Article:

Hess, S and 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. ISSN 1755-5345

https://doi.org/10.1016/j.jocm.2013.12.005

A mixed random utility - random regret model linking the choice of decision rule to latent character traits

Stephane Hess* Amanda Stathopoulos†

December 2, 2013

Abstract

An increasing number of studies are concerned with the use of alternatives to random utility maximisation as a decision rule in choice models, with a particular emphasis on regret minimisation over the last few years. The initial focus was on revealing which paradigm fits best for a given dataset, while later studies have looked at variation in decision rules across respondents within a dataset. However, only limited effort has gone towards understanding the potential drivers of decision rules, i.e. what makes it more or less likely that the choices of a given respondent can be explained by a particular paradigm. The present paper puts forward the notion that unobserved character traits can be a key source of this type of heterogeneity and proposes to characterise these traits through a latent variable within a hybrid framework. In an empirical application on stated choice data, we make use of a mixed random utility-random regret structure, where the allocation to a given class is driven in part by a latent variable which at the same time explains respondents' stated satisfaction with their real world commute journey. Results reveal a linkage between the likely decision rule and the stated satisfaction with the real world commute conditions. Notably, the most regret-prone respondents in our sample are more likely to have aligned their real-life commute performance more closely with their aspirational values.
Keywords: random regret, random utility, latent class, decision rules, hybrid models

1 Introduction

Significant effort has gone into understanding and modelling variability in human behaviour across different sectors, ranging from health to transport. In recent years, there has been particular interest in the use of alternatives to the random utility maximisation paradigm as a decision rule in discrete choice models. While some work has looked again at long standing alternatives such as elimination by aspects (EBA) or other non-compensatory decision rules, a key area of investigation has been the potential use of regret minimisation rather than utility maximisation (see e.g. Chorus, 2012).

While early studies focussed solely on establishing which specific decision paradigm offered the best performance on a given dataset, the assumption that all people in a sample use the

*Institute for Transport Studies, University of Leeds, Leeds LS2 9JT, s.hess@its.leeds.ac.uk, Tel: +44 (0)113 34 36611, Fax: +44 (0)113 343 5334

†Transport and Mobility Laboratory, ENAC, EPFL, amanda.stathopoulos@epfl.ch

same decision process was soon identified as a potentially restrictive assumption. As a result, novel approaches have been put forward where several decision processes are allowed to co-exist in the same population. It was in this context that Hess et al. (2012) proposed the use of latent class structure where different decision rules are used in different classes. Each respondent falls into every class according to an estimated probability, where this can vary across respondents. In a similar vein, a mixed model has been proposed by Chorus et al. (2013) where instead of distinguishing sub-groups of respondents, the attributes are allowed to be processed according to either regret or utility.

Any model allowing for multiple decision paradigms clearly offers improvements in flexibility and potential further insights into behaviour. Studies comparing the use of different decision paradigms on different datasets can of course never determine whether a given decision rule was actually 'used' by the respondents in the data, and can only establish which specific rule works best in explaining the observed choices. Similarly, studies allowing jointly for multiple decision paradigms can only ever come to conclusions as to a specific paradigm having been a more likely approach for a given respondent, but not determine which one, if any of those included in the model, was actually used.

