Utility maximisation vs. regret minimization in health choice behaviour: evidence from four datasets

Abstract

Choice models in health are almost exclusively based on the neoclassical economic paradigm of utility maximization. Recently developed choice models have captured and shown empirical support for regret minimization as an alternative decision rule. In health economics, recent applications of RRM models indicate that individuals making health-based choices may exhibit regret minimization-type behaviour. In this paper, we build on this research using a more flexible model that allows for heterogeneous decision rules, separately from preference heterogeneity, and comparing it to models that assume single decision rules. We use four datasets from diverse settings in which individuals make health choices: tobacco markets, genomic testing, and HIV prevention. We found that, if a one-size-fits-all rule is applied, then utility maximization was preferable to regret minimization for these datasets. However, we also find that individuals apply varying decision rules in similar proportions in these health settings, suggesting that for heterogeneous decision rules were needed to capture these behaviours in these settings.

Key words: utility maximisation; regret minimization; decision rule; health choices

JEL codes: C35; I12; I18

Conflicts of interest: none to declare

Introduction

Behavioural evidence has emerged as an important form of evidence which can inform health policy. Data from discrete choice experiments (DCEs) and non-experimental data, analysed using choice models, have proven popular for understanding many health behaviours, from purchasing consumer goods, to clinical decision making and lifestyle choices (de Bekker-Grob et al., 2012; Clark et al., 2014; Soekhai et al., 2019). Choice models provide rich information on preferences related to an array of health decisions, including important heterogeneity in those preferences. They allow for estimating metrics such as welfare measures and willingness-to-pay (if cost is included as an attribute), and can also provide behavioural forecasts of how changes in attributes might impact product market shares (Regier et al., 2020). Given the value of these metrics to policymakers[1] it is critical that these models yield reliable evidence.

Analysing choice data requires assumptions about behavioural decision rules, on which individuals evaluate the available options and reach their decision(s). In health, choice data are almost exclusively treated in a manner consistent with neoclassical utility-maximization using random utility maximization (RUM; see e.g. McFadden, 1974) models. The RUM has served as the workhorse for health-based choice models for several decades. It has proven useful for understanding a range of health behaviours and has produced credible results that accord with underlying theories and match patterns of behaviour observed in reality, e.g. recovering tobacco behaviours that accord with those observed with large-scale, secondary data sources (Marti et al., 2019).

More recently, however, some economists have questioned the appropriateness of neoclassical assumptions such as utility-maximization. Empirical evidence on smoking, for example, suggests that individuals may exhibit alternative behaviours to utility maximization (Cawley and Ruhm, 2011). Several authors have suggested that the emotion of regret may play a role across a range of health behaviours (Colenda et al., 1995; Sorum et al., 2004; Frank, 2007; Brewer et al., 2016; Groopman and Hartzband, 2017). If it is the case that regret drives consumers' decisions, in place of or alongside utility, then evidence derived from choice models may be enhanced by considering regret-based models.

Regret minimization involves the pairwise comparisons of the available options. Regret arises if an alternative has more of a desirable attribute (or less of an undesirable attribute) than another. Individuals then choose options in a way to minimise this regret. The theory was first acknowledged in the economic literature by Loomes and Sugden (1983) and subsequently developed into analytically tractable choice models (Hey and Orme, 1994; Chorus, 2012); the latter termed random regret minimization (RRM) models. In the context of choice experiments, the RRM embodies the idea that individuals are willing to accept poor performance on some attribute(s) to achieve better performance on other attribute(s). RRM models are highly reference-dependent given that they are formulated based on attribute level comparisons across alternatives. As such, they are well-paced to capture choice set heuristics such as compromise effects (Chorus and Bierlaire, 2013) and decoy effects (Guevara and Fukushi, 2016).

These models have had recent applications in health (see literature review below). The emerging evidence suggests that in some cases, RRM performs better than RUM. Other studies suggest that hybrid RUM-RRM models, treating some attributes as RUM and others as RRM, can improve model fit (de Bekker-Grob and Chorus, 2013). Following from this research, we examine whether the reason for these mixed results is due to individuals adopting different

[1] E.g. the UK's national institute for health and care excellence (NICE) recently recognized the value of these studies (https://www.nice.org.uk/news/blog/patient-preferences-studies-how-nice-could-make-use-of-them)

decision rules when making choices. Understanding this is critical because substandard behavioural assumptions could result in suboptimal evidence to inform health policy.

Our main contribution is to compare models that allow for differing decision rules in health settings. These are RUM and RRM[2]. As a starting point, we review existing RRM studies in health settings. We then conduct a modelling exercise using a range of choice models, including developing a flexible choice model that allows individuals to adopt either RUM-type behaviour, RRM-type behavior, or some mix of these decision rules, i.e. heterogeneous decision rules. This builds on health literature that investigates alternative decision rules to RUM (Arana et al., 2008; Boeri et al., 2013; Erdem et al., 2014; Biondi et al., 2019). The model allows for multiple decision rules by specifying latent classes, one each for each decision rule. This allows us to consider whether all individuals adopt a particular decision rule, or different individuals adopt different decision rules. Thus, we relax the one-size-fits-all assumption imposed by modelling a single decision rule (though this model would detect this behaviour if it was present). We test this model against its simpler counterparts to understand if, along with its behavioural appeal, it can better explain choice data.

We also use the models to predict choice shares and compare these to real-world market shares (where data are available to do this). From these, we make inferences about the implied behaviours of individuals.

Our model follows a growing number of studies in the choice modelling literature that recognize that alternative decision rules to RUM are at play in individuals' choice behaviour (Hess et al., 2012a; Leong and Hensher, 2012; Adamowitz and Swait, 2012; Gonzalez-Valdez and Raveau, 2018; Balbontin et al., 2019). Using latent classes to model different decision rules has been set out in the economics literature and more recently adopted in choice modelling (c.f. Stahl and Wilson, 1995; Hess et al., 2012a). The latter work raised important questions about possible confounding between decision rule heterogeneity and preference heterogeneity; failure to account for these differences can result in misleading inferences on the decision rules that individuals use (Hess et al., 2012b). Subsequently, methods have been developed specifying preference heterogeneity within latent classes by using (further) latent classes (Hess and Stathopoulos, 2013), random heterogeneity (Boeri et al., 2013), and latterly model averaging (Hancock and Hess, 2020), to separate out decision rule heterogeneity from preference heterogeneity. In this paper, we opt for a latent class within latent class approach; our specification is discussed in later sections.

We apply these models to four health datasets of both stated preference (SP) and (separately) revealed preference (RP) data, reflecting a range of geographical and socioeconomic diversity. They comprise both online and face-to-face methods of collecting data; and sampling from either the general population or specific populations (e.g. smokers). These data sets were chosen to reflect the varied circumstances in which health-based choices are studied and made. The first is an SP dataset of US smokers' choices of tobacco products, for an application of addiction-related purchasing behaviours. The second is a corresponding RP dataset of these smokers' actual smoking behaviours. The third is an SP dataset of Canadians making choices about genomic testing to inform clinical decision making and cancer screening. In the fourth, we explore if patterns hold in middle-income settings, a SP dataset of choices of HIV prevention product in South Africa.

Review of Regret Minimization choice models in health

[2] In preliminary modelling, we also estimated the muRRM model (van Cranenburg et al., 2015), which is a generalisation of the preceding two models, allowing for either of the two behaviours, something in between, or what is termed a "pure" regret minimization behaviour. We did not find evidence that this model was able to better explain the data than the RUM or RRM and so discarded this model.

A review of studies in health that use RRM models was conducted (March 2020). We included any study of health that applied RRM models to choice data. Both economics and medical databases were searched (econlit, medline, pubmed). We also used reference mining of studies that met our selection criteria. In total, our search yielded five studies that use RRM models in health (Boeri et al., 2013; de Bekker-Grob et al., 2013; Paul et al., 2018; Ryan et al., 2018; Biondi et al., 2019).

