Title: The role of regret minimization in lifestyle choices affecting the risk of coronary heart disease

Authors: Marco Boeri, Alberto Longo, José M. Grisolía, W. George Hutchinson, Frank Kee

PII: S0167-6296(12)00168-3
DOI: doi:10.1016/j.jhealeco.2012.10.007
Reference: JHE 1624

To appear in: Journal of Health Economics

Accepted date: 11-10-2012

Please cite this article as: Boeri, M., Longo, A., Grisolía, J.M., Hutchinson, W.G., Kee, F. The role of regret minimization in lifestyle choices affecting the risk of coronary heart disease, Journal of Health Economics (2010), doi:10.1016/j.jhealeco.2012.10.007

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
The role of regret minimization in lifestyle choices affecting the risk of coronary heart disease

Marco Boeri, Alberto Longo, José M. Grisolía, W. George Hutchinson, Frank Kee

Queen's University of Belfast, Gibson Institute for Land, Food and Environment, School of Biological Sciences, 97 Lisburn Road, Medical Biology Centre, Belfast BT9 7BL, United Kingdom

E-mail: m.boeri@qub.ac.uk

Research Highlights:
This paper introduces the discrete choice paradigm of Random Regret Minimization (RRM) to the field of health economics
The combined use of RRM and Random Utility Maximization (RUM) models provide useful behavioural insights on choice
Whilst the RUM is suitable for calculating welfare estimates, the RRM highlights how anticipated regret affects choices
We find that the choices of overweight or obese respondents and smokers are more likely to conform to the RUM approach
We also find that the choices of people in good health and with higher education are more likely to reflect the RRM approach

Abstract
This paper introduces the discrete choice model-paradigm of Random Regret Minimization (RRM) to the field of health economics. The RRM is a regret-based model that explores a driver of choice different from the traditional utility-based Random Utility Maximization (RUM). The RRM approach is based on the idea that, when choosing, individuals aim to minimize their regret – regret being defined as what one experiences when a non-chosen alternative in a choice set performs better than a chosen one in relation to one or more attributes. Analysing data from a discrete choice experiment on diet, physical activity and risk of a fatal heart attack in the next ten years administered to a sample of the Northern Ireland population, we find that the combined use of RUM and RRM models offer additional information, providing useful behavioural insights for better informed policy appraisal.

*Keywords*: Random Regret Minimization; Random Utility Maximization; dietary choices; physical activity; coronary heart disease risk; behavioural economics.
1. Introduction

Discrete choice experiments (DCE) are a survey-based technique used to investigate the trade-offs that people are prepared to make between different hypothetical goods or services, where respondents are shown alternative variants of the good or service described by a set of attributes and are asked to choose the most preferred one (Ryan and Hughes, 1997; Vick and Scott, 1998). Since the first applications of DCE in health economics to value patient experiences (Ryan and Hughes, 1997; Ryan, 1999), this technique has become widely used to investigate a wide range of policy questions (Ryan et al., 2008). De Bekker-Grob et al. (2010), in a recent systematic review, found an increase in the number of DCE health economics studies, starting with about three applications per year in the 90s to fourteen in 2008. These studies base their analysis on the linear-in-parameters Random Utility Maximization (RUM) model (McFadden 1974; Train 2009), which assumes that respondents’ choices are driven by utility maximization - e.g. respondents choose the option in a choice set that maximizes their expected utility. The RUM’s popularity is mainly due to its strong economic foundations, its conceptual elegance and its formal tractability. Many RUM-based models have closed-form formulations for choice probabilities, and can be easily coded and estimated using standard discrete choice-software packages. The simplest and most popular RUM is the Multinomial Logit Model (RU-MNL), developed by Luce (1959), Marschak (1960) and McFadden (1974).¹

¹ Other RUMs have been developed to overcome some limitations of the RU-MNL. For example, the mixed logit, the nested logit and the multinomial probit model (Train, 2009) relax the RU-MNL assumption of the Independence of Irrelevant Alternatives (IIA). Other RUMs have added behavioural realism to the model (e.g. Swait, 2001; Arentze and Timmermans, 2007; Kivetz et al. 2004, and Zhang et al. 2004, among others). These models are, however, more complicated to estimate and their results are often more difficult to interpret than the output of the RU-MNL.
Recently, Chorus (2010) introduced an approach based on Random Regret Minimization (RRM), as a complement to the RUM for the analysis of DCE data. The Random Regret Minimization Multinomial Logit Model (RR-MNL) is built on the idea that, when choosing, individuals aim to minimize their regret rather than to maximize their utility – regret being defined as what one experiences when a non-chosen alternative performs better than a chosen one, on one or more attributes.\(^2\) The interesting aspect of the RR-MNL is its ability to translate the notion of regret minimization into an operational, easily estimable, discrete choice model that can be employed as a complement to the RU-MNL to analyse risky and riskless choices. Despite these advantages, the departure from the utility maximization choice paradigm makes the RR-MNL not particularly suitable for welfare analysis (e.g., Chorus & de Jong, 2011). This contrasts with RUM-based models, which are firmly rooted in welfare-economics.

Evidence from previous studies suggests that regret can be an important factor in both medical decision making (Smith, 1996; Djulbegovic et al. 1999 and Sorum et al. 2004) and personal healthcare decisions (Brehaut et al., 2003 and Ziarowski et al., 2009), particularly when the outcome of the choice may be unfavourable. Smith (1996) explored the concept of regret in healthcare decision making from the patient’s viewpoint, arguing that valuation techniques based on the utility theory, such as QALYs, will not necessarily reflect the true preferences of the individual. Ziarowski et al. (2009) found that anticipated regret played an important role in the decision to vaccinate adolescent girls against HPV. Furthermore, several studies highlighted that regret can be experienced by patients before or after cancer-related decisions (e.g., Hu et al., 2003, Montgomery et al., 1999, Clark et al. 2001, 2003, Payne et al. 2000, and Joseph-Williams et al., 2010).

\(^2\) The idea that regret is an important determinant of choice behaviour is not new and is well established theoretically and empirically in many fields, including marketing (e.g. Simonson, 1992; Zeelenberg and Pieters, 2007), microeconomics (e.g. Loomes and Sugden, 1982; Sarver, 2008), psychology (e.g. Zeelenberg, 1999; Connolly, 2005), the management sciences (e.g. Savage, 1954; Bell, 1982), transportation (e.g., Chorus et al., 2006, 2009) and health (Ziarowski et al., 2009).
Therefore when analysing DCE data, it may be worthwhile, especially when researchers suspect that both utility maximization and regret minimization may play important roles in choosing, to approach the data from these two different behavioural perspectives, and estimate both RU-MNL and RR-MNL. In this paper, we use both models to investigate individuals’ lifestyle choices, described in terms of dietary habits, physical activity, risk of a fatal coronary heart disease (CHD) event in the next ten years and cost. The data were collected surveying a sample representative of the population of Northern Ireland aged 40-65. To further explore the characteristics of respondents that are mainly associated with a utility maximization or a regret minimization approach, we run a binary logit model that explains the likelihood of the two decision choice paradigms using respondents’ characteristics as explanatory variables. To our knowledge, this is the first application that employs both RU-MNL and RR-MNL in health economics. Furthermore, this is the first attempt to describe the profile of “regret minimizers” and “utility maximizers” in the context of individuals’ lifestyle choices.

2. Modelling DCE: Regret vs. Utility

2.1 Modelling Utility and Regret

The RU-MNL is based on the utility maximization theory (Thurstone, 1927; Manski, 1977) and assumes that, when choosing, a respondent maximizes his/her utility function:

\[ U_{ni} = V_{ni} + \lambda_{ni}, \]  

(1)

where \( U_{ni} \) is the utility function maximized by respondent \( n \) when choosing alternative \( i \), \( V \) is the observed part of utility and \( \lambda \) is the unobserved part of the utility Extreme Value Type I-
The observed linear-in-parameter utility function is \( V_{ni} = \beta' X_{ni} \), where \( X \) is a vector of \( m \) attributes describing alternative \( i \), and \( \beta \) is a vector of \( m \) parameters to be estimated. Each parameter \( \beta_m \) captures the slope of the utility function for the attribute \( m \). A positive and significant parameter estimate means that individuals consider the attribute as positively affecting their utility; a negative and significant parameter estimate suggests that respondents dislike alternatives characterised by higher levels of the corresponding attribute \( m \). In this context the probability for individual \( n \) of choosing alternative \( i \) over any other alternative \( j \) in the choice set is represented by a RU-MNL (McFadden, 1974):

\[
Pr_{ni} = \frac{e^{V_{ni}}}{\sum_{j=1}^{J} e^{V_{nj}}}.
\]  

(2)

The RRM approach postulates that, when choosing between a set of alternatives, decision-makers aim to minimize anticipated regret:

\[
\beta_{ni} = R(\theta, X_{ni}) + \beta_{ni}
\]  

(3)

where \( \beta_{ni} \) is the regret function minimised by respondent \( n \) when choosing alternative \( i \), \( R \) is the observed part of regret, \( \theta \) is a vector of parameters to be estimated and \( \beta \) is the unobserved part of regret Extreme Value Type I-distributed. The observed part of the regret function as described by Chorus (2010), is:

---

3 The assumption that the error terms are independently and identically distributed extreme value Type I is well rooted in choice modelling (see for example Ben-Akiva nd Lerman, 1985, Train 2009). The mean of the extreme value distribution is not zero; however, with DCE, researchers are interested in the difference between alternatives, hence the difference between two error terms that have the same mean is zero. The difference between two extreme value distributions is distributed logistic. Therefore, using the extreme value distribution for the errors, which implies the logistic distribution for the error differences, is almost the same as assuming that the errors are independently normal. As Train (2009, page 35) points out, “the extreme value distribution gives slightly fatter tails than the normal, which means that it allows for slightly more aberrant behavior than the normal. Usually, however, the difference between extreme value and independent normal errors is indistinguishable empirically.”
In equation (4), the observed part of regret is defined as the sum of all so-called binary regrets associated with bilaterally comparing alternative \( i \) with all the other alternatives \( j \) in the choice set. This comparison is done for all attributes \( m \). The parameter \( \theta_m \) captures the slope of the regret-function for attribute \( m \). It reflects the upper bound of the extent to which a unit increase in the relative performance of an attribute influences the level of regret that is associated with a comparison of another alternative. In other words, the estimated coefficients reflect the potential contribution of an attribute to the regret associated with that alternative. A positive coefficient for an attribute suggests that regret increases when the difference in that attribute between a chosen and a non-chosen alternative increases. A negative coefficient for an attribute implies that regret increases when a considered alternative is compared to another alternative with a decreasing value on that attribute. In other words, decreasing attribute-values of a competing alternative lead to increases in regret associated with the considered alternative when the attribute has a negative sign, such as is the case with the cost attribute.

Acknowledging the fact that minimizing the random regret is mathematically equivalent to maximizing the negative of the random regret, the probability for individual \( n \) of choosing alternative \( i \) over any other alternative \( j \) in the choice set in this context is represented by a RR-MNL (Chorus, 2010):

\[
P_{R_{ni}} = \frac{e^{R_{ij}}}{\sum_{j=1}^{N} e^{R_{ij}}} \tag{5}
\]

Equation (5) is similar to equation (2). Both models result in logit-choice probabilities and are equally parsimonious: each parameter estimated from a RR-MNL has a counterpart in
a RU-MNL; note that, when there are only two alternatives to consider, the RR-MNL and the RU-MNL generate the same choice probabilities.\(^4\)

It is worth emphasizing that, while the RU-MNL assumes that respondents use a fully compensatory behaviour, the RR-MNL implies semi-compensatory decision makers. A fully compensatory behaviour implies that individuals, when trading off attributes defining an alternative, follow a decision rule that allows a positive evaluation of an attribute to compensate for a negative evaluation of another attribute. A semi-compensatory behaviour, on the contrary, assumes that individuals adopt a decision rule that does not allow for a bad performance of an alternative with respect to an attribute to be directly compensated by a good performance of another attribute. This means that improving an alternative in terms of an attribute on which it already performs well relative to other alternatives generates only small decreases in regret, whereas a corresponding deterioration of the performance on another equally important attribute on which the alternative has a poor performance relative to other alternatives may generate substantial increases in regret. As a result, the extent to which a strong performance on one attribute can compensate for a poor performance on another depends on the relative position of each alternative in the choice set. This difference in compensatory behaviour between the two models emerges directly from the underlying structure of the two choice paradigms. Whether respondents use a fully-compensatory or a semi-compensatory decision making strategy is a matter for empirical investigation that the analyst can uncover by estimating both RU-MNL and RR-MNL.

One of the strengths of the RU-MNL is its suitability in retrieving welfare measures such as Willingness to Pay (WTP). Following the formula described in McFadden (1999), WTP

---

\(^4\) The RR-MNL, following its recent introduction in transportation (Chorus, 2010), has so far been employed for analysing choices among shopping destinations (Chorus, 2010), parking lots (Chorus, 2010), road pricing policies (Chorus et al., 2011), departure times (Chorus and de Jong, 2011), travel mode (Pathan, 2010), site choice for recreational activities (Thiene et al., 2012 and Boeri et al., 2012) and online dating-profiles (Chorus and Rose, 2011).
can indeed be computed as the negative of the ratio between the coefficient estimate of an attribute and the coefficient estimate for the cost:

\[ WTP_{m} = -\frac{\beta_{m}}{\beta_{cost}} \] (6)

On the other hand, the RR-MNL does not offer the researcher such convenient welfare measures. However, the ratios of attribute parameter estimates can be used to explore the relative effect on regret of hypothetical policy changes.

2.2 The characteristics of utility maximizers and of regret minimizers

After estimating the RU-MNL and the RR-MNL, we are interested in examining which respondent characteristics are more likely to describe either a “regret minimizing” or a “utility maximizing” decision maker. We therefore compute the contribution to the value of the Log-likelihood function for each respondent’s sequence of choices under both the RU-MNL and the RR-MNL. We then create a dummy variable equal to one when the respondent’s sequence of choices is better described by the regret minimization approach, i.e. when the Log-likelihood of the RR-MNL outperforms the Log-Likelihood of the RU-MNL, and zero otherwise. Next, we run a logit regression on this variable where the characteristics of the respondents are used as explanatory variables:

\[ P(d|y_{n}) = \frac{1}{1+\exp(-d \cdot \beta \cdot Z_{n})}. \] (7)

In Equation (7), \( d \) is a dummy variable equal to 1 when the sequence of choices, \( y_{n} \), faced by respondent \( n \) is better described by a regret minimization strategy, and 0 otherwise, \( \beta \) and \( Z \) are parameters of the logit regression on this variable and \( Z \) is a vector representing the characteristics of respondent \( n \).
3. The case study

The survey instrument

Given the recent government concern for the rising levels of overweight and obesity in Northern Ireland, we administered a DCE questionnaire to assess the hypothetical trade-offs that a representative sample of the adult population of Northern Ireland is willing to make to change their lifestyles to improve their health. The questionnaire was administered through Computer Assisted Personal Interviews (CAPI), which allowed us to develop respondent’s tailored DCE questions built using some answers given by the respondent in the initial part of the questionnaire. The questionnaire was finalized after extensive focus groups, one-on-one interviews with members of the general public and discussions with stakeholders.

The questionnaire began with general questions about the respondent’s age, gender, health, weight, height and family history of cardiovascular diseases. These data were incorporated into the QRISK1 prediction algorithm developed by the University of Nottingham for cardiovascular diseases risk in the British population (Hippisley-Cox et al., 2008) to estimate the respondent’s own cardiovascular diseases risks. The outcome of the algorithm, in terms of the percentage risk of having a fatal heart attack caused by CHD in the next ten years, was then shown to the respondent and later reported in the “current choice” of the DCE questions. Next, we elicited the respondent’s engagement with moderate-intensity and vigorous-intensity activities, using the UK National Health Service version of the International Physical Activity Questionnaire (IPAQ) (Craig et al, 2003). We then collected information about the respondent’s current eating habits using the Block Questionnaire (Block et al, 2000), a tool developed in the nutritional literature, that offers a snapshot of an individual’s

---

5 In Northern Ireland, 59% of adults are either overweight or obese, 25% are sedentary (a person who has not performed any activity of at least a moderate level, lasting 20 minutes, on at least one occasion in the last 7 days) and only 38% meet the recommended level of physical activity (30 minutes of moderate activity on at least 5 days a week) (DHSSPS, 2011).
levels of energy intake through questions about the eating frequency and the portion size for
selected items. We adapted the Block Questionnaire to the Irish diet considering the main
sources of energy (Joyce et al, 2007). Respondents were asked the frequency of consumption
of these items from ‘never’ to ‘five or more times a week’ and the portion size. Each item
Chiabai (2007), we provided the respondent with a risk tutorial using visual aids to explain
the concept of probability. Familiarity with the concept of probability would later help the
respondent when answering the DCE questions.

The questionnaire then presented ten DCE questions.\footnote{Caussade et al. (2005) showed that setting between 6 to 13 choice situations seems to be optimal because this range minimizes the error variance of the estimates.} Each DCE question entailed three
alternatives: the respondent’s current lifestyle and two alternative hypothetical lifestyles.
Each alternative was described by a diet, an amount of physical activity, a risk of a fatal heart
attack in the next ten years and a cost to the respondent. An example of DCE question is
shown in figure 1.

![Figure 1 about here.]

Physical exercise was defined as minutes spent doing moderate-intensity physical
exercise per day. The levels of this attribute were: the current level of physical exercise, and
increases by 10, 20, 30, or 40 minutes per day from an individual’s current level. The cardio
vascular risk was defined as the percentage risk of a fatal heart attack caused by CHD in the
next ten years. The level for the current lifestyle was the one resulting from the QRISK1
prediction algorithm. The levels for the other alternative scenarios were calculated as a
reduction in such a risk by 40%, 50%, 60%, 75% and 85%. Therefore, for a respondent
whose current risk was equal to 5%, a 50% reduction would result in a risk of 2.5%. Cost was described as an increase in the money spent on food and physical exercise per week compared to the current level. The levels for this attribute were set to 0 for the current lifestyle and increases by 2, 5, 7, 10, 15, and 18 GBP\textsuperscript{8} per week for the other hypothetical alternatives. The diet attribute was the most complex to define. In focus groups we ruled out the idea of using a hypothetical food basket described in terms of an abstract nutritional content, as such description would not convey well the information of the ‘taste’ of food and the ‘sacrifice’ resulting from reducing the consumption of favourite food products and from increasing the consumption of fruit and vegetables. We also discarded the possibility of using flagship unhealthy food items, such as pizza, chicken curry, fish and chips, or an “Irish Breakfast”, as they might not have been relevant to all respondents, and might have reduced the credibility of the DCE and questioned the incentive compatibility of the survey instrument. Therefore, we used the information collected from the Block Questionnaire. For each respondent, we selected the five food items most frequently consumed.\textsuperscript{9} This information was presented to respondents under the “current choice.” The alternative hypothetical scenarios were described in terms of reduction in the consumption of these five items and an increase in fruit and vegetables. We selected four levels for the diet attribute defined in terms of overall fat content. Considering the current diet as the reference value, we defined light, medium, high and restricted diets, corresponding to reductions in fat intake by 10\% (light), between 20\% and 30\% (medium), between 40\% and 50\% (high) and between 60\% and 75\% (restricted) from the current diet respectively. We then used these levels of fat in our econometric analysis. This approach allowed us to compare diets across respondents.

\textsuperscript{8} During the time of the survey, 1 GBP (£) was worth about 1.60 US Dollars.
\textsuperscript{9} In five focus groups we found that individuals were able to answer choice experiments questions described by five food items, and were struggling when additional food items were included. The CAPI automatically selected for each respondent the five items that were mostly and most frequently consumed.
and build a variable expressed in terms of reduction in grams of fat from the current diet.\textsuperscript{10} Table 1 shows the attributes and their levels used in the DCE.

[Table 1 about here.]

Different respondents received different choice scenarios that were generated through a Bayesian efficient design, based on the minimization of the $D_b$-error criterion (Ferrini and Scarpa, 2007). The design was implemented in two waves, in February and July 2011.

The data

A survey of 493 respondents was administered to randomly selected households providing a representative sample of the Northern Ireland population aged between 40 and 65 years old. These age boundaries were set to keep the survey credible since it was found very difficult for younger people to consider their risk of dying for reasons linked to cardiovascular disease in the next 10 years.\textsuperscript{11} In addition, the actual risk of a fatal heart attack caused by CHD for people younger than 40 is close to zero (Conroy et al. 2003). As the actual risk was used as the baseline scenario from which reductions in risk were calculated, these improvements would have looked negligible to the eye of potential respondents younger than 40 years.

Table 2 provides descriptive statistics for the sample. Average net annual income per household was £29,051. Women were slightly over represented, with 57\% of individuals in the sample. Regarding Body Mass Index (BMI), the mean was 26.85, with 31.03\% of overweight and 26.37\% of obese respondents. This is comparable with the population data of

\textsuperscript{10} We are aware that this approach may lead to a researcher bias, as in our econometric model we assume that respondents trade off grams of fat when choosing different lifestyles. However, we are unaware of a more efficient approach to investigate comparable dietary choices across respondents in a tractable way for a DCE survey. Our approach leads to comparable choices, and choices meaningful to respondents. In focus groups we tried adding the information of ‘grams of fat’, but this information appeared to convey a wrong message, as respondents would only grasp the unhealthy message conveyed by ‘grams of fat’ and were unable to consider the ‘taste’ and ‘sacrifice’ elements of reducing the consumption of their favourite food items.

\textsuperscript{11} Krupnick et al, 2002 sampled a population of similar age in the USA for valuing mortality risk reduction from public programs aimed at reducing air pollution.
Northern Ireland, which includes 36% of overweight and 23% of obese people (DHSSPS, 2011).

4. Results

Table 3 presents the results from the RU-MNL and the RR-MNL. Both models were estimated in Biogeme 2.2 (see Bierlaire, 2003, 2009). In both models, the dependent variable takes on a value of 1 if a respondent chooses alternative $i$ and zero otherwise.

The table shows that for both models all parameters are highly statistically significant and have the expected signs. Looking firstly at the RU-MNL, we observe that respondents are inclined to follow a diet that entails a sacrifice in terms of reduced fat intake: the positive sign of the coefficient means that, ceteris paribus, the more restricted the diet of an alternative is in terms of fat intake, the more likely respondents are to choose that alternative. The parameter on physical activity is positive, which means that individuals consider physical activity as positively affecting utility. The coefficient estimates for both risk and cost are negative and significant, suggesting that respondents dislike (all the rest been equal) more expensive alternatives, as well as alternatives with higher risk of a fatal heart attack caused by CHD. Table 4 presents WTP computed for each attribute using the output of the RU-MNL, following Equation (6). We find that respondents are willing to pay £0.01 per week for one additional minute of physical activity per week, which means that, on average, respondents are ready to pay about £1 a month for increasing their physical activity by 25 minutes per week. We further find that respondents have a positive WTP for reduction of fat intake, which is equal to about £0.03 per gram of fat. This equates to a WTP of £1.37 per week for a 50 grams reduction of fat intake in a weekly diet, which corresponds to a reduction of about
60 grams of butter in a week. It is not surprising that people are willing to pay to reduce their fat intake. This is consistent with the attitude of the food industry that applies a price premium to low fat products. Respondents also have a negative WTP for increase in risk of dying. This means that they have a positive WTP for risk reduction. Respondents are, on average, willing to pay about £0.80 per week for reducing their risk of dying from a fatal heart attack by one point in the next 10 years. This implies a value of £41.34 per year, and, considering a 3.5% discount rate, over 10 years, this is equal to £355.84.

[Table 4 about here.]

Turning our attention to the RR-MNL, we notice that the Log Likelihood function is similar to the one of the RU-MNL, suggesting that both the RU-MNL and the RR-MNL do a similar job in modelling respondents’ preferences. We also notice that the estimated coefficients from the RR-MNL maintain the same signs as in the RU-MNL. However, the interpretation of the coefficients from the two models is not directly comparable. In fact, a positive and significant coefficient in the RR-MNL suggests that regret increases as the level of that attribute increases in a non-chosen alternative, compared to the level of the same attribute in a chosen one. For example, the positive coefficient estimate for physical activity indicates that regret increases as the level of physical activity increases in a non-chosen alternative, compared to its level in a chosen option. Similarly, for the diet attribute, described in terms of reduction in fat intake, the positive coefficient indicates that regret increases as the difference in fat reduction between a chosen and a non-chosen alternative increases. On the other hand, the negative coefficient for risk (and for cost) suggests that regret decreases as the difference in risk (cost) between the chosen and the non-chosen alternative increases, because the non-chosen alternative has become riskier (more expensive). To further explore the RR-MNL results, we can compute the ratio between two estimated coefficients to investigate the relative importance of the attributes on anticipated
regret. For example, by observing that the ratio of the estimates of the coefficients for risk and cost is close to 1 (0.87), we conclude that the impacts on regret would be similar from policies that changed non-chosen alternatives by either decreasing by one point the risk of dying from a fatal heart attack in the next 10 years or decreasing weekly expenditures by £1. Similarly, we can compute the ratio of the estimates of the coefficients of physical activity and risk. This ratio is equal to -0.0152, suggesting that the impact on regret would be similar from policies that changed non-chosen alternatives by either reducing the risk of dying by 1 point or by increasing the weekly amount of physical activity by about 66 minutes. We can also compare the relative impact on regret of reducing fat content in diet and increasing physical activity levels. The ratio of the estimates of the coefficients for weekly increase in moderate-intensity physical activity and weekly grams of fat reduction is equal to 0.48. This indicates that the impacts on regret would be similar from policies that changed non-chosen alternatives by either increasing by one minute the level of weekly physical activity or reducing by two grams the weekly intake of fat.

Another interesting aspect is the investigation of the drivers of regret minimization and of utility maximization. We report the results from the binary logit model described in section 2.2 to examine which respondents’ characteristics are more likely to explain a regret minimization approach, rather than a utility maximization approach. Table 5 presents the results from this exploration. The dependent variable is equal to one when the Log-likelihood value at the individual level is better for the RR-MNL than for the RU-MNL, and zero otherwise. The independent variables used in this model include: a dummy variable ‘male’ equal to one if the respondent is a male and zero if the respondent is a female; three dummy variables – ‘underweight’, ‘overweight’ and ‘obese’ – equal to one if the respondent is underweight, overweight or obese respectively (normal weight being the reference dummy); a dummy variable ‘good and very good health’ equal to one if the respondent declares to be in
good or very good health, and zero otherwise; a dummy variable ‘high education’ equal to one if the respondent has at least a graduate degree, and zero otherwise; a dummy variable ‘smoker’ equal to one if the respondent is a smoker and to zero if the respondent is a non-smoker. Finally, respondent’s age and its squared value are also included in the model.

[Table 5 about here.]

Table 5 shows that the intercept is positive and significant, suggesting that, ceteris paribus, the variation in respondents’ behaviour not captured by other independent variables in the binary logit model mainly captures regret minimization. When we look at the coefficient estimates of the variables used in Table 5, we notice that there is no statistically significant difference between male and female respondents: we cannot claim that the behaviour of one group or the other is better described by either utility maximization or regret minimization. Looking at the BMI dummy variables, we notice that the coefficient for underweight is not statistically different from the one for normal weight, the reference dummy, captured by the intercept, suggesting that the behaviour of these two groups of respondents is mainly driven by regret minimization. On the other hand, choices made by overweight and obese respondents are more likely to be driven by utility maximization. This result suggests that overweight or obese persons will accord less weight to the negative consequences of not choosing an option scoring better in one or more attributes. We interpret this result considering that those with normal BMI levels seem to consider the long term effects of excessive food intake: for this group of respondents the regret from consuming additional food is an important driver of their choices. On the opposite, those with high BMI levels are more likely to consider the utility from food intake, and are less worried about the long term effects of their choices.
We also find that highly educated persons and individuals in good or very good health are more likely to be regret minimizers. A highly educated person, as well as a person in good or very good health, is probably more concerned about, and less willing to trade off, the features of different healthy and non-healthy alternatives. Indeed, a person in good or very good health aims to minimize the regret of choosing a lifestyle that may be worse in certain features compared to the foregone alternative. It is possible that such people are concerned about the long term effects of making poor choices.

A further interesting outcome of the binary Logit model is that the choices of smokers are better described by the utility maximization paradigm, as the coefficient estimate for smoker is negative and significant. It appears that smokers’ choices are not driven by the regret of making a poor choice, rather, their decisions are driven by the utility they can obtain from their choice.

Finally, the explanatory variables capturing the age of the respondent and its squared value \( \text{Age}^2 \) show that respondents of all ages are more likely to minimize their regret, and after the age of 42 the likelihood to be regret minimizer increases, suggesting that respondents are generally concerned with making a wrong choice, and that this concern tends to increase, all else being equal, the older a person becomes.

5. Conclusions and discussion

This study presented a novel approach for the analysis of DCE health economics data, by estimating both the RU-MNL and the RR-MNL to investigate the drivers of choice behaviour. We estimated the two models on DCE data about individuals’ choices for hypothetical lifestyles defined in terms of diet, physical exercise, risk of a fatal heart attack in the next 10 years and cost. Both models convey useful information. The RU-MNL informs how utility increases if respondents choose alternatives described by stricter diets, increased in physical exercise levels and reductions in risk of a fatal heart attack. The model also tells
us how much respondents are willing to pay for hypothetical programs that increased their expected utility. The RR-MNL gives complementary information. It tells us that regret decreases if the levels of non-chosen alternatives – with respect to chosen alternatives – increase for risk of a fatal heart attack and cost and decrease for physical exercise and diet. We therefore argue that, given the ease with which the RR-MNL can be estimated using standard econometric approaches, analysts could consider applying both the RR-MNL and the RU-MNL to DCE data. This approach would capture the behavioural influences on choices more accurately than assuming in all instances that individuals always make choices within a utility maximization framework or within a regret minimization choice paradigm. Indeed, the RR-MNL ‘tells a story’ that is fundamentally different from the RU-MNL storyline, and, as such, it may lead to new behavioural insights and new policy implications. The importance of the RR-MNL in the field of health economics is yet to be determined, but findings from the field of social-psychology suggest that it may play an important role. In a recent review, Zeelenberg and Pieters (2007) state that regret-minimization has been found to be a particularly important choice behaviour when choices are perceived as important, difficult or uncertain, when the decision-maker expects to receive feedback about chosen and non-chosen options, and when the decision-maker believes that he or she will be held accountable for his or her choices. However, the RR-MNL alone, given its limitations for estimating welfare measures, should not replace the RU-MNL, but used in conjunction with it.

Following this idea, we find interesting insights on both the behavioural premises of respondents’ choices in the context of reducing the risk of dying from coronary heart disease. The RU-MNL informs us about respondents’ WTP for hypothetical programs affecting the amount of fat intake in the diet, the levels of moderate intensity physical exercise, and the risk of a fatal heart attack in the next ten years. In addition, the RR-MNL results tell us how
hypothetical programs would affect anticipated regret. Moreover, exploring which choice paradigm is mostly associated with different respondents’ characteristics provides useful information for policymakers. A policy aimed at improving the lifestyle of the population by promoting physical activity and healthy food baskets should take into consideration that the effects of the interventions may be different according to the choice paradigm used by decision makers. If the goal is to improve the wellbeing of less advantaged individuals, such as smokers, people with high BMI levels, or less well educated persons, then policies that use the leverage offered by a utility maximization paradigm, such as incentives, may have a better chance of success than those premised on regret minimization, such as disincentives or penalties for non-salutary behaviour.

Our research opens different avenues for future research. This is the first study that explores the RRM choice paradigm along with the RUM in health economics. Other studies should explore the use of both models in related contexts, for example, in medical decision making, to explore patients’ preferences for different treatment options or to examine real as opposed to hypothetical choice situations. In addition, whilst our research explored which respondents’ socioeconomic characteristics are more likely to be associated with each of the two choice paradigms, further research should examine the extent that the social-psychology findings are supported by economics findings. Finally, in this paper we assumed that preferences are homogenous across respondents, an assumption that may be relaxed using mixed logit models. We leave this for future research, however, as noted by Chorus (2012), the extension towards RRM-based Mixed Logit models is as straightforward as in the case of RUM-based Mixed Logit models.
References


Table 1: Attributes and levels

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diet (reduction in the consumption of the respondent’s five most energy intensive food items)</td>
<td>Current, light, medium, high and restricted diet</td>
</tr>
<tr>
<td>Cost (GBP per week)</td>
<td>0,2,5,7,10,15,18</td>
</tr>
<tr>
<td>Physical activity (increase in daily minutes)</td>
<td>0,10,20,30,40</td>
</tr>
<tr>
<td>Percentage risk reduction from respondent’s actual risk</td>
<td>40, 50, 60, 75, 85</td>
</tr>
</tbody>
</table>
Table 2: Socioeconomic statistics for the sample

<table>
<thead>
<tr>
<th>Income (annual income per household) (£ sterling)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 3120</td>
<td>10.02%</td>
</tr>
<tr>
<td>3121-4160</td>
<td>9.83%</td>
</tr>
<tr>
<td>4161-5200</td>
<td>10.40%</td>
</tr>
<tr>
<td>5201-6240</td>
<td>7.90%</td>
</tr>
<tr>
<td>6241-7280</td>
<td>7.90%</td>
</tr>
<tr>
<td>7281-10400</td>
<td>9.06%</td>
</tr>
<tr>
<td>10401-15600</td>
<td>13.10%</td>
</tr>
<tr>
<td>15601-20000</td>
<td>13.49%</td>
</tr>
<tr>
<td>20000-40000</td>
<td>42.77%</td>
</tr>
<tr>
<td>More than 40000</td>
<td>0.00%</td>
</tr>
<tr>
<td>Mean</td>
<td>29,051</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>40-50</td>
<td>47.06%</td>
</tr>
<tr>
<td>50-55</td>
<td>21.91%</td>
</tr>
<tr>
<td>More than 55</td>
<td>31.03%</td>
</tr>
<tr>
<td>Mean</td>
<td>50.73</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BMI</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Underweight BMI &lt; 18.5</td>
<td>4.67%</td>
</tr>
<tr>
<td>Normal BMI 18.5-25</td>
<td>37.93%</td>
</tr>
<tr>
<td>Overweight BMI 25-30</td>
<td>31.03%</td>
</tr>
<tr>
<td>Obese BMI &gt; 30</td>
<td>26.37%</td>
</tr>
<tr>
<td>Mean</td>
<td>26.85</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sex</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>43%</td>
</tr>
<tr>
<td>Female</td>
<td>57%</td>
</tr>
</tbody>
</table>

# of respondents: 493
Table 3: Model estimates for RU-MNL and RR-MNL (4,930 observations)

<table>
<thead>
<tr>
<th>Attribute</th>
<th>RUM</th>
<th></th>
<th>RRM</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost (increase in weekly expenditures)</td>
<td>-0.0985</td>
<td>-15.48</td>
<td>-0.0616</td>
<td>-17.66</td>
</tr>
<tr>
<td>Increase in Physical Activity (minutes per week)</td>
<td>0.00134</td>
<td>9.34</td>
<td>0.000816</td>
<td>9.91</td>
</tr>
<tr>
<td>Fat reduction (grams per week)</td>
<td>0.0027</td>
<td>5.63</td>
<td>0.0017</td>
<td>5.32</td>
</tr>
<tr>
<td>Risk of fatal heart attack in next 10 years</td>
<td>-0.0783</td>
<td>-5.34</td>
<td>-0.0537</td>
<td>-5.27</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-5,280.37</td>
<td></td>
<td>-5,275.37</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>4,930</td>
<td></td>
<td>4,930</td>
<td></td>
</tr>
</tbody>
</table>
### Table 4: WTP from the RU-MNL estimates

<table>
<thead>
<tr>
<th>Attribute</th>
<th>WTP</th>
<th>Std. err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase in Physical Activity (minutes per week)</td>
<td>0.0136</td>
<td>0.00094</td>
</tr>
<tr>
<td>Fat reduction (grams per week)</td>
<td>0.0274</td>
<td>0.00483</td>
</tr>
<tr>
<td>Risk of fatal heart attack in next 10 years</td>
<td>-0.795</td>
<td>0.148</td>
</tr>
</tbody>
</table>
Table 5: Logit model to explain determinants of RR-MNL outperforming RU-MNL for each person (493 observations).†

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th></th>
<th>t-stat</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>9.402</td>
<td>6.525</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male (dummy variable)</td>
<td>-0.092</td>
<td>1.473</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight (dummy variable)</td>
<td>-0.197</td>
<td>1.341</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overweight (dummy variable)</td>
<td>-0.319</td>
<td>4.323</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obese (dummy variable)</td>
<td>-0.264</td>
<td>3.324</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good and very good health (dummy variable)</td>
<td>0.495</td>
<td>11.659</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High education (dummy variable)</td>
<td>0.038</td>
<td>3.475</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoker (dummy variable)</td>
<td>-0.315</td>
<td>4.208</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.426</td>
<td>7.542</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age^2</td>
<td>0.005</td>
<td>8.246</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

†The dependent variable is equal to 1 if RR-MNL outperforms RU-MNL and 0 otherwise. The reference dummy variable for the BMI dummy variables is ‘Normal weight’.
Figure 1: Example of a choice card