The empirical evidence in the literature does suggest that different decision rules work differently well in different datasets, while further gains in explanatory power can be obtained by relaxing the assumption of within sample homogeneity. Focussing particularly on the case of RUM *vs* RRM, the main approach has indeed been to use a population-wide treatment where all respondents are assumed to use *either* regret or utility based decisions (e.g. Thiene et al., 2012; Chorus et al., 2009; Hensher et al., 2011; Chorus and Rose, 2011). These applications have revealed to date that regret-minimisation favours compromise alternatives with average performance for each attribute (Chorus et al., 2008). Importantly, a regret driven choice process is shown to fit the data better than utility maximisation in about half of the existing studies Chorus (2012). The differences in fit, however, are typically small. Other studies have gone further by allowing for mixtures of decision rules, for example Hess et al. (2012); Chorus et al. (2013); Bekker-Grob and Chorus (2013). A key issue in this context was highlighted by Hess et al. (2012) who observe substantial confounding between taste heterogeneity and decision rule heterogeneity, an issue we will return to later on in the paper. A first observation to be made is that work on comparisons between decision at the sample levels or work allowing for within sample heterogeneity has arguably put too much emphasis on statistical fit and not on the reasonableness of the behavioural implications or economic outputs. More importantly however, there has been a comparative lack of investigation as to the likely drivers of decision rule heterogeneity, i.e. the question of what makes a specific person more or less likely to make choices that can be explained by a given paradigm. It is this gap that the present paper seeks to address.

Most existing work does not offer an explanation as to why one behavioural paradigm might be more suitable than another, whether working at the sample level or the level of individual respondents. When such explanatory factors for decision rule heterogeneity have been investigated, they mainly relate to observable features of the decision-maker or the context of the choice. For instance, the approach of subsetting data reveals that men's choices appear to be governed more by regret-minimisation (Chorus and Rose, 2011). Other work has found that instances of superior performance of regret models compared to RUM models align with a lack of experience with the decision setting (Boeri et al., 2012).

In the present paper, we take a somewhat different approach. In particular, we argue that there are inherent and unobserved character traits which influence the way in which given people make their choices. While observable characteristics such as age or gender may play a role in forming these traits, there is substantial scope for idiosyncratic differences across people. Examples of such character traits may include the fact that some individuals are more risk averse than others. Similarly, some respondents will be more driven towards compromise solutions while others are more relentless in their pursuit of maximum gain. It is this latter example that our empirical work relates to and which forms the basis of our methodological developments, although they are easily transferable to other context. In particular, we look at the specific case of a model where we allow for the two most widely used types of paradigms, namely random utility maximisation (RUM) and random regret minimisation (RRM). We argue that individuals who are driven to select the best overall outcome are more likely to make choices that can be explained by RUM while those who seek to avoid frustration, disappointment or regret are more likely to make choices that can be explained by RRM.

A crucial component of our work is the assumption that it is inherent character traits that determine which way a person makes their choices, rather than focussing on the influence of the specific choice setting. The issue then remains how such character traits can be captured in our models. The vast majority of applications looking at decision rule heterogeneity make use of data from stated choice (SC) surveys. The choices captured in such data provide a snapshot of the preferences by a person in a very controlled setting and may not be well suited on their own for identifying underlying factors that drive the decision process. Instead, we put forward the combined use of SC data and data relating to real world behaviour, and in particular a respondent's stated satisfaction with their real world choices.

Regret can be said to relate to prevention-based decision making where individuals are concerned with avoiding unsatisfactory outcomes (Crowe and Higgins, 1997). This is in contrast with a promotion-based goal direction where decisions are driven by a desire to approach a specific end-state. In line with these insights we use data on aspirations to explain regret minimisation efforts. Our specific hypothesis in this context is that a respondent who has an underlying tendency to avoid outcomes that potentially lead to regret, frustration or disappointment is likely to have aligned his/her real world choices accordingly. To test this hypothesis, we collect information on respondents' satisfaction with their real world commute journeys in the context of a survey which first presented them with a set of stated choice scenarios for journeys to work. We then develop a hybrid structure which explains both the stated choice behaviour and the stated satisfaction with their real life situation. Our choice model allows jointly for utility maximisation and regret minimisation within a latent class framework, where the probability of a given respondent falling into either decision rule class is a function of a latent variable which relates to the specific character traits discussed above. This latent variable is used at the same time to model respondents' satisfaction with their real world choices. A strong pattern emerges which shows that those respondents who in our hypothetical scenarios are found to make choices which are better explained by regret minimisation are also more likely to have expressed a higher level of satisfaction with their real life commute. We interpret this as such respondents having an underlying tendency to avoid regret and have used this over time to align their real world commute with their aspirations.

Our work links up well with other developments. Indeed, early explorations of latent variable models (e.g. Ben-Akiva et al., 1999) discuss the impact of latent constructs in influencing decision processes. Similarly, the seminal paper by Swait (2001) on non-compensatory preference cutoffs suggests as a future area of research the use of behavioural indicators to improve the identification of the cutoffs. These insight are echoed also for the specific context or regret minimisation modelling in Chorus and Bierlaire (2013) who underscore the need to explore to what extent personality can trigger the decision-process under study.

The remainder of this paper is organised as follows. Section 2 gives an outline of the methodology, while 3 describes the survey instrument and model specification. Results and discussions are given in Sections 4 and 5.

2 Methodology

Following Hess et al. (2012), a general specification of a model allowing for different decision rules within a latent class framework is given by:

$$LC_n\left(\beta_1,\ldots,\beta_S,\pi_{n,1},\ldots,\pi_{n,S}\right)=\sum_{s=1}^{S}\pi_{n,s}LC_{n,s}\left(\beta_s\right). \tag{1}$$

With this notation, LC_n is the contribution to the likelihood function of the observed choices for respondent n (out of N). This probability of observed choices is given by a weighted average over S different types of models, where $LC_{n,s}$ is the probability of the observed sequence of choices for person n if model s is used, and $\pi_{n,s}$ is the weight attached to model s (representing a specific decision process), where $\sum_{s=1}^{S}\pi_{n,s}=1$, $\forall n$. In the above specification, β_s is the vector of parameters (e.g. utility coefficients) used in model s.

Hess et al. (2012) use a specification of this type for several different combinations of behavioural paradigms, including a mixture of regret minimisation and utility maximisation. The approach has two key shortcomings. First, there is a risk of confounding between heterogeneity in sensitivities and heterogeneity in decision rules. Second, there is limited insight into the factors determining the choice of decision rule.

Hess et al. (2012) observe clear evidence of the first of the above problems, with for example the random regret class in their RUM-RRM mixture primarily capturing behaviour that exhibits strong fare sensitivity. When allowing for additional random heterogeneity within each decision rule class, they observe substantially different patterns of heterogeneity in decision rules, with a notable decrease in the weight for the RRM class. The inclusion of additional random heterogeneity was dealt with by Hess et al. (2012) in a continuous manner, i.e. using integration over the distribution of parameters for the given model. This would allow us to rewrite Equation 1 as:

$$LC_n\left(\Omega_1,\ldots,\Omega_S,\pi_{n,1},\ldots,\pi_{n,S}\right)=\sum_{s=1}^{S}\pi_{n,s}\left[\int_{\beta_s}LC_{n,s}\left(\beta_s\right)f\left(\beta_s\mid\Omega_s\right)\mathrm{d}\beta_s\right], \tag{2}$$

where we now have that $\beta_s\sim f\left(\beta_s\mid\Omega_s\right)$. In essence, we now replace the models in each class of the overall structure by continuous mixture equivalents of those used in the specification in Equation 1. The use of a continuous specification within each class (e.g. a mixed RUM model in one class, a mixed RRM model in another class, etc) imposes substantial demands in terms of computational complexity as well as empirical identification. Additionally, the results of continuous mixture models are strongly influenced by the assumptions in terms of distributions for each parameter. In the present paper, we instead put forward the use of an additional layer of latent classes to accommodate the within model heterogeneity. Specifically, we now use:

$$LC_n\left(\beta^{(1)},\ldots,\beta^{(S)},\pi_{n,1},\ldots,\pi_{n,S},\varpi_n^{(1)},\ldots,\varpi_n^{(S)}\right)=\sum_{s=1}^{S}\pi_{n,s}\sum_{k=1}^{K_s}\varpi_{n,s,k}LC_{n,s}\left(\beta_{s,k}\right). \quad (3)$$

In this notation, we now have that $\beta^{(1)}=\langle\beta_{1,1},\ldots,\beta_{1,K_1}\rangle$, where this corresponds to one vector for each of the K_1 classes using model 1, and where we allow for the possibility that K_s varies across models, e.g. we might have more RUM classes than RRM classes. We additionally have that $\varpi_n^{(1)}=\langle\varpi_{n,1,1},\ldots,\varpi_{n,1,K_1}\rangle$ gives the weights for the K_1 classes conditional on using model 1, where $\sum_{k=1}^{K_s}\varpi_{n,s,k}=1,\ \forall s$. This model now uses K_s different classes for model s, where, if $K_s=1\,\forall s$, the model collapses back to the specification in Equation 1. As can be noted from Equation 3, the averaging across classes is performed at the level of individual respondents (i.e. sequences of choices), recognising the repeated choice nature of the data. It is clear that the use of additional random heterogeneity in sensitivities conditional on a given model type cannot completely eliminate the risk of confounding between this type of variation and heterogeneity in decision rules. It can only ever aim to reduce this risk, where the use of a latent class approach can have potential advantages over a continuous mixture kernel in terms of the flexibility of the resulting distribution in sensitivities within a model.

The other shortcoming of the structure thus far is the lack of explanation as to what drives the likelihood of a given paradigm being more appropriate for one specific respondent than another. One possibility would be to make the class allocation probabilities π a function of respondent specific characteristics, i.e. writing $\pi_{n,s}=g\left(\gamma_s,z_n\right)$ where z_n is a vector of characteristics of respondent n and γ_s is a vector of estimated parameters. Alternative, we could make the class allocation probabilities a function of the settings of the choice task, which would also involve repositioning the weighted averaging across classes to the level of individual tasks rather than individual respondents.

In the present paper, we take a different approach. In particular, we hypothesise that underlying respondent-specific character traits can be used to explain the allocation to the different decision paradigm classes. Working with a single such trait for the sake of exposition, let us refer to it as α_n for respondent n. Simplifying our overall structure further to the case of just two decision paradigms (i.e. $S=2$, which is consistent with our later empirical setting), we now make the probability of a given respondent n being allocated to either of the two paradigm classes a function of α_n by writing:

$$\pi_{n,1}=\frac{1}{1+e^{\delta_{\pi,2}+\tau\alpha_n}} \quad (4)$$

$$\pi_{n,2}=\frac{e^{\delta_{\pi,2}+\tau\alpha_n}}{1+e^{\delta_{\pi,2}+\tau\alpha_n}}. \quad (5)$$

In this specification, $\delta_{\pi,1}$ is a constant that allows us to capture the sample level weight for decision rule 2 (using 1 as the base). The estimated parameter τ determines whether respondent n is more or less likely to be allocated to class 2 than the average respondent in the sample, depending on the value of α_n.

The key issue at this stage is that the character traits likely to drive decision rule heterogeneity cannot be 'observed' by an analyst, i.e. α_n is latent. A structural equation is used to model the value of α_n as:

$$\alpha_n = \gamma' z_n + \eta_n, \tag{6}$$

where z_n is a vector of socio-demographic characteristics of respondent n, γ is a vector of estimated parameters and η_n is a random disturbance which follows a standard Normal distribution across individuals. In practice, it will likely be very difficult to find meaningful socio-demographic explanators for underlying character traits, as these are more likely to be intrinsic to a person and shaped by experience and lifestyle, either of which are difficult to capture in data. This is a very similar situation to the difficulties inherent to to explaining attitudes on the basis of socio-demographics, as highlighted recently by Abou-Zeid and Ben-Akiva (2014), drawing also on Anable (2005).

Thus far, this model would simply allow for random (through η_n) and deterministic (through $\gamma' z_n$) variations across respondents in the class allocation probabilities. The model would be able to estimate the relationship between the latent character traits and the likely decision rules only on the basis of the data on hypothetical choices. As highlighted in the introduction, these however only provide a snapshot of preferences in a very controlled settings at a particular point in time and arguably do not permit us to make the full link to what we regard as person specific character traits which are constant over a longer time horizon. For this reason, we make use of additional information relating to other manifestations of these character traits.

Let us assume that our data contains additional variables at the level of each individual which we hypothesise to be a function of the same latent character traits that also drive the allocation to different decision paradigm classes in our latent class structure. The identification of such variables is a difficult task and could encompass a range of different formats, be it answers to questions on attitudes and perceptions, or descriptors of lifestyle and past experiences. Of crucial importance within the behavioural concept at the heart of our approach is that they need to relate to long term traits rather than short term feelings.

Under the assumption that these additional measures, referred to hereafter as *indicators*, say $I_{n,1}$ to $I_{n,M}$ grouped together into a vector I_n, are linked to the same underlying character traits α_n, we model their values as:

$$I_{n,m} = \delta_{I,m} + \zeta_m \alpha_n + \upsilon_{n,m}, \tag{7}$$

where $\delta_{I,m}$ is a sample level constant for indicator m, ζ_m captures the impact of the latent variable α_n on this indicator, and $\upsilon_{n,m}$ is an error term. The distributional assumptions made for the error term (e.g. normal, logistic, etc) determine the functional form of the measurement model and are thus a reflection of the nature of the indicators (e.g. continuous, ordinal, etc). As an example, if $I_{n,m}$ is continuous, we can for example use a normal density function to explain its value, and the likelihood $LI_{n,m}$ of the observed value for indicator m for person n is then given by:

$$LI_{n,m} = \frac{1}{\sigma_m} \phi \left(\frac{I_{n,m} - \zeta_m \alpha_n - \delta_{I,m}}{\sigma_m} \right) \tag{8}$$

where $\phi\left(\right)$ is the standard Normal density function. If the indicators are zero centred by subtracting the population mean, the estimation of $\delta_{I,m}$ becomes redundant.

In estimation, we now jointly maximise the likelihood of the observed choices and the observed values of the indicators. With both components being a function of the latent variables, we enable the model to create a link between the behaviour in the short term context (i.e. stated choice) and the longer term character traits. The resulting structure is an example of a hybrid model (Ben-Akiva et al., 1999) and falls into a growing body of work that exploit such models for the analysis of a range of behavioural traits such as attitudes, perceptions and future plans (see e.g. Bolduc et al., 2008; Choudhury et al., 2010; Abou-Zeid et al., 2010; Daly et al., 2012). We believe this to be the first specification of such a model within the context of allowing long term character traits to explain decision rule heterogeneity in choice data.

The combined model specification now gives the joint probability of the observed choices and the observed values of the indicators, both of which depend on α_n. Owing to the random component η_n in α_n, this joint probability does not have a closed form solution, and integration over the distribution of η_n is thus needed. Specifically, we have:

$$L_n = \int_{\eta_n} \left[\sum_{s=1}^{S} \pi_{n,s}\left(\alpha_n\right) \sum_{k=1}^{K_s} \varpi_{n,s,k} LC_{n,s}\left(\beta_{s,k}\right) \right] \left[\prod_{m=1}^{M} LI_{n,m}\left(\alpha_n\right) \right] \phi\left(\eta_n\right) \mathrm{d}\eta_n \tag{9}$$

where we make use of S different decision paradigm classes, with K_s classes for additional heterogeneity specific to a given decision paradigm class s. The first component gives the probability of the observed choices for respondent n, where this is given by a latent class structure with two layers of classes (for decision paradigm heterogeneity and within paradigm heterogeneity), as in Equation 3. The latent term α_n enters this part of the model in the decision paradigm class allocation weights $\pi_{n,s}$, as illustrated in Equation 4. The second component of the model gives the probability of the observed set of values for the indicators, given by a product across the individual measurement model components, each time explaining the value of an indicator as a function of α_n, as in Equation 7.

3 Data and model specification

3.1 Survey work

For our empirical application, we made use of data from an online survey conducted on rail and bus commuters in the UK in 2010. The survey gathered 3,680 observations from 368 respondents with information on experienced trip features (averaging trip attributes across 10 regular trips corresponding to a week of commuting). The baseline data was used to generate stated choice scenarios presenting three commuting options: one reference scenario representing each respondent's current situation, kept invariant across the 10 tasks, and two hypothetical, pivoted, alternatives. Six attributes were selected to characterise the commute; travel time in minutes, fare in £, the rate of crowded trips (out of ten trips), the rate of delays (out of ten trips), the average length of delays (across delayed trips), and the provision of a delay information service with different pricing. The survey was designed using NGene (ChoiceMetrics, 2012) with a D-efficient experimental design alongside appropriate conditions to avoid dominant alternatives. A total of 60 choice scenarios were generated and these were blocked into 6 sets of 10 tasks, minimising correlation with the blocking variable.

After completion of the stated choice tasks, data was also collected on aspirations relating to commute journeys. In particular, respondents were asked to provide information on *"acceptable"* and *"ideal"* conditions for the four core attributes attributes of their reference trip, namely travel time, fare, the rate of crowding and the rate of delays. Respondents were solicited for these values with the following instruction: "Taking into account technical constraints as well as the high rate of usage of the public transport network, could you for your current commute trip indicate the ideal values and realistic/acceptable values for the following attributes".

Table 1 provides some summary statistics on the answers to these questions. Similar to findings by Redmond and Mokhtarian (2001), the ideal values are observed to be lower than current conditions in the vast majority of cases, with the highest rate across attributes of ideal values being higher than current values arising for travel time and crowding, at only 3.26%, potentially a result of respondent mistakes. The fact that acceptable values can be higher than current values is not counter to intuition. Figure 1 explores the response patterns for the disparities between stated ideal and acceptable values and the levels currently experienced further. It is clear that when moving from *ideal* to *acceptable*, the distribution of differences with current values shifts towards the left, that is towards smaller, even negative differences. Overall, there is a clear pattern where *acceptable* values form an intermediate value between current and ideal fare or time. Further details on the relationship between declared aspirations and current trip performance can be found in Stathopoulos and Hess (2012).

Table 1: Comparing aspirations to current experience

	travel time		fare	
	ideal	acceptable	ideal	acceptable
equal to current	20.92%	31.79%	9.51%	17.12%
higher than current	3.26%	10.33%	0.82%	5.98%

	crowding		rate of delays	
	ideal	acceptable	ideal	acceptable
equal to current	35.05%	38.32%	21.74%	30.43%
higher than current	3.26%	13.04%	1.63%	11.96%

3.2 Specification for empirical example

Two different models were used in our empirical example, a simple latent class structure and the hybrid equivalent incorporating the role of a latent variable. In addition, for model fit comparisons, we also estimated simple RUM and RRM models as well as a latent class structure with one class per paradigm, i.e. $K_s = 1, \forall s$. All models were coded and estimated using Ox 6.2 (Doornik, 2001), making use of Halton draws for the continuous random component in the hybrid model, with simultaneous estimation of both model components in Equation 9. Given the highly non-linear form of the log-likelihood function, multiple estimation runs were carried out, each time using an initial search with $2,000$ sets of different randomly generated starting values.

In the specific empirical application conducted for the present paper, we make use of two

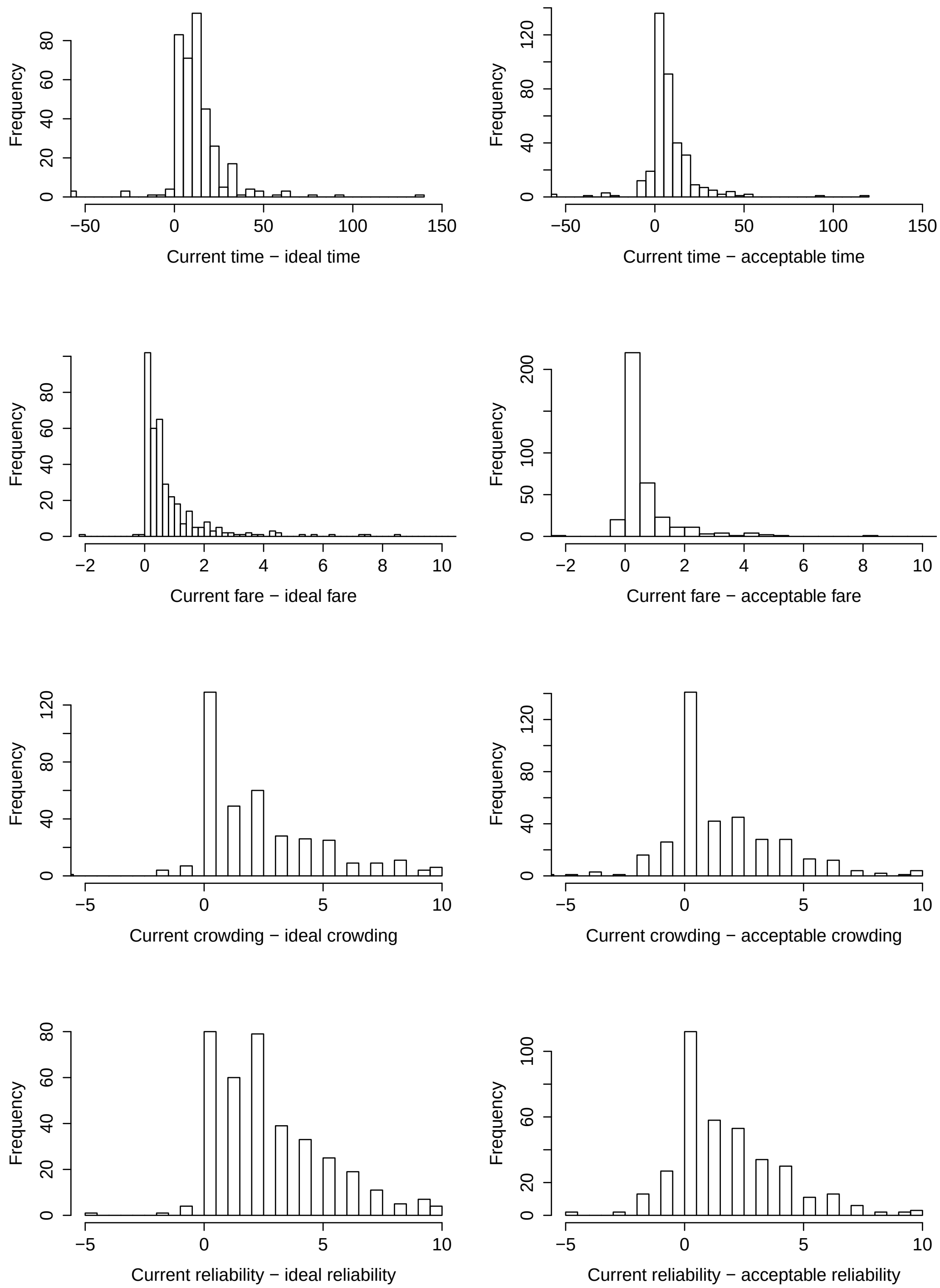

Figure 1: Differences between current and acceptable or ideal values for main journey attributes

decision paradigms within one model, such that $S = 2$, where $s = 1$ corresponds to a RUM structure, while $s = 2$ corresponds to a RRM structure. To allow for additional heterogeneity within the paradigm specific classes, