic and those who don't. Figure 1 summarises all choices for organic in relation to socio-economic characteristics, behaviour and lifestyle of respondents, including only statistically significant difference in proportion of responses. Pro-organic behaviours are associated with the highest proportion of respondents choosing organic, relative to the rest of the sample, ranging from 4.3% (those who perceive organic as environmentally friendly) to 9.3% (those who perceive organic as healthier). A quite strong impact have those who buy organic to support local production (9.1%), those who perceive organic as being fresher (8.5%) and because it is not genetically manipulated (8.3%). Only generation "baby boomers" and those who don't buy organic because they perceive it as expensive have a negative effect on the proportion of individuals choosing organic.'Gen Y - Millennials' are born between 1980 and 1994 and form 21% of the sample. They are responsible for 24% of organic chicken purchases (compared to other generations), thus with a positive effect on consumption of organic chicken. 'Gen Baby Boomers' are respondents born between 1946 and 1964 and form 43% of the sample. They buy relatively less organic having a negative effect on organic consumption. As more detailed descriptive behaviour statistics and further analyses in the Appendix shows, age has in general a negative impact on organic consumption.

[17] A similar 'scale' was constructed also with 7 reasons for not buying organic such as high price and low availability. This scale was called con-organic.

[18] The first scale answers the question 'How satisfied are you with your life as a whole these days?' with answers from 'very dissatisfied to very satisfied' on a five point scale and the second question is a comparison with other people that seem to enjoy life regardless of what is going on in their life with answers scaled from 1 to 5. The questions can be found in the Appendix.

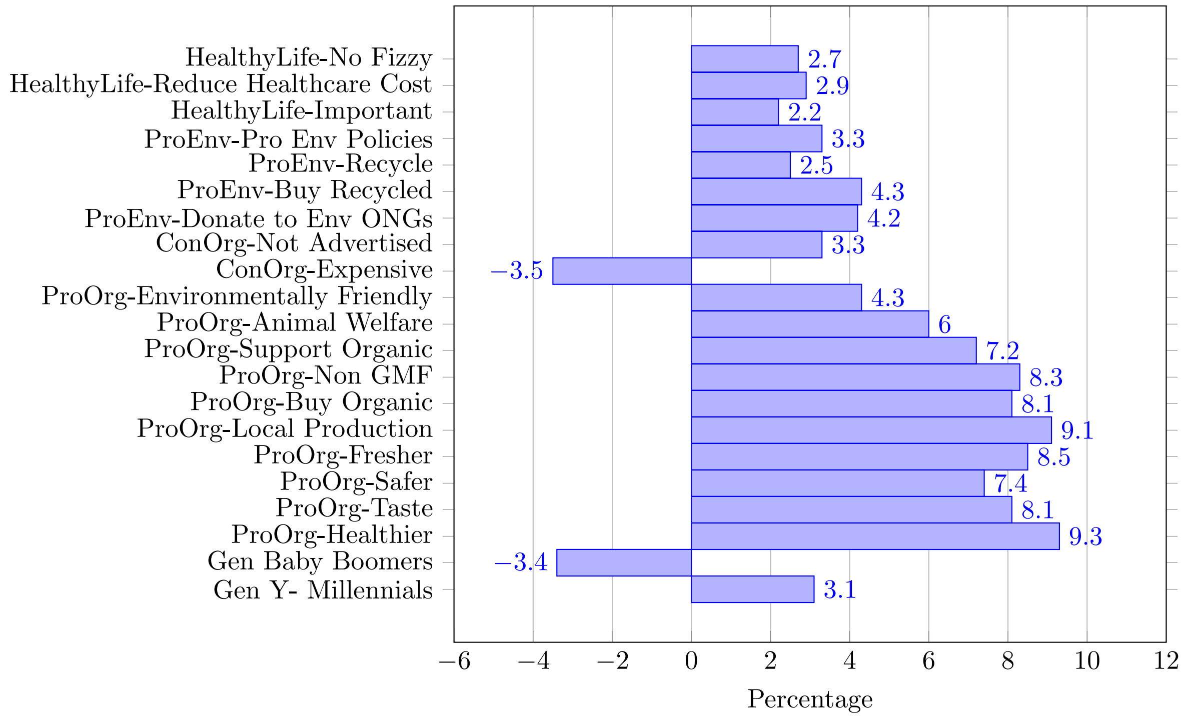

Figure 1: Choice for organic: deviation from sample average

4.2 RP Data and impact of socio-economic characteristics

The RP dataset contains 336,970 chicken purchases from 26,658 households from the Kantar Worldpanel in 2016. Only purchases of 400g chicken breast were included, so the RP data is consistent with the type of product presented to respondents in the choice experiment, and the SP and RP datasets can be pooled together. As result the number of households finally included in the analysis is reduced to 9,948, contributing to 58,170 observations.

Table 3 shows the socio-economic characteristics of the households from the RP sample. For representativeness, some of these should be checked against the characteristics of the household member responsible for the shopping, rather than the average indicators of the UK population. The average age of the household member responsible for the grocery shop in the sample is 48.3 years old, 78% of which are female. These are different from the national figures in 2016, but are not counter-intuitive when considering the average grocery shopper, as previously discussed. These are not significantly different from figures in table 2, i.e. average age of the respondents around 50, and most shoppers being women, which is expected in a sample of supermarket shoppers. In contrast, there is a significant difference in the proportion of married couples between the RP (29.2%) and SP data (66.7%), both figures deviating from the national average in 2016 (50.9% in England and Wales (ONS, n.d.), and 47% in Scotland (*Scottish household survey 2016: annual report* n.d.)), which would be close to the average of the two samples, thus a joint data would be more representative in this regard.

Households are on average formed by three people (2.9) and have one child (0.7) - similar to the SP data. Further family structure amongst households are fairly distributed, but highest proportion of households are “pre-family” (housewife below 45 years old, with no children) and “empty-nest” (45-65 years old with no children in the household), forming 19% and 18% of the sample, respectively.

Table 3: Summary statistics RP data: Households

Variable	Description	Obs	Mean	Std.Dev.	Min	Max
Female	dummy = 1 if female	9948	0.783	0.412	0	1
Age	age of respondent	9948	48.3	13.9	18	95
Married	dummy = 1 if married	9948	0.292	0.455	0	1
Adults	adults in the household	9948	2.205	0.830	1	8
Children	children in the household	9948	0.725	0.995	0	7
Household size	individuals in the household	9948	2.930	1.256	1	10
Pre-Family	age < 45 with no children	9948	0.123	0.328	0	1
Young-Family	youngest child 0 to 4 years old	9948	0.189	0.391	0	1
Middle-Family	youngest child 5 to 9 years old	9948	0.117	0.322	0	1
Older-Family	youngest child > 10 years old	9948	0.115	0.320	0	1
Older Dependents	44+ years old and 3+ adults	9948	0.134	0.340	0	1
Empty nest	45 to 65 years old, 1 to 2 adults	9948	0.184	0.388	0	1
Retired	65+, no children, 1 to 2 adults	9948	0.176	0.381	0	1
Employment FT	full-time employment (30h+)	9948	0.430	0.495	0	1
Employment PT	part-time employment (<30h)	9948	0.234	0.424	0	1
Unemployed	dummy = 1 if unemployed	9948	0.021	0.142	0	1
Income level	from 1 (>£10K) to 8 (£70k+)	8312	4.226	1.944	1	8

The smallest number (11%) are "middle-family" (families with youngest child between 5 and 9 years old) and "older-family" (families with youngest child above 10 years old).

With respect to employment and income, central indicators of WTP, the RP sample has an average gross income within category 4 (£30,000 to £39,999 per annum), which was above the national average income in 2016 (£26,300) similarly to the SP sample, and the proportion of unemployed people in sample was 2% (below half the 4.8% rate for the UK in 2016). Regarding employment, 43% of respondents were full-time while the 23% of part-time workers (respectively, 1% below and 7% above the national rates).

Table 4 shows the impact of some socio-economic characteristics on the organic attribute. The results show that most of them have a significant impact, justifying the use of socio-economic characteristics in the regressions. There is a large literature showing how various characteristics impact on organic consumption. Griffith and Nesheim (2010) explore heterogeneity wtp for organic products across different family structures; Costanigro, Kroll, and Thilmany (2012), Gschwandtner (2018), Yue, Alfnes, and Jensen (2009); and Wong et al. (2010) are just few of the many studies that show the contribution of individuals' characteristics such income, gender, family structure, employment status, etc as on organic food consumption. It is well acknowledged that organic consumption is highly correlated with income for example. One interesting result is the negative correlation with BMI suggesting a positive correlation between weight and the quality of nutrition. The fact that consumers characteristics are are associated with consumption in general has led to them being used as instrumental variables in hedonic pricing regressions sometimes using datasets similar to the present ones (Follain and Jimenez, 1985; Bishop and Timmins, 2011; Ribeiro, Gschwandtner, and Revoredo-Giha, 2021). This is also the reason why they will be used as interaction terms in the regressions.

Table 4: *Impact of socio-economic characteristics on the organic attribute*

VARIABLES	Probit	OLS
Income	0.062***	0.001***
	(0.005)	(0.000)
Female	0.024	0.000
	(0.022)	(0.000)
BMI	-0.027***	-0.000***
	(0.002)	(0.000)
PreFamily	0.159***	0.002***
	(0.042)	(0.001)
YoungFamily	0.204***	0.003***
	(0.037)	(0.001)
MiddleFamily	0.294***	0.005***
	(0.038)	(0.001)
OlderFamily	-0.024	-0.000
	(0.045)	(0.000)
EmptyNest	0.263***	0.004***
	(0.038)	(0.001)
Retired	0.383***	0.006***
	(0.040)	(0.001)
Constant	-2.289***	0.009***
	(0.077)	(0.001)
Observations	238,860	238,860
R-squared		0.002

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

5 Results

5.1 Joint estimation

As previously discussed, although stated preferences methods offer insight not yet available in the market, revealed preference methods are usually regarded as more reliable, thus making a strong case for joining stated and revealed datasets. In this study, the choice experiment and its survey offered valuable information about how consumer socio-economic characteristics and their behaviours affect their demand for some socially desirable attributes. However, for the WTP estimates, a joint approach offers more valid results.

