Front of Pack Food Labels and dietary choice determinants: what works and for whom?

Mara Thiene*, Riccardo Scarpa, Alberto Longo, George Hutchinson

Mara Thiene[a] (mara.thiene@unipd.it)

Riccardo Scarpa[b]

Alberto Longo[c]

George Hutchinson[c]

(forthcoming in Food Policy, accepted on Sept. 14th, 2018)

[a]Dpt. of Land, Environment, Agriculture and Forestry (TESAF)
University of Padua
Viale dell'Università, 16
35020 Legnaro (PD) Italy
Tel. 049-8272760
Fax 049-8272750

[b]Durham University Business School, UK,
[b]University of Verona, Italy
[b]University of Waikato, New Zealand

[c]School of Biological Sciences, Queen's University Belfast, Northern Ireland

Acknowledgements

Funding for this research was provided by the UKCRC, grants number RES-590-28-0001 and MR/K023241/1, by the Department for Employment and Learning, Northern Ireland, and by the Spanish Ministry of Economy and Competitiveness through Grant No. ECO2014-52587-R.

Abstract

The introduction of an effective Front of Pack food labelling (FoPL) system is at the forefront of the food policy debate. Nutritional information is seen as an effective tool to help fight obesity and its associated co-morbidities, such as cancer and cardiovascular disease, for which unhealthy diet represent a major preventable risk factor. This paper explores the influence of FoPL formats on consumer's stated choice of weekly food baskets using data from a discrete choice experiment carried out in Northern Ireland in 2011. Two of the three baskets were experimentally designed while the third represented the respondent's actual current food choice (or status-quo basket). Four nutritional attributes were used: (*i*) total fat, (*ii*) saturated fat, (*iii*) salt, and (*iv*) sugar. Baskets were portrayed at different price levels to elicit the sensitivity of choice to price and to derive marginal willingness to pay estimates. Results from random utility models with various forms of heterogeneity reject the null of no association between preference classes and healthier food baskets and also the null of no effect of the nutritional information described. We find that the influence of the FoPL format used to convey nutritional information combines with selected socio-demographic covariates to determine membership to preference classes. A sensitivity analysis is used to validate the preferred model and the response sensitivity of selection probabilities to potential policy levers, such as a more realistic appreciation of self-body image and the habit of reading labels.

Key words: food choice, dietary habits, discrete choice experiment, Front of Pack food labels

1. Introduction

The UK and the Republic of Ireland, along with Luxemburg and Finland, are the four EU countries in the top 10 nations in the world for prevalence of obesity (WHO, 2015). In the UK, according to the *"cost of living and food survey"* the average adult body weight increased by 5.1kg between 1993 and 2014, when it reached 77.5 kg (The Economist 2016, August 13th). A high prevalence of overweight people is associated with a high incidence of a variety of serious life-style related non-transmissible diseases, such as type two diabetes, many types of cancer and cardiovascular conditions. The incidence of overweight is higher in older people. So, countries heading towards a larger share of aging population are expected to suffer more. Recent estimates from the U.K. National Health Service, for example, project the cost of direct treatment for diabetes to balloon over the next 25 years, moving from 10% of the NHS budget to 17% (NHS, 2012).

The growth of human body weight is not only a developed world problem, but it is a global phenomenon. A recent study by the NCD Risk Factor Collaboration (AAVV, 2016, Lancet) used over 19 million body measurements to compute body mass index (BMI) across 186 countries. Data was collected over the period 1975-2014 and shows that if current trends continue *"by 2025, global obesity prevalence will reach 18% in men and surpass 21% in women; severe obesity will surpass 6% in men and 9% in women"*.

At the national level, the UK official statistics (HSCIC 2015) predicts the current obesity trends to continue, showing increases with age, greater prevalence in men than women and among the lower-middle social class These statistics show that the causes are to be found in excessive energy intake, decreased rates of intense physical activity and more widespread sedentary lifestyles; all of which are further exacerbated by a generally unbalanced diet (especially outside the London area), at least when compared to the government recommended *"eat-well plate"* guidelines. All this reflects negatively on the national health care bill, which is already extremely high. Widespread preventive action is now urgently needed. The use of potentially useful market-based instruments, such as taxes on calorie-rich foods (fat-tax, sugar-tax, etc.), is still being debated. Which ways are effective to provide information to those consumers who most need it in order to nudge them towards healthier food choices remains a mostly unanswered issue, yet an answer is badly needed as labeling is still seen as the dominant tool in the policy arena.

To revert the weight gain tendency and in order to encourage healthier eating, the UK food and health authorities have embarked on a joint effort to promote nutritional information via adequate front of pack labels (FoPLs). Consumers' nutritional choices play a causative role in weight gain. Coupled with increasing consumer education, lowering the cost of information and interpretation of the nutritional consequences of food choices is seen by many as an essential component of any policy directed to stem and possibly revert the current trend. The information content of back of pack labels have been the subject of much regulation and studies, but the switch in emphasis to placing nutrition information on Front of Pack Labels (FoPLs) is mostly due to the perceived necessity to more forcefully attract consumer's attention to the health consequences of food choice. In the USA in 2011, FoPLs recommendations were published by the Institute of Medicine and also by the Grocery Manufacturers Association and Food Marketing Institute, who started their own labelling scheme. In October 2012, the UK FSA announced a voluntary scheme for FoPLs, which was to be put in place by 2014.

Since December 2016 nutritional information have become mandatory on back of pack labels of pre-packed food in the UK. Such information may be repeated in the FoPLs, but this is still a voluntary initiative, which complements the already mandatory labelling information required by the EU Food Information Consumer regulations 1924/2006 and 1169/2011. To promote adoption, a guidance document for creating FoPLs for pre-packed food sold by retail outlets was published in June 2013 by the Department of Health. This was collated following several studies conducted between 2001 and 2013 designed to understand what particular form of FoP labelling is most fit for purpose. The document is part of a series of policy actions taken to encourage voluntary adoption by the UK food industry. Such actions started in 2014, and it is hence still too early to draw conclusions on their effects on health or weight change in the population. Will these voluntary initiatives affect dietary habits and, for example, decrease obesity and other diet-based non-communicable diseases? Will the evidence constitute a legitimate base for compulsory policy in the UK and possibly elsewhere? Epidemiological studies will provide an answer to such important questions in the years to come. But some preliminary evidence can be gleaned from patterns of choices using experimental choice design, as done in the present study.

A whole body of research from nutritionists dictates the nutritional categories that provide salient dietary information to consumers, such as sugar, fat, saturated fat and salt contents of each food package relative to the guideline daily amounts (GDA). Several experimental cognitive studies in food consumer research have explored the communication effectiveness of labels. Results have supported the use of specific types of FoPL, on the basis of their ability to attract consumers' visual attention better than others. For example, by comparing mandated nutritional information (the nutritional Facts Panel, NFP) in the US and FoP nutritional labels, Becker *et al*. (2015) found that FoPL were attended earlier, more often and that the use of colours increased attention to labels.

Consensus seems to indicate that FoPL should have chromatic elements and it might work best if combined with other succinct recognizable signals, such as health certificates (see Bialkova *et al.* 2013, Hersey *et al*. 2013). While the effect of socio-economic covariates have also been studied, these focussed on the use of nutrition information from food labels during meal planning (Nayga 1996, 1997) at home or when comparing brands when shopping (Nayga *et al.* 1998). In general, these studies showed the importance of education, along with other factors. However, fewer studies explored whether specific FoPLs affect how healthy consumers' food choices are. Fewer still have done so while accounting for age, perceived weight, education, marginal utility of income and other consumer characteristics relevant for the evaluation of social impact of policy. Yet, this information seems crucial in the overall evaluation of a mandatory FoPL policy, or even of a voluntary labelling initiative. With this study we try to fill this research gap. We recognise that the range of factors affecting food choice is ample and articulated and that these have been the subject of investigation for a long time within several disciplines (see for example Pollard et al., 2002 and Raghunathan et al. 2006).

The hypothesis we investigate here is that, faced with alternative types of nutritional signals in FoPLs, consumers will be affected differently depending on their latent taste segment and on their body weight status. Such latent segmentation and differential effects on choice would provide some insight with respect to the effectiveness of nutritional signals in FoPLs.

While awaiting clearly interpretable clinical data from randomised trials, which can be persuasively used to drive and design the food policy for FoPLs in the UK and elsewhere, some interim insight can be derived from hypothetical food choice studies. In this paper we present results of a survey using discrete choice experiment data. We extend the findings reported in the original Food Standard Agency 2012 report, the results of which were used to issue guidelines by the Department of Health (2013). In fact, the original report documented extensively the degree of comprehension of alternative FoPLs (text only, traffic light systems, GDAs and mixtures thereof), but fell short of establishing the link to healthier food choice by those who most need to make them. Our study provides results that corroborate the original report by systematically linking FoPL types to specific consumer profiles, and to healthier food choice. Our results further show that relevant self-reported factors such as self-image perception, BMI, gender, frequency of reading labels and age are differentially associated with preference groups and with healthy food choice. The main shortcoming of this study is that with the exception of the status quo basket it relies on quite abstract and hypothetical rather than real food choices. Yet, the results are sufficiently strong to motivate further experimental research on real food choice behaviour of alternative FoPLs thereby informing evidence-based policy design.

The rest of the paper is articulated as follows. Section 2 reports on the state of knowledge and on the underlying research in FoPL, highlighting the research gaps that our study fills, with an emphasis on defining the broader research strategy enabling the design of an effective labelling policy. Section 3 reports the survey design, the data and the methods of analysis used in our study. We use a mixed logit design that layers discrete and continuous mixing and explore 4 separate FoPLs. Section 4 provides a thorough discussion of the findings and of model validation, while Section 5 concludes by indicating the way forward in research design to inform policy actions.

2. Front of Pack Nutritional Food Labelling: a summary of relevant research

Starting from the seminal work by Asam and Bucklin (1973), the use of food nutritional labels by consumer has been the focus of literally hundreds of consumer studies. Several reviews on the issue are available, both for the US and the EU (Balcombe *et al*. 2010, Hawley *et al.* 2012, Soederberg Miller and Cassady 2015). Therefore the following review is quite selective. An early review of six studies (Jacoby *et al*. 1977) concluded that *"most consumer neither acquire such information when making a purchase decision nor comprehend most nutrition information once they receive it"*. In response to this and several other studies that showed very low use of nutritional labels by consumers (as low as 20% in the US), Klopp and MacDonald (1981) asked why this should be the case to a sample of Wisconsin shoppers. They found that less educated consumers tended to make significant lower use of labels and spent shorter time in food planning. So did consumers with lower self-assessment of nutrition knowledge.

Over thirty years later, Nørgaard and Brunsø (2009) reached similar conclusions in a study of families; they state that: *"Parents seldom use nutritional information when they seem to sense an overflow of information, information that is too technical and a problematic presentation of energy distribution, and/or when their health consciousness is limited"*, suggesting that *"parents [are] more likely to prefer food labels with concise information and more visual aspects"*. Such need for simplification had also emerged from a review of 58 studies conducted between 2003-2006 in the EU-15 by Gruner and Wills (2007). Given the importance of visualization of nutritional elements to guide healthy diets, and the necessity to provide such information to consumers in a succinct, yet clear manner, interventions have been devised to place these on FoPLs, which is in the immediate field of vision (i.e. FoPLs), rather than relegating them to the back of the pack labels. In 2012, according to the UK Food Standard Agency (FSA), approximately 80% of pre-packed processed food products sold carried nutrition information on FoPLs. Previous work by Malam *et al.* (2009) found that UK consumers were to some degree confused and distracted by the diversity of existing FoPLs, due to the difference of interpretive elements. In an analysis of the information impact of such elements they concluded that using a text scale (high, medium, low) had the greatest impact on comprehension. They further recommended that combining text with traffic light colour coding and percent of guideline daily amounts (GDAs) enabled more consumers to make healthier food choices, partly because the normative signal was more reinforced by traffic light colours. The study did not elaborate as to whether or not those in most need to correct their diets (e.g. overweight subjects) were differently affected by the various FoPLs. Based on this and other studies, in March 2010 the FSA board encouraged food businesses to use all three elements to signal nutritional amounts: (1) colours from the traffic light system (red, amber and green) or TLS, (2) text signals (high, medium or low) or TXT and (3) percentage Guideline Daily Amounts (% GDAs) in order to enable UK consumers to interpret nutritional information (FSA 2010). Furthermore, the board highlighted that the FSA does not support FoPLs using only % GDAs, but that these should be combined with either traffic light colours or text, and should ideally have all three elements. Finally, consumers seem to value FoPLs, as results from a willingness to pay survey across EU countries shows (Gregori *et al.* 2015).

The two most common FoPL elements currently adopted in the UK market place are GDAs—developed by the food industry—and TLS, developed by the FSA. But combinations of the two styles are commonplace and often include basic text signals too. These two most common labelling formats are discussed further below, but it is worth noting that there are other initiatives more specifically directed at fighting the problem of an increasingly overweight population. For example, the "activity equivalent calorie labelling" recently promoted by the Royal Society for Public Health (RSPH), which claims that nutrition information signalled by using equivalence of physical activities are best understood by most.

i) Traffic Light System (TLS Format)

Independent research by the FSA has investigated FoPL extensively and produced a large body of literature (see Synovate, 2005). Following reviews published in 2005, the FSA concluded the Traffic Light System (TLS) to be the most effective FoPL label to enable consumers to make informed dietary choices about food products. The TLS is a FoPL which informs and warns consumers on the nutritional content of processed foods indicating the amount of calories, fat, saturated fat, salt and sugar of processed foods per 100gr by assigning colour-coded levels: high content is something to be warned about, and hence is red; medium content is less worrisome and it is amber; and low content is the way to go, and hence is green.

Early studies based on eye-tracking experiments (Jones and Richardson 2007) showed TLS to be relatively more effective at attracting attention. Some literature (Hodgings *et al*. 2012) classify this system as a semi-directive system, as it provides behavioural normative content rather than neutral information as opposed to nutritional table of content, for example. TLS labels have been shown to perform well in attracting attention, even when consumers have limited time and have specific goals (van Herpen and van Trijp 2011). Recent neurological investigation using MRI scan on subjects during choice with different FoPLs provided evidence that *"salient traffic light labels influence the valuation of food products by [activating] a [brain] region implicated in endogenous and exogenous self-control and its connectivity"* (Enax *et al.* 2015).

Other research supports the use of colour indicators. For example, research by Feunekes *et al.* (2008) support findings by the FSA in that the multiple TLS was the easiest FoPL to comprehend. Epstein *et al.* (1998) also provide evidence that diets based on the TLS can help reduce levels of obesity. Andrews *et al.* (2011) found that the combination of TLS-GDA is more desirable in terms of food choice outcomes than the single summary indicator "Smart choices" used in the US. Thorndike *et al.* (2012) found that a simple colour coded labelling intervention increased sales of healthy items and decreased those of unhealthy ones. More recently, Crosetto *et al.* (2016) found that GDA performs better than TLS when subjects do not face time constraints, but when time is limited TLS outperforms GDA with an increasing number of nutritional goals.

However, there exists conflicting evidence suggesting that the TLS is not the most accurate or desirable information format to convey nutrient levels in food (Grunert and Willis 2007; Hodgkins *et al.* 2012). The objection is linked to the red colour being potentially interpreted as "no go" signal, which might lead to systematic under-supply of some important nutrient groups, such as important fat categories.

ii) Percentage Guideline Daily Amounts (GDA Format)

The GDA scheme typically shows the fat, saturated fat, sugar and salt per portion of the food and indicates the percentage the portion contributes to GDA. It is important to note that GDAs are a guide, not a target, to how much energy and key nutrients the average healthy person needs in order to achieve a balanced diet. They are based on the 'average' adult. However, physically active people will have higher requirements, and smaller people, like children, will have lower ones. Note that similar acronyms exist. For example, RDAs (recommended daily amounts) were set by the Department of Health in 1979 for nutritional requirements for different population subgroups. In 1991 the Department of Health replaced these with DRVs (dietary reference values), which was a comprehensive term covering criteria for nutritional and energy intakes. DRVs are only to be used as guidelines and are for healthy people. DRVs are commonly reported as recommended daily intakes or recommended daily amounts. Current nutrient recommendations are given in FSA Nutrient and food based guidelines for the UK (2007).

2.1 Studies on the effect of FoPLs and food choice

Discrete choice experiments (DCEs) have a recent and successful history in evaluating consumer preferences for food labels and their content. Gracia *et al.* (2009) employ DCE data and found that consumers were willing to pay more for a nutritional facts panel than a simple nutritional claim. Balcombe *et al.* (2010, 2015) design a DCE based on the TLS to examine the relationship between nutritional food labels (with colour indicating level of nutritional content) and price. Their results seem to indicate that utility is improved more when moving from red to amber (i.e. when remedying potential loss) than when moving from amber to green (i.e. when achieving potential health gains), which suggests a form of gain-loss asymmetry, also apparent in our results, albeit in different form.

Empirical studies of effects of FoPLs on food choice while monitoring eye-tracking have also shown that *"Adding both health marks and traffic light colours (v. traffic lights only) to numeric nutritional information* *produces favourable outcomes from the perspective of public health"* (Koenigstorfer *et al.* 2013), thereby providing grounds for the study of interaction effects on choice, which we undertake here. This is important because there is a tenuous line between striking the right balance with a synergistic combination of displays and over-cluttering, as shown in visual search studies (Bialkova *et al.,* 2013).

Aschemann-Witzel *et al.* (2013) also studied the effect on healthy food choices of nutritional label format in Poland and Germany, but in the context choice sets of varied size. Their results show that colour coding is more effective than simple text in inducing healthy choices when the choice set is large. Consumers perceived that colour coding was enabling them to make healthier food choices when asked to do so, but label format had no effect when consumers were asked to choose only on the basis of their personal preferences.

Effects of coloured and monochrome GDA labels on healthy choices were investigated in an eye-tracking study by Bialkova *et al.* (2014). They found an effect of nutrition labels on choice via consumer attention, which was attracted most by colour GDA. The effect of monochrome GDA FoPLs on consumer choice has recently been assessed (Boztug *et al.* 2015) using scanner data. The study concludes that *"the GDA label introduction reduces attraction of unhealthier products in terms of market share but does not affect product choice behaviour"*, as a consequence the authors *"agree that GDA labels are generally insufficient to adjust consumer behaviour towards healthier alternatives"*.

In closing this review we briefly touch upon studies on the segmentation of food consumers into types and their reaction to alternative nutritional label information. While it is well-established in the literature that antecedent volition (i.e. pre-established goals) (Swait 2014a, 2014b) is a natural driver of the influence of additional information on choice, relatively few studies have looked at latent segments and how they related to nutritional values and health in food choice. Visschers *et al.* (2013) conducted a cluster analysis of nutrition information use from nutrition tables in labels in relation to consumer's health and nutrition interest. They identify 4 segments, but conclude pessimistically with regards to the outlook with which improvement of nutrition labels is likely to stimulate nutrition information usage among consumer types.

From our literature review the issues of interaction effects between label formats that can be jointly used, their effect on latent consumer segments, and especially on obese consumers, all emerge as research topics worthy of further investigation. Our study was designed to cast some light on these issues by an adequate use of DCEs data.

3. Survey and Data

To facilitate the development of the methods section we first illustrate the survey with which we generated the food choice data. In a discrete choice experiment (DCE) respondents are faced with the task of choosing between several experimentally designed alternatives. Using the recorded choices from the experimental design analysts retrieve the underlying preference structure using adequate behavioural theories and statistical models. This method was chosen for this study as it most closely replicates real food choices in a hypothetical setting. In a grocery shop consumers buying their weekly food basket continually compare and evaluate food items on the basis of their taste, previous experience and label information.

3.1. Survey details

The development of the DCE survey instrument followed a lengthy, systematic process, consistent with the recommendations from the literature. The various stages involved a literature review, expert consultation, focus group research and pilot study, prior to fielding the main questionnaire to collect the final data (full details in Brown, 2014).

Three preliminary focus groups were held to understand the role of FoPLs in food choice. Early versions of the questionnaire were tested in further focus groups and individual interviews. These were followed by an in-depth test of the questionnaire with a pilot study of 32 respondents. Information was collected on respondents' attitudes towards food and on their personal characteristics to help explain responses to the choice experiment exercise.

In order to elicit the effect of price on food choice, price was also a descriptor of the alternative food baskets evaluated in each choice task, which included two differently priced baskets of weekly food shopping to be compared with the current status-quo food basket, self-reported by each respondent. The focus on the weekly packaged food basket (i.e. a collection of packaged foods bought in a regular week of grocery shopping) was dictated by the fact that limiting the attention to a single product would inevitably restrict the external validity of the results across food products. This choice imposes its own cost in the form of diminished realism of the hypothetical choice scenario, which in our eyes seems the lesser of two evils. Nutritional contents were conveyed in terms of four types of front of pack nutritional food labels. The use of an individual-specific status-quo alternative follows recommendations from recent studies (e.g. Marsh *et al.,* 2011; Boeri *et al.*, 2013; Grisolia *et al*. 2013, 2015). Since baseline diets differ across respondents, it would be arbitrary to present all respondents with an identical status quo. The individual elicitation of the status-quo food basket was achieved by presenting respondents with a visual aid based on food cards from which the assortment of the usual packaged foods bought by the respondent was identified. Such cards were designed based on a protocol developed with assistance from experts in food nutrition and psychology. A systematic approach was taken to ensure consistency and accuracy. Extensive testing was carried out in individual interviews and further tests were conducted during the formal pilot study. Prior to fielding the main survey, example food cards were checked by health professionals (these included registered NHS dieticians and nutritionists working in an academic capacity) to ensure satisfactory representation of foods and nutritional levels from an expert perspective. An example food card was created for each nutritional attribute. Each card displayed a range of foods in categories of high, medium and low according to the content of the nutrient in question in a wide range of food products (See examples in the Appendix). These were displayed to respondents at the moment of the identification of the individual usual weekly basket (status-quo basket), and used to assign to the reference baskets their respective nutritional classifications. See the appendix for examples.

3.2 Sample and survey

The sampling frame included all residents of Northern Ireland. The sample was drawn using stratified quota sampling using wards within electoral districts in Northern Ireland. Specifically, a two stage sampling process was used. Stage one involved a random selection of wards in Northern Ireland within geographic areas. These were selected so as to provide both urban and rural sub-samples. Samples drawn from each ward were proportional to the overall population in the ward. Stage two involved a quota sample within each of the selected wards. Quotas were assigned according to age, gender, socio-economic classification so as to match known demographics based on Census data and mid-year population estimates from the Northern Ireland Statistics and Research Agency. The survey was administered between December 2010 and March 2011, using face-to-face computer assisted personal interviews (CAPI). It was conducted by professionally trained and experienced market-research interviewers.

3.3 Alternatives and choice tasks

The discrete choice experiment consisted of a panel of 16 choice tasks per respondent. In the choice tasks alternatives were presented as “your current weekly basket” (the status quo weekly basket as described by the respondent), “Food Basket A” or “Food Basket B”. Given our concern with an individual's whole diet, we found it desirable to frame the alternatives in terms of “your weekly food basket”. Findings from focus groups and individual interviews confirmed that presenting the alternatives in terms of a weekly shopping basket was easily conceptualised by respondents. Indeed, the concept of a basket has been used successfully in previous food choice studies (Balcombe *et al.*, 2010). The Integrated Household Survey (IHS) includes a section known as the Living Costs and Food (LCF), which records weekly consumption and expenditure for each item of food in the average UK food basket (DEFRA 2010). Previous data from DEFRA surveys has been used in economic analysis regarding food choice. For example, Pretty *et al.*, (2005) carried out an assessment of the full cost of the weekly food basket in relation to farm costs and food miles.

3.4 Packaged Food Basket Attributes

Selection of relevant attributes to describe the alternative FoPLs is important in the design of the DCE survey. Care should be taken to reduce the cognitive burden on respondents (Powe *et al.*, 2005). Attributes selection was based on expert consultations, literature review and findings from our focus groups. Apart from the price attribute, four nutritional attributes were selected, specifically: sugar, fat, saturated fat and salt. The attributes and their levels are described in Table 1.

The four nutritional attributes had common reasons for inclusion in the survey: (*i*) all are typically reported on back of pack nutritional food labels; (*ii*) there are associated health implications with a diet exceeding guideline daily amounts (GDAs) in any one, some or all of these nutritional attributes; (*iii*) healthy eating advice from the UK government groups these nutrients together—saturated fat, fat, salt and sugar—stating that all healthy individuals should consume a diet that contains ‘moderate’ amounts of each of them; (*iv*) all can be used as indicators for taste, which typically has a strong influence on food choice.

The price attribute was specified for each basket and presented as a percentage increase, decrease or no change to the respondent’s defined current weekly food basket, which acted as a subjective reference point. Percentage changes were 50% and 20% from the price of the current food basket in each direction. The pre-testing results indicated that respondents' found this to be acceptable in terms of both payment vehicle and amount. The price range variation was informed by the report by the UK office of national statistics on family expenditures (Family Spending 2009).

3.5 Experimental Design

As in many choice experiment applications, our number of attributes and their levels result in a full factorial with too large a number of choice set combinations to have them all evaluated by respondents, let alone to have sufficient replicates to assess taste heterogeneity across respondents. So, an experimental design criterion is used to assign specific fractions of the full factorial to each respondent in a manner that all the effects with a-priori relevance are identified. Apart from identification, the design typically generates an allocation plan such that the choice data ensure a statistically efficient estimate of a random utility model (Ferrini and Scarpa 2007). That is, under a-priori assumptions the design produces estimates minimizing expected variance of estimates. However, several other criteria aside from efficiency are possible (see, for example Rose and Scarpa 2008).

Efficient experimental designs have come to the fore in recent years. Bayesian efficient designs, as employed in this study, can be used to accommodate uncertainty associated with assumed prior parameter values. Various criteria are used to determine the efficiency of the design. D_b error minimization is the most common criteria and the one used here. In a Bayesian efficient design the efficiency of a design is evaluated over a number of different draws taken from the prior parameter distributions assumed in generating the design (Ferrini and Scarpa, 2007; Scarpa *et al.*, 2007; Bliemer *et al.*, 2008). The efficient experimental design was generated using the software package Ngene, which is a standard in this field.

3.6 Nutritional label treatments

To uncover the differential effects due to the accumulation of the four nutritional signals in the label formats, respondents were randomly assigned to the following treatments: (*i*) *FoP label with text only* (TXT) (high, medium or low). For example, if a basket of goods is labelled "high" for the respective nutrient (fat, saturated fat, salt or sugar) this means that it is considered to have high levels of the respective nutrient per 100gr servings; "high" is interpreted as most unhealthy while "low" is considered the healthiest, with "medium" in between; (*ii*) *FoP label using multiple traffic lights* (MTL) adds a chromatic signal (red for high, amber for medium and green for low) to the text signal for each nutrients in the basket; (*iii*) *FoP label using Guideline Daily Amount* (GDA) rather than traffic light colours, this format adds to the text the GDA percentages; (*iv*) *Integrated FOP label format* (HYB). Both traffic light colours and GDA percentages are combined into a hybrid signal for each nutrient, on top of the text. Examples of food baskets are reported in Figure 1. Respondents had already defined their status quo level of these nutrients from their actual food purchase (See show cards in the Appendix) In terms of information load one expects HYB to be superior to all others, and TXT to be inferior to all others, with MTL and GDA to have intermediate effects, possibly different in size according to whether chromatic or percentage information result as most effective. The impact on healthy choice may, or may not correlate to information load, and this issue is part of our investigation.

3.7 Socio-economics covariates

Given our intention to test the role of a number of socio-economic variables in explaining taste latencies and sensitivity to FoPLs types by weight sub-samples, several covariates were also collected to be used in estimation of the choice probability model. The first two are age and gender as they are well-known determinants of food choice. These were followed by two additional variables related to individual body mass index (BMI) and self-body image. BMI was calculated based on data each respondent provided in terms height and weight. With regards to self-body perception, respondents were asked the following question: *"When you think of your ideal body weight, would you say you are currently: a lot over, a little over, about ideal, a little under, a lot under."* A last question investigated the level of engagement in terms of acquiring information; respondents were asked to answer the following question *"How often do you read* *these front of pack food labels when you are buying food: never, rarely, occasionally, usually, always, don't know/can't remember".*

4. Research questions, theory and methods

In this empirical study we set out to answer the following policy-relevant research questions:

1) Do food basket choices relate to latent preference classes with different propensity to select healthy food baskets?
2) Do FoPL formats determine probabilistic membership to such classes?
3) Is there a residual heterogeneity within classes which can further explain within-class taste variation for some food attributes?
4) Are choice predictions valid from the viewpoint of their plausibility with self-reported height/weight data (BMI) and other socio-economic variables in the sample data?
5) Are there policy-relevant differences in the way FoPLs formats affect the propensity to select healthy food basket? In other words, do various FoPLs affect the propensity of subjects to abandon a reference basked to select a healthy food basket? If so, how?

More specifically, the aim of the study is to account for the role of FoPL formats on packaged food basket choice via the existing latent differences across respondents' taste and ability to discriminate between alternatives (latent taste and scale classes). So, to simultaneously account for preference heterogeneity and varying levels of multiplicative correlation (often defined as error scale) in a tractable manner, we use both forms of preference mixing, continuous and discrete. To do so we specify choice probabilities using a latent class (LC) logit model, but a subset of taste coefficients, after testing, are also assumed to be continuosly random within preference classes. We name this a latent class random parameter logit model (LC-RPL) (amongst others Bujosa *et al*. 2010, Hess *et al*. 2012, Franceschinis *et al*. 2017) .

We denote the latent preference classes with c and the latent multiplicative correlation classes with s. Conditional on belonging to a specific c,s-latent class combination, a consumer's chooses the favorite food basket i from a set of $j \in J$ mutually exclusive alternatives, with J = 3. The probability of this choice is characterized by different profiles for nutritional attributes (weekly food baskets) and types of information display in the FoPL. Nutritional attributes report high, intermediate and low levels of fat, sugar, saturated fat and salt, and include the cost of the food basket.

Respondent n is asked to choose her favorite food basket in a panel of T=16 experimentally designed choice tasks nt. Following the conventional random utility (RU) maximization approach (Thurstone 1927, Manski 1977), each respondent n is assumed to select the utility-maximizing food basket from the set. For a respondent n with a particular combination of preference-class c and scale-class s, the indirect utility of alternative i in choice task t is denoted by $V(\lambda_s, \boldsymbol{\beta}_c, \mathbf{x}_{nit})$, and the overall total utility includes a random component ε i.i.d. Gumbel:

$$U_{nit|gc} = V(\lambda_s, \boldsymbol{\beta}_c, \mathbf{x}_{nit|gc}) + \varepsilon_{nit|sc}, \qquad (1)$$

where $\mathbf{x}_{nit|sc}$ is the vector of five food attributes, described by their respective levels; $\boldsymbol{\beta}_c$ is a vector of preference-class utility coefficients to be estimated and λ_s is the scale-class specific value for the scale parameter[1] (multiplicative correlation factor).

Because of the assumption on the stochastic component, the probability for a consumer n belonging to latent class combination s,c of choosing alternative i over alternative j in the choice set nt is given by a multinomial logit model (McFadden 1974):

$$\mathrm{Pr}_{nit|sc} = \frac{\exp(\lambda_s \boldsymbol{\beta}_c' \mathbf{x}_{nit})}{\sum_{j=1}^{J} \exp(\lambda_s \boldsymbol{\beta}_c' \mathbf{x}_{njt})} \qquad (2)$$

The RUM latent class choice model is characterized by a discrete mixture of choice probabilities, over a finite number of c preference classes and s scale-classes, each of which shows a homogenous choice behavior (Provencher et al. 2002, Boxall and Adamowicz 2002, Hensher and Greene 2003, Scarpa and Thiene 2005). It follows that the mixing distribution $f(\boldsymbol{\beta})$ is discrete, with a random parameter vector $\boldsymbol{\beta}_c$ denoting a finite set of c different vector values. There is a fairly active debate on how to adequately account for the potentially confounding role of the scale/multiplicative correlation parameter of the Gumbel error (Burton *et al.,* 2016).

[1] There has been a debate addressing the potential confounding between scale and taste heterogeneity (Hess and Rose, 2012). Since the use of the term "scale parameter" has become established in the literature, we also use it here, but warn the reader to interpret it as a factor able to capture multiplicative correlation, and direct readers to the recent clarification note by Hess and Train (2017) for further details on its correct interpretation.

The importance of the scale parameter was first raised by Swait and Louviere in their seminal paper (1993), who argued that respondents do not necessarily display the same level of certainty when making choices. Louviere and Eagle (2006) pointed out that ignoring the scale factor may confound heterogeneity in preferences with heterogeneity in error variance, thereby potentially obtaining biased estimates. Recently, various approaches were implemented to address variation in taste and its correlations via the scale parameter (Keane 2006, Fiebig *et al.* 2010, Scarpa *et al.* 2012, Hess and Rose 2012, Thiene *et al.* 2015; Hess and Train, 2017).

The probability of observing a choice sequence, conditional on being in scale class *s* (i.e. on a given degree of discrimination) and preference class *c* is:

$$\Pr(y_n|s,c) = \prod_{t=1}^{T_n} \frac{\exp(V_{nit|sc})}{\sum_{j=1}^{J}\exp(V_{njt|sc})} = \prod_{t=1}^{T_n} \frac{\exp(\lambda_s \boldsymbol{\beta}_c' \mathbf{x}_{nit})}{\sum_{j=1}^{J}\exp(\lambda_s \boldsymbol{\beta}_c' \mathbf{x}_{njt})} \qquad (3)$$

We hypothesize that for each latent class significant food attributes effects are estimated in the class specific utility function. Formally, this implies λ_s and $\boldsymbol{\beta}_c$ be different from zero for all scale classes *s* and taste classes *c*. Rejecting the null implies a positive answer to part of research question 1) above. The other part (i.e. whether they relate to healthier food choice) depends on the specific value estimates for $\boldsymbol{\beta}_c$.

For each latent preference class *c* and scale class *s*, membership probabilities are defined via a multinomial logit approach, with class-specific constant α_c:

$$\pi_{c,s} = \left[\frac{\exp(\alpha_c+\alpha_s+\boldsymbol{\gamma}_c'\boldsymbol{z}_n)}{\sum_{c=1}^{C}\sum_{s=1}^{S}\exp(\alpha_c+\alpha_s+\boldsymbol{\gamma}_c'\boldsymbol{z}_n)}\right] \qquad (4)$$

where $\boldsymbol{z}_n$ is a vector of covariates of respondent *n*, $\boldsymbol{\gamma}$ the vector of associated parameters, α_c and α_s are class-specific constants and must sum to zero for identification. In our investigation, key determinants of preference class membership are types of FoPLs, along with the individual characteristics, especially those related to health issues and the conventional socio-demographics.

We hypothesize that for each latent class significant membership determinants are estimated in the class specific membership probability function. Formally this implies that the elements of the vector $\boldsymbol{\gamma}_c$, as well as the preference and scale-specific intercepts α_c, α_s be different from zero for some scale classes *s* and taste classes *c*. Rejecting the null implies a positive answer to part of research question 2) above. The other part (i.e. which specific determinants relate to healthier food choice) depends on the specific value estimates for $\boldsymbol{\gamma}_c$.

The unconditional probability of a sequence of choices over all classes is:

$$\Pr(y_n) = \sum_{c=1}^{C} \sum_{s=1}^{S} \pi_{c,s} \prod_{t=1}^{T_n} \frac{\exp(\lambda_s \boldsymbol{\beta}_c' \mathbf{x}_{nit})}{\sum_{j=1}^{J} \exp(\lambda_s \boldsymbol{\beta}_c' \mathbf{x}_{njt})} \quad (5)$$

Previous studies using finite mixture of preference classes found that allowing for further heterogeneity within each preference class, by means of continuously varying random parameters, produced significant increases in model fit (Bujosa *et al*. 2010, Hess *et al*. 2012, Greene and Hensher 2013, Campbell *et al.* 2014, Boeri *et al.* 2014, Farizo *et al.* 2014, Yoo and Ready 2014, Franceschinis *et al*. 2017). There is no *a-priori* strong rationale for negating this occurrence in our data. On the contrary, respondents belonging to the same preference class are expected to show some continuous form of variation in preference for some sub-set of attributes with random coefficients $\widetilde{\boldsymbol{\beta}}$, while maintaing the shared values within the class for the other coefficients. So, we estimate a latent class model that accommodates in the vector of utility coefficients some continuously random coefficients. This allows for continuous heterogeneity of tastes across respondents within the same preference class. The unconditional choice probability than becomes:

$$\Pr(y_n) = \pi_{c,s} \prod_{t=1}^{T_n} \int_{\beta} Pr_{nit} f(\widetilde{\boldsymbol{\beta}}) d\beta \quad (6)$$

Specifically, in our case, an extensive specification search showed that the utility coefficients for the current food basket (i.e. the status quo), high level of fat and high level of salt are best specified as continuously random within each preference class[2]. Normal distributions are assumed for such random parameters in each preference class, such that $\widetilde{\boldsymbol{\beta}} \sim N(\boldsymbol{\mu}, \boldsymbol{\Omega})$ and $\boldsymbol{\mu}, \boldsymbol{\Omega}$ are the subject of estimation from the DCE data.

We hypothesize that at least some of the taste parameters within classes have specific hyperparameters $\boldsymbol{\Omega}$ of their continuous distribution that are significantly different from zero. Rejecting the null implies a positive answer to research question 3) above.

[2] We engaged in a specification search exploring all sets of random utility coefficients. The reported model is the one with best improvement in model fit. A mixed logit with all random coefficients (normally distributed) except for price and full correlation gives an AIC of 22,643 which is much higher than what found in our favorite model: 17,002.

From the normative viewpoint the question we hope to answer relates to whether specific FoPL associate themselves with preference patterns (i.e. latent classes) more or less likely to induce healthy food choices. For example, a preference structure systematically favouring selection of tastier food baskets with high levels of salt, fat and sugar is bad for health. Given the broad heterogeneity documented in the food taste literature, we must account for other systematic differences associated with individual-specific variables. For example, standard socio-economics (age and sex), self-perception of body weight (how this departs from the ideal) and more objective body weight measures (BMI) and their correlation with self-image.

In the model validation section, the effects of systematic exposure to specific FoPL is explored, at the individual respondent level, in terms of differences in predicted marginal probabilities of membership to classes with differing propensity to select healthier food baskets. This analysis highlights what FoPL formats increase membership to given taste classes and hence the propensity of healthier food choice; and from what other preference classes these increases are drawn. This provides an answer to research question 4) and to part of question 5).

Finally, to specifically answer research question 5), exposure effects to FoPL formats are also explored in a more direct form by comparing the differences in predicted choice probabilities when the choice task contains two alternatives: the status quo basket of each respondent and the basket with the healthiest attribute profile across FoPL (the one with lowest levels of sugar, salt, fat and saturated fat) when both are offered at the same price[3]. A larger positive absolute value difference between the two predicted probabilities implies a propensity to stay with either the SQ basket, or the healthier basket, whichever has the largest probability. OLS regressions can be used to ascertain the significance of the marginal effects of FoPL formats on these propensities, while accounting for other background variables to avoid omitted variable bias.

5. Results and discussion

5.1 Description of sample characteristics.

Forty percent of our sample of 797 respondents are men, while the average age of respondents is 48. Personal annual gross income has an average of about £13,800. In terms of education, 33% of respondents holds a high

[3] We are grateful to an anonymous reviewer for suggesting this line of investigation that we found to be persuasive and well corroborated by our data.

school diploma, 10% of them holds a post school diploma and 10% a university degree or above. In terms of employment status, 52% has either a full time or a part time job, 10% is unemployed and 35% of the sample is retired, student or homemaker. The average weekly expenditure for food shopping is £40.95. The large majority of respondents shop for food at the supermarket (96%), but a substantial fraction also shops for food at local shops (68%) and at the butcher (47%). A small fraction shops on line (5%). In terms of Body Mass Index, almost 33% of the sample have weight in the normal range, 25% are overweight and 18% are obese. 37% of respondents perceive their body weight as a little or a lot over, 40% as about ideal and 4% as a little or a lot underweight. The Health Survey of Northern Ireland in 2010-11 (DHSS&PS), instead reports only 7% as with normal weight, 36% as overweight and 18% as obese. These sample statistics hence denote some degree of under-reporting in terms of weight and/or over-reporting in terms of height. An issue to take into account in the policy implications of this study.[4]

28% of the sample never or rarely read labels, 23% do so occasionally and 36% usually or always. Importantly for this study to be used in the policy arena, computed BMI values correlate positively with attributes of the self-reported status-quo food basket, such as price (ρ=0.23) and high levels of key nutrients (high sugar 0.17, high fat 0.22, high salt 0.19 and high saturated fat 0.21).

5.2 Choice models

5.2.1 *Specification search*

All 11,628 food basket choices from the 797 complete panels are used in our choice analysis[5]. As it has become customary in taste heterogeneity studies, we benchmark our model specification search on the conditional logit specification with fixed utility coefficients, in which all respondents are restrictively assumed to be “preference clones”. We then run a specification search to explore the dimensions of preference heterogeneity over a range of 2-8 preference classes. Given the non-nested nature of the various specifications, we use information criteria (IC) (Bayesian, Akaiki, Akaiki-3 and corrected-AIC) to guide us to the optimal number of latent preference classes to fit the data, even though this method has its limitations (see discussion in McLachlan and Peel 2000,

[4] We are grateful to an anonymous reviewer for point this out.

[5] Estimation of parameters was via maximization of the sample log-likelihood and it was conducted with Latent Gold Choice version 5.0 using the expectation-maximization algorithm from an adequately large number of random starting points, to minimize the probability of local maxima.

Thacher *et al*. 2005, Morey and Thiene 2012, 2017). In our search, the IC values decrease as the number of classes increases throughout. The best model was hence selected based on two combined criteria: the plausibility of parameter estimates and the plateauing of the marginal improvement of IC values as a new class is added. This combined approach suggests a four preference-class model is best. Incidentally, four segments were also found by a similar segmentation study on use of nutrition information in Switzerland (see Visschers *et al.* 2013) and on another study on perception of FoPLs in France (Méjean *et al.* 2013). Altogether it is comforting to see that the latent preference classes clearly separate into groups with distinct propensities to healthy food choice. We then explore the effect of scale/multiplicative correlation classes and find that the fit does not significantly improve by adding more than a second class for this factor. The latent scale-preference classes are therefore eight in total.

Once ascertained that preference classes can map into healthy food choice, the next step of the specification search involves the crucial testing of whether the FoPLs treatments and the individual-specific variables systematically act as determinants of class membership probabilities for both coefficient and scale heterogeneity. Statistical evidence is found in favor of such covariates influencing preference-class membership probabilities, but not for effects on scale-class, which therefore remains unconditional. A final step in the specification search concerns the testing for the presence of continuous residual heterogeneity within preference-classes. This leads to a final model including both discrete and continuous mixing preference variation. Taste distributions for high level of fat, high level of salt and for the status quo are assumed to be distributed independent normal within each preference class, whereas all the remaining attribute coefficients are kept fixed within each preference class.

To summarize the analytics of the above narrative on the specification search, Table 3 reports the information criteria statistics for a selection of the estimated models: *i*) conditional logit model (MNL); *ii*) four-class preference model (LCM); *iii*) four-class preference and two-class scale model (LCM and scale); *iv*) four-class preference and two-class scale model with covariates (LCM and scale); *v*) four-class preference and two-class scale model with covariates and random parameters (LC-RPL and scale). By inspecting Table 3, one notes a gradual improvement in terms of model fit moving from the basic MNL model, which is used as a benchmark, to the rather articulated latent class with within-class continuous random parameters. Importantly, one notes a substantial improvement (more than 210 points) moving from the latent class model to the LC-RPL model specification, which allows for three continuously random parameters. In what follows we then focus on results description from the LC-RPL model specification.

5.2.2 *Fixed preference* ($\widetilde{\boldsymbol{\beta}}$)

We start by looking at results from the fixed coefficient conditional logit model (Table 4), which is used as a benchmark. The SQ reveals a positive and significant effect on utility coefficients, thereby implying that respondents show a preference for their current food shopping basket over the other alternatives, everything else equal. The price coefficient is negative and significant, as expected. The estimated coefficients for nutritional attributes (except for low saturated fat and low salt) are all significantly different from those for the intermediate level, which was kept as baseline. Importantly, attribute coefficient estimates conform to prior expectations in that they appear to be monotonic with negative preferences towards high levels of unhealthy nutrient attributes, denoting possibly more palatable but unhealthier food baskets; and positive preferences for low levels, denoting healthier but less palatable food baskets. Overall this seems to suggest that people, tend to give up palatability to obtain healthier food options as a result of their understanding of nutritional levels information portrayed in the FoPL. These findings seem in line with the literature (e.g. Balcombe *et al.*, 2010).

The conditional logit model fails to retrieve the latent structure of variation in taste preference and its relation with healthy food choice. Some subjects may prefer food higher in some nutrient level (say fat or salt) because of their individual preference in taste. Others may dislike high levels of a nutrient because they perceive them as unhealthy or simply do not like the taste. This implies that the coefficients of the nutritional attributes may display estimated values of diverse magnitude or sign. Effects of FoPL treatments and socio-economic covariates can be investigated with a fixed coefficient model using adequate interactions with FoPL attributes, but this approach hides latent preference structures (results of a logit model with interactions are available from the authors upon request), which instead are allowed to emerge in our random coefficient latent class approach as acting on class membership probabilities equations.

5.2.3 *Class preference* ($\widehat{\boldsymbol{\beta}}_c$)

Latent class specifications allows analysts to capture different preference structures according to the nature and number of classes in the population of respondents and answer research question 1). In interpreting these models it is customary to try and associate each class with a specific preference profile. In our case we seek to emphasize how class differences relate to healthy food choice. Then, using membership probability estimates, the individual-specific determinants of class membership are discussed in terms of propensity of different subjects to belong to each preference class. We also add a scale-class discussion that separates food consumers in highly and moderately discriminating (i.e. high and low choice determinacy) because we find evidence of continuous random utility coefficients within each class.

Parameters estimates of the four-class model are reported in Table 5. In terms of membership probabilities regarding preference classes, respondents show an averaged 38% probability of belonging to preference class 1, 32% of belonging to class 2, 20% to class 3 and 10% to class 4. Turning to classes with different multiplicative correlation, we note that the scale parameter for scale class 1 (the one with highest scale) is set to one for identification purposes. The relative value of the scale parameter for scale class 2 (averaged probability of 0.593) is 0.16 that of scale class 1, thereby suggesting that respondents have higher likelihood to act as they belong to this scale class, which displays a choice behavior with much lower multiplicative correlation than those in class 1. This implies a much smaller signal to noise ratio than in scale class 1.

Taste parameter estimates of preference classes, with only few exceptions, are statistically significant, suggesting that the preference profile of each class is quite well identified. Second, the coefficient for low saturated fat (*stfat_L*), which was insignificant in the fixed effect model, is now significant across all classes, although but it displays different signs. So, this food basket feature matters differently across preference latent structures.

5.2.3.1 *Class 1 (healthy all-rounders)*

With 38% probability, collects people that tend to healthy food choice along all nutrient dimensions. The coefficient signs have negative preferences for high levels and positive preferences for low ones. Importantly, respondents with these preferences tend to comparatively dislike their current food basket, as signaled by the negative sign of the *SQ* coefficient, which implies a propensity to modify their current diet behavior, corroborating research question 1). Interestingly, research question 3) is also answered as the estimates of standard deviations for *SQ*, *fat_H* and *sug_H* are significant: despite the negative means, the effects on utility of these high nutrient levels vary greatly within this otherwise homogenous preference class. This is of particular relevance as it provides evidence of heterogeneity beyond that of latent classes, by allowing for extra taste variation within the same class. Specifically, they imply that within this class, only 7.6% are attracted by baskets with high sugar content in the label, even a smaller share of 1.5% by high fat and about one fifth would tend to stick to their status quo basket.

Respondents with class 1 preferences display the lowest sensitivity to cost for healthy nutrient attributes, as validated by the marginal willingness to pay estimates (WTP) reported in Table 6. They are willing to pay between £35-£46/week more for a weekly food basket with low level attributes, with largest WTP for low sugar doses. On the other side of the spectrum we find baskets with high doses of fat, to avoid which they are willing to pay as much as £88.2/week. As a consequence, they are inclined to spend a substantial amount of money to move towards healthier food baskets from medium nutrient dosed ones. Because of their inclination to lower the doses of all unhealthy nutrients, the prototype respondents of this class are named here the "*healthy all-rounders*".

5.2.3.2 *Class 2 (high fat lovers)*

With 32% probability, this class shows little residual heterogeneity: the only coefficient found to be significantly random in this class is that for the SQ basket. Its large standard deviation estimate implies an 85% probability of having a propensity to stay with their SQ food choice. Consumers with these preference significantly favour both low and high sugar levels to medium ones as well as medium level of salt and saturated fat. The only nutrient they seem to appreciate in high doses is fat, perhaps for its taste. For want of a better term, we call this class "*high fat lovers*", but altogether it does seem to be inclined towards a moderately unhealthy food choice in our experiment.

5.2.3.3 *Class 3 (selectively focussed*)

We named class 3, with 20% probability, *"selectively focussed"* as their choice is affected only by a few nutritional attributes: low salt and low saturated fat, for which they are willing to pay £52.3/week (the large value across classes) and £32.9/week, respectively. They show the largest WTP estimates to avoid all high nutritional levels (more than £120/week). Interestingly, the high aversion towards high doses of fat is characterized by a significant variation in preference, as suggested by the value of the standard deviation of this parameter, but with most coefficient values in the negative range. Similar to class 1, on average, they are mostly inclined to change their current food basket. The estimated distribution indicates that only 14.4% in this class has a propensity to stay with their SQ food basket.

5.2.3.4 *Class 4 (moderately interested*)

The 4th class is the lowest probability one (about 10%) and we named it *"moderately interested"*. As in class 2, the only random coefficient is for the SQ and it shows a negative mean, but with a large standard deviation, which implies, like in class 1, that about 20% has a propensity to stay with their SQ food basket. Its member seem to only partially compromise taste with health as their choices are associated positively with intermediate doses of nutritional FoPL values. In fact, for all four nutrients coefficient signs for both high and low levels are negative, suggesting moderate amounts being the favourite. Respondents in this class display the highest sensitivity to cost, which induces low values of WTP estimates. In other words, these people are often unhappy with their current food basket and would sometime like to change it, but they do not seem to be strongly affected by nutritional labels. As a consequence, they are unwilling to spend money to secure such change.

5.2.4 *Class determinants* ($\hat{\boldsymbol{\gamma}}$)

Having identified the sizes and the salient effects of FoPL nutrient messages on propensity to healthy food choice in latent groups with homogeneous preferences, we now turn our attention to exploring their statistical association with individual specific policy relevant social covariates, and to answer question 2). Socio-economic effects on food choice have been found before. So, although not novel, these effects are interesting for model validation. We separate these variables into a first set with three FoPL formats (HYD, GDA and MTL, since TXT is the baseline), the set of conventional socio-economic variables (income, education attainment, age, sex, etc.) and the final set of food choice context self-reports (perceived departure from ideal body weight, BMI, propensity to read food labels, etc.).

FoPL formats are known to convey different amount of information by means of various visual features. A key policy question that can be asked to endorse a given FoPL format over others is whether it significantly affects class membership probabilities, and if so how it associates with more or less healthy food choice.

5.2.4.1 *FoPL formats*

In our model, all effects refer to the baseline probability of belonging to the highest probability class 1 (*healthy all rounders*). All else being equal, compared to TXT, the hybrid FoPL (HYB)—the most informative label format—significantly increases membership probability to class 3 (*selectively focussed*). From a policy perspective this is an interesting and positive finding, as the preference features of this class provide scope for designing and implementing a tailored policy to increase the role of nutrient information in food purchase involvement for saturated fat and salt.

The GDA format is the second most informative as it only differs for lack of the colour signals from the HYB. This treatment is never significant at conventional level, but has the highest asymptotic z-value for a negative effect on membership to class 2 (*high fat lovers*) and for positive effect on class 3. The negative effect lowers the probable membership to class 2 in favour to the healthier class 1 and increases that of class 3. For both the significance is just outside the customary levels, but in light of the more recent recommendation to interpret p-values (Wasserstein and Lazar, 2016) it makes sense to highlight this result regardless of conventional level of significance.

In terms of visual signal, the traffic light in text format (MTL) is only just more informative than the least informative FoPL (TXT) as it only adds colors to the TXT display. Compared to the latter it only shows a significant and negative effect on membership probability to class 2 (*high fat lovers*), denoting by default a positive role in determining association with groups making healthier food choices. For memberships to classes 3 and 4 its effect has low significance. Overall our data provide a positive answer to research question 2) and 3), since the matrix $\mathbf{\Omega}$ is significantly different from zero, and its structure varies plausibly across preference classes.

5.2.4.2 *Socio-economic covariates*

Moving to the socio-economic covariates, we see that older age significantly affects only membership to class 2; it makes sense that elderly people are more likely to be in this group because they are often less inclined to collect new information from FoPL and to use it to improve their knowledge about food products: this might require comparative higher cognitive effort or accrue comparatively lower perceived benefits. Being a woman significantly increases membership to class 3, which is the *selectively focussed* class. Women might have more familiarity with food choices as they often shop for food for the whole household.

Self-reports on the frequency of reading FoPLs have a negative association with memberships probabilities to classes 2 and 4, which by default implies they are positively associated (with high significance) to the other two healthier food choice classes. This is definitely an interesting piece of information for policy, as both classes 2 and 4 involve respondents who are either moderately affected by nutritional details (class 4) or only partly affected (class 2). So, those who read FoPL details frequently are associated with healthier food choices. We cannot state causation, although this is obviously very plausible, so a campaign aiming at increasing the frequency of reading such details might steer consumers towards healthier food baskets. This obvious link can be used as a validation of the robustness of the model. Causation could be explored in future research with field experiments based on randomised treatments.

A salient feature, in the context of stemming the growth of overweight prevalence, is the association between self-reported perception of having an "ideal body weight" and class membership, as well as its association with the more objective BMI values. Perceiving oneself as having an ideal body weight is significantly and positively associated only with membership to class 2. These people do not perceive to have weight-related reasons to steer away from high fat baskets and indulge in tasty meal selections. On the other hand, having a high BMI has a negative and significant association with class 3, which implicitly makes it positively associated with the baseline class of healthy food choosers. At least in this hypothetical choice context, those with a weight problem, objectively measured or perceived, seem to pay attention to FoPL and to use them for healthier choice. This suggests that the choice experiment reached out to its target audience.

5.3 Sensitivity analysis and determinants of membership probabilities

Discussing signs and relative magnitude of structural coefficients $\hat{\boldsymbol{\gamma}}$ of probability models offers some insight on the direction and intensity of associations between preference groups and their drivers. However, further insight on model validity can be gleaned by a sensitivity analysis. So, in this section the estimates of the coefficients determining class membership probabilities are used to perform a sensitivity analysis. The aim is to describe changes in class membership probabilities, and hence on degree of healthy food choice, as a consequence of changes in their determinants. The ultimate goal is, in fact, to draw a selection of scenarios that can provide useful suggestions for policy design, which in this case must be tailored on the characteristics of the target population.

Figure 2 shows how class membership probabilities change as age increases. The baseline is defined by the profile for a male respondent who decided the favourite food basket using the TXT format for FoPL, and who reports to never read food labels, a normal body weight (BMI group 3) and who perceives their own body weight as about ideal. Young males with such individual traits display a high probability of belonging to class 4, the *moderately interested.*

As age increases within this profile a major shift in membership probability takes place from class 4 to class 2. That is, from *moderately interested* to *high fat lovers.* From a policy perspective, this is important as it suggests a policy addressing older people, or educating middle age people to be more attentive about food choices. If one is prepared to assume that the change is age-induced, rather than being a feature associated to the specific age cohort, then one may conclude that without a tailored action, young males with 15% probabilities of belonging to class 2 may see this probability grow to nearly 50% by the time they are 60 years old guys: a three-fold increase. Clearly, more research is necessary to establish this causal dependency.

One may wonder what effect would have to change some elements of this profile on the age range. Figure 3 describes this effect on a woman reporting to “always read the label” (except for the first set of bars), and who decides based on a HYB label, i.e. the label format conveying the richest amount of information. The combined effect on membership probability of sex and of label type change (from TXT to HYB) can be seen by comparing the first set of bars on the left between Figure 2 and 3. The effect is strong and positive for class 2 membership, and negative for class 1. Focusing on the first two sets of bars in Figure 3 shows the effect of moving from "never" to "always" reading FoPLs, everything else being equal, for an 18 year old woman. As can be seen "always reading FoPL" is strongly associated with classes with healthier food choices. Specifically, we note a two-fold decrease in membership probability for class 2 (*high fat lovers*) and a drop from 50% to 3% in class 4 (*moderately interested*).

Turning the attention to the five blocks of bars on the right of Figure 3 allows us to explore the effect of age increase on class membership. We note that, as expected, being older makes it more likely to belong to class 2, a relatively unhealthy food choice group, with a probability change from 10% to 26%, which draws mostly from class 4 (the *moderately interested*). From a policy perspective, there is obvious scope to target older women, even when they read FoPL and correctly think of themselves as of ideal weight, to improve their diet habits. This needs doing with action beyond food labeling. Perhaps with an information campaign directed to the personalized interpretation of the information content of labels.

Let us now turn to Figure 4 which investigates the interesting effect of the five BMI categories (from normal BMI to the highest obesity of class III) on class membership probabilities. The baseline in this case are 30 years old women who never read FoPLs, are shown a HYB format, and perceive own weight as "about ideal". Let us ignore for the moment the rightmost block of bars and focus on the first five. From these comparisons, there emerges a quite clear picture: all else equal, increasing BMI (that is, *effective* weight, not the perceived one) redistributes membership probabilities from class 4 to class 2. That is from the *moderately interested* group to the *fat lovers*, which for highest BMI ends up with a 61% membership probability. Hence, there is clear evidence for the need to target food choice policies to this group of effectively overweight and obese people, who despite having objective issues in terms of own weight (as shown by reported BMI), incorrectly perceive their body weight class and hence discount their health risks.

How much does a realistic perception of own body weight combined with reading FoPL affect class membership in an extreme case? To answer this question let us now focus on the two very last groups of bars on the right side of Figure 4. The last set of bars to the right shows how class membership probabilities change with respect to the second to the last set when these conditions are imposed, i.e. when own weight perception is correct (a lot over-weight for a class III obese woman) and reading FoPL is imposed. The two effects combined produce a major redistribution in the class membership probabilities: class 1 (the healthy food choice) increases from 10% to 65%, followed by a smaller increase in class 3 (that also chooses quite well), whereas class 2 and class 4 show a drastic decrease, moving from 61% to 13% and from 24% to 3%, respectively. This suggests that a policy promoting a *realistic* body weight image and a regular reading of FoPL details is associated with potentially *strong* health benefits from the adoption of healthier diet. Similar results are found also with label formats different from HYB. A proposition worth exploring further in field experiments.

5.4 Distributions of individual marginal WTP estimates and taxation targeting

The literature has often discussed the cross effect of price-based instruments to discourage the dietary intake of unhealthy nutrients. Taxing one nutrient—for example fat—can, by statistical association, discourage the uptake of other complementary nutrients—for example salt. One way to inform policy design is to explore the degree of association between individual-specific marginal willingness to pay (mWTP) implied by the sequences of choice data of each respondent. mWTPs can be computed in our sample, conditional on the pattern of the 16 observed choices, for high (and therefore unhealthy) levels of nutrients in the weekly food baskets. Figure 5 shows the quantile contours of a bivariate kernel density of mWTP for a weekly diet high in fat and high in salt. The north-east quadrant delimited by the dashed line shows the density of those in the sample with positive mWTPs for both, while those in the south-west quadrant show the densities for those with negative values. In this quadrant we recognize a group with strong adversity to a diet with high values in salt and fat (less than £-150/week) and a group with medium aversion (around £-50/week). The highest density is found along the dashed line (£=0/week) for high fat, but around £-15/week for high salt.

The north-west quadrant collects those that have positive view of high fat, but negative for high salt. These respondents would not adjust their high salt diet as a consequence of a tax on high fat, since they already dislike high salt, but those in the north-east quadrant would. Although the latter group has smaller density. The south-east quadrant collects those with positive view of high salt, but negative for high fat. A similar reasoning applies here for a tax on high salt—it would not reduce the consumption of high fat in this group.

The policy implication is that the segment in the north-east quadrant is the only segment that would be subject to cross effects in case a tax was exclusively imposed on high levels of either salt or fat. This segment is a low density one and hence cross tax effects are likely to be small. Similar policy directions can be derived for other levels or other nutrients. Some of these are available from the authors upon request.

5.5 Effects of FoPL types on class membership

Figure 6 illustrates the marginal effects on (posterior) predicted class membership probabilities for each of the three FoPL formats, using TXT as baseline. Values are separated by BMIs computed from self-reported measures (on the right obese respondents with a BMI>30) to emphasize differences between the two target groups. The effects are plotted in increasing order so as to illustrate the sample distribution at the various level of response.

For example, focussing on the effect of HYB for non obese, it can be noticed that exposure to this FoPL draws prevalently from membership of classes 3 (selectively focussed) and 2 (high fat lovers) to contribute mostly to membership of class 4 (moderately interested), class 1 (healthy all-rounders) and class 2 (high fat lovers). However, this layout demonstrates that the membership density lost by class 1 is small compared to the density gained, so that class 1 has a net gain, as does (more evidently) class 4.

A comparison across the not obese and obese plots shows that, while the change in both groups draws prevalently from class 3 (selectively focussed on low salt and on low saturated fat) and is directed mostly to class 4 (moderately interested), the densities of the contribution varies: the contribution to class 4 is much higher in the non obese sub-sample. This implies that HYB labels affect the target population (obese people) by making them relatively more aware across the board of nutrition information, and not only of low salt and saturated fat.

The overall effect of the specific MTL label shows little difference across sub-samples, but it is of particular interest because it draws from class 2 membership (high fat lovers) and contributes to classes 3 (selectively focussed). This suggests that traffic light colours are effective across both weight groups.

5.6 Effects of FoPL types on healthy choice

Figure 7 reports the predicted differences between the probability of selection of the status quo food basket and the healthiest (i.e. lowest content of sugar, salt, fat and saturated fat) food basket profile on offer. Sample predictions are obtained from the model in Table 5. As evident from the plot, the pattern of positive predicted differences (those with propensity to choose the SQ-basket on the upper part of the graph) differ substantially from that of negative ones (those with propensity to select the healthy basket in the lower part of the graph). The effects of moving from TXT to other FoPL formats is best evidenced in Figure 8 where we plotted the sub-sample *differences* in predicted probabilities of sticking to the SQ basket computed for the most basic TXT labels and those predicted with other labels makes the effect more apparent. Such values are nearly always negative, because TXT shows the highest propensity not to change. Also, they have a much narrower range, as the effect is only due to change of FoPL. Interestingly though, this plot shows clearly how the non-obese respondents are more affected by GDA than MTL, while to obese respondents the two FoPLs are equivalent in terms of this specific effect relative to TXT. However, the latter group shows a smaller difference, indicating lower responsiveness to all FoPLs, but particularly to HYB.

We formally investigate the statistical significance of FoPLs on these differences with regards to various subgroups of respondents. The hypothesis is that, once accounted for background variables to avoid omitted variable bias, the marginal effects of FoPL formats and their interactions be significant and have plausible signs. A Chow test of structural stability across signs of the dependent variable is rejected, consistently with gain-loss asymmetry. In Table 7 we report OLS results for two separate regressions, one for respondents with predicted propensity to change to the SQ basket and the other to the healthy basket. The dependent variables are the two sets of absolute values of the differences (positive and negative) in predicted posterior choice probabilities or |Pr(sq)-Pr(healthy)|. Positive effects of independent variables indicate larger absolute value differences (i.e. less uncertainty in choice), or stronger propensity. The effect of different types of FoPL is measured using TXT or HYB as a baseline and positive effects are to be interpreted as producing stronger propensity. Interaction effects of interest are those with groups of respondents that are in need to correct their current food choice. So, we use dummy variables indicating exposure to FoPLs, on their own as well as interacted with indicators of subgroups, which are also used on their own as background variables. These subgroups of interest are being a *woman*, self-reporting body measures indicating *obesity* (BMI>30) and a dummy variable indicating *misperceiving* one's own body weight while being obese (1 if one manifests this misperception). Additional background variables include *age* and *age squared*, index of frequency to *read labels* and self-perception of an *ideal own body weight*. The variables used have good explanatory power for the two propensities to change (adj. R^2 0.87 for those with SQ propensity and 0.52 for those with propensity to move to the healthy basket).

The results of the single coefficients offer much ground for discussion, we limit our comments here to the significant effects of FoPL formats when they are interacted with obesity, gender and self-image misperception.

5.6.1 *Explaining propensity for status-quo baskets*

With respect to the move from TXT or HYB, moving to GDA or to MTL reduces the propensity to stay with the status-quo basket. This effect is exacerbated for women for GDA (with borderline significance) and for obese respondents exposed to MTL, while for obese people who mis-perceive their own body weight the effect is similar and significant for both GDA and MTL. Being woman, obese and having reported a higher score for ideal body image significantly increase propensity for the SQ basket, and so does being older (with a peak extrapolated at age 91), while the self-reported frequency score for reading labels decreases this propensity.

5.6.2 *Explaining propensity for healthy baskets*

For this type of propensity the pattern of significance and the directions of the effects are somewhat different. Compared to the move from TXT or HYB, moving to GDA significantly *increases* the propensity to select a healthy basket. This effect is less significant and less than half the magnitude estimated for a move from TXT to MTL; the latter effect (on the margin) is nullified for non obese women. Being obese significantly reduces the propensity to healthy food baskets, especially for those obese respondent that self-report a perception of a normal weight. Being older increases propensity to healthy food baskets, but this effect decreases at squared speed with age. The marginal effect of frequency of reading labels is highly significant and positive, that of being a woman is also positive, but only marginally significant. Self-reporting a higher ideal body image score decreases this propensity significantly.

6. Implications for future research and for policy

Deriving strong policy recommendations of immediate applicability to the field of food labeling from a stated preference study with limited external validity as the present one is obviously unwarranted without further field testing, which we advocate. A further limitiation is that we did not address how consumers can substitute unhealthy food items with healthy ones to achieve a *satisficying* level of healthiness in the overall mixture of packaged foods in the basket. This because doing so would require a prohibitively expensive experimental design and be impractical.

We nevertheless derive some potentially important policy suggestions from our study, which further validate and extend the evidence supporting the recommendation to use GDA by Malam *et al.* (2009). The overall picture depicted by our analysis of the Northern Irish food consumers is quite articulated. They display good sensitivity to nutritional labels for the most part (classes 1 and 3 represent together nearly 60 percent) with about 10 percent displaying moderate interest. About one third of the total (class 2) represents a hard core of relatively insensitive users of FoPL information. However, significant differences exists across determinants of memberships to the four preference groups with regards to both, label formats and socio-economic covariates. A significant residual of within-class preference heterogeneity is present, as shown by both continuously random preferences as well as differences in choice determinism (or ability to discriminate). These technical issues should be born in mind in future by choice analysists operating in this area and by those wishing to develop future field tests.

6.1 Policy implications

A policy-salient result is that FoPLs induce respondents of different self-reported weight categories to respond differently. FoPL based on traffic light systems (MTL) and daily amount guidelines (GDA) induce stronger responses towards healthier baskets in self-reported obese respondents, compared to the baseline text only or hybrid FoPLs. When the alternative to the status-quo basket is the healthiest food basket, the propensity to select the healthy food shows different sensitivity to determinants, depending on whether the propensity is positive or negative. This suggests potential for different policy targets: one, for example, for nudging FoPLs that portray a visual colour enhancement with respect to the basic text. This because they emerge as comparatively more effective at increasing membership probabilities into preference classes associated with healthier food choice. Choices made under the most visually informative label format (HYB), have higher membership of the preference structure that appears *selectively focused* (class 3) on specific nutritional factors (salt and saturated fats), and it does so in our sample for a large proportion of respondents, even though it shows a markedly lower impact on obese ones (see Figure 6). But, it seems to be effective mostly on already nutritionally sensitized food consumers. How valuable its use can be will hence depend on how large a share of the population this preference class represents, bearing in mind that even though it mostly draws from the "*fat lovers*", it also draws in part from "*healthy all rounders*".

The marginally less informative FoPL format GDA appears as a determinant in the membership of larger preference classes, detracting from class 2 (*high fat lovers*) and adding to class 3 (*selectively focused*), mostly drawing from class 1 (*healthy all rounders*). Once again, GDA appeals positively to the already nutritionally sensitized food buyers, but in our sample it induces to a class change a smaller sample proportion than HYB and it has similar drawbacks. However, in the propensity to choose healthier baskets when compared to the SQ, our simulation shows the GDA label as having the strongest effect on non-obese respondents, and as strong as the MTL for obese ones. This is a result contrary to that by Botzug *et al.* (2015) who conclude that *"GDA labels are generally insufficient to adjust consumer behaviour towards healthier alternatives"*. Altogether these results point the finger to the role of nutrition education as a means to sensitize customers as a necessary precursor of FoPL effectiveness, when these contain more information.

What clearly emerges in the sensitivity analysis we conducted to validate the model is the role of other drivers behind preference, such as gender, the perception gap between BMI and self-body image and age, with being obese at the forefront. This points the finger to the potential scope for methods other than alternative forms of FoPLs formats, and towards information programs specifically tailored to specific sub-groups of consumers, a form of individualised labeling. While much emphasis and past research work has been focused only on FoPL formats, the wider policy picture seems to require a much broader multi-dimensional intervention, mostly based on education and directed to specific groups.

6.2 Further research

Given the small space available to convey information in FoP food labels, the search remains for a succinct prescription for information on nutritional content that can be broadly effective. Direction for further research might include labeling initiatives directed towards specific groups for specific foods (individualized information). Information directed to younger age groups and groups with low nutritional education might rely on labelling signals based on physical activity caloric equivalency. Interpreting these messages does not require knowledge of suggested daily caloric intake or pre-existing sensitivity to specific nutrition factors. For example, recent research in the USA (Bleich *et al.* 2012 and Bleich *et al.* 2014) demonstrates that at least black youth are more inclined to heed and act upon activity equivalent calories metrics than they are on simple caloric amounts. The effect has also been shown to be mediated by parents' choices for their children fast food meals (Viera and Antonelli, 2014). Admittedly, caloric intake does not provide as full a nutritional picture, but in a fight against obesity and overweight it might be more relevant to encourage consumer to consider both lowering intake and increasing physical activity, rather than expecting to act upon complex multi-dimensional nutritional messages.

Official UK statistics on caloric intake are problematic. For example, a recent report (Harper and Hallsworth, 2016) showed that official statistics on food expenditures (the National Diet and Nutrition Survey data and the Living Costs and Food Survey data) are systematically under-estimating caloric consumption when compared to other survey statistics from the same population (e.g. Kantar Worldpanel) and from evidence derived from other objective measurements. The reduction in the average physical activity necessary to produce the observed average body weight increase cannot be reconciled with the reported intake. A conclusion supported also by Doubly Labelled Water, which indicates calorie under-reporting of about 32 percent. On the other side of the equation, self-reports on physical activity in England in 2008 showed that "data indicated that 39% of men and 29% of women met the Chief Medical Officer's minimum recommendations for physical activity; the data from accelerometers indicated that only 6% of men and 4% of women had done so" (Harper and Hallsworth, 2016, page 11). These skewed self-reports are possibly due to an increased awareness of being overweight, the need for dieting and increased physical exercise in order to lose weight.

The above measures, once combined with GDA or MTL FoPLs might work better than alternative combinations, at least for certain target groups. A view recently supported also by the Royal Society for Public Health chief executive (Cramer 2016). More research is needed in this area, which can move from the basis of relatively weak evidence from hypothetical choice under experimental conditions to more persuasive evidence from field tests based on real choice. Randomised control trials in the dimensions suggested by this study may offer the way forward in this field.

In response to our initial question, whether obese care about FoPL, our result show that they do, but differently from other consumers. For example the effects of MTL and GDA formats in selecting healthy food baskets, using TXT as a baseline, are predicted to be identical for obese, but not so for others.

7. References

AAVV (2016), Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19·2 million participants, *Lancet*, 387, 1377–96.

Andrews, J., Burton, S., & Keys, J. (2011). Is simpler always better? Consumer evaluations of front-of-package nutrition symbols. *Journal of Public Policy & Marketing*, 30(2), 175–190.

Asam E.H. and Bucklin L.P. (1973) Nutrition labeling for canned goods: A study of consumer response. *J. Marketing* 37, 32-37.

Aschemann-Witzel J., K.G. Grunert, H. van Trijp, S. Bialkova, M.M. Raats, C. Hodgkins, (2013) Effects of nutrition label format and product assortment on the healthfulness of food choice, *Appetite*, 71, 63–74.

Balcombe K., I. Fraser, S. di Falco, (2010) Traffic lights and food choice: A choice experiment examining the relationship between nutritional food labels and price. *Food Policy*, 35 (3), 211-220

Balcombe, K.; Fraser, I., McSorley, E. (2015), 'Visual Attention And Attribute Attendance In Multi-Attribute Choice Experiments', *Journal of Applied Econometrics*. 30:447-467

Becker, M. W., Bello, N. M., Raghav, S. P., Peltier, C., & Bix, L. (2015). Front of pack labels enhance attention to nutritional information in novel and commercial brands. *Food Policy*, 56, 76-86.

Bialkova, Svetlana, Klaus G. Grunert a, Hans van Trijp. (2013) Standing out in the crowd: The effect of information clutter on consumer attention for front-of-pack nutrition labels. *Food Policy*, 41:65–74

Bialkova S., K.G. Grunert, H.J. Juhl, G. Wasowicz-Kirylo, M. Stysko-Kunkowska, H.C.M. van Trijp (2014) Attention mediates the effect of nutrition label information on consumers' choice. Evidence from a choice experiment involving eye-tracking, *Appetite*, 76, 66–75

Bleich SN, Herring BJ, Flagg DD, Gary-Webb TL. (2012) Reduction in purchases of sugar-sweetened beverages among low-income Black adolescents after exposure to caloric information. *Am J Public Health*, 102 (2):329-335.

Bleich SN, CL Barry, TL Gary-Webb, BJ Herring (2014) Reducing sugar-sweetened beverage consumption by providing caloric information: how Black adolescents alter their purchases and whether the effects persist *Am J Public Health*, 104, 2417–2424

Brown, H. (2014) *A comparison of Front of Pack Nutritional Food Labelling formats in Northern Ireland using a discrete choice experiment*, Ph.D. Thesis, Queen's University Belfast, Northern Ireland

Bliemer M., Rose J. M., and S. Hess. (2008) Approximation of Bayesian efficiency in experimental choice designs. *Journal of Choice Modelling*, 1:98–127.

Boeri, M., Longo, A., Grisolía, J. M., Hutchinson, W. G., & Kee, F. (2013). The role of regret minimisation in lifestyle choices affecting the risk of coronary heart disease. *Journal of health economics*, 32(1), 253-260.

Boeri, M., Scarpa, R. & Chorus, C.G. (2014) Stated choices and benefit estimates in the context of traffic calming schemes: Utility maximization, regret minimization, or both? *Transportation Research Part A: Policy and Practice*, 61, pp. 121-35.

Boztuğ Y, Juhl HJ, Elshiewy O, Jensen MB (2015) Consumer response to monochrome Guideline Daily Amount nutrition labels. *Food Policy*.;53:1–8.

Boxall, P. C., V. L. Adamowicz (2002) Understanding Heterogeneous Preferences in Random Utility Models: The Use of Latent Class Analysis. *Environmental and Resource Economics*, 23 (4): 421-46.

Bujosa, A., Riera, A., Hicks, R.L. (2010). Combining discrete and continuous representations of preference heterogeneity: a latent class approach. *Environmental Resource Economics*, 47: 477–493.

Burton, M., Davis K., Kragt M. E. (2016) Interpretation issues in heteroscedastic conditional logit models, Working Paper, University of Western Australia, School of Agricultural and Resource Economics, http://purl.umn.edu/235296

Campbell, D., Hensher, D.A. and Scarpa, R. (2014). Bounding WTP distributions to reflect the 'actual' consideration set. *Journal of Choice Modelling*, 11:4-15.

Cramer S. (2016) Food should be labelled with the exercise needed to expend its calories, *Bmj*. 353, i1856.

Crosetto, P.;Muller, L.;Ruffieux, B. (2016). Helping consumers with a front-of-pack label: Numbers or colors? Experimental comparison between guideline daily amount and traffic light in a diet-building exercise, *Journal of Economic Psychology* , 55, August, 30-50.

DEFRA (2010) Family food. a report on the family food module of the living costs and food survey 2008. A National Statistics Publication.

Department of Health (2013) Guide to creating a front of pack (FoP) nutrition label for pre-packed products sold through retail outlets, Food Standards Agency.

Enax, L., Hu, Y., Trautner, P., & Weber, B. (2015). Nutrition labels influence value computation of food products in the ventromedial prefrontal cortex. *Obesity*, 23(4), 786–792

Epstein L. H., M. D. Myers, H. A. Raynor, and B. E. Saelens. (1998) Treatment of pediatric obesity. *Pediatrics*, 101, 554–570

Farizo B.A., J. Louviere, M. Soliño, (2014) Mixed integration of individual background, attitudes and tastes for landscape management, *Land Use Policy*, 38 (2014), pp. 477–486

Ferrini, S. and Scarpa, R. (2007). Designs with a-priori information for nonmarket valuation with choice-experiments: a Monte Carlo study, *Journal of Environmental Economics and Management* 53, 342–363.

Feunekes, G., Gortemaker, I., Willems, A., Lion, R., & van den Kommer, M. (2008). Front-of-pack nutrition labelling: Testing effectiveness of different nutrition labelling formats front-of-pack in four European countries. *Appetite*, 50, 57–70.

Fiebig DG, Keane MP, Louviere J, Wasi N (2010) The generalized multinomial logit model: accounting for scale and coefficient heterogeneity. *Mark Sci* 293:393–421

Food Standards Agency. (2007) FSA nutrient and food based guidelines for UK institutions, Available online at https://www.food.gov.uk/sites/default/files/multimedia/pdfs/nutrientinstitution.pdf

Food Standards Agency. (2010) Front of pack (fop) nutrition labelling. Available online at http://www.food.gov.uk/multimedia/pdfs/board/fsa100307.pdf

Food Standard Agency (2012), Consultation on front of pack nutrition labelling, Available online at http://www.food.gov.uk/multimedia/pdfs/consultation/consult-fop-ni.pdf.

Franceschinis, C., Thiene, M., Scarpa, R., Rose, J., Moretto, M., Cavalli, R. (2017) Adoption of renewable heating systems: An empirical test of the diffusion of innovation theory, *Energy*, Article in Press

Gracia A., M.L. Loureiro, and R. Nayga Jr. (2009), Consumer's valuation of nutritional information: A choice experiment study. *Food Quality and Preference*, 20 (7):463–471.

Greene, W.H. and Hensher, D.A. (2013). Revealing additional dimension of preference heterogeneity in a latent class mixed multinomial logit model. *Applied Economics*, 45(14): 1897-1902.

Gregori Dario, Simonetta Ballali BA, Claus Vögele PhD, Francesca Galasso BS, Kurt Widhalm, Paola Berchialla PhD & Ileana Baldi PhD (2015) What Is the Value Given by Consumers to Nutritional Label Information? Results from a Large Investigation in Europe, *Journal of the American College of Nutrition*, 34(2):120-125.

Grisolía, J. M., Longo, A., Boeri, M., Hutchinson, G., & Kee, F. (2013). Trading off dietary choices, physical exercise and cardiovascular disease risks. *Social Science & Medicine*, 93, 130-138.

Grisolía, J. M., Longo, A., Hutchinson, G., & Kee, F. (2015). Applying Health Locus of Control and Latent Class Modelling to food and physical activity choices affecting CVD risk. *Social Science & Medicine*, 132, 1-10.

Grunert, K. and J. Wills (2007) A review of European research on consumer response to nutrition information on food labels *Journal of Public Health*, 15(5), 385–399.

Harper H, Hallsworth M. (2016) Counting Calories: how under-reporting can explain the apparent fall in calorie intake. Report, Behavioral Insights Team, 1-43.

Hawley, Kristy L., Christina A. Roberto, Marie A. Bragg, Peggy J. Liu, Marlene B. Schwartz and Kelly D. Brownell. (2012) The science on front-of-package food labels. *Public Health Nutrition*, 16(3):430–439

Health Survey Northern Ireland 2010/11: Obesity Analysis (2011), Public Health Information & Research Branch, Information & Analysis Directorate, Department of Health, Social Services & Public Safety, Bulletin 5

Hensher D. A. and W. H. Greene. (2003), The mixed logit model: the state of practice. *Transportation*, 30:133–176.

Hensher, D., and W. Greene (2003) A Latent Class Model for Discrete Choice Analysis: Contrasts with Mixed Logit. *Transportation Research, Part B* 37:681-98.

Hersey, J., Wohlgenant, K., Arsenault, J., Kosa, K., & Muth, M. (2013). Effects of front-of-package and shelf nutrition labeling systems on consumers. *Nutrition Reviews*, 1-14.

Hess S. and Rose JM (2007) A latent class approach to modelling heterogeneous information processing strategies in SP studies. In: Oslo workshop on valuation methods in transport planning, Oslo

Hess S. and Rose JM (2012) Can scale and coefficient heterogeneity be separated in random coefficients models? *Transportation* 39(6):1225–1239.

Hess, S., Stathopoulos, A., Daly, A. (2012) Allowing for heterogeneous decision rules in discrete choice models: an approach and four case studies. *Transportation* 39 (3), 565–591.

Hess, S and Train, K. (2017) Correlation and scale in mixed logit models, *Journal of Choice Modelling*, 23, 1-8.

Hodgkins, C., Barnett, J., Wasowicz-Kirylo, G., Stysko-Kunkowska, M., Gulcan, Y., Kustepeli, Y., et al. (2012). Understanding how consumers categorise nutritional labels. A consumer derived typology for front-of-pack nutrition labelling. *Appetite*, 59, 806–817.

HSCIC (2015), Statistics on Obesity, Physical Activity and Diet, Report by Lifestyles Statistics Team, Health and Social Care Information Centre, England.

Jacoby J., Chestnut R.W. and Silberman W. (1977) Consumer use and comprehension of nutrition information. *J. Con. Res*. 4(2), 119-128.

Jones G, Richardson M. (2007) An objective examination of consumer perception of nutrition information based on healthiness ratings and eye movements. *Public Health Nutr*;10, 238-244.

Keane M (2006) The generalized logit model: preliminary ideas on a research program. In: Presented at Motorola-CenSoC Hong Kong meeting, October 22

Klopp P, MacDonald M. (1981) Nutrition labels: an exploratory study of consumer reasons for non-use. *Journal of Consumer Affairs*, 15, 301 – 16.

Koenigstorfer, J., Wasowicz-Kiryło, G., Stysko-Kunkowska, M., Groeppel-Klein, A., (2013) Behavioural effects of directive cues on front-of-package nutrition information: the combination matters! *Public Health Nutrition*, 17(9), 2115–2121

Louviere J. and Eagle T (2006) Confound it! That pesky little scale constantmesses up our convenient assumptions. In: Proceedings of the sawtooth software conference, Sawtooth Software, Sequem, Washington, DC, USA, pp 211–2.

Malam, S.; Clegg, S.; Kirwan, S. & McGinigal., S. (2009), 'Comprehension and use of UK nutrition signpost labelling schemes. Prepared for the Food Standards Agency, *FSA*.

Manski, C., (1977) The structure of random utility models. *Theor. Decis.* 8, 229–254.

Marsh, D. Mkwara, L. Scarpa, R. (2011), Do respondents' perceptions of the Status Quo matter in non-market valuation with choice experiments? An application to New Zealand freshwater streams, Sustainability, 3 (9): 1593-1615.

McFadden D (1974) Conditional logit analysis of qualitative choice-behaviour. In: Zarembka P (ed) Frontiers in econometrics. Academic Press, New York

McLachlan G, Peel D (2000) Finite mixture models. Wiley, New York

Mejean C., Macouillard P., Peneau S., Hercberg S., Castetbon K. (2013) Perception of front-of-pack labels according to social characteristics, nutritional knowledge and food purchasing habits. *Public Health Nutr*. 16:392–402

Morey E., Thiene M., (2012) A parsimonious, stacked latent-class methodology for predicting behavioral heterogeneity in terms of life-constraint heterogeneity, *Ecological Economics*, 74, pp.130–144.

Morey E., Thiene M., (2017) Can personality traits explain where and with whom you recreate? A latent-class site-choice model informed by estimates from a mixed-mode LC cluster models with latent-personality traits, *Ecological Economics*, 138, 223-237.

Nayga R. M. (1996). Determinants of Consumers' Use of Nutritional Information on Food Packages, *Journal of Agricultural and Applied Economics*, 28(2), 303-312.

Nayga R. M. (1997) Impact of Sociodemographic Factors on Perceived Importance of Nutrition in Food Shopping. *The Journal of Consumer Affairs*, 31(1) 1-9.

Nayga R. M., Lipinski D., Savur N. (1998) Consumer's use of nutritional labels while food shopping at home, *The Journal of Consumer Affairs*, 32(1), 106-120.

NHS (2012) http://www.nhs.uk/news/2012/04april/Pages/nhs-diabetes-costs-cases-rising.aspx, accessed 29/10/2016

Nørgaard M.K. and Brunsø K. (2009). Families' Use of Nutritional Information on Food Labels. *Food Quality and Preference*, 20, 597-606.

Pollard, J., Kirk, S.L. and Cade, J.E., 2002. Factors affecting food choice in relation to fruit and vegetable intake: a review. *Nutrition research reviews*, 15(2), 373-387.

Powe N. A., G. D. Garrod, and P. L. McMahon (2005), Mixing methods within stated preference environmental valuation: choice experiments and post-questionnaire qualitative analysis. *Ecological Economics*, 52:513–526.

Pretty J.N., A.S. Ball a, T. Lang, and J.I.L. Morison (2005), Farm costs and food miles: An assessment of the full cost of the uk weekly food basket. *Food Policy*, 30:1–19.

Provencher, B., K. Barenklau, R. C. Bishop (2002) A Finite Mixture Logit Model of Recreational Angling with Serially Correlated Random Utility, *American Journal of Agricultural Economics* 84 (4), 1066-75.

Raghunathan, R., Naylor, R.W. and Hoyer, W.D., 2006. The unhealthy=tasty intuition and its effects on taste inferences, enjoyment, and choice of food products. *Journal of Marketing*, 70(4), pp.170-184

Rose J. M. and R. Scarpa. (2008), Designs efficiency for non-market valuation with choice modelling: how to measure it, what to report and why. *Australian Journal of Agricultural and Resource Economics*, 52(3), 253-282.

Scarpa, R, Campbell, D, & Hutchinson, W 2007, 'Benefit Estimates for Landscape Improvements: Sequential Bayesian Design and Respondents' Rationality in a Choice Experiment', *Land Economics*, 83 (4), 617-634

Scarpa, R., Thiene, M. (2005) Destination choice models for rock climbing in the Northeastern Alps: a latent-class approach based on intensity of preferences, *Land economics*, 81(3), pp. 426-44

Scarpa R, Thiene M, Hensher D (2012) Preferences for tap water attributes within couples: an exploration of alternative mixed logit parameterizations. *Water Resour Res* 48:W01520

Soederberg Miller L. M. and Cassady D. L. (2015) The effects of nutrition knowledge on food label use. A review of the literature, *Appetite*, 92(1), 207–216.

Swait J and Louviere JJ (1993) The role of the scale parameter in the estimation and comparison of multinomial logit models. *J Mark Res* 30(3):305–314

Swait, Joffre, Neil Brigden and Richard Johnson (2014a) "Categories Shape Preferences: A Model of Taste Heterogeneity Arising From Categorization of Alternatives," Special Issue on Antecedent Volition, J. Swait and W. Adamowicz, Guest Editors, *Journal of Choice Modelling*, http://dx.doi.org/10.1016/j.jocm.2014.05.003, 13:3-23.

Swait, Joffre and Wiktor Adamowicz (2014b), "Choosing how best to choose: Antecedent Volition and decision process representation in discrete choice models", *Journal of Choice Modelling*, doi:10.1016/j.jocm.2015.01.003, 13:1-2.

Synovate (2005), Quantitative evaluation of alternative food signposting concepts, report prepared for the f.s.a. Technical report

Thacher J, Morey E, Craighead WE (2005) Using patient characteristics and attitudinal data to identify treatment preference groups: a latent-class model. *Depress Anxiety* 212:47–54

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

Thorndike A. N., L. Sonnenberg, J. Riis, S. Barraclough, D. E. Levy (2012) A 2-phase labeling and choice architecture intervention to improve healthy food and beverage choices. American Journal of Public Health, 102(3):537–533.

Thurstone L. (1927), A law of comparative judgment. *Psychological Review*, 34, 273–286.

Train K. E. (2003), Discrete choice methods with simulation. Press Syndicate of the University of Cambridge, Cambridge.

Van Herpen E., van Trijp H.C.M. (2011) Front-of-pack nutrition labels. Their effect on attention and choices when consumers have varying goals and time constraints. *Appetite*, 57:148-160.

Viera A.J., R. Antonelli (2015) Potential effect of physical activity calorie equivalent labeling on parent fast food decisions, *Pediatrics*, 135(2), 376–382

Visschers V.H.M., C. Hartmann, R. Leins-Hess, S. Dohle, M. Siegrist (2013) A consumer segmentation of nutrition information use and its relation to food consumption behavior, *Food Policy*, 42, 71–80

Wasserstein R. L., Lazar N. A. (2016) The ASA's Statement on p-Values: Context, Process, and Purpose, *The American Statistician*, 70:2, 129-133, DOI:10.1080/00031305.2016.1154108

WHO (2015), Obesity Facts, *The European Journal of Obesity*, Vol. 8, Supplement 1, May 2015, 1-272.

Yoo, James & Ready, Richard C., (2014). Preference heterogeneity for renewable energy technology, *Energy Economics*, Elsevier, vol. 42(C), pages 101-114.

Table 1 - Attributes and levels

Attributes	Levels
Sugar	High, Medium, Low
Fat	High, Medium, Low
Saturated	High, Medium, Low
Salt	High, Medium, Low
Price	+50% , +20%, 0, -20% , -50%

Table 2 – Description of nutritional label treatments

Description	Sample	Abbreviation
Text only	High, Medium, Low Text	TXT
Text, Colour	Multiple Traffic Light	MTL
Text, % GDA	% Guideline Daily Amount	GDA
Text, Colour, % GDA	Hybrid	HYB

Table 3 – Summary statistics of estimated models

Model Specification	Log*L*	BIC	AIC	AIC3	CAIC	N. par
MNL model	-11,952.1	23,971.0	23,924.2	23,934.2	23,981.0	10
4-Class model (LCM)	-8,961.7	18,210.7	18,009.5	18,052.5	18,253.7	43
4-Class model (LCM) 2-scale	-8,700.5	17,701.6	17,490.9	17,535.9	17,746.6	45
4-Class model (LCM) 2-scale with Covariates	-8,638.3	17,737.5	17,414.6	17,483.6	17,806.5	69
4-Class model (LC-RPL) 2-scale with Covariates	-8,420.2	17,381.6	17,002.4	17,083.4	17,462.6	81

Table 4 – Estimates from Multinomial Logit Model

Attributes	Coeff.	|z-value|
price	-0.01	-14.61
sug_Low	0.11	3.37
sug_High	-0.26	-7.60
fat_Low	0.17	5.25
fat_High	-0.26	-7.65
stfat_Low	0.03	0.85
stfat_High	-0.46	-13.43
slt_Low	0.07	1.97
slt_High	-0.36	-10.63
SQ	0.32	16.38
Pseudo-R^2		0.0408
Log-likelihood		-11,952.1

Table 5 – Estimates from Latent Class Model

Attributes	Healthy all rounders		High fat lovers		Selectively Focussed		Moderately interested			
	Class 1		Class 2		Class 3		Class 4		Wald	*p*-value
	Coeff.	|z-value|	Coeff.	|z-value|	Coeff.	|z-value|	Coeff.	|z-value|	(Likel. Ratio)	
Class size (Preference)	38.2		31.8		19.6		10.5			
Food choice attributes:										
price	-0.01	4.2	-0.04	5.9	-0.06	3	-0.64	7.3	98.26	<0.01
sug_Low	0.6	4.6	1.08	4.1	-0.59	1.3	-1.13	2.2	43.76	<0.01
Mean: *sug_High*	-0.96	6	0.91	3.9	-7.07	6.5	-1.15	2.6	84.73	<0.01
St. dev.: *sug_High*	0.67	4.4	0	0	1.42	1.7	0	0	17.14	<0.01
fat_Low	0.46	3.9	0.15	0.9	-0.16	0.4	-0.57	1.2	94.67	<0.01
Mean: *fat_High*	-1.15	6.5	0.34	1.8	-10.3	7.4	-1.53	3.3	50.59	<0.01
St. dev.: *fat_High*	0.53	2.7	0	0	3.08	4.1	0	0	106.01	<0.01
stfat_Low	0.5	3.9	-0.62	3.1	1.84	4.5	-1.23	2.6	60.03	<0.01
stfat_High	-1.09	7.1	-1	4.9	-9.67	6.9	-0.9	1.8	91.51	<0.01
slt_Low	0.6	3.9	-1.18	5.1	2.93	5.2	-0.27	0.5	74.53	<0.01
slt_High	-0.74	5	-0.54	3.2	-10.15	7.4	-1.14	2.2	79.10	<0.01
Mean: *SQ*	-7.41	6.4	20.38	7.3	-2.58	5.9	-7.57	5.3	24.69	<0.01
St. dev.: *SQ*	8.83	7.6	19.73	7.1	2.43	6.2	8.74	5.9	21.72	<0.01
Membership Equations:										
Intercept	0	--	-0.92	0.2	0.19	0.2	3.63	2.62	(92)*	<0.01
HYB	0	--	0.11	0.3	0.83	2.3	0.3	0.7	(92)*	<0.01
GDA	0	--	-0.6	1.7	0.57	1.6	-0.44	0.9	11.01	0.01
MTL	0	--	-0.74	2.2	-0.11	0.3	-0.2	0.4	(92)*	<0.01
Age (48)	0	--	0.03	3.7	0	0.5	-0.01	1.4	29.72	<0.01
Woman (60)	0	--	0.37	1.5	0.57	2	0.27	0.8	66.53	<0.01
How often read FoPL (2.8)	0	--	-0.61	5.7	-0.08	0.6	-1.08	7	8.15	0.04
Perceived ideal body weight (2.5)	0	--	0.43	2.2	0.04	0.2	-0.19	0.7	12.74	0.01
BMI class (3.8)	0	--	0.09	0.7	-0.34	2.6	-0.2	1.2	(82)**	<0.01
Scale parameter classes	Scale class 1		Scale class 2							
Class size (Scale)	40.7		59.3							
Scale parameter	1 (fixed)		0.16	16.93						
N. respondents		797		N. obs.	11,628	Pseudo R-squared 0.34				
Log-likelihood(AIC)	-8,420.2(17,002)									

* Jointly tested using likelihood ratio test; ** tested across three membership equations using the likelihood ratio test.

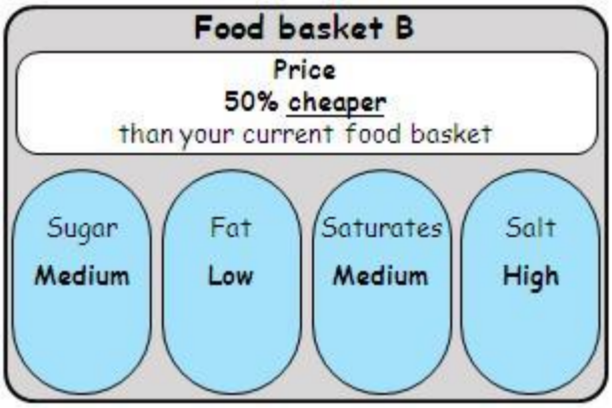
Table 6 – Willingness to Pay estimates (marginal)

Attributes	Class1	Class2	Class3	Class4
sug_Low	46.5	30.7	-10.6	-1.8
sug_High	-74.1	26.0	-126.2	-1.8
fat_Low	35.7	4.2	-2.9	-0.9
fat_High	-88.2	9.8	-183.8	-2.4
stfat_Low	38.6	-17.8	32.9	-1.9
stfat_High	-83.7	-28.5	-172.6	-1.4
slt_Low	46.0	-33.5	52.3	-0.4
slt_High	-56.9	-15.2	-181.3	-1.8

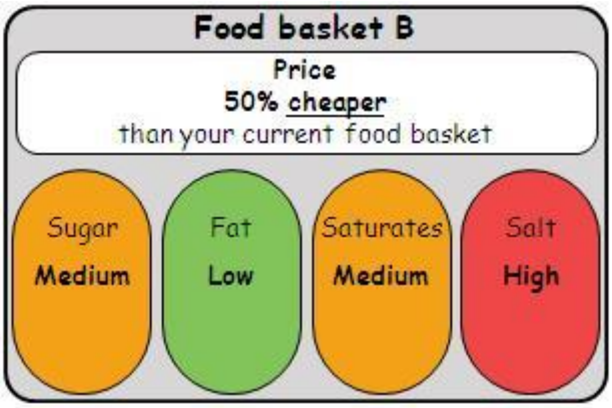
Table 7. OLS results for positive and negative choice probability differences between SQ and healthy basket

Propensity	Status quo basket		Healthy basket	
y=|Pr(sq)-P(healthy)|	y|y>0		y|y<0	
	Estimate	|*t* value|	Estimate	|*t* value|
(Intercept)	0.20510	10.85	0.63310	28.73
GDA from TXT or HYB	-0.06039	6.94	0.03761	4.29
GDA x *Woman*	-0.01705	1.78	-0.01001	1.00
GDA x *Obese*	-0.00801	0.77	0.00770	0.69
GDA x *Misperceived Obese*	-0.04447	4.36	-0.00089	0.09
MTL from TXT or HYB	-0.06259	7.20	0.01590	1.82
MTL x *Woman*	-0.00231	0.24	-0.01853	1.86
MTL x *Obese*	-0.02298	2.22	0.02081	1.85
MTL x *Misperceived Obese*	-0.03209	3.18	0.01162	1.17
Obese	0.05558	6.99	-0.04114	4.67
Obese|Perceived Normweight	0.00482	0.55	-0.03099	3.31
Age	0.00850	13.11	0.00010	0.13
Age^2	-0.00005	7.46	-0.00003	3.58
How often read FoPL	-0.07176	56.45	0.02693	20.72
Ideal Body Image	0.05721	19.39	-0.01030	3.38
Woman	0.02005	2.91	0.01276	1.79
Adjusted R-squared:	0.8741		0.5211	
F-statistic:	426.2 d.f. 15,904		75.21 d.f. 15,1008	
p-value:	< 2.2e-16		< 2.2e-16	

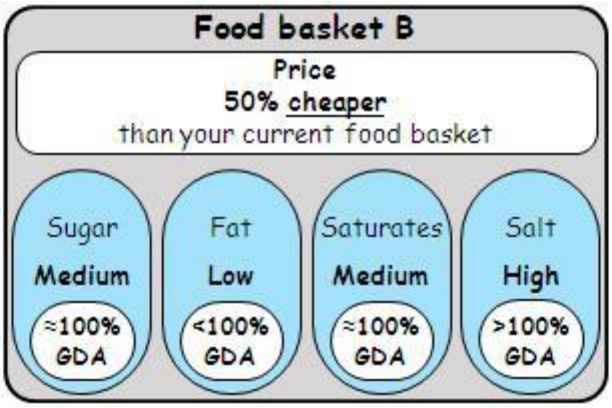
Figure 1 – Examples of Food baskets (choice tasks)



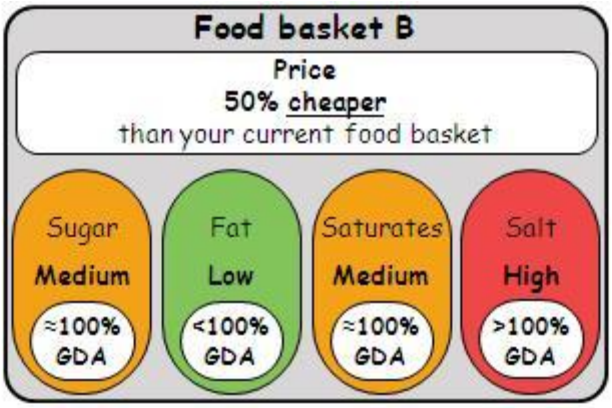
i) Text only



ii) Multiple Traffic Light



iii) % Guideline Daily Amount



iv) Hybrid

Figure 2 – Class membership probabilities by age increase for a baseline respondent described as male, MTL label format, perceived own body weight as ideal and with normal BMI.

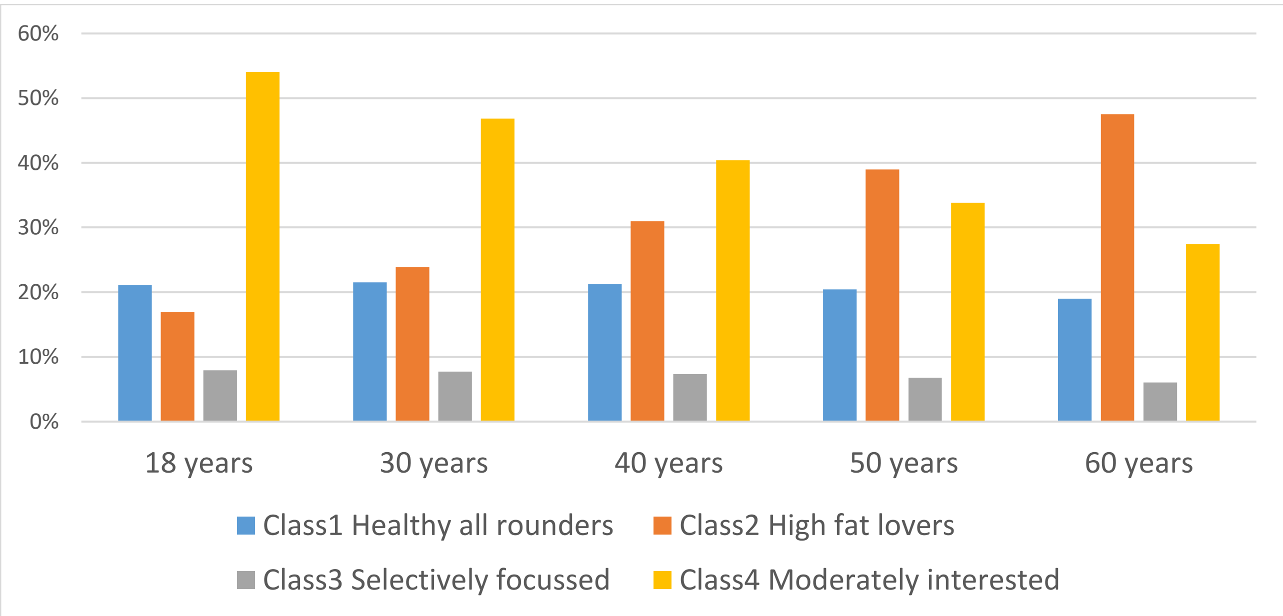


Figure 3 - Class membership probabilities by age increase and by reading or not nutritional information on FoPL. Baseline respondent: woman, HYB label format, perceived own body weight as ideal and with normal BMI.

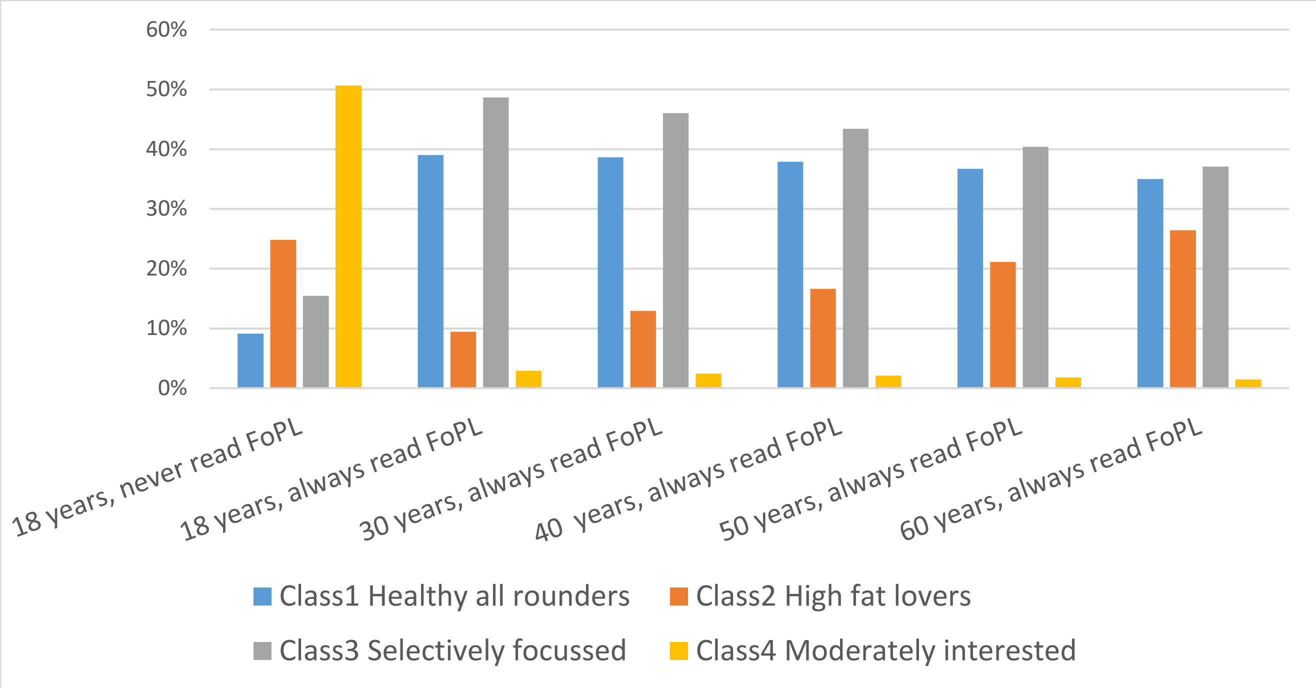


Figure 4 - Class membership probabilities by bodyweight increase and by reading or not FoP labels. Baseline respondent: 30 years old women, normal BMI, perceive their body weight as ideal, and have HYB label format.

Figure 5 - Distributions of individual marginal WTP estimates for high fat and high sugar level.

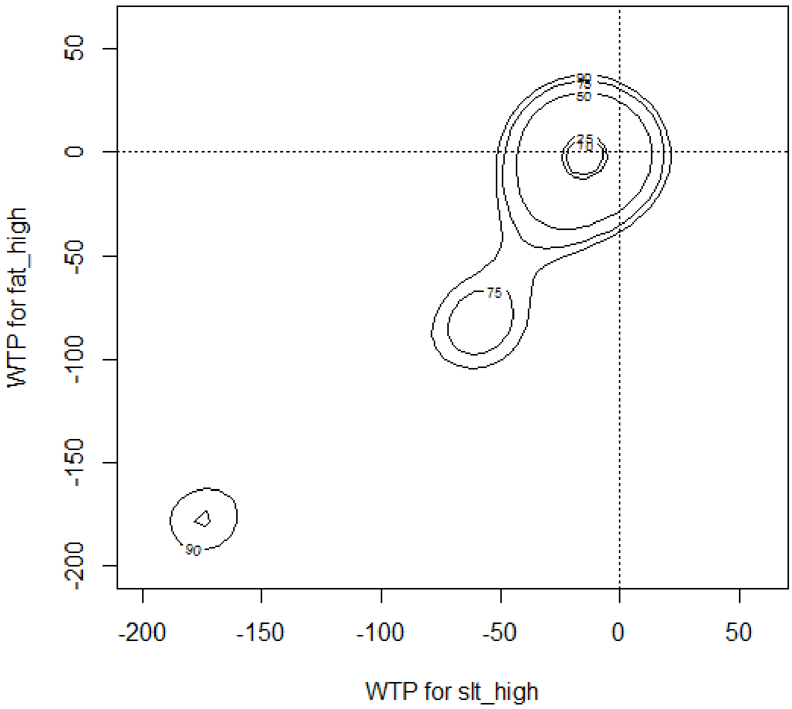
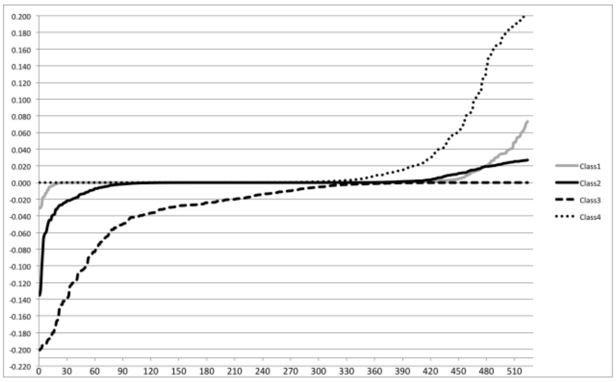
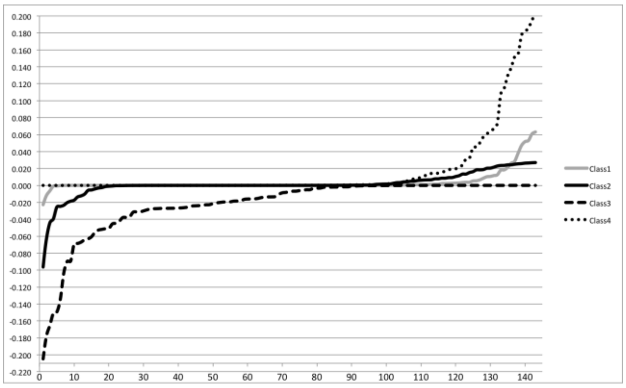


Figure 6 – Marginal effects of FoPL types on predicted class membership posterior probabilities (TXT as a baseline).

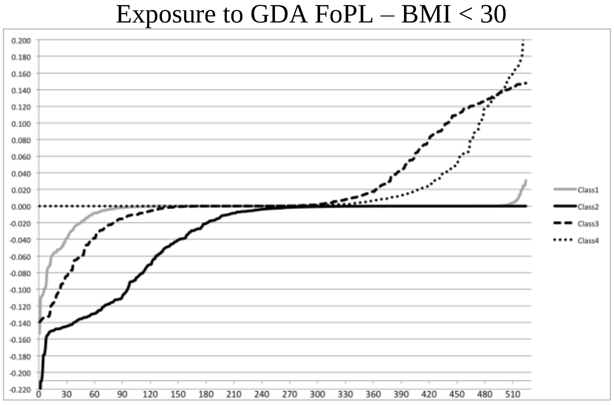
Exposure to HYB FoPL – BMI < 30



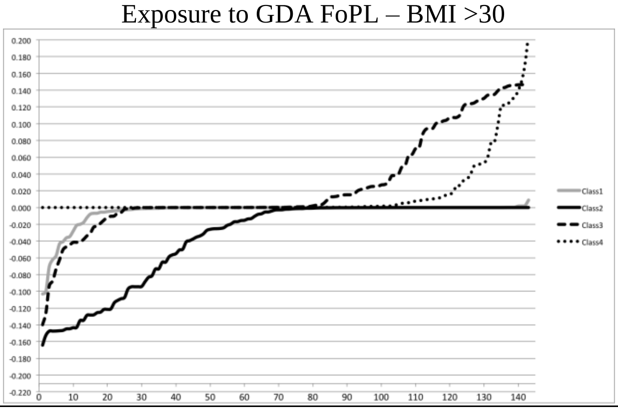
Exposure to HYB FoPL – BMI > 30



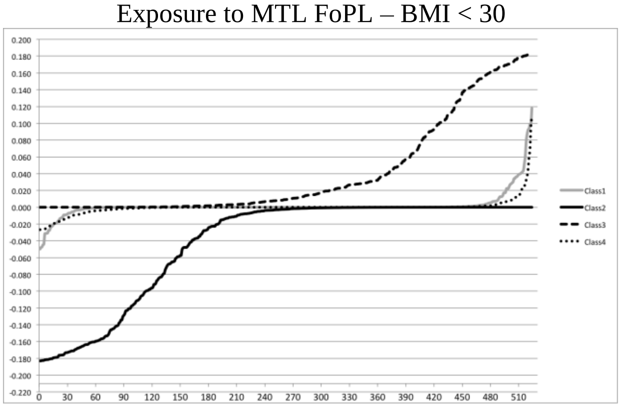
Exposure to GDA FoPL – BMI < 30



Exposure to GDA FoPL – BMI >30



Exposure to MTL FoPL – BMI < 30



Exposure to MTL FoPL – BMI > 30

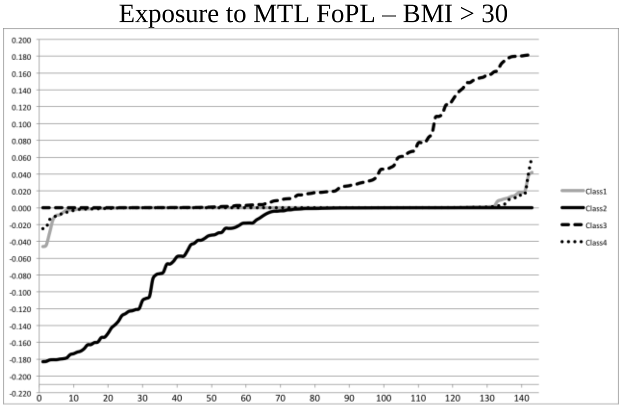


Figure 7 – – Effect of FoPL types predicted choice between SQ and healthy baskets by BMI groups

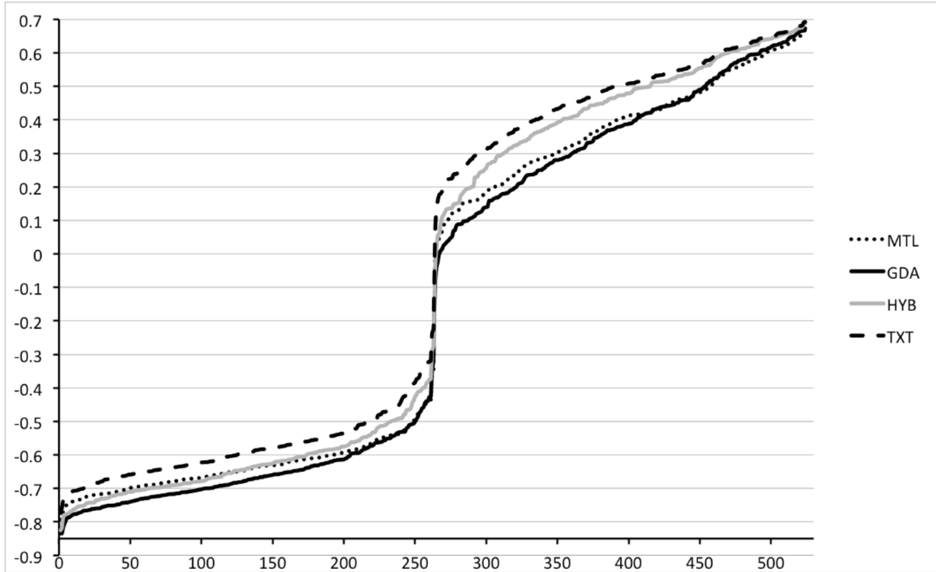
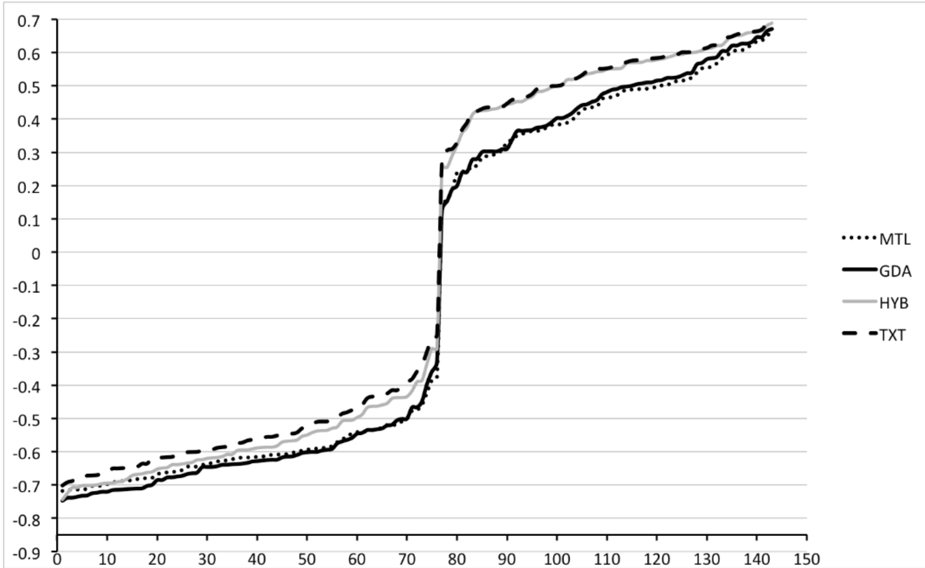
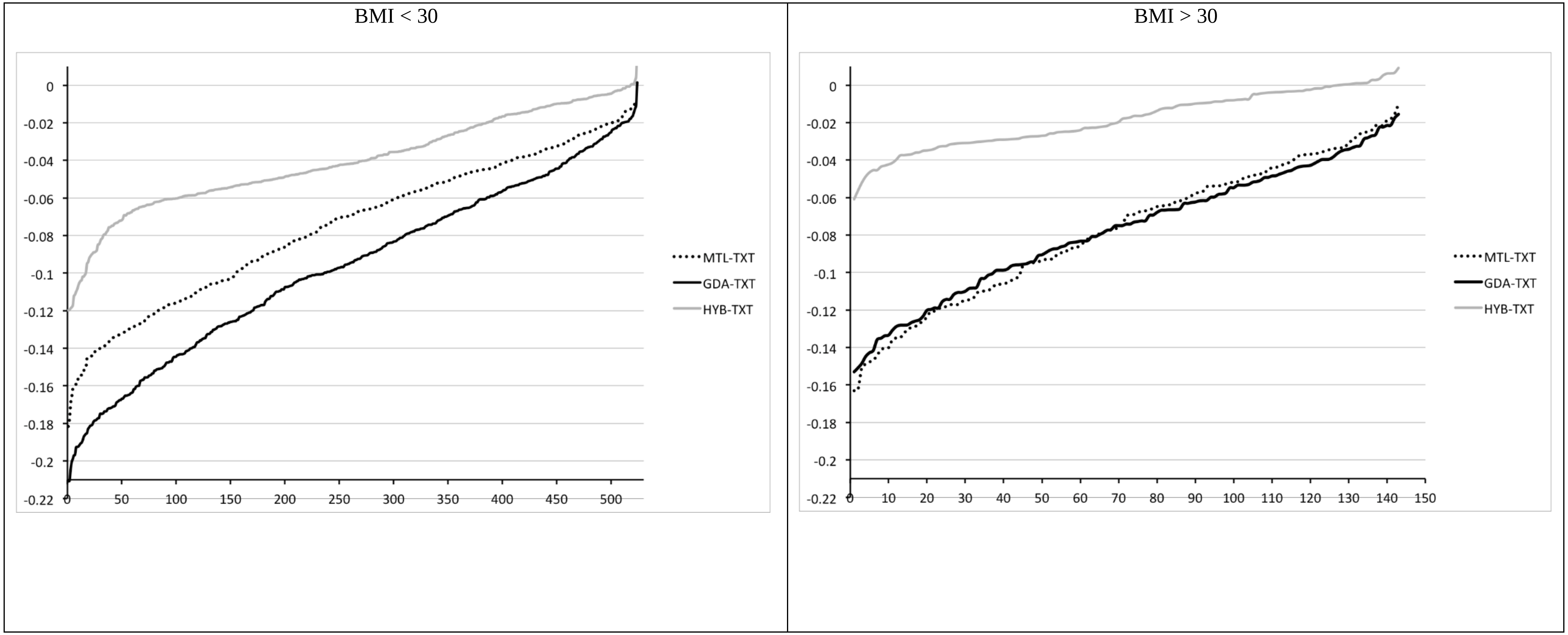
Differences of predicted probabilities between SQ and healthy food basket (BMI < 30)	Differences of predicted probabilities between SQ and healthy food basket (BMI > 30)
	

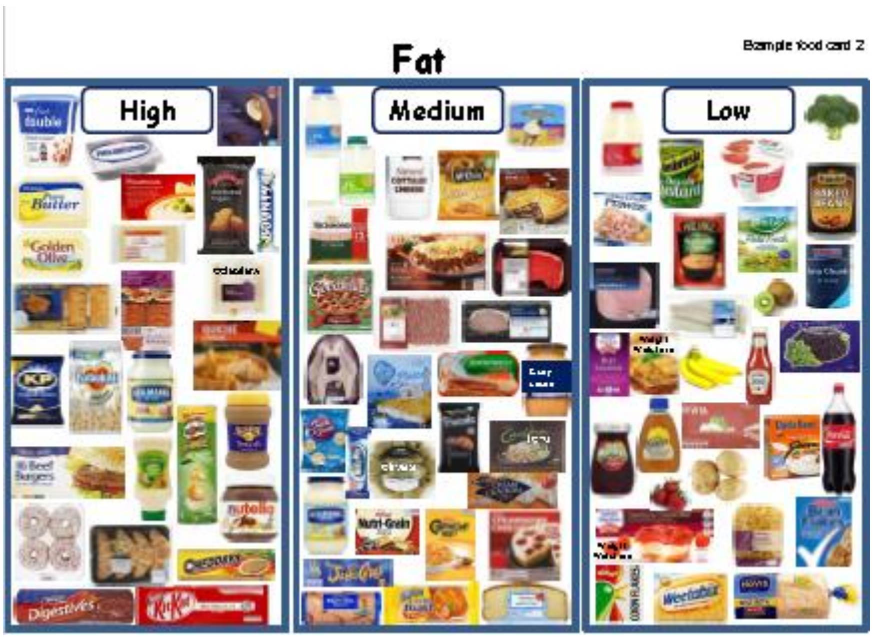
Figure 8 – Selection of the SQ probabilities differences between other FoPL and TXT by BMI groups



Appendix

Example of food card for sugar

Example of food card for fat



Correlation of BMI with SQ basket attributes' levels

	bmi	***sug_l***	***fat_l***	***stfat_l***	***slt_l***	***sug_h***	***fat_h***	***stfat_h***	***slt_h***	***price***
bmi	1.00									
sug_l	-0.04	1.00								
fat_l	-0.13	0.64	1.00							
stfat_l	-0.15	0.63	0.82	1.00						
slt_l	-0.08	0.57	0.58	0.60	1.00					
sug_h	0.17	-0.70	-0.64	-0.61	-0.51	1.00				
fat_h	0.22	-0.53	-0.76	-0.68	-0.47	0.74	1.00			
stfat_h	0.21	-0.50	-0.67	-0.76	-0.48	0.71	0.84	1.00		
slt_h	0.20	-0.48	-0.56	-0.59	-0.70	0.65	0.66	0.70	1.00	
price	0.23	0.02	-0.05	-0.07	-0.02	-0.03	0.04	0.07	0.09	1.00