These studies reported RUM-RRM comparisons across different health settings, without finding consistent conclusions on the favoured decision rule. Boeri et al. (2013) use RUM and RRM models using SP data on diet and exercise behaviours. They find that the RRM model improves fit, with a higher log-likelihood than the RUM model, but do not find substantial differences between the two. They also note that deriving welfare measures (e.g. WTP) from RRM models is problematic, but recent work has provided some solutions (Dekker and Chorus, 2018). de Bekker-Grob and Chorus (2013) use RUM, RRM, and hybrid models (some attributes are RUM; others are RRM) on two datasets in clinical decision making contexts. Their results indicate that on different data, different models perform better or worse. They also show that differences between models are fairly modest in both cases. Paul et al. (2018) use revealed preference data on hospital choices to test RUM vs RRM. They find that for multinomial logit (MNL) models, RUM outperforms RRM. They also use heteroskedastic MNL models in which case the RRM performs better than RUM. Ryan et al. (2018) relate RUM and RRM to visual attention using eye-tracking technology in their experiment of lifestyle choices amongst students. Their results favour RRM, but statistically significant differences were not demonstrated. Finally, Biondi et al. (2019) compare RUM and RRM models on consumer food choices. Their sample level results show similar fit and predictive power between the two approaches (though they use hit rates for prediction, which ignore the error term of the model (Hess and Palma, 2019)). However, subgroup analyses based on factor analyses of personality traits, indicate that the different approaches each perform better on different subgroups.

Overall, RRM shows promise as an alternative to RUM in explaining health choices. Although in most cases both approaches performed similarly, different approaches performed better and worse according to personality traits. These findings are consistent with studies in the non-health literature (cf. Chorus et al., 2014). One explanation is simply that individuals are using decision heuristics that lie between RUM and RRM, a finding partly supported by the hybrid findings of de Bekker Grob and Chorus (2013). If this is the case, models that generalise the two behaviours could provide an improved/enhanced account of choice behaviours. Another explanation is that different individuals, or groups of individuals, are applying different decision rules when making choices. This idea is supported by the findings in subgroup analyses by Biondi et al. (2019). If this is the case, then more flexible models may be required to capture this group-wise decision making behaviour.

These review findings motivated our empirical study, where we explored in greater depth whether different individuals in health settings are applying different decision rules. We use more sophisticated choice models which could more accurately explain health choices and provide greater insight into decision making processes for policymaking.

Aside from the models, we also consider other aspects of the evidence available. These studies are conducted in individual settings (though some use two rather than one dataset). Here we use more datasets for comparisons, and also across multiple settings in multiple countries. We also have data that is both SP and RP; and that has been collected using different approaches,

i.e. online and face-to-face. These features are not present collectively in other studies. Considering these together allows us greater insights into individuals' decision making.

Methods

Sampling and datasets

1.1 Smokers' stated preference tobacco product choice data, USA

Data were taken from an online DCE on 2,031 US adult smokers conducted in 2017 (1531 current smokers; and 500 self-reported recent quitters) (Buckell et al., 2019). Sampling was based on quotas derived from the Behavioral Risk Factor Surveillance System (BRFSS) data in 2013/14, comprising gender, age, education and region to make the sample representative. The sample size is well in excess of minimum sample size calculations (de Bekker-Grob et al., 2015). A series of exercises were conducted to promote the quality of the data (e.g. attention checks in the survey, minimum time threshold, and removing duplicates individuals). Descriptive statistics are presented in table 7.

The DCE was based on a review of the literature and a pilot study. The literature review comprised prior DCEs in tobacco (e.g. a systematic review by Regmi et al., 2018); market data on tobacco product prices (Cuomo et al., 2015); and scientific literature on the harms of tobacco products (Jha et al., 2013; McNeill et al., 2015). In the study individuals chose between cigarettes, e-cigarettes and opt-outs. Respondents were presented with 2 of each product and made two choices in each choice task. Attributes (levels) were price ($4.99, $7.99, $10.99, $13.99), flavours (tobacco, menthol, fruit, sweet), level of nicotine (none, low, medium, high) and health harm expressed in life years lost to the average smoker (2 years, 5 years, 10 years, unknown); attributes and levels are shown in Appendix. Some levels were omitted to make choices realistic (e.g. fruit/sweet cigarettes are not on the market in the US). This design is based on a review of the literature and a pilot study.

A Bayesian D-optimal design was used (Hensher et al., 2015). Priors were obtained from a MNL model in analysis of pilot study data on 87 respondents. 3 blocks of 12 had individuals randomized to them. Each individual answered 12 choice sets, balancing concerns of learning and respondent fatigue (Hess et al., 2012). A practice choice scenario was given to all respondents to ensure that they understood how the choice scenarios worked.

1.2 Smokers' revealed preference tobacco product choice data, USA

Smoking behaviours and information on products' attributes were collected from the 2,031 sampled individuals. Individuals were categorized as either a smoker (uses cigarettes only, n=1038), a dual user (uses cigarettes and e-cigarettes, n=619), a vaper (exclusive use of e-cigarettes, n=148) or a recent quitter (has recently stopped smoking, n=226). Data was collected on products' prices and flavours. These data are used to build choice models, extending previous work in Buckell and Hess (2019).

2. Genomic testing choice data, Canada

A DCE evaluated individuals' preferences for the following 9 attributes: number (of 100) of people identified with genetic marker, medical expert agreement on getting tested, information on your risk of getting other diseases, cost of testing, number of people (of 100) with genetic marker who respond to treatment, number of family members (of 100) with genetic marker who get the disease, quality of life change from individualized treatment, length of life change from individualized treatment, length of life change for your family from screening. Attributes and levels were developed from a scoping review (Regier et al., 2018), followed by three focus

groups in metropolitan Vancouver. Three focus groups with 13 members of the public (3 to 7 per group) guided attribute identification. Inclusion criteria were: healthcare experience within the past 6 months; English-speaking adults; >=18 years of age. A semi-structured interview guide was created from a published literature review of discrete choice experiments in precision medicine (Regier et al., 2018), and a facilitator led discussion. Transcripts from each focus group were produced and analyzed using thematic analysis (Strauss, 1987). The set of identified themes were collapsed into a core list, with the final attributes determined by the research team. An initial choice task was constructed based on team input. The identified attributes were refined and respondent burden assessed through pre-test think aloud interviews with 14 members of the public.

In the resultant questionnaire, a Bayesian, D-optimal design (using Ngene software) with enforced overlap between three attributes was constructed using informative priors generated from a MNL model in a pilot study (n=100). The final experimental design resulted in 144 choice questions. In the final questionnaire, each respondent was randomly assigned to one of 9 blocks that included 16 choice questions. The final questionnaire was administered to 1140 individuals utilizing the largest probability-based online panel that is representative of the adult Canadian population. Descriptive statistics are presented in table 7.

3. HIV prevention product choices among a general population sample, South Africa

In 2015, 367 HIV negative women (199 aged 16-17 and 168 aged 18-49) were interviewed in a randomised face-to-face household survey conducted in a peri-urban township on the outskirts of Johannesburg, South Africa (Quaife et al., 2017). Descriptive statistics are presented in table 7.

The DCE was developed through an analysis of a previous DCE and focus groups discussions carried out in previous research (Terris-Prestholt et al., 2014), specifically identifying important characteristics of prevention products and exploring optimal ways to present these in a clear and relatable manner to participants. This was supplemented by a scoping literature review to identify new products and additional attributes which could be important to respondents, which was added to and refined through piloting. We opted to show three alternatives of new products in each task using an unlabelled design where each alternative represents a generic product within which all characteristics can change as prescribed by the statistical design. In this experiment, respondents chose between three unlabelled alternatives of new HIV prevention products and an opt-out. Products were described by product type (oral pill, injectable, reusable diaphragm, vaginal gel, and vaginal ring), HIV prevention efficacy (55%, 75%, 95%), contraceptive ability (yes, no), STI protection (yes, no), frequency of use (coitally, daily, weekly, monthly, every three months, every six months, annually), and side-effects (nausea, stomach cramps, dizziness, none). A Bayesian D-optimal design was generated using priors estimated on a MNL model in a pilot using a sequential orthogonal design.

Choice Models

1. Utility Maximization (RUM)

Based on McFadden (1974), the RUM has been used overwhelmingly for choice models in health. In this formulation, the individual reconciles their product/attribute preferences for each of the available options and chooses that which maximizes their utility. Respondents' utility is a linearly-additive function of attribute/product preferences and the product-attribute

combinations available. For each alternative, the individual is assumed to choose the option that delivers the highest utility.

$$U_{ni} = V_i + \varepsilon_{ni} = \sum_m \beta_m \cdot x_{im} + \varepsilon_{ni} \quad (1)$$

Where U_{ni} is the utility for decision maker n for product i, comprising deterministic and random utility. V_i is the deterministic component of utility; and ε_{ni} is the random component of utility. x_{im} is the mth attribute/product; β_m are preference parameters to be estimated. Estimation is operationalized by assuming a type-I extreme value error distribution on ε_{ni} and estimating choice probabilities for each product with a multinomial logit (MNL) model.

$$P_{rum,ni} = \prod_{t=1}^{T} \sum_{i=1}^{I} (c_{nt} = i) \frac{\exp(V_i)}{\sum_{j=1\dots J} \exp(V_j)} \quad (2)$$

Where $P_{rum,ni}$ is the RUM probability of respondent n's sequence of choices over T choice sets. I is the set of alternatives and $c_{nt} = i$ takes the value if 1 if alternative i is chosen in choice set t; 0 otherwise.

2. Regret minimization (RRM)

Discrete choice models that capture regret minimization were developed by Chorus (2012). In this setting, a regret function is defined, in which the pairwise comparisons of attributes/products between the available options are specified. Regret for a given attribute is generated if an alternative has more of a desirable attribute (or less of an undesirable attribute), than the option at hand. If the current option has more of a desirable attribute (or less of an undesirable attribute), no regret is generated. For each available option, overall regret is determined as the sum of pairwise regrets, and the individual chooses the alternative which minimizes this regret.

$$RR_{ni} = R_i + \varepsilon_{ni} = \sum_{j \neq i} \sum_m ln\left(1 + \exp\left[\beta_m \cdot \left(x_{jm} - x_{im}\right)\right]\right) + \varepsilon_{ni} \quad (3)$$

Where RR_{ni} is the regret function for individual n for product i. R_i is deterministic regret and ε_{ni} is the random component of regret. x_{im} is the value of the mth attribute of the considered option; and x_{jm} is the value of the mth attribute of the option against which the considered option is being compared. β_m is the preference parameter for the mth attribute to be estimated. The operator $ln(1 + \exp[\dots])$ imposes that regret is generated when an attribute of an alternative is more desirable than the option at hand [3] (Chorus, 2012). As per the RUM, choice probabilities can be estimated using a MNL, but now the negative of the regret function is substituted into the equation. The sign of the equation relative to RUM is reversed because regret is a negative emotion (whereas utility is a positive emotion).

$$P_{rrm,ni} = \prod_{t=1}^{T} \sum_{i=1}^{I} (c_{nt} = i) \frac{\exp(-R_i)}{\sum_{j=1\dots J} \exp(-R_j)} \quad (4)$$

[3] As pointed out by Hess and Chorus (2015), if the operator is modified to $ln(0 + \exp[\dots])$, the RUM behaviour is obtained. As a specification check that RRM is its RUM counterpart, we make this modification to the RRM model and see that the log-likelihood is identical to that obtained from the RUM. If any parameters are treated as RUM in the RRM model, then these should be identical to RUM estimates.

Where $P_{rrm,ni}$ is the RRM probability of respondent *n*'s sequence of choices over *T* choice sets. *I* is the set of alternatives and $c_{nt} = i$ takes the value if 1 if alternative *i* is chosen in choice set *t*; 0 otherwise.

3. A latent class, decision rule heterogeneous choice model (LCDRH)

A latent class model is specified that accommodates the preceding decision rules (Hess and Stathopoulos, 2013; Hess and Chorus, 2015; Gonzalez-Valdez and Raveau, 2018; Nielsen and Jacobsen, 2020). Each of the latent classes corresponds to a particular decision rule. Each individual belongs to a class, and so applies a given decision rule, up to a (estimated) probability. These class membership probabilities indicate which proportions of individuals in the sample applied the possible decision rules when making their choices. The model thus allows for any of the decision rules to be applied by all individuals, or some mixture of decision rules. The model combines the respective MNL formulations from each of the above models and adds class membership probability terms.

$$P_{lcdrh,ni} = \sum_{dr=1}^{D} \pi_{dr} \, . \, P_{dr,ni} \quad (5)$$

Where

$$\sum_{dr=1}^{D} \pi_{dr} = 1 \text{ and } 0 \leq \pi_{dr} \leq 1 \; \forall dr \quad (6)$$

To impose the restrictions in (9), a multinomial logit formulation is used,

$$\pi_{dr} = \frac{\exp(\theta)}{\sum_{dr=1 \ldots D} \exp(\theta)} \quad (7)$$

Here, D is the set of possible decision rules that individuals apply (RUM or RRM). π_{dr} is the class membership probability for class dr, with θ to be estimated. $\exp(\theta)$ can be further parameterized to allow for individual-specific, deterministic class membership probability; it is also possible to treat class membership as a random variable (Gonzalez-Valdez and Raveau, 2018). $P_{dr,ni}$ is the choice probability of respondent *n*'s sequence of choices over *T* choice sets in class dr (as per eqns. (2) and (4)). .

Next, we estimate within-class latent classes, to accommodate unobserved preference heterogeneity as distinct from decision rule heterogeneity (Hess and Stathopoulos, 2013; Nielsen and Jacobsen, 2020). Then, the choice probability becomes,

$$P_{lcdrh,ni} = \sum_{dr=1}^{D} \pi_{dr} \, . \sum_{k=1}^{K_{dr}} \omega_{dr,k} \, . \, P_{dr,k,ni} \quad (8)$$

Where K are additional classes that allow β_m to vary within decision rules. $\omega_{dr,k}$ are within-decision-rule (i.e. conditional) weights for the K classes, where $\omega_{dr,k} = 1 \; \forall dr$. The latter restriction can again be imposed by using logit formulation. K can be allowed to vary across classes.

With 2 decision rules and 2 classes within each rule, we estimate 4-class, decision rule heterogeneous models. The choice probability is as below.

$$P_{lcdrh,ni} = \pi_{rum}\left(\omega_{rum1} . P_{rum1,ni} + \omega_{rum2} . P_{rum2,ni}\right) + \pi_{rrm,n}\left(\omega_{rrm1} . P_{rrm1,ni} + \omega_{rrm2} . P_{rrm2,ni}\right) \quad (9)$$

Where decision rule class membership is estimated as,

$$\pi_{rrm} = \frac{e^{\delta_{rrm}}}{1+e^{\delta_{rrm}}} \quad (10)$$

$\delta_{rrm}$is a parameter to be estimated; and $\pi_{rum} = 1 - \pi_{rrm}$.

And within-rule class membership is estimated as,

$$\omega_{rrm2} = \frac{e^{\delta_{rrm2}}}{1+e^{\delta_{rrm2}}} \quad (11)$$

$\delta_{rrm2}$is a parameter to be estimated; and $\omega_{rrm1} = 1 - \pi_{rrm2}$. (Within-rule class memberships for RUM are derived analogously.)

Comparisons of model fit

From estimation, the log-likelihoods, AIC, and BIC are reported. In addition, the Vuong test of non-nested models is applied (Vuong, 1989; Hensher, 2015); simulation suggests this is preferred for experimental data (Strazzera et al., 2013). LCDRH models were tested against a model that supresses decision rules – a standard 4-class latent class model - as the recommended comparator against which to test the specification (Hancock and Hess, 2020). This helps to establish the presence of decision rule heterogeneity over and above preference heterogeneity if the 4-class decision rule heterogeneous choice model can better explain the data than the 4-class preference heterogeneity model.

Forecasts

We use sample enumeration for forecasting (Train, 2009), where the probabilities across all possible alternatives are added across all individuals in the sample. This involves first computing the predicted choice probabilities for each alternative for each observation from the choice models (as in equations 2 and 4). For each product, we then take the mean of the predicted probabilities over all observations to yield predicted choices shares of products in the sample. These are the reference against which forecasts are compared. To make the forecasts, attributes in the utility functions are adjusted (e.g. altering nicotine levels in the tobacco SP dataset). With adjusted attribute levels, the choice probabilities are recalculated as are the choice shares. These are the forecasts and compared with the base choice shares. We forecast only for the tobacco datasets as these are the only datasets in our empirical analysis with real-world market shares for comparing predictions.

Limitations

We note several drawbacks of our approach. First, while the data that we used span a broad set of health settings and populations, we acknowledge that four is a reasonably small number of datasets, especially as the RP and SP tobacco datasets are drawn from the same sample. This is limits the conclusions we are able to draw. In addition, these were the only data for which market share data were available against which to compare forecasts. Therefore, for other datasets, we could see potential differences in predicted choice shares, but we were unable to ascertain which of them was better able to predict reality. Second, while we were able to examine the performance of our models in terms of fit, we were limited in being able to assess their performance in terms of external validity for a single setting. Third, while we have controlled for within-decision-rule heterogeneity, we recognise that 2 latent classes may be

insufficient to absorb all of this heterogeneity. We also describe in our sensitivity analyses our attempts to model this heterogeneity differently. Fourth, even with pared down versions of model (3), where non-significant parameters are removed, we note the considerable additional complexity relative to its simpler constituents. Indeed, we were unable to estimate models on two of four datasets, even using procedures to aid estimation by searching for improved initial starting values. This is a practical limitation of this approach. Lastly, we note that three of four of our datasets are designed based on RUM. While we are not aware of evidence that links the design of an experiment to induced behaviour of that type, we cannot reject the possibility that the RUM design could be linked to RUM behaviours.

Results

Table 1 presents the diagnostic information for the sets of choice models applied to the four datasets. We first consider the comparison between RUM and RRM; that is, between models (1)-(2). In all datasets, we see that the fit of these models is similar. In all four cases, the log-likelihood of the RUM is higher than that of the RRM. Vuong tests prefer RUM over RRM for tobacco SP, genomic testing and HIV prevention. For tobacco RP, the Vuong test is inconclusive.

[Insert table 1 here]

Turning to the LCDRH models, (3), we first report that these models failed to estimate on the genomic testing data and the tobacco RP data. This is perhaps unsurprising for the tobacco RP data due to the fact that we only observe each individual once. For genomic testing SP, each additional class requires estimating 18 additional parameters, which we suspect is the reason for issues in estimation. We next consider the fit of model (3) across the tobacco SP and HIV prevention datasets. Compared to their constituent comparators, the fit of the data is considerably improved in the LCDRH models; Vuong tests indicate these models are preferred. We tested the LCDRH against 4-latent class preference heterogeneity models (one for RUM and one for RRM) that supress decision rules; Vuong tests in all cases supported the decision rule heterogeneity.

We next consider the class shares, where we again see some divergence in which decision rule applies across datasets. For tobacco SP, we see that the RRM decision rule is dominant, with a 52% class share compared to 48% of RUM. We test this formally by observing if *classprob_rrm*=0[4]. We do not reject the null in this case and so we do not detect an unequal class share. This suggests an equal of RUM and RRM in this sample. In contrast, it is RUM, at 56%, that is, dominant for the HIV prevention data, leaving 44% of RRM. Again, the parameter on the class membership parameter is not different from zero, suggesting equal balance between RUM and RRM. We summarize these findings from LCDRH models across datasets graphically in Fig.1. For HIV prevention, the class shares match the patterns of log-likelihoods of models (1)-(2) in that RUM was preferred to RRM. For tobacco SP, however, this is not the case. In the simple MNL models, RUM has a better fit than RRM, though in the LCDRH model, it appears that RRM is more prevalent than RUM amongst respondents.

[Insert Fig.1 here]

[4] This derives from the specification of the model: when the parameter is zero, the class shares are 50:50, since $classprob_{rrm} = \exp(0)/(1 + \exp(0)) = 0.5$

Tables 2 and 3 show the forecasts from the tobacco SP and RP data. In both datasets, the predicted market shares are almost identical. This is unsurprising as the fits of the models, too, are very close.

Across datasets, we observe consistent preference parameter estimates for the choice models, presented in Tables 2-5 with one for each dataset. For models (1)-(2) across all of the datasets, very similar preference structures are observed. That is, both the directions and order of preferences for attribute levels are the same (though for a variety of reasons – scale, model specifications – it is inappropriate to compare the magnitudes of these coefficients). For model (3), within-decision-rule heterogeneity in preferences is seen by comparing parameters across classes with the same decision rules. For tobacco SP, preferences for flavoured tobacco products vary across classes (the directions of the coefficients are reversed), with both RUM and RRM sub-classes. Likewise, for HIV prevention, preferences for vaginal rings vary in the RUM class (preferences reverse) and the RRM class (significant vs. non-significant parameter). These results suggest that the model is indeed capturing preference heterogeneity within classes.

[Insert tables 2-5 here]

Decision rules and study design features

In table 6, we analyse how study design features are related to the implied decision rules from our estimates. The datasets related to each setting are listed. 5 features of the study designs are considered. The setting, whether clinical or health behaviours, is examined. Whether studies were fielded in more or less developed countries was considered. Features of the data were considered, namely SP vs RP and online vs face-to-face data collection. Finally, the sampling – general population vs specific population – was examined. Across all of the study design features, the dominant decision rule was RUM. However, we also point out some key caveats in some cases, namely that in the LCDRH models, some RRM behaviour was observed; and that in other cases, more complex models were inestimable. Of course, another key caveat is the limited number of datasets available here to perform this type of analysis. This notwithstanding, the general results seem to be consistent with the analyses presented above.

Sensitivity analyses

We estimated mixed logit versions of all of our models. For models (1)-(2) we found that, for each of the datasets, the overall patterns were consistent with those of the MNL models. The RUM models were preferred in all cases, though the gain in fit of the RUM was more pronounced. The patterns of preferences for attribute levels were in the same directions and in the same order across models. So whilst these models allowed for additional flexibility in preferences, they did not materially impact on the main findings. (Though in passing, we note that we are not primarily interested in preferences.) There were also challenges with estimation. These models took a long time to estimate, even with the modest number of draws used in estimation. Results from Czajkowski et al. (2019) indicate that more draws may be needed to achieve stability in estimation. For these reasons, we are treating the mixed logit models as sensitivity analyses, rather than as for our main results. We present the mixed models (1)-(2) in the Tables A1-A4 in the appendix.

For the mixed LCDRH models, estimation verged on being prohibitively slow, and using fewer draws still (as few as 100). We do not report the results of these models because 100 draws seems insufficient for mixed logit models with many (in this case 33) dimensions (Czajkowski et al., 2019).

Whilst it is common to specify deterministic heterogeneity in the class membership probabilities, we chose not to for two reasons. First, that we wish to keep the exposition and

comparison of models as clear as possible. Second, that due to the different data collections, we would not have been able to specify the same set of covariates across datasets, meaning that comparisons become more complicated. To address the concern that class shares could be impacted by the addition of deterministic heterogeneity, we estimated this model for tobacco SP data. We present this model in the appendix Table A5. The class shares were very similar to those presented in the main analyses; though the share is slightly in favour of RUM in this case.

Summary and Conclusions

In this paper, we have sought to provide empirical evidence on the decision rules that individuals adopt when making choices in health. Specifically, we made comparisons between utility maximisation and regret minimisation.

We found that, in models that assume a single decision rule, utility maximisation was the dominant decision rule applied by individuals in all of our datasets. This is different from literature reviews in other fields that find support for RRM or hybrid models of RUM and RRM (Chorus et al., 2014). This finding is also different to previous RRM applications in health, that have, broadly speaking, not found substantial differences between the two approaches (c.f. literature review section). We caveat this finding by noting that RUM and RRM models had similar log-likelihoods in both of the tobacco datasets. Given that we have a wider array of data sources to draw on than previously, this general finding may be supportive in light of the RUM assumptions imposed routinely by researchers in health.

However, we also found evidence of decision rule heterogeneity in cases where we were able to estimate LCDRH models. For tobacco, when the fit of RUM and RRM models were similar, class share for RRM was slightly higher than for RUM. This means that the finding compared to the RUM models (where RUM would be slightly preferred) is overturned. For HIV prevention, where the fit of RUM and RRM diverged, the latent class model gave more weight to the dominant class. These findings, particularly when the fits of one-rule models are similar, suggest that researchers in health should think carefully about specifying decision rules in choice models. That is, where the RUM-RRM balance may be similar, as is the case here, more sophisticated models seem better able to capture behaviours.

Another key finding was that accounting for unobserved preference heterogeneity at the same time as decision rule heterogeneity made a substantive difference to the results (which is in keeping with previous results, e.g. Hess et al., (2012b)). Using latent classes potentially leads to the confounding of decision rule and preference heterogeneity. The LCDRH models outperformed comparator models that modelled only preference heterogeneity, thought this may not be the case in other datasets. This testing is critical in attempting to avoid confounding preference and decision rule heterogeneity (though if this can ever be fully achieved is open to question).

A third important finding was the difficulty in estimating more complex models. For two of four datasets, 4-class LCDRH models were inestimable. Further, models that used mixing distributions, rather than latent classes, to capture unobserved heterogeneity were not a feasible option given the processing time needed to estimate models, even using a low number of draws; in cases where better computer power is available, models as per Boeri et al. (2013) may be feasible. For these reasons, modelling decision rule heterogeneity may be difficult in some

settings, particularly when sample sizes are limited or models have many parameters, as was the case here.

Where we were able to compare forecasts to their real-world market shares, we found that forecasts were almost identical. However, these were in cases where the fits of those models were extremely similar also. Finally, it is notable that our analyses of choices made around HIV prevention, from a DCE undertaken with a population of young adults in peri-urban South Africa, are not notably different to those of datasets from high-income settings. This consistency suggests that decision rules may not vary substantially across different populations, or by income setting, reinforcing the applicability of choice models to understand human behaviours in different contexts.

In conclusion, based on our findings, if a one-size-fits-all rule is applied, then utility maximization appears preferable to regret minimization. Ultimately, though, a given dataset may be more akin to RRM than RUM, and the specification of decision rules remains an empirical question and testing should be conducted to determine the presence of decision rule/preference heterogeneity. We also found that a substantial proportion of individuals in all of the considered samples exhibit regret minimization. This suggests that individuals do indeed apply varying decision rules in health settings, and more complex choice models may be required to capture these behaviours. Moreover, approaches to accounting for preference heterogeneity may confound preference and decision rule heterogeneity. We anticipate that these findings will be useful to support researchers and policymakers when trying to understand how individuals are making choices and when designing policies in health settings.

Acknowledgements

This work was supported by the NIHR Oxford Biomedical Research Centre. It was also supported by Genome British Columbia (grant no: GEN-001).

References

Adamowicz, W. L., & Swait, J. D. (2012). Are Food Choices Really Habitual? Integrating Habits, Variety-seeking, and Compensatory Choice in a Utility-maximizing Framework. *American Journal of Agricultural Economics, 95*(1), 17-41. doi:10.1093/ajae/aas078

Araña, J. E., León, C. J., & Hanemann, M. W. (2008). Emotions and decision rules in discrete choice experiments for valuing health care programmes for the elderly. *Journal of Health Economics, 27*(3), 753-769. doi:https://doi.org/10.1016/j.jhealeco.2007.10.003

Balbontin, C., Hensher, D. A., & Collins, A. T. (2019). How to better represent preferences in choice models: The contributions to preference heterogeneity attributable to the presence of process heterogeneity. *Transportation Research Part B: Methodological, 122*, 218-248. doi:https://doi.org/10.1016/j.trb.2019.02.007

Biondi, B., Van der Lans, I. A., Mazzocchi, M., Fischer, A. R. H., Van Trijp, H. C. M., & Camanzi, L. (2019). Modelling consumer choice through the random regret minimization model: An application in the food domain. *Food Quality and Preference, 73*, 97-109. doi:https://doi.org/10.1016/j.foodqual.2018.12.008

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. doi:https://doi.org/10.1016/j.jhealeco.2012.10.007

Brewer, N. T., DeFrank, J. T., & Gilkey, M. B. (2016). Anticipated regret and health behavior: A meta-analysis. *Health Psychology, 35*(11), 1264-1275. doi:10.1037/hea0000294

Buckell, J., & Hess, S. (2019). Stubbing out hypothetical bias: improving tobacco market predictions by combining stated and revealed preference data. *Journal of Health Economics, 65*, 93-102. doi:https://doi.org/10.1016/j.jhealeco.2019.03.011

Buckell, J., Marti, J., & Sindelar, J. L. (2019). Should flavours be banned in cigarettes and e-cigarettes? Evidence on adult smokers and recent quitters from a discrete choice experiment. *Tobacco Control, 28*(2), 168. doi:10.1136/tobaccocontrol-2017-054165

Cawley, J., & Ruhm, C. J. (2011). Chapter Three - The Economics of Risky Health. In M. V. Pauly, T. G. McGuire, & P. P. Barros (Eds.), *Handbook of Health Economics* (Vol. 2, pp. 95-199): Elsevier.

Clark, M. D., Determann, D., Petrou, S., Moro, D., & de Bekker-Grob, E. W. (2014). Discrete Choice Experiments in Health Economics: A Review of the Literature. *PharmacoEconomics, 32*(9), 883-902. doi:10.1007/s40273-014-0170-x

Colenda, C. C., Poses, R. M., Rapp, S. R., Leist, J. C., & Reboussin, D. M. (1995). Anticipated Regret Associated With Treatment Decisions for Agitated Dementia Patients. *The American Journal of Geriatric Psychiatry, 3*(1), 68-74. doi:https://doi.org/10.1097/00019442-199524310-00009

Chorus, C.G., 2012. Random Regret-based Discrete Choice Modeling: A Tutorial. Springer, Heidelberg, Germany.

Chorus, C. G. (2014). A Generalized Random Regret Minimization model. *Transportation Research Part B: Methodological, 68*, 224-238. doi:https://doi.org/10.1016/j.trb.2014.06.009

Chorus, C. G. and M. Bierlaire (2013). "An empirical comparison of travel choice models that capture preferences for compromise alternatives." Transportation 40(3): 549-562.

Chorus, C., van Cranenburgh, S., & Dekker, T. (2014). Random regret minimization for consumer choice modeling: Assessment of empirical evidence. *Journal of Business Research, 67*(11), 2428-2436. doi:https://doi.org/10.1016/j.jbusres.2014.02.010

Cuomo RE, Miner A, Mackey TK. (2015) Pricing and sales tax collection policies for e-cigarette starter kits and disposable products sold online. Drug Alcohol Rev 015;35:110–4.

Czajkowski, M., Zagórska, K., & Hanley, N. (2019). Social norm nudging and preferences for household recycling. *Resource and Energy Economics, 58*, 101110. doi:https://doi.org/10.1016/j.reseneeco.2019.07.004

de Bekker-Grob, E. W., Ryan, M., & Gerard, K. (2012). Discrete choice experiments in health economics: a review of the literature. *Health Economics, 21*(2), 145-172. doi:10.1002/hec.1697

de Bekker-Grob, E. W., & Chorus, C. G. (2013). Random Regret-Based Discrete-Choice Modelling: An Application to Healthcare. *PharmacoEconomics, 31*(7), 623-634. doi:10.1007/s40273-013-0059-0

de Bekker-Grob, E. W., Donkers, B., Jonker, M. F., & Stolk, E. A. (2015). Sample Size Requirements for Discrete-Choice Experiments in Healthcare: a Practical Guide. *The Patient - Patient-Centered Outcomes Research, 8*(5), 373-384. doi:10.1007/s40271-015-0118-z

Dekker, T., & Chorus, C. G. (2018). Consumer surplus for random regret minimisation models. *Journal of Environmental Economics and Policy, 7*(3), 269-286. doi:10.1080/21606544.2018.1424039

Erdem, S., Campbell, D., & Thompson, C. (2014). Elimination and selection by aspects in health choice experiments: Prioritising health service innovations. *Journal of Health Economics, 38*, 10-22. doi:https://doi.org/10.1016/j.jhealeco.2014.06.012

Frank, R. G. (2007). Behavioral Economics and Health Economics. In P. Diamond & H. Vartiainen (Eds.), *Behavioral Economics and Its Applications* (pp. 195-234): Princeton University Press.

Gonzalez-Valdes, F., & Raveau, S. (2018). Identifying the presence of heterogeneous discrete choice heuristics at an individual level. *Journal of Choice Modelling, 28*, 28-40. doi:https://doi.org/10.1016/j.jocm.2018.05.001

Groopman, J., & Hartzband, P. (2017). The Power of Regret. *New England Journal of Medicine, 377*(16), 1507-1509. doi:10.1056/NEJMp1709917

Guevara, C. A. and M. Fukushi (2016). "Modeling the decoy effect with context-RUM Models: Diagrammatic analysis and empirical evidence from route choice SP and mode choice RP case studies." Transportation Research Part B: Methodological 93: 318-337.

Hancock, T. & Hess, S., What is really uncovered by mixing different model structures: contrasts between latent class and model averaging. Working paper, University of Leeds. http://www.stephanehess.me.uk/papers/working_papers/Hancock_Hess_2020_model_averaging.pdf

Hensher, D. A., Rose, J. M., & Greene, W. (2015). *Applied Choice Analysis*. Cambridge: Cambridge University Press.

Hess, S., Hensher, D. A., & Daly, A. (2012b). Not bored yet – Revisiting respondent fatigue in stated choice experiments. *Transportation Research Part A: Policy and Practice, 46*(3), 626-644. doi:https://doi.org/10.1016/j.tra.2011.11.008

Hess, S., Stathopoulos, A., & Daly, A. (2012a). Allowing for heterogeneous decision rules in discrete choice models: an approach and four case studies. *Transportation, 39*(3), 565-591. doi:10.1007/s11116-011-9365-6

Hess S, Stathopoulos A. (2013). A mixed random utility — Random regret model linking the choice of decision rule to latent character traits. Journal of Choice Modelling. 9:27-38.

Hess, S. & Chorus, C. (2015), Utility Maximisation and Regret Minimisation: A Mixture of a Generalisation, in Rasouli, S. & Timmermans, H. (eds), Bounded Rational Choice Behaviour: Applications in Transport, chapter 2, pp. 31-48, Emerald, Bingley.

Hess, S., & Palma, D. (2019). Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application. *Journal of Choice Modelling, 32*, 100170. doi:https://doi.org/10.1016/j.jocm.2019.100170

Hey, J. D. and C. Orme (1994). "Investigating Generalizations of Expected Utility Theory Using Experimental Data." Econometrica 62(6): 1291-1326.

Jha, P., C. Ramasundarahettige, V. Landsman, B. Rostron, M. Thun, R. N. Anderson, T. McAfee and R. Peto (2013). "21st-Century Hazards of Smoking and Benefits of Cessation in the United States." New England Journal of Medicine 368(4): 341-350.

Leong, W., & Hensher, D. A. (2012). Embedding Decision Heuristics in Discrete Choice Models: A Review. *Transport Reviews, 32*(3), 313-331. doi:10.1080/01441647.2012.671195

Loomes, G. and R. Sugden (1983). "A Rationale for Preference Reversal." The American Economic Review 73(3): 428-432.

Marti J, Buckell J, Maclean JC, Sindelar J. TO "VAPE" OR SMOKE? EXPERIMENTAL EVIDENCE ON ADULT SMOKERS. Economic Inquiry. 2019;57(1):705-25.

McFadden, D. "Conditional Logit Analysis of Qualitative Choice Behavior," in Frontiers in Econometrics, edited by P. Zarembka. New York: Academic Press, 1974, 105–42.

McNeill, A., L. S. Brose, R. Calder, S. C. Hitchman, P. Hajek, and H. McRobbie. E-Cigarettes: An Evidence Update. A Report Commissioned by Public Health England. London: Public Health, 2015.

Nielsen, & Jacobsen, J. (2020). Decision rules matter in choice experiments about hunting and trading bushmeat. *Conservation Biology, Forthcoming*.

Paul, P., Berlin, C., Maessen, M., & Valtonen, H. (2018). A comparison of regret-based and utility-based discrete choice modelling – an empirical illustration with hospital bed choice. *Applied Economics, 50*(40), 4295-4305. doi:10.1080/00036846.2018.1444260

Quaife, M., Eakle, R., Cabrera Escobar, M. A., Vickerman, P., Kilbourne-Brook, M., Mvundura, M., . . . Terris-Prestholt, F. (2017). Divergent Preferences for HIV Prevention: A Discrete Choice Experiment for Multipurpose HIV Prevention Products in South Africa. *Medical Decision Making, 38*(1), 120-133. doi:10.1177/0272989X17729376

Regier, D. A., Weymann, D., Buchanan, J., Marshall, D. A., & Wordsworth, S. (2018). Valuation of Health and Nonhealth Outcomes from Next-Generation Sequencing: Approaches, Challenges, and Solutions. *Value in Health, 21*(9), 1043-1047. doi:https://doi.org/10.1016/j.jval.2018.06.010

Regier, D. A., Veenstra, D. L., Basu, A., & Carlson, J. J. (2020). Demand for Precision Medicine: A Discrete-Choice Experiment and External Validation Study. *PharmacoEconomics, 38*(1), 57-68. doi:10.1007/s40273-019-00834-0

Regmi, K., D. Kaphle, S. Timilsina and N. A. A. Tuha (2018). "Application of Discrete-Choice Experiment Methods in Tobacco Control: A Systematic Review." PharmacoEconomics - Open 2(1): 5-17.

Ryan, M., Krucien, N., & Hermens, F. (2018). The eyes have it: Using eye tracking to inform information processing strategies in multi-attributes choices. *Health Economics, 27*(4), 709-721. doi:10.1002/hec.3626

Soekhai, V., de Bekker-Grob, E. W., Ellis, A. R., & Vass, C. M. (2019). Discrete Choice Experiments in Health Economics: Past, Present and Future. *PharmacoEconomics, 37*(2), 201-226. doi:10.1007/s40273-018-0734-2

Sorum, P. C., Mullet, E., Shim, J., Bonnin-Scaon, S., Chasseigne, G., & Cogneau, J. (2004). Avoidance of Anticipated Regret: The Ordering of Prostate-Specific Antigen Tests. *Medical Decision Making, 24*(2), 149-159. doi:10.1177/0272989X04263163

Stahl, D. O. and P. W. Wilson (1995). "On Players′ Models of Other Players: Theory and Experimental Evidence." Games and Economic Behavior 10(1): 218-254.

Strauss AL. Qualitative analysis for social scientists. Cambridge Cambridgeshire ; New York: Cambridge University Press; 1987

Strazzera, E., Contu, D. and Ferrini, S. (2013), Check it out! a monte carlo analysis of the performance of selection criteria and tests for models fitted to choice experiments data, in `Proceedings of the International Choice Modelling Conference', Citeseer.

Terris-Prestholt, F., K. Hanson, C. MacPhail, P. Vickerman, H. Rees and C. Watts (2014). "How Much Demand for New HIV Prevention Technologies Can We Really Expect? Results from a Discrete Choice Experiment in South Africa." PLOS ONE 8(12): e83193.

Train, K. E. Discrete Choice Methods with Simulation. Cambridge: Cambridge University Press, 2009.

van Cranenburgh, S., Guevara, C. A., & Chorus, C. G. (2015). New insights on random regret minimization models. *Transportation Research Part A: Policy and Practice, 74*, 91-109. doi:https://doi.org/10.1016/j.tra.2015.01.008

Vuong, Q. H. (1989). "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses." Econometrica 57(2): 307-333.

Tables and Figures

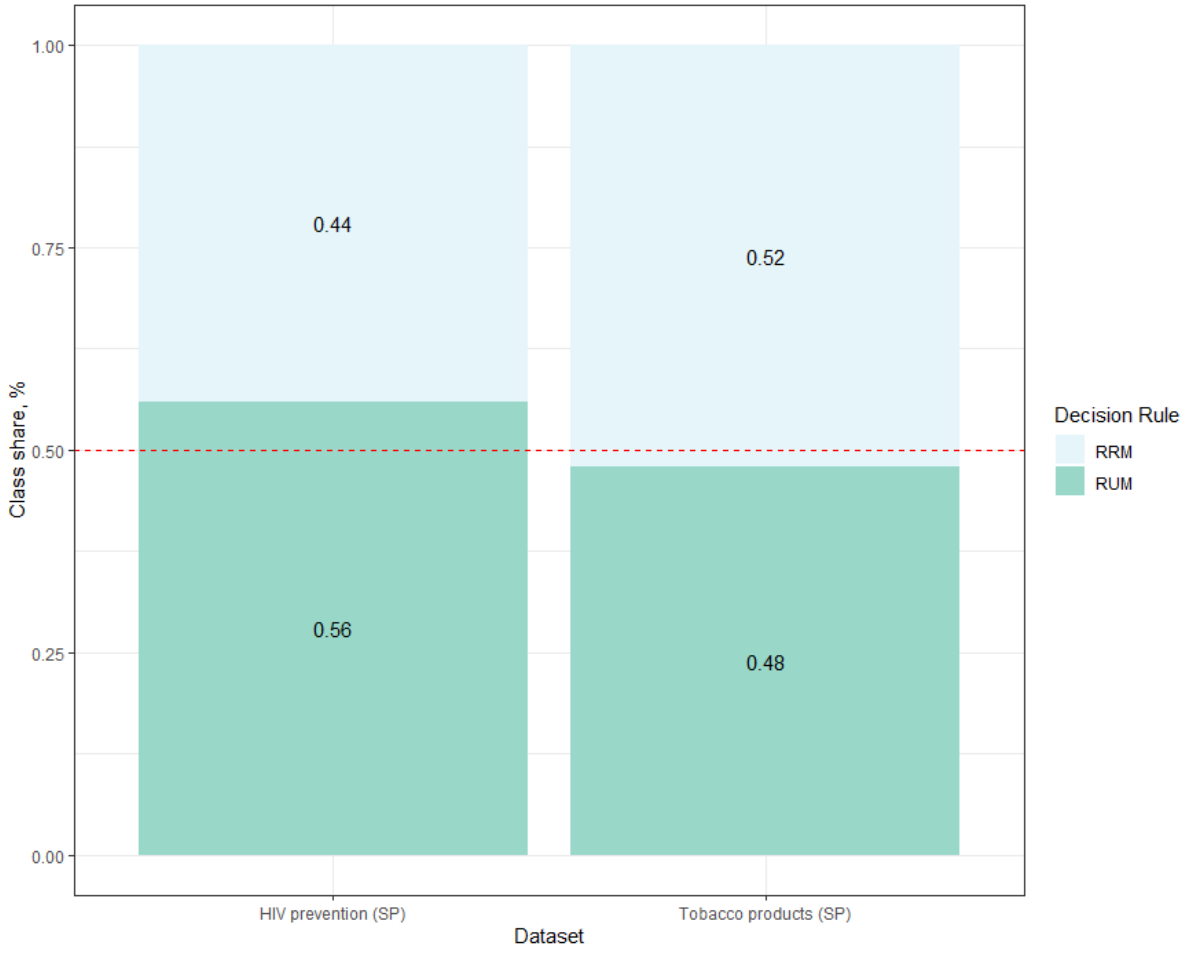

Figure 1: Class shares from latent class, decision rule heterogeneous models.

	Tobacco Products, SP			Tobacco Products, RP		Genomic Testing		HIV Prevention		
	Model 1: RUM	Model 2: RRM	Model 3: LCDRH	Model 1: RUM	Model 2: RRM	Model 1: RUM	Model 2: RRM	Model 1: RUM	Model 2: RRM	Model 3: LCDRH
No. of individuals	2031	2031	2031	2031	2031	1140	1140	367	367	367
No. of observations	24372	24372	24372	2031	2031	18240	18240	3658	3658	3658
LL (null)	-43668.76	-43668.76	-43668.76	-2815.56	-2815.56	-25286.01	-25286.01	-5071.07	-5071.07	-5071.07
LL (final)	-39413.25	-39419.77	-32329.68	-2202.05	-2202.53	-21457.20	-21487.81	-4612.44	-4629.35	-3960.98
Estimated parameters	8	8	22	7	7	17	17	11	13	17
AIC	78842.49	78855.53	64703.36	4418.11	4419.05	42948.40	43009.61	9246.87	9284.69	7955.95
BIC	78907.30	78920.34	64881.59	4457.42	4458.37	43081.19	43142.40	9315.12	9365.35	8061.43
Class Share RUM			0.48							0.56
Class Share RRM			0.52							0.44
Testing between models										
Vuong test: (1) vs (2)	5.96			1.34		8.48		3.65		
Vuong test: (1) vs (3)	-35.43							-13.21		
Vuong test: (2) vs (3)		-35.50							-13.43	
Specification test of LCDRH										
LR test statistic vs 4-Latent Class RUM			35.43							13.21
LR test statistic vs 4-Latent Class RRM			35.49							13.43

Table 1: Summary table for mixed logit models. LL – log-likelihood, AIC – Akaike Information Criteria, BIC – Bayesian Information Criteria. Vuong test – test statistic for Vuong test, values >1.96 favor the left-hand model, values <-1.96 favour the right-hand model, values in between are inconclusive.

	Model 1: Utility Maximisation (RUM)		Model 2: Regret Minimisation (RRM)		Model 3: LCDRH							
					Class 1: RUM		Class 2: RUM		Class 3: RRM		Class 4: RRM	
	Beta estimate	Rob. t-ratio (0)	Beta estimate	Rob. t-ratio (0)	Beta estimate	Rob. t-ratio (0)	Beta estimate	Rob. t-ratio (0)	Beta estimate	Rob. t-ratio (0)	Beta estimate	Rob. t-ratio (0)
Product constants:												
Cigarettes	2.44	43.69	2.38	44.00	4.39	20.32	2.09	11.36	0.71	4.08	4.33	25.11
E-cigarettes	1.52	23.06	1.47	22.64			3.35	24.19			2.58	14.02
Attributes:												
Price	-0.10	-29.88	-0.03	-31.12	-0.12	-9.44	-0.10	-15.29	-0.07	-12.27	-0.03	-16.13
Nicotine: none											-0.09	-3.44
Nicotine: low												
Nicotine: high												
Flavor: menthol	-0.34	-9.97	-0.11	-10.05	-4.34	-21.97					0.13	5.83
Flavor: fruit/sweet	-0.18	-5.46	-0.06	-5.46			0.27	6.36	-0.18	-3.28		
Life years lost: 2	0.68	10.49	0.24	13.16	1.19	7.09			0.20	3.73		
Life years lost: 5	0.18	13.65	0.06	4.02								
Life years lost: unknown	0.49	4.08	0.17	10.30	1.32	6.58			0.19	4.44		
Class Membership Probability:												
RRM									0.08		1.46	
RUM2							0.72	8.84				
RRM2											1.02	11.85
Forecasts:												
Cigarettes		51.12		51.12								51.14
E-cigarettes		35.88		35.88								35.48
Opt-out		13.00		13.00								13.38

Table 2: Parameter estimates for Tobacco SP MNL models. Omitted attribute levels (attributes/products): opt-out (products), medium (nicotine), tobacco (flavour), 10 years (life years lost). Parameters that are missing are so as they were not found to be statistically significantly different from zero and removed from the model.

	Model 1: Utility Maximisation (RUM)		Model 2: Regret Minimisation (RRM)	
	Beta estimate	Rob. t-ratio (0)	Beta estimate	Rob. t-ratio (0)
Product constants:				
Smoker				
Dual User	-0.70	-8.52	-0.67	-9.11
Vaper	-3.60	-16.83	-3.58	-17.07
Recent Quitter	-1.71	-16.01	-1.68	-17.04
Attributes:				
Price	-0.02	-2.64	-0.01	-2.72
Flavor: fruit/sweet	2.82	11.56	1.75	9.31
Flavor: other	0.85	2.41	0.45	2.28
Flavor: none	2.29	7.63	1.36	6.38
Forecasts				
Smoker	51.11		51.11	
Vaper	7.29		7.29	
Dualuser	30.48		30.48	
Recent Quitter	11.13		11.13	

Table 3: Parameter estimates for Tobacco RP MNL models. Omitted attribute levels (attributes/products): smoker (products), tobacco (flavour).

	Model 1: Utility Maximisation (RUM)		Model 2: Regret Minimisation (RRM)	
	Beta estimate	Rob. t-ratio (0)	Beta estimate	Rob. t-ratio (0)
Treatment change	-1.98	-27.04	-1.98	-27.88
Treatment change & Hereditary Screening	-1.38	-18.16	-1.38	-18.73
Treatment change	-2.05	-28.41	-2.06	-29.48
Most experts agree	0.57	17.47	0.28	17.07
All experts agree	0.91	24.76	0.46	24.44
Percent genetic marker	1.02	19.14	0.51	17.45
Sftod	0.30	11.62	0.15	11.39
Sfaod	0.45	17.00	0.23	16.51
Cost of test	0.00	-23.93	0.00	-24.24
Percent responsive	1.03	19.51	0.48	19.77
QoL: poor to fair	-0.09	-2.32	-0.05	-2.47
QoL: fair to excellent	0.57	15.46	0.29	14.77
QoL: poor to excellent	0.75	19.48	0.39	18.31
Life year gain: 1	0.15	4.33	0.07	4.00
Life year gain: 2	0.41	11.16	0.20	10.58
Life year gain: 5	0.83	21.14	0.44	19.82
Family year gain: na	-0.23	-2.81	-0.12	-2.89
Family year gain: 5				

Table 4: Parameter estimates for Genomic Testing SP MNL models. PGM - number (of 100) of people identified with genetic marker, Experts agree - medical expert agreement on getting tested, OD: treatable - information on your risk of getting other treatable diseases, OD: all - information on your risk of getting all other diseases, Cost - cost of testing, Percent responsive - number of people (of 100) with genetic marker who respond to treatment, QOL - quality of life change from individualized treatment, Life year gain - length of life change from individualized treatment, Family Year Gain - length of life change for your family from screening. Omitted attribute levels (attributes/products): No test (products), few (experts agree), no information (OD), no change (quality of life), 0 (life years gained), none (family year gain). Parameters that are missing are so as they were not found to be statistically significantly different from zero and removed from the model.

	Model 1: Utility Maximisation (RUM)		Model 2: Regret Minimisation (RRM)		Model 3: LCDRH							
					Class 1: RUM		Class 2: RUM		Class 3: RRM		Class 4: RRM	
	Beta estimate	Rob. t-ratio (0)	Beta estimate	Rob. t-ratio (0)	Beta estimate	Rob. t-ratio (0)	Beta estimate	Rob. t-ratio (0)	Beta estimate	Rob. t-ratio (0)	Beta estimate	Rob. t-ratio (0)
Optout	3.18	13.75	2.80	13.86	1.70	2.84	2.49	2.96	4.86	15.15	9.84	5.49
Prep	-0.45	-5.69	-0.16	-3.60								
Silcs	-0.38	-2.87	-0.16	-2.50								
Gel	-0.35	-3.48	-0.12	-2.36								
Ring	-0.57	-6.76	-0.29	-7.32	-0.25	-3.23						
HIV protection	0.03	13.20	0.01	12.26	0.04	8.03					0.04	5.88
Pregnancy prevention	0.71	11.37	0.32	9.92	0.54	6.67	2.38	4.51	1.61	11.83		
STI protection	0.82	10.94	0.36	9.80	0.14	1.98	5.78	4.13	1.31	9.69		
Use: Daily	0.21	2.54	0.13	3.16								
Use: Weekly												
Use: Monthly	0.29	4.41	0.07	5.80								
Use: 3 Months			0.21	3.87								
Use: 6 Months			0.16	3.34								
Effects: dizzy	-0.19	-3.51	-0.09	-3.47								
Effects: stomach												
Effects: nausea												
Class Membership Probability:												
RRM									-0.24		-1.44	
RUM2							-2.06	-8.13				
RRM2											-0.24	-1.00

Table 5: Parameter estimates for HIV Prevention SP MNL models. Prep - Oral pre-exposure prophylaxis, Gel - Microbicide gel, Slics - Microbicide gel with SILCS diaphragm, Ring - Vaginal ring. Use – frequency of use. Effects – side effects. Omitted attribute levels (attributes): injection (product), no (protection), no (pregnancy protection), coitally (use), none (effects). Parameters that are missing are so as they were not found to be statistically significantly different from zero and removed from the model.

Study Design Feature		Tobacco RP	Tobacco SP	HIV Prevention	Genomics	Dominant DR	Caveats
Setting	Clinical	0	0	0	1	RUM	complex models inestimable
Setting	Health behavior	1	1	1	0	RUM	some evidence of RRM in tobacco SP and HIV prevention LCDRM models
Economy	More Developed	1	1	0	1	RUM	some evidence of RRM in tobacco SP LCDRM model
Economy	Less Developed	0	0	1	0	RUM	some evidence of RRM in HIV prevention LCDRM model
Data type	SP	1	0	1	1	RUM	some evidence of RRM in tobacco SP and HIV prevention LCDRM models
Data type	RP	0	1	0	0	RUM	complex models inestimable
Data collection	Online	1	1	0	1	RUM	some evidence of RRM in tobacco SP LCDRM model
Data collection	Face-to-face	0	0	1	0	RUM	some evidence of RRM in HIV prevention LCDRM model
Sampling	General population	0	0	1	1	RUM	some evidence of RRM in HIV prevention LCDRM model
Sampling	Specific population	1	1	0	0	RUM	some evidence of RRM in tobacco SP LCDRM model

Table 6: Decision rules and study design features.

	Tobacco, n=2,031			HIV prevention, n=408			Genomic Testing, n=1,140		
	Mean	n	% sample	Mean	n	% sample	Mean	n	% sample
Age	40.0			23.5					
Age Catergory: 18-34								310	27.19%
Age Catergory: 35-44								185	16.23%
Age Catergory: 45-54								203	17.81%
Age Catergory: 55-64								199	17.46%
Age Catergory: 65+								230	20.18%
Prefer not to answer								13	1.13%
Female		1101	54.21%		408	100.00%		584	51.23%
White		1759	86.61%						
Black		183	9.01%						
Asian		51	2.51%						
Hispanic		166	8.17%						
Education: College or higher		972	47.86%		34	8.33%			
Cigarettes per day	12.5								
E-cigarette use		767	37.76%						
Employment: Full time					43	10.54%			
Employment: Part time					42	10.29%			
Employment: Student/scholar					213	52.21%			
Employment: Work seeker/unemployed					96	23.53%			
Employment: Other					11	2.70%			
In a stable relationship					233	57.11%			
Cohabiting					60	14.71%			
Household monthly income: <R5,000					271	66.42%			
Household monthly income: R5,000-15,000					111	27.21%			
Household monthly income: <R5,002					17	4.17%			
HIV positive					40	9.80%			
Not sexually active					271	66.42%			
Age at first sex				17					
Number of lifetime partners				5.7					
Number of partners in previous year				2.3					
Current regular sexual partner					383	93.87%			
Condom use at last sex with regular partner					134	32.84%			
Currently using any form of contraception					185	45.34%			
Ever received an HIV test					68	16.67%			
External sexual partner in prior 3 months					14	3.43%			
British Columbia								146	12.81%
Alberta								136	11.93%
Prairie								89	7.81%
Ontario								398	34.91%
Quebec								254	22.28%
Atlantic								112	9.82%
Prefer not to answer								5	<1%

Table 7: descriptive statistics of individual characteristics from each of the datasets.