For the observations from the RP data to be consistent with the choice respondents faced in the choice experiment, a set of alternatives needed to be constructed from the RP data. Firstly, only purchases of chicken breast were kept. A purchase (i.e. choice made) was selected by tracking the household in a given week from a supermarket in a given postcode. Only single purchases in a week were considered. As there was a large number of observations, only occasions when there were two options of chicken breast were included. Although this tried to be more consistent with the experiment, and facilitates the construction of the alternatives, it is not required to limit the number of options to two. As with the choice experiment, each product and its attributes were randomly assigned as either alternative one or two for each week in every shop. The third option was again no buy (status quo - SQ), thus a choice in which all attribute dummies receive zero value, and only household characteristics are kept. Again, to be consistent with the choice cards from the choice experiment households who chose not to buy (alternative 3) were tracked from the data as they bought a different product in the respective shop in the same week. The alternatives (choice card equivalent) used were available in the shop, but the consumer chose neither alternative. The 'status quo' (SQ) outcome has to be incorporated in the alternative construction so that it captures the preferences of individuals in the population that don't consume the product, so that potential estimates of total welfare are more accurate. Failure to do so would lead to an over-estimation of the WTP.

Table 5: Summary Statistics (joint estimation with socio-economic characteristics)

	SP			RP			RPSP		
Variable	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Price (£/Kg)	4.011 (3.492)	0	10	6.426 (1.521)	0	10	6.043 (2.159)	0	10
Volume	0.400 (0.000)	0.4	0.4	0.669 (0.396)	0	9	0.626 (0.377)	0	9
Organic	0.172 (0.377)	0	1	0.000 (0.021)	0	1	0.028 (0.164)	0	1
Freedom_Food	0.314 (0.464)	0	1	0.002 (0.042)	0	1	0.051 (0.221)	0	1
Quality	0.353 (0.478)	0	1	0.007 (0.081)	0	1	0.062 (0.240)	0	1
LesswChemicals	0.177 (0.381)	0	1	0.000 (0.000)	0	0	0.028 (0.165)	0	1
EnvFriendly	0.331 (0.470)	0	1	0.000 (0.000)	0	0	0.052 (0.223)	0	1
Healthy	0.000 (0.000)	0	0	0.000 (0.015)	0	1	0.000 (0.013)	0	1
Offer	0.000 (0.000)	0	0	0.476 (0.499)	0	1	0.400 (0.490)	0	1
Age	50.519 (15.627)	18	80	46.518 (13.078)	18	95	47.153 (13.593)	18	95
Income	3.334 (1.962)	1	8	4.385 (1.952)	1	8	4.218 (1.991)	1	8
Unemployed	0.117 (0.322)	0	1	0.019 (0.137)	0	1	0.035 (0.183)	0	1
Married	0.682 (0.466)	0	1	0.279 (0.448)	0	1	0.343 (0.475)	0	1
Children	0.623 (0.971)	0	5	0.815 (1.010)	0	7	0.785 (1.006)	0	7
PreFamily	0.077 (0.266)	0	1	0.126 (0.331)	0	1	0.118 (0.322)	0	1
YoungFamily	0.005 (0.067)	0	1	0.210 (0.407)	0	1	0.177 (0.382)	0	1
Observations	10,632			56,362			66,994		

Standard errors in parentheses

Table 5 shows the summary statistics from the stated preference (SP), revealed preference (RP) and combined (RPSP) data. In choice experiments, the number of alternatives is smaller, but form also a representative sample and was designed to capture enough variation of attributes across alternatives. The RP data is larger and presumably more representative of the UK consumer population. Although this is not the main motivation of this joint estimation study, RP results can be used as an external validity of SP results.

Joining the two datasets provides us with 66,994 observations (each alternative counts as an observation). To join the two datasets the SP and RP need to be consistent. In other words the joint C-logit model needs to accept the parameter restrictions imposed by the joint data, and this can be tested with a log-likelihood ratio (LLR) test. Joining the whole RP data with the SP data provided a p value = 0.000, thus one can reject the null hypothesis that the coefficients in the SP and the RP models are the same. After investigating the source of inconsistencies, the main variables were status quo (SQ), price and volume. Mismatch in price was due the fact that the RP data includes values up to £26 per kilogram, while in the SP the highest value for price was £10. Volume also differs between the two datasets as the choice experiment gave options of 400g chicken breast packs only[19] Therefore, the price had to be truncated to £10 per kilogram, and volume was included as explanatory variable, and limited to up to 1 kilogram. The status quo option in the samples has also been found a significant source of scale inconsistency. This could be driven by difference in unobserved factors influencing consumers to not choose a product from the given alternatives. One example is that when faced with the alternatives, the respondent might feel more inclined to make a choice, not only for the hypothetical characteristic of the experiment (they don't have to pay for their choice), but also because the attributes are more clearly available than in a real shopping situation. Availability may also have an effect on status quo, e.g. respondents when shopping should be less encouraged to buy from a choice set with which they are not totally happy or familiar, so they can choose a different shop (or different set of alternatives) which is not the case in the experiment where they are encouraged to chose an alternative in every round.

Table 6 shows regression results from the SP, RP and joint RPSP conditional logit models (columns 1 to 3, respectively), and results from the joint RPSP heteroscedastic conditional logit (CLHet) in Column 4. After accounting for the inconsistencies in price, volume and SQ, new LLR tests give p-values of 0.000 for the RPSP-Clogit model (column 3), and 0.1718 for the CLHet model (column 4), thus one cannot reject the null that the coefficients in in columns (1) and (2) are the same, once the difference in scale between the SP and RP datasets is addressed. Independent analyses of the SP and RP results would fall into the joint estimation comparison category of external validity, and the LLR test that indicates that the SP and RP parameters are equivalent. This is already a contribution, but to proceed with a joint estimation, this should be a condition. Nonetheless, J. C. Whitehead, Pattanayak, et al. (2008) argue that even if the convergence validity test fails, joining both data would mitigate the bias in both SP and RP estimations, thus the joint estimation is still encouraged.

Comparing the two joint models, the scale term (last row) of 0.873 is significant, thus to make the SP and RP data equivalent one needs to control for scale differences. These are likely driven by the characteristics of sample and, as discussed this study includes interaction terms to account for within sample heterogeneity, which also helps to control for these variations. In other words, the estimates can be driven by preferences for the attributes and the variance of unobservable elements of utility (Vass et al., 2018). The heteroscedastic conditional logit model would address this issue. In fact, post regression loglikelihood test shows an improvement in the model (+117.8), associated with a p value = 0.000, thus there is a statistically significant improvement in the model. Therefore, results from the CLHet (Column 4) are the most reliable.

[19] However, keeping only 400g package for the RP data did not offer enough values and variation of attributes hene the RP contains values up to 1 kg.

The scale parameter (RP) is positive, which indicates that the RP data has a statistically significant greater scale parameter. CLHEt parametises λ_n as $exp(X_{n\lambda})$, where X_n is a vector of individual characteristics. Therefore, the scale parameter is the exponential the term (RP), i.e. 2.39 is the scale of difference of the RP data compared to the SP. The scale parameter also shows decreased error variance in the RP data, which is intuitive. For the attributes the coefficients, which indicate the change in the probability that an alternative would be chosen by the individual, can differ substantially between the two joint models, notably price, organic, animal welfare and volume. These indicate that not taking the scale difference into account would offer misleading results.

Another issue is a second source of heterogeneity: heterogeneity within samples. The CLHet model assumes individuals within the samples have the same preferences. The inclusion of interaction terms mitigates this. Therefore, tables 6 presents results from the four models with interaction terms.

When interpreting the coefficients it is important to clarify that in the CLHet model (column 4) the scale parameter for the SP group is set to one and hence the estimated parameters are for this group of people and the ones for the RP group are scaled from those (Davis, Burton, and Kragt, 2019). It can be observed that price has a negative impact on the choice probability and that the impact appears to be twice as large in the RP data than in the SP data. Joining the two datasets mitigates for this and the result of the joint estimation yields a much lower coefficient than for RP alone. It might be worth remembering this when looking at the WTP results. The impact of volume is difficult to interpret as the SP data includes no variation (just 400g packs). However, the larger the volume the lower the choice probability which is intuitive when considering that consumers might prefer smaller packages. When looking at the organic coefficient, this has a positive and significant impact on the choice probability and is similar between SP and RP. It increases significantly when joining the two datasets but decreases and becomes similar to (or slightly higher than) the SP results as expected. The impact of 'Animal Welfare' is positive and significant as expected and appears to be much larger in the RP than in the SP dataset, which is mitigated when joining the two datasets. However, the model accounting for heterogeneity yields a much smaller coefficient than the one without, closer to the results for the SP. The impact of 'Quality' is also positive and significant as expected but not very different between SP and RP, and the joint estimation. However, in the RP estimation this important attribute is insignificant. It only becomes significant in the joint estimation, as found in the literature, which suggests that the joint estimation results are more reliable (Gschwandtner and Burton, 2020). Using less chemicals and environmental friendliness are attributes measured only in the SP data and therefore, the joint data results are similar to the results for the SP estimation, especially when accounting for heterogeneity. Low chemical usage has a counter-intuitive negative coefficient whilst environmental friendliness is positive and significant as expected. 'Healthy' is an attribute measured only in the RP data and does not have a significant impact on the choice probability. This is probably due to the way this is measured in the RP dataset where the label wasn't very clearly visible to the consumers. A better label would probably be needed to measure the impact of this important attribute. The SQ has as expected in both separate estimations and in the joint estimation a negative and significant impact on the choice probability.

For the attributes with interaction terms (price, organic and quality), the interpretation of the coefficients differs substantially from the ones without interactions, as now the effect of these attributes on the choice probability needs to take into consideration all variables containing them. In general, for a linear model, the effect of Z on P_i would be $P_i = \alpha Z + \beta X * Z$. Thus, for the organic attribute (oganic = 1) it would be $1.009 - 0.020 * Age$. Similarly, for price $-0.234 + 0.004 * Income$, and for quality $0.330 + 0.399 * Unemployed + 0.056 * Income$. Using the values from table 5, the average values for organic, price and quality would be respectively 0.066, -0.217, and 0.580. The interaction terms offer valuable insight about the heterogeneity within the sample. All variables of consumer characteristics were tested, but these were either not relevant or incompatible (did not passed the LLR test of restricted parameters). As the choice experiment and scanned data were collected by

Table 6: Joint Regression (individuals' interaction terms)

VARIABLES	(1) SP_CLogit	(2) RP_CLogit	(3) RPSP_CLogit	(4) RPSP_CLHet
Price	-0.243***	-0.554***	-0.481***	-0.234***
	(0.023)	(0.018)	(0.013)	(0.017)
Volume		-0.760***	-0.747***	-0.318***
		(0.029)	(0.029)	(0.026)
Organic	0.963***	0.995	1.762***	1.009***
	(0.234)	(1.831)	(0.236)	(0.219)
AnimalWelfare	0.295***	0.958***	0.898***	0.318***
	(0.069)	(0.263)	(0.055)	(0.057)
Quality	0.408***	0.385	0.389***	0.330***
	(0.086)	(0.395)	(0.084)	(0.072)
LessChemicals	-0.209**		-0.401***	-0.206**
	(0.088)		(0.081)	(0.082)
EnvFriendly	0.183***		0.214***	0.180***
	(0.041)		(0.042)	(0.040)
Healthy		-0.376	-0.400	-0.157
		(0.837)	(0.820)	(0.350)
Offer		-0.220***	-0.222***	-0.092***
		(0.023)	(0.023)	(0.012)
SQRP		-2.809***	-2.688***	-1.173***
		(0.041)	(0.039)	(0.085)
SQSP	-2.112***		-3.162***	-2.096***
	(0.100)		(0.077)	(0.088)
AgeOrganic	-0.020***	0.004	-0.023***	-0.020***
	(0.004)	(0.041)	(0.005)	(0.004)
UnemploymentQuality	0.387***		0.376***	0.399***
	(0.119)		(0.122)	(0.117)
IncomeQuality	0.030	0.226***	0.070***	0.056***
	(0.021)	(0.075)	(0.020)	(0.016)
IncomePrice	0.007*	0.008**	0.003	0.004***
	(0.004)	(0.004)	(0.003)	(0.001)
Scale Parameter (RP)*				0.873***
				(0.072)
LLR	-3235.137	-14195.506	-17552.434	-17434.680
LLR Test			235.508	8.074
			(0.000)	(0.172)
Observations	10,632	56,362	66,994	66,994
Number of groups			22,470	22,470

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1
except for the row below the LLR Test where there are p-values in parentheses.
*The scale parameter is the exponential the term (RP)=2.39

different companies, from different samples, most individual characteristics and criteria were different, thus limiting the scope. However, age, unemployment and income were found to be both relevant to explain consumers choices, and consistent across the SP and RP samples. Age again is central in defining relatively a lower marginal utility from the organic label, while income is positively correlated with quality and price and has a positive impact on the choice probability, as expected. The joint estimation also shows an increase in marginal utility from the quality attribute for unemployed people, which might seem counter-intuitive, but is consistent with the descriptive analysis. As explained, a large part of unemployed people consist of young people that are only temporally unemployed and hence might still be able to search for a higher quality in food. In summary, the coefficient of the interaction variable tells us by how much the influence of the attribute on the choice probability will change when the consumer characteristic changes by 1 unit. The marginal effects from age on organic, and of income on quality and price are respectively, -0.020, 0.056 and 0.004. These show that an increase in 1 year in age would decrease the probability to choose organic by 2%, while a 1 unit higher income increases the probability to choose quality by 5.6%, and to chose a product with a 1 unit higher price by 0.4% . For dummy variables, the analysis is clearer. In the joint results, unemployed individuals would increase the quality coefficient by approx. 0.4.

Altogether, these have an effect on WTP estimations. Firstly, one can also check whether the WTP values for the attributes across the two samples are the same. The WTP values are estimated as the ratio between the attribute coefficient and the price coefficient (both from explanatory variables in the models). WTP estimates are not affected by difference in variance across SP and RP samples, i.e. the heteroscedastic errors (Vass et al., 2018). If normalising $\lambda_{SP} = \lambda_{RP} = 1$ is not true, i.e. heteroscedastic, the scale hetoregeneity is elimated in the WTP estimation:

$$\frac{\lambda\beta_{attribute}}{\lambda\beta_{price}} = \frac{\beta_{attribute}}{\beta_{price}}$$

Willingness to pay estimates are given by the formula in equation (15) are shown in table 7.

$$WTP = -\frac{(Attribute \quad Coefficient)}{(Price \quad Coefficient)} \tag{15}$$

Estimates from table 7 control for the heterogeneity of preferences and are a more diligent approach to interpret the WTP. They show that consumers are willing to pay less per unit for larger packages (volumes) which is intuitive. It might be interesting to observe that the WTP of the joint model not accounting for heterogeneity would be significantly lower than the one from the RP estimation (the only one that has variation in volume). However, when accounting for heterogeneity (column 4) the WTP is not significantly different to the one from the RP estimation. The interesting result in this table is that the WTP for the attribute 'Organic' is around £4 from the SP estimation, around £2 from the RP estimation and an average of £3.66 in the joint estimation. However, when accounting for heterogeneity, the WTP for the organic attribute is even larger than from SP, RP and RPSP which is mainly driven by a smaller coefficient (in absolute value) for price in this estimation. Nevertheless, the difference is not big and the amount is plausible (Ribeiro, Gschwandtner, and Revoredo-Giha, 2021). Regarding 'AnimalWelfare', consumers appear to want to pay £1.21 more from the SP estimate, £1.73 from the RP estimates and £1.87 from the joint estimate for a higher animal welfare. However, when accounting for heterogeneity the WTP for higher animal welfare decreases to £1.36, less than half than for 'Organic' which might seem intuitive as 'Organic' is perceived to imply higher animal welfare as well. A bit more surprising maybe is the fact that the WTP for a higher quality is £1.68 from the SP estimates, only £0.7 from the RP estimates and £0.81 from the joint estimation. As with organic, when controlling for heterogeneity, the WTP increases to £1.41, closer to the SP WTP. It is still much lower than for 'Organic' maybe because consumers usually associate 'Organic' also with a higher quality among other attributes. The HC-logit model estimates a negative effect from the "low chemical usage" label, although counter-intuitive and different from results from the joint c-logit, this is consistent with the SP results (the RP data does not have the attribute), re-inforcing the notion that scale heterogeneity has to be addressed, as there is unlikely to have any other reason why the joint estimation would deviate from the SP in this case. Results suggest that IID is not a major source of bias in our joint estimation, but scale differences can have a significant impact on WTP, thus if not addressed, the SP and RP offer better results if estimated separately. 'EnvFriendly' has a positive WTP of £0.75 in the SP estimates where it is available. The joint estimation would lead to a decrease of the WTP to £0.44. However, when accounting for heterogeneity, as in the case of 'AnimalWelfare' and 'Organic' the WTP increases to a level slightly higher than the SP estimates (£0.77). The WTP for 'Healthy' is not significantly different from zero which has probably to do with the way this attribute was measured in the RP dataset. The attribute was not available in the CE and very few products from the revealed sample had the healthy label, thus for future studies one would need more information for better analysis, most likely by including as an attribute in a CE, as the attribute is not commonly available from real data. Products being on offer reduce the WTP in the RP estimates (the only dataset where this attribute is available) by £0.4. This reduction is slightly higher in the joint estimation (£0.46) but decreases again to approx £0.4 when heterogeneity is accounted for. Therefore. it appears that only for 'environmental attributes' such as 'Organic', 'AnimalWelfare' and 'EnvFriendly' the joint estimation accounting for heterogeneity increases slightly the WTP compared with SP results but as explained above, we do expect values to be similar to the results for the SP group. What the results clearly show is that using just SP or RP data might lead to misleading results. Definitively, in the case of 'Organic' it would lead to an underestimation of the WTP and to a miss-allocation of resources for organic production. Using the joint estimation leads to more reliable results.

Table 7: WTP Comparison - with interaction terms (£)

	(1)	(2)	(3)	(4)
Attributes	SP	RP	RPSP	RPSP (CLHet)
Volume		-1.37 (0.13)	-1.55 (0.14)	-1.36 (0.13)
Organic	3.96 (1.91)	1.80 (6.49)	3.66 (0.97)	4.31 (1.83)
AnimalWelfare	1.21 (0.48)	1.73 (0.94)	1.87 (0.22)	1.36 (0.40)
Quality	1.68 (0.69)	0.70 (1.39)	0.81 (0.33)	1.41 (0.61)
LessChemicals	-0.86 (0.79)		0.83 (0.32)	-0.88 (0.75)
EnvFriendly	0.75 (0.35)		0.44 (0.17)	0.77 (0.35)
Healthy		-0.68 (2.96)^{i}	-0.83 (3.34)^{i}	-0.67 (2.93)^{i}
Offer		-0.40 (0.09)	-0.46 (0.10)	-0.39 (0.08)

Lower (-) and upper (+) limits in parentheses
i: Statistically insignificant

6 Conclusions

The aim of the present paper was to estimate consumer preferences for environmental attributes related to organic chicken meat. Traditionally, the WTP for organic food has been estimated using either stated or revealed preferences techniques. Stated preferences techniques have the advantage that they can elicit preferences for attributes or combination of attributes that do not exist yet in the market, which is often the case for so called 'non-use values' such as higher animal welfare or environmentally friendly production typically associated with organic meat. However, they have the strong disadvantage that consumers do not have to actually behave like this in a real situation and therefore the estimates suffer from hypothetical and other types of biases. Revealed preferences techniques rely on actual market transaction and are best suited to elicit preferences for so called 'use-values' such as perceived better health, better taste or higher quality in general. As they are based on real choices, they do not suffer from hypothetical bias and hence offer more valid estimates. However, they suffer from collinearity and from the difficulty to isolate the effects of a specific attribute which is arguably what is mostly required in policy applications (Adamowicz, J. Louviere, and Williams, 1994). They are also riddled with other problems such as heteroskedasticity and endogeneity which are often not trivial to address. To our knowledge this is the first paper that uses a joint approach which addresses the shortcomings of both methods in order to elicit the WTP for various attributes in organic chicken meat. This is especially important for organic products as they contain both 'use-values' and 'non-use values' that would be difficult to elicit in a robust manner using the approaches separately. Moreover, by joining the stated and the revealed data we obtain a larger dataset with more observations that enable more robust WTP estimates.

The stated preference (SP) data was collected using a choice experiment, arguably the most advanced SP method to date as it is designed to enable trade-offs between alternatives and between attributes. Most importantly, it enabled us to elicit the preference for specific environmental attributes typically associated with organic such as 'animal welfare' and 'environmentally friendly' and combinations between them that might not exist in a single product on the market. Moreover, it enabled us to elicit a rich set of socio-economic variables and behavioural attitudes of consumers towards organic products, the environment and towards a healthy lifestyle that would have been very difficult to obtain in revealed data. These variables helped the analysis in three different ways: for validity checks, for behavioural insights related to organic consumption, and as interaction terms in the regressions in order to help with heterogeneity and to relax the assumption of independently and identically distributed variables (IID). The study contributes by enriching the revealed preference data coming from scanned consumers shopping's with a rich set of variables obtained from a choice experiment regarding organic products.

Using just the stated preference dataset would yield a relatively low number of observations that might leave doubt about the validity of results. Even though the choice experiment contained two different treatments against hypothetical bias and potentially managed to reduce it, stated preferences are subject to variety of different other biases such as anchoring bias, response bias, non-response bias etc. The revealed preference (RP) data helps to ground the choices made in the experiment in actual choices made by consumers in supermarkets. Moreover, the Kantar World-panel used for the construction of the dataset for the RP estimations offers a number of individuals that is 50 times larger than the ones used in the choice experiment leading to a significantly larger number of observations in total. Even though both samples were constructed to be representative for the UK consumer population this confers not only more certainly with respect to representativeness but also with respect to the robustness of the results.

The separate SP and RP methods used both conditional logit, and the main joint estimation model applied a heteroscedastic Conditional logit (CLHet). Both c-logit and CLHet models included interaction terms, providing information about consumer characteristics and behaviours when selecting the product and their attributes, while addressing the IID strong assumption of these models.

Joint estimations have their advantages, but the problems associated with such studies can also not be ignored. These include, inter-alternative error structures, unobserved heterogeneity effects, state-dependence and scale difference. This study demonstrates that heteroscedastic c-logit (CLHet), as a panel data study, with interaction terms applied to different SP and RP samples offers a viable way to offset these issues simultaneously. To our knowledge this is the first study to apply CLHet with interaction terms to address the assumption of homogenous preferences and other problems associated with joint estimations. The results show that after accounting for heterogeneity and taking into consideration the scale effect, the preferences of the two datasets are similar and can be meaningful combined.

The data collected through the choice experiment show that consumers state to buy organic food products primarily because they are healthier, fresher and in order to support local production. The main stated reason for not buying organic meat appears to be its higher price which is intuitive as organic meat is significantly more expensive than conventional one. Amongst the variables that are positively and significantly correlated with a 'pro-organic' attitude are income and gender (female) and among the ones that are negatively correlated age is the most prominent. More generally a 'pro-organic' attitude appears to be correlated with 'green behaviour' and a 'healthy lifestyle' and these attitudes appear to be also correlated with being 'happy'.

The results of the joint regression, indicate that environmental friendliness increases the choice probability of buying chicken meat by only 18%. This is significantly below the influence from the attribute 'animal welfare' (32%) and from the attribute 'quality' (33%) which might encompass being fresher and having a better taste. The attribute that contributes mostly to the choice probability is 'organic' potentially because it summarizes all the other attributes. Without combining the SP and RP data it would not have been possible to estimate in a robust manner how much each of these attributes contributes to the probability to buy organic meat individually.

The interaction terms available in the main model enlighten the consumer characteristics driving WTP for the organic label, and show that age, income and employment status are the main source of heterogeneity, with very large variations. The joint estimation, results show that the WTP for organic label would decrease by 2% per year of age in the stated and joint preference approaches. These reveal market and intervention opportunities, e.g. sellers can target younger generations, while policy making in line with information on the benefits of organic food, especially targeting older generation might help increase organic sales. Better marketing in terms of the attributes incorporated into the organic label could be also beneficial as results indicate that consumers do not necessarily understand these.

The results related to the willingness to pay (WTP) indicate that individuals are willing to pay £1.36 (22.7%) for higher animal welfare, and £0.77 (12.8%) for environmentally friendliness per 1 kg unit. The fact that environmentally friendliness has a relatively low WTP, and less chemical usage has a negative effect on WTP, indicates that meat consumers are more concerned about the welfare of the animals, thus WTP for a more 'humane' treatment of the chicken in the production process. At the same time, consumers appear to be willing to pay an average £1.41 (23.5%) for better quality, indicating that their own welfare is even more valued. The results also indicate that consumers are willing to pay a premium of £4.31 (71.8%) for the attribute organic that encompasses several of these attributes simultaneously.

Finally, the results of the present study show that the common problems risen from pooling SP and RP data can be mitigated with the application of a heteroscedastic conditional logit model, and the inclusion of interaction terms between attributes and consumer characteristics. They suggest that the IID assumption does not significantly change results, but failure to accommodate for scale differences can have severe impact on the estimates of coefficients of the choice probability, as the SP and RP can have different parameters. For this study, the parameter restriction was only accepted in the CLHet model, and failed in the c-logit model. Concluding we can say that the joint estimation offers more reliable results when estimating the WTP for various attributes in chicken meat and that the use of interaction terms in joint estimation studies promises to be a fruitful avenue for future research.

References

Adamowicz, Wiktor, Jordan Louviere, and Michael Williams (1994). "Combining revealed and stated preference methods for valuing environmental amenities". In: *Journal of environmental economics and management* 26(3), pp. 271–292.

Adamowicz, Wiktor, Joffre Swait, et al. (1997). "Perceptions versus objective measures of environmental quality in combined revealed and stated preference models of environmental valuation". In: *Journal of environmental economics and management* 32(1), pp. 65–84.

Arbindra, R and M Wanki (2005). "Perceived Risks of Agrobiotechnology and Organic Food Purchase in the United States". In: *Selected Paper prepared for presentation at the Southern Agricultural Economics Association Annual Meetings*, pp. 5–9.

Arnoult, Matthieu, Alexandra Lobb, and Richard Tiffin (2010). "Willingness to pay for imported and seasonal foods: A UK survey". In: *Journal of International Food & Agribusiness Marketing* 22(3-4), pp. 234–251.

Assembly, General (2015). "Sustainable development goals". In: *SDGs Transform Our World* 2030.

Balcombe, Kelvin et al. (2016). "Information customization and food choice". In: *American Journal of Agricultural Economics* 98(1), pp. 54–73.

Ban, Ki-moon (2016). "Sustainable Development Goals". In:

Ben-Akiva, Moshe, Mark Bradley, et al. (1994). "Combining revealed and stated preferences data". In: *Marketing Letters* 5(4), pp. 335–349.

Ben-Akiva, Moshe and Takayuki Morikawa (1990). "Estimation of switching models from revealed preferences and stated intentions". In: *Transportation Research Part A: General* 24(6), pp. 485–495.

Bishop, Kelly C and Christopher Timmins (2011). *Hedonic prices and implicit markets: Estimating marginal willingness to pay for differentiated products without instrumental variables.* Tech. rep. National Bureau of Economic Research.

Brooks, Kathleen and Jayson L Lusk (2010). "Stated and revealed preferences for organic and cloned milk: combining choice experiment and scanner data". In: *American Journal of Agricultural Economics* 92(4), pp. 1229–1241.

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

Burton, Michael, Abbie Rogers, and Claire Richert (2017). "Community acceptance of biodiversity offsets: evidence from a choice experiment". In: *Australian Journal of Agricultural and Resource Economics* 61(1), pp. 95–114.

Cameron, Trudy Ann (1992). "Combining contingent valuation and travel cost data for the valuation of nonmarket goods". In: *Land Economics*, pp. 302–317.

Carlsson, Fredrik, Peter Frykblom, and Carl Johan Lagerkvist (2005). "Using cheap talk as a test of validity in choice experiments". In: *Economics letters* 89(2), pp. 147–152.

Carson, Richard T et al. (1996). "Contingent valuation and revealed preference methodologies: comparing the estimates for quasi-public goods". In: *Land economics*, pp. 80–99.

Combris, Pierre, Sebastien Lecocq, and Michael Visser (1997). "Estimation of a hedonic price equation for Bordeaux wine: does quality matter?" In: *The Economic Journal* 107(441), pp. 390–402.

Costanigro, Marco, Stephan Kroll, and Dawn D Thilmany (2012). *Local, organic, conventional—Asymmetric effects of information and taste on label preferences in an experimental auction.* Tech. rep.

Cummings, Ronald G (1986). "Valuing environmental goods". In: *An assessment of the contingent valuation method*, pp. 104–107.

Cummings, Ronald G and Laura O Taylor (1999). "Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method". In: *American economic review* 89(3), pp. 649–665.

Cushing, Brian et al. (2007). "Conditional logit, IIA, and alternatives for estimating models of interstate migration". In:

Davis, Katrina J, Michael Burton, and Marit E Kragt (2019). "Scale heterogeneity and its implications for discrete choice analysis". In: *Land Economics* 95(3), pp. 353–368.

DeShazo, JR and German Fermo (2002). "Designing choice sets for stated preference methods: the effects of complexity on choice consistency". In: *Journal of Environmental Economics and management* 44(1), pp. 123–143.

Eiswerth, Mark E et al. (2000). "The value of water levels in water-based recreation: A pooled revealed preference/contingent behavior model". In: *Water Resources Research* 36(4), pp. 1079–1086.

Ellickson, Paul B, Mitchell J Lovett, and Bhoomija Ranjan (2019). "Product launches with new attributes: a hybrid conjoint–consumer panel technique for estimating demand". In: *Journal of Marketing Research* 56(5), pp. 709–731.

Farrell, Joseph and Matthew Rabin (1996). "Cheap talk". In: *Journal of Economic perspectives* 10(3), pp. 103–118.

Feit, Eleanor McDonnell, Mark A Beltramo, and Fred M Feinberg (2010). "Reality check: Combining choice experiments with market data to estimate the importance of product attributes". In: *Management science* 56(5), pp. 785–800.

Follain, James R and Emmanuel Jimenez (1985). "Estimating the demand for housing characteristics: a survey and critique". In: *Regional science and urban economics* 15(1), pp. 77–107.

Gopalakrishnan, Sathya et al. (2011). "The value of disappearing beaches: a hedonic pricing model with endogenous beach width". In: *Journal of Environmental Economics and Management* 61(3), pp. 297–310.

Griffith, Rachel and Lars Nesheim (2010). *Estimating households' willingness to pay.* Tech. rep. CEMMAP Working paper.

Griffith, Rachel and Lars Nesheim (2013). "Hedonic methods for baskets of goods". In: *Economics Letters* 120(2), pp. 284–287.

Grijalva, Therese, Alok K Bohara, and Robert P Berrens (2003). "A seemingly unrelated Poisson model for revealed and stated preference data". In: *Applied Economics Letters* 10(7), pp. 443–446.

Gschwandtner, Adelina (2018). "The organic food premium: a local assessment in the UK". In: *International Journal of the Economics of Business* 25(2), pp. 313–338.

Gschwandtner, Adelina and Michael Burton (2020). "Comparing treatments to reduce hypothetical bias in choice experiments regarding organic food". In: *European Review of Agricultural Economics* 47(3), pp. 1302–1337.

Hausman, Jerry (2012). "Contingent valuation: from dubious to hopeless". In: *Journal of economic perspectives* 26(4), pp. 43–56.

Hensher, David, Jordan Louviere, and Joffre Swait (1998). "Combining sources of preference data". In: *Journal of Econometrics* 89(1-2), pp. 197–221.

Hensher, David A (2010). "Hypothetical bias, choice experiments and willingness to pay". In: *transportation research part B: methodological* 44(6), pp. 735–752.

Hensher, David A and Mark Bradley (1993). "Using stated response choice data to enrich revealed preference discrete choice models". In: *Marketing Letters* 4(2), pp. 139–151.

Hindsley, Paul et al. (2021). *Joint Estimation of Revealed Preference Site Selection and Stated Preference Choice Experiment Recreation Data Considering Attribute NonAttendance.* Tech. rep.

Hoffman, Saul D and Greg J Duncan (1988). "Multinomial and conditional logit discrete-choice models in demography". In: *Demography* 25(3), pp. 415–427.

Hole, Arne Risa et al. (2006). "Small-sample properties of tests for heteroscedasticity in the conditional logit model". In: *Economics Bulletin* 3(18), pp. 1–14.

Horrace, William C and Ronald L Oaxaca (2006). "Results on the bias and inconsistency of ordinary least squares for the linear probability model". In: *Economics Letters* 90(3), pp. 321–327.

Huang, Ju-Chin, Timothy C Haab, and John C Whitehead (1997). "Willingness to pay for quality improvements: should revealed and stated preference data be combined?" In: *Journal of environmental economics and management* 34(3), pp. 240–255.

Jacquemet, Nicolas et al. (2011). "Social psychology and environmental economics: a new look at ex ante corrections of biased preference evaluation". In: *Environmental and resource economics* 48(3), pp. 413–433.

King, Steven and Iain Fraser (2013). "Divestment of the English Forestry Estate: An economically sound choice?" In: *Ecological economics* 88, pp. 25–31.

Kling, Catherine L (1997). "The gains from combining travel cost and contingent valuation data to value nonmarket goods". In: *Land Economics*, pp. 428–439.

Landry, Craig E, J Scott Shonkwiler, and John C Whitehead (2020). "Economic Values of Coastal Erosion Management: Joint Estimation of Use and Existence Values with recreation demand and contingent valuation data". In: *Journal of Environmental Economics and Management* 103, p. 102364.

Larson, Douglas M and Daniel K Lew (2013). "How do harvest rates affect angler trip patterns?" In: *Marine Resource Economics* 28(2), pp. 155–173.

Laughland, Andrew S et al. (1996). "Construct validity of averting cost measures of environmental benefits". In: *Land Economics*, pp. 100–112.

Layman, R Craig, John R Boyce, and Keith R Criddle (1996). "Economic valuation of the Chinook salmon sport fishery of the Gulkana River, Alaska, under current and alternate management plans". In: *Land Economics*, pp. 113–128.

Lea, Emma and Tony Worsley (2005). "Australians' organic food beliefs, demographics and values". In: *British food journal*.

Lew, Daniel K and Douglas M Larson (2011). "A repeated mixed logit approach to valuing a local sport fishery: the case of Southeast Alaska salmon". In: *Land Economics* 87(4), pp. 712–729.

Lew, Daniel K and Douglas M Larson (2012). "Economic values for saltwater sport fishing in Alaska: a stated preference analysis". In: *North American journal of fisheries management* 32(4), pp. 745–759.

Lew, Daniel K and Douglas M Larson (2014). "Is a fish in hand worth two in the sea? Evidence from a stated preference study". In: *Fisheries Research* 157, pp. 124–135.

List, John A and Craig A Gallet (2001). "What experimental protocol influence disparities between actual and hypothetical stated values?" In: *Environmental and resource economics* 20(3), pp. 241–254.

Loomis, John B (1997). "Panel estimators to combine revealed and stated preference dichotomous choice data". In: *Journal of Agricultural and Resource Economics*, pp. 233–245.

Louviere, Jordan J, David A Hensher, and Joffre D Swait (2000). *Stated choice methods: analysis and applications*. Cambridge university press.

McFadden, Daniel et al. (1973). *Conditional logit analysis of qualitative choice behavior*. Institute of Urban and Regional Development, University of California.

Mitchell, Robert Cameron and Richard T Carson (2013). *Using surveys to value public goods: the contingent valuation method*. Rff Press.

ONS (n.d.). *Unemployment by age and duration (seasonally adjusted)*. URL: https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment.

Padel, Susanne and Carolyn Foster (2005). "Exploring the gap between attitudes and behaviour: Understanding why consumers buy or do not buy organic food". In: *British food journal*.

Penn, Jerrod M and Wuyang Hu (2018). "Understanding hypothetical bias: An enhanced meta-analysis". In: *American Journal of Agricultural Economics* 100(4), pp. 1186–1206.

Phaneuf, Daniel J, Laura O Taylor, and John B Braden (2013). "Combining revealed and stated preference data to estimate preferences for residential amenities: A GMM approach". In: *Land Economics* 89(1), pp. 30–52.

Ribeiro, Jose Eduardo, Adelina Gschwandtner, and Cesar Revoredo-Giha (2021). "Estimation of a Hedonic Price Equation for Chicken Meat in the UK: Does the Organic Attribute Matter?" In:

Rimal, Arbindra P, Wanki Moon, and Siva Balasubramanian (2005). "Agro-biotechnology and organic food purchase in the United Kingdom". In: *British Food Journal*.

Rosen, Sherwin (1974). "Hedonic prices and implicit markets: product differentiation in pure competition". In: *Journal of political economy* 82(1), pp. 34–55.
Scarpa, Riccardo et al. (2003). "Valuing indigenous cattle breeds in Kenya: an empirical comparison of stated and revealed preference value estimates". In: *Ecological Economics* 45(3), pp. 409–426.
Scottish household survey 2016: annual report (n.d.). (2016). Accessed: 2021-08-20, from https://www.gov.scot/publications/scotlands-people-annual-report-results-2016-scottish-household-survey/pages/2/.
Stobbelaar, Derk Jan et al. (2007). "Adolescents' attitudes towards organic food: a survey of 15-to 16-year old school children". In: *International Journal of Consumer Studies* 31(4), pp. 349–356.
Swait, Joffre, Jordan J Louviere, and Michael Williams (1994). "A sequential approach to exploiting the combined strengths of SP and RP data: application to freight shipper choice". In: *Transportation* 21(2), pp. 135–152.
Thomas, Gregory Owen et al. (2019). "The English plastic bag charge changed behavior and increased support for other charges to reduce plastic waste". In: *Frontiers in Psychology* 10, p. 266.
Tonsor, Glynn T and Robert S Shupp (2011). "Cheap talk scripts and online choice experiments:"looking beyond the mean"". In: *American Journal of Agricultural Economics* 93(4), pp. 1015–1031.
Vass, Caroline M et al. (2018). "Scale heterogeneity in healthcare discrete choice experiments: a primer". In: *The Patient-Patient-Centered Outcomes Research* 11(2), pp. 167–173.
Von Haefen, Roger H and Daniel J Phaneuf (2008). "Identifying demand parameters in the presence of unobservables: a combined revealed and stated preference approach". In: *Journal of Environmental Economics and Management* 56(1), pp. 19–32.
Wardman, Mark (1988). "A comparison of revealed preference and stated preference models of travel behaviour". In: *Journal of transport economics and policy*, pp. 71–91.
Whitehead, John, Tim Haab, and Ju-chin Huang (2012). *Preference data for environmental valuation: combining revealed and stated approaches*. Vol. 31. Routledge.
Whitehead, John C, Christopher F Dumas, et al. (2008). "Valuing beach access and width with revealed and stated preference data". In: *Marine Resource Economics* 23(2), pp. 119–135.
Whitehead, John C and Daniel K Lew (2020). "Estimating recreation benefits through joint estimation of revealed and stated preference discrete choice data." In: *Empirical Economics* 58(4).
Whitehead, John C, Subhrendu K Pattanayak, et al. (2008). "Combining revealed and stated preference data to estimate the nonmarket value of ecological services: an assessment of the state of the science". In: *Journal of Economic Surveys* 22(5), pp. 872–908.
Whitehead, John C, Daniel J Phaneuf, et al. (2010). "Convergent validity of revealed and stated recreation behavior with quality change: a comparison of multiple and single site demands". In: *Environmental and Resource Economics* 45(1), pp. 91–112.
Willis, Ken G and Guy D Garrod (1991). "An individual travel-cost method of evaluating forest recreation". In: *Journal of agricultural Economics* 42(1), pp. 33–42.
Wong, Jonathan et al. (2010). "Consumer premiums for environmentally friendly grass-fed and organic milk in the Southeast". In: *Journal of Agribusiness* 28(345-2016-15277), pp. 75–88.
Wooldridge, Jeffrey M (1996). "Estimating systems of equations with different instruments for different equations". In: *Journal of Econometrics* 74(2), pp. 387–405.
Yue, Chengyan, Frode Alfnes, and Helen H Jensen (2009). "Discounting spotted apples: Investigating consumers' willingness to accept cosmetic damage in an organic product". In: *Journal of Agricultural and Applied Economics* 41(1), pp. 29–46.

Appendix

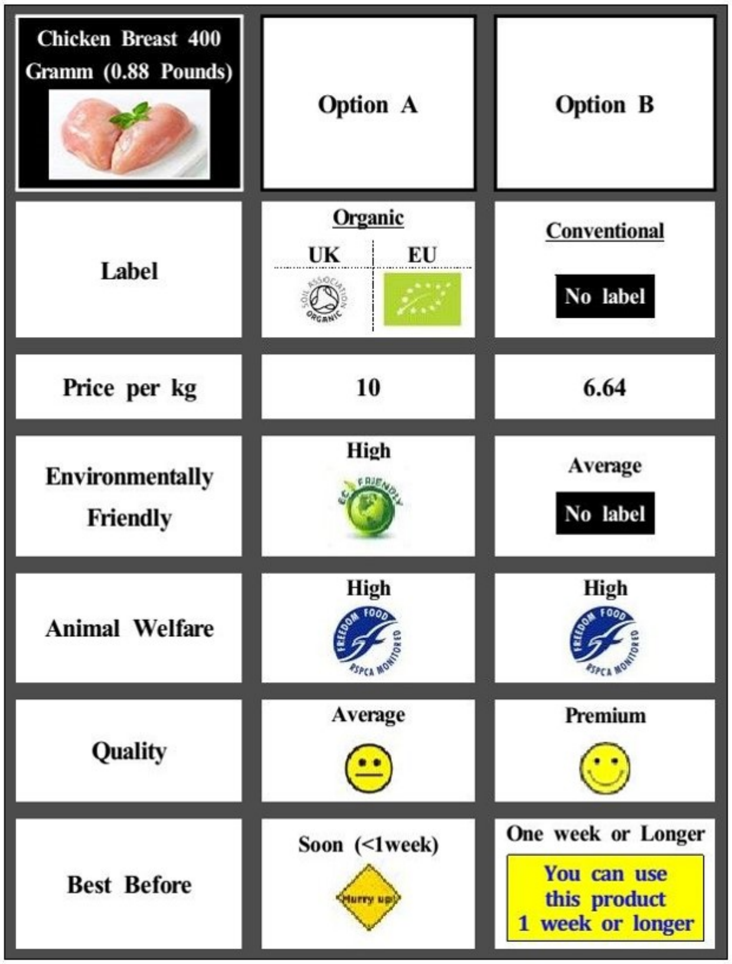

Figure 2: Example of choice card

Table 8: Summary statistics: choice cards (12,120 observations)

Variable	Description	Mean	Std.Dev.	Min	Max
Organic	dummy for organic label	0.172	0.377	0	1
Chemical Usage	dummy = 1 if "average" or = 0 if "low"	0.177	0.382	0	1
Env Friendly	dummy for Eco Friendly label	0.332	0.471	0	1
Animal Welfare	dummy for animal welfare label	0.315	0.464	0	1
Best Before	=1 if $\geq$ 1 week or = 0 if $<$ 1 week	0.334	0.472	0	1
Quality	=1 if "premium" or = 0 if "average"	0.352	0.478	0	1
Price	£3/ £3.5/ £5.75/ £6.64/ £8.32/ £10	4.015	3.496	0	10
Status quo	alternative C, no purchase	0.333	0.471	0	1
HBT0	no HBT	0.253	0.435	0	1
HBT1	dummy for "cheap talk" HBT	0.277	0.448	0	1
HBT2	dummy for "honesty priming" HBT	0.234	0.423	0	1
HBT3	dummy for both HBT	0.236	0.424	0	1

ProEnvir = Green Behavior	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	neutral (4)	Somewhat agree (5)	Agree (6)	Strongly agree (7)
I think that the TV/PC should not be left on standby overnight (1)	❍	❍	❍	❍	❍	❍	❍
It is important to switch off lights in empty rooms (2)	❍	❍	❍	❍	❍	❍	❍
I think that one should not let the tap run while brushing teeth (3)	❍	❍	❍	❍	❍	❍	❍
I think that becoming a member of and/or donating money to an environmental organization like the World Wide Fund is important (4)	❍	❍	❍	❍	❍	❍	❍
I think one should buy recycled products (5)	❍	❍	❍	❍	❍	❍	❍
I think that one should use own bags while shopping whenever possible (6)	❍	❍	❍	❍	❍	❍	❍
I think that preserving the environment should be a high political priority (7)	❍	❍	❍	❍	❍	❍	❍
I think one should use public transport rather than a car whenever possible (8)	❍	❍	❍	❍	❍	❍	❍
I think that better insulated houses will reduce the heating costs and therefore environmental damage (9)	❍	❍	❍	❍	❍	❍	❍
I think that waste separation and recycling is useful and important (10)	❍	❍	❍	❍	❍	❍	❍

ProOrganic = Reasons for buying organic	Strongly Disagree (1)	Disagree (2)	Somewhat Disagree (3)	neutral (4)	Somewhat Agree (5)	Agree (6)	Strongly Agree (7)
They are healthier (1)	❍	❍	❍	❍	❍	❍	❍
They taste better (2)	❍	❍	❍	❍	❍	❍	❍
They are fresher (3)	❍	❍	❍	❍	❍	❍	❍
They are safer (4)	❍	❍	❍	❍	❍	❍	❍
They involve higher animal welfare (5)	❍	❍	❍	❍	❍	❍	❍
They are produced in a more environmentally friendly way (6)	❍	❍	❍	❍	❍	❍	❍
To support local producers (7)	❍	❍	❍	❍	❍	❍	❍
I have bought organic products before and I was satisfied with them (8)	❍	❍	❍	❍	❍	❍	❍
I fear that conventional products may be genetically manipulated and may be less safe than organic products (9)	❍	❍	❍	❍	❍	❍	❍
I want to support the Organic Movement (10)	❍	❍	❍	❍	❍	❍	❍

ConOrganic = Reasons for not buying organic	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	neutral (4)	Somewhat agree (5)	Agree (6)	Strongly agree (7)
They are too expensive (1)	❍	❍	❍	❍	❍	❍	❍
I do not trust the organic label (2)	❍	❍	❍	❍	❍	❍	❍
I cannot easily find them on the shelf (3)	❍	❍	❍	❍	❍	❍	❍
It is not easy to recognize them (4)	❍	❍	❍	❍	❍	❍	❍
I do not find them in the same variety as conventional products (5)	❍	❍	❍	❍	❍	❍	❍
They do not look as good as conventional products (6)	❍	❍	❍	❍	❍	❍	❍
They do not last as long as conventional products (7)	❍	❍	❍	❍	❍	❍	❍
Organic products are not sufficiently advertised (8)	❍	❍	❍	❍	❍	❍	❍
I have always bought conventional products and it is hard to change the habit (9)	❍	❍	❍	❍	❍	❍	❍
I do not know much about organic products (10)	❍	❍	❍	❍	❍	❍	❍

Healthy = Questions about Lifestyle	Strongly Disagree (1)	Disagree (2)	Somewhat disagree (3)	neutral (4)	Somewhat agree (5)	Agree (6)	Strongly agree (7)
I consider that leading a healthy lifestyle is important (1)	❍	❍	❍	❍	❍	❍	❍
I consider that eating healthy food is important (2)	❍	❍	❍	❍	❍	❍	❍
I consider that doing enough exercise is important (3)	❍	❍	❍	❍	❍	❍	❍
I consider spending many hours in front of the TV/PC is unhealthy (4)	❍	❍	❍	❍	❍	❍	❍
I consider that eating fast food is unhealthy (5)	❍	❍	❍	❍	❍	❍	❍
In general I prefer to eat low fat and low sugar food products (6)	❍	❍	❍	❍	❍	❍	❍
I usually get upset when I eat too much fattening food (7)	❍	❍	❍	❍	❍	❍	❍
I consider that drinking fizzy drinks is unhealthy (8)	❍	❍	❍	❍	❍	❍	❍
I consider that smoking is unhealthy (9)	❍	❍	❍	❍	❍	❍	❍
I consider that every amount invested in a healthy lifestyle decreases healthcare costs (10)	❍	❍	❍	❍	❍	❍	❍

Happy1: Some people are generally very happy. They enjoy life regardless of what is going on, getting the most out of everything. To what extent does this characterization describe you?

- ❍ Not at all 1 (1)
- ❍ 2 (2)
- ❍ 3 (3)
- ❍ 4 (4)
- ❍ A great deal 5 (5)

Happy2: How satisfied are you with your life as a whole these days?

- ❍ Very Dissatisfied (1)
- ❍ Dissatisfied (2)
- ❍ Neither (3)
- ❍ Satisfied (4)
- ❍ Very Satisfied (5)

A Descriptive Statistics SP Data

A.1 Scales

Table 9: Correlation check: multivariate regression - respondent characteristics and pro-organic behaviour

	Age		Children		Married		Female		Income	
Pro-Org Reason	Coef.	t-val.	Coef.	t-val.	Coef.	t-val.	Coef.	t-val.	Coef.	t-val.
Health	-4.376	-10.31	0.104	3.82	-0.043	-3.32	0.031	2.22	-38.227	-0.71
Taste	2.301	6.08	0.034	1.39	-0.089	-7.59	0.138	11.21	275.954	5.75
Safe	-2.277	-5.62	0.090	3.44	0.070	5.58	0.004	0.31	54.619	1.06
Fresh	-5.826	-14.94	0.170	6.76	0.062	5.13	-0.004	-0.35	262.040	5.29
An. Welfare	2.663	6.74	-0.043	-1.70	-0.028	-2.31	0.007	0.56	115.923	2.31
Env. Friendly	1.058	2.48	-0.177	-6.46	0.068	5.14	-0.046	-3.31	31.630	0.58
Local Production	0.693	1.86	-0.104	-4.33	0.058	5.00	-0.049	-4.09	80.565	1.70
Buy Org	-0.348	-0.90	0.133	5.33	-0.011	-0.91	0.041	3.31	-35.953	-0.73
Non GMF	-0.971	-2.50	-0.055	-2.20	0.064	5.36	0.074	5.90	32.883	0.67
Support Org	-2.201	-5.13	-0.001	-0.04	-0.037	-2.76	-0.020	-1.42	61.202	1.12
_cons	53.414	192.28	0.607	33.93	0.618	72.11	0.517	57.46	2085.840	59.14

Table 9 indicates the correlation between the main characteristics of individuals and the pro-organic reasons for buying organic. One can observe that age is negatively and significantly correlated with most of these, excluding taste, animal welfare, environmental friendliness and local production, which indicate more concerns about the environment and availability. In fact, table 9 shows the opposite for females, households with more children and higher income, i.e. they are more concerned with variables with potentially higher direct impact on their quality of life. It is worth mentioning that 67.9% of respondents indicated that they buy organic because it is environmentally friendly, the highest proportion, followed by those who state they buy them because of animal welfare (59.6%), and those who bought organic before and were satisfied with the product (56.4%). This can be seen in the last column of Table 10.

More generally, table 15 indicates that 'pro-organic attitude' is positively correlated with income, female, being married, and with children, whilst negatively correlated with age. Respondents with children would generally be 'pro-organic', with exception of pro-animal welfare (negative correlation) and pro-environmentally friendly (not significant) motivations. In addition, table 16 shows that individuals who stated that they are happy, who also are less likely to be on any special diet, do not buy organic because it is safe, and are less concerned about GMF. However, "happy individuals" are positively correlated with all other pro-organic behaviours. Respondents who have a healthy lifestyle, and those who are pro environment, also showed positive correlations with 'pro-organic behaviour'.

Respondents were also asked why they do not buy organic, similar to pro-organic, their responses were aggregated in order to identify the 'con-organic attitudes', i.e. negative perception about organic products. The experiment asked individuals ten questions about the reasons why they don't buy organic (Table 11). The main reason given for not buying organic is because they are perceived as expensive (77%) followed by not being able to recognize organic products in the shop (51%) and by a low variety (41%). Hence, the reasons for not buying organic are more associated with price and availability rather than the characteristics of organic product per se. Furthermore, most justifications for not buying organic do not show significant difference in the proportion of times in which respondents had chosen organic, as opposed to non-organic and no choice. The only significant differences were for those who stated that they don't buy organic because it is expensive, (4% lower buying frequency

Table 10: Buy Organic Characteristic: pro-organic

Why buy organic?	NonOrg	Organic	dif	St_Err	t_value	p_value	Sample %
Health	0.423	0.516	0.093	0.018	5.10	0.000	43.0
Taste	0.424	0.505	0.081	0.018	4.40	0.000	43.0
Safe	0.482	0.556	0.074	0.018	4.00	0.000	48.7
Fresh	0.339	0.424	0.086	0.018	4.85	0.000	34.5
Local Production	0.517	0.608	0.090	0.018	4.90	0.000	52.3
Buy Organic	0.559	0.640	0.080	0.018	4.40	0.000	56.4
Non GMF	0.422	0.505	0.083	0.018	4.55	0.000	42.8
Support Organic	0.403	0.475	0.072	0.018	4.00	0.000	40.8
Animal Welfare	0.592	0.652	0.060	0.018	3.30	0.001	59.6
Environmentally Friendly	0.676	0.719	0.042	0.018	2.45	0.015	67.9

in column ‘dif’.).

Table 11: Don’t Buy Organic Characteristic: con-organic

Why don’t buy	NonOrg	Organic	dif	St_Err	t_value	p_value	Sample %
Expensive	0.774	0.739	-0.035	0.015	-2.30	0.022	77.2
Not Advertised	0.360	0.393	0.033	0.018	1.80	0.069	36.2
Low Variety	0.414	0.435	0.020	0.018	1.10	0.267	41.6
Not Available	0.259	0.275	0.017	0.016	1.00	0.305	25.9
Cannot Recognise	0.515	0.503	-0.013	0.018	-0.70	0.477	51.5
Do Not Trust	0.266	0.257	-0.009	0.017	-0.55	0.586	26.5
Not Attractive	0.185	0.192	0.006	0.015	0.40	0.673	18.6
Habit	0.389	0.383	-0.006	0.018	-0.35	0.733	38.8
Not Known	0.317	0.321	0.004	0.017	0.25	0.797	31.7
Short Expiring Date	0.348	0.348	-0.001	0.018	-0.05	0.967	34.9

Similarly, table 12 shows results for the proportion of individuals who chose organic based on their approach to the environment, referred as ‘green behaviour’. Only four types of attitudes show significant difference in the proportion of individuals who choose organic, ranging from 2.5% to 4.2% (column ‘dif’). These are those who donate to environmental institutions, recycle, buy recycled items, and think that environmental policies should be prioritised. What may be also interesting to observe is that the behaviors mostly associated with ‘green behaviours’, are switching the lights off in empty rooms (92.9%) and using own bags while shopping (91.3%). The latter is supported by evidence as the introduction of a plastic bag fee in 2015 in the UK appears to have been very successful in changing behavior (Thomas et al., 2019).

Table 16 also indicates that pro-environmental behaviour is positively correlated with all pro-organic attitudes.

Table 13 shows the difference in responses from individuals with regards to their lifestyle, and indicates any differences in the proportion of those who chose organic in comparison with those who don’t. Only those who answered that a healthy lifestyle is important, those who take such an approach to reduce healthcare costs, an those who don’t drink fizzy drinks are suggested to show significant difference in the proportion of their choices for organic, all with positive effect on the choice for organic product. It might be interesting to observe that most people appear to associate a healthy lifestyle with healthy eating (90.3%) and not smoking (91.3%).

Table 12: Buy Organic Characteristic: green-behaviour

One should:	NonOrg	Organic	dif	St_Err	t_value	p_value	Sample %
Donate the Env ONG's	0.320	0.362	0.042	0.018	2.40	0.015	32.3
Buy Recycled	0.666	0.709	0.042	0.018	2.40	0.016	66.9
Recycle	0.888	0.913	0.025	0.011	2.20	0.028	88.9
Prioritise Env Policies	0.790	0.823	0.033	0.015	2.15	0.032	79.2
Use Public Transport	0.512	0.533	0.022	0.018	1.20	0.238	51.3
Close Tap When Brushing	0.799	0.810	0.011	0.015	0.80	0.438	80.0
Use Own Bags	0.912	0.918	0.005	0.011	0.55	0.588	91.3
Switch Lights Off	0.928	0.931	0.003	0.009	0.25	0.802	92.9
Unplug Devices	0.723	0.727	0.004	0.017	0.25	0.818	72.3
Use Better Insulation	0.880	0.879	-0.001	0.012	-0.05	0.951	87.9

Table 13: Buy Organic Characteristic: healthy lifestyle

Healthy Lifestyle	NonOrg	Organic	dif	St_Err	t_value	p_value	Sample %
Health is Important	0.882	0.904	0.022	0.012	1.90	0.059	88.3
HLS to Reduce Healthcare Cost	0.723	0.752	0.029	0.017	1.75	0.079	72.5
No Fizzy	0.743	0.770	0.027	0.016	1.70	0.094	74.5
Exercises are Important	0.860	0.880	0.019	0.013	1.55	0.123	86.1
Upset with Fast Food	0.426	0.454	0.028	0.018	1.55	0.125	42.8
Healthy Food is Important	0.902	0.918	0.016	0.011	1.50	0.138	90.3
No Fast Food	0.752	0.768	0.016	0.016	1.00	0.322	75.2
Low Fat/Sugar	0.637	0.647	0.010	0.018	0.55	0.571	63.8
Non Smoker	0.913	0.910	-0.003	0.011	-0.25	0.807	91.3
More Active (Less TV)	0.695	0.698	0.004	0.017	0.20	0.843	69.5

A.2 Correlation tables

The correlation tables in this section reinforce the results above. Table 14 shows that being female correlates positively and significantly with almost all reasons pro-organic with exception of local production. Age on the other hand correlates negatively and significantly with all pro-organic reasons. Elderly people, as opposed to young ones, do not seem to be equally enthusiastic about organic products. This might seem surprising but is something that has been also found in previous literature on organic consumption (Padel and Foster, 2005; Rimal, Moon, and Balasubramanian, 2005; Gschwandtner, 2018). Income correlates positively and highly significantly with all pro-organic reasons which is intuitive as a higher income enables purchase of organic food which is more expensive. The correlations for 'Married' and 'Children' are more mixed with both having positive and negative negative significant correlations. The negative correlations in the case of children might seem surprising but is in line with literature which shows that the impact of children on organic consumption is mixed. On one hand people with children would like to offer them a better nutrition, on the other hand the increase in number of family members reduces the available budget per person in the household (Arnoult, Lobb, and Tiffin, 2010). These results are also corroborated by the ones in Table 15 which shows correlations between socio-economic characteristics and 'pro-organic behaviour'.

Table 14: Multivariate regression: pro-organic behaviour (Y) and household characteristics (X)

Y:	Healthy		Taste		Safe		Fresh		An. Welfare		Env. Frindl.		Local Production		Buy Org	
	Coef.	t-val.	Coef.	t-val.	Coef.	t-val.	Coef.	t-val.	Coef.	t-val.	Coef.	t-val.	Coef.	t-val.	Coef.	t-val.
Female	0.061	6.17	0.149	15.05	0.045	4.56	0.038	4.00	0.052	5.17	0.031	3.32	0.016	1.63	0.072	7.20
Married	0.042	4.00	-0.035	-3.34	0.102	9.65	0.068	6.75	0.047	4.46	0.091	9.04	0.095	8.81	0.041	3.90
Age	-0.007	-20.62	-0.002	-6.40	-0.006	-19.10	-0.007	-21.34	-0.002	-5.57	-0.003	-9.76	-0.004	-12.73	-0.004	-11.26
Children	-0.005	-0.98	0.006	1.16	-0.009	-1.79	0.003	0.55	-0.022	-4.24	-0.040	-8.11	-0.032	-6.07	0.004	0.79
Income	0.000	6.94	0.000	13.63	0.000	6.36	0.000	9.62	0.000	9.00	0.000	7.44	0.000	7.30	0.000	6.58
_cons	0.668	31.59	0.393	18.47	0.674	31.56	0.558	27.57	0.583	27.23	0.732	36.15	0.631	29.16	0.649	30.38

Residual 10,626; overal significance: p=0.000

Results in Table 16 sow that being happy (columns 1 and 2) correlates positively and significantly with almost all pro-organic reasons except for safety and non-gmo which are negative but not significant. Being on a special diet correlates negatively and significantly with both happiness measures which is intuitive. Reasons against organic (ConOrganic) correlates negatively and significantly with 'HealthyLife' and 'ProEnvironment' which suggests that an 'pro-organic attitude' is related to a healthy lifestyle and a pro environmental attitude. Leading a healthy lifestyle on the other hand (column 3) correlates positively and significantly with most pro-organic reasons. All these correlations are intuitive and confirm the results obtained previously in the study and can be seen as robustness checks.

Table 15: Correlation: socio-economic and pro-organic behaviour

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Age	1.000											
	(0.000)											
(2) GenderFem	-0.239	1.000										
	(0.000)											
(3) Married	0.095	-0.007	1.000									
	(0.000)	(0.412)										
(4) Child	-0.319	0.188	0.181	1.000								
	(0.000)	(0.000)	(0.000)									
(5) Children	-0.271	0.167	0.185	0.846	1.000							
	(0.000)	(0.000)	(0.000)	(0.000)								
(6) HIncomeMid	-0.127	-0.023	0.199	0.141	0.155	1.000						
	(0.000)	(0.016)	(0.000)	(0.000)	(0.000)							
(7) ProOrgHealth	-0.19	0.101	0.019	0.109	0.069	0.099	1.000					
	(0.000)	(0.000)	(0.041)	(0.000)	(0.000)	(0.000)						
(8) ProOrgTaste	-0.088	0.150	0.002	0.084	0.085	0.133	0.515	1.000				
	(0.000)	(0.000)	(0.859)	(0.000)	(0.000)	(0.000)	(0.000)					
(9) ProOrgSafe	-0.175	0.088	0.091	0.123	0.093	0.102	0.587	0.435	1.000			
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)				
(10) ProOrgFresh	-0.233	0.104	0.061	0.151	0.133	0.136	0.465	0.532	0.461	1.000		
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)			
(11) ProOrgAnWe	-0.069	0.048	0.053	-0.023	-0.007	0.095	0.413	0.323	0.471	0.342	1.000	
	(0.000)	(0.000)	(0.000)	(0.012)	(0.421)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
(12) ProOrgEnFr	-0.099	0.048	0.082	0.013	0.015	0.088	0.477	0.365	0.492	0.32	0.611	1.000
	(0.000)	(0.000)	(0.000)	(0.165)	(0.100)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
(13) ProOrgLocProd	-0.114	0.033	0.057	0.037	0.032	0.093	0.461	0.365	0.384	0.409	0.393	0.448
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
(14) ProOrgBuyOrg	-0.130	0.085	0.049	0.092	0.094	0.084	0.529	0.448	0.473	0.402	0.441	0.457
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
(15) ProOrgNonGMF	-0.166	0.122	0.067	0.121	0.079	0.093	0.487	0.406	0.511	0.401	0.369	0.405
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
(16) ProOrgSupportOrg	-0.178	0.082	0.039	0.076	0.072	0.105	0.557	0.427	0.529	0.441	0.437	0.467
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Standard errors in parenthesis

Table 16: Correlation: lifestyle and pro-organic behaviour

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Happy1	1.000											
	(0.000)											
(2) Happy2	0.736	1.000										
	(0.000)											
(3) HealthyLife	0.052	0.063	1.000									
	(0.000)	(0.000)										
(4) ProEnvironm	0.057	0.053	0.587	1.000								
	(0.000)	(0.000)	(0.000)									
(5) ConOrganic	-0.035	0.011	-0.122	-0.167	1.000							
	(0.000)	(0.216)	(0.000)	(0.000)								
(6) SpecialDiet	0.106	-0.039	0.094	0.038	-0.007	1.000						
	(0.000)	(0.000)	(0.000)	(0.000)	(0.473)							
(7) ProOrgHealth	0.029	0.011	0.244	0.194	-0.166	0.045	1.000					
	(0.001)	(0.245)	(0.000)	(0.000)	(0.000)	(0.000)						
(8) ProOrgTaste	0.033	0.006	0.218	0.142	-0.235	0.045	0.515	1.000				
	(0.000)	(0.484)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)					
(9) ProOrgSafe	-0.015	-0.046	0.317	0.253	-0.133	0.082	0.587	0.435	1.000			
	(0.089)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)				
(10) ProOrgFresh	0.047	0.038	0.188	0.161	-0.132	0.089	0.465	0.532	0.461	1.000		
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)			
(11) ProOrgAnWe	0.007	-0.031	0.287	0.285	-0.093	0.021	0.413	0.323	0.471	0.342	1.000	
	(0.453)	(0.001)	(0.000)	(0.000)	(0.000)	(0.019)	(0.000)	(0.000)	(0.000)	(0.000)		
(12) ProOrgEnFr	0.012	0.032	0.288	0.322	-0.127	0.016	0.477	0.365	0.492	0.32	0.61	1.000
	(0.18)	(0.000)	(0.000)	(0.000)	(0.000)	(0.084)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
(13) ProOrgLocProd	0.097	0.063	0.264	0.291	-0.103	0.073	0.461	0.365	0.384	0.409	0.393	0.448
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
(14) ProOrgBuyOrg	0.069	0.047	0.205	0.225	-0.211	0.016	0.529	0.448	0.473	0.402	0.441	0.457
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.087)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
(15) ProOrgNonGMF	-0.005	-0.012	0.272	0.228	-0.145	0.09	0.487	0.406	0.511	0.401	0.369	0.405
	(0.600)	(0.182)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
(16) ProOrgSupportOrg	0.006	0.022	0.33	0.338	-0.219	0.108	0.557	0.427	0.529	0.441	0.437	0.467
	(0.478)	(0.016)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Standard errors in parenthesis

Attributes

Attribute	Description	Image	coding
Label	Organic label	EU UK	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

Figure 3: Attributes used in the choice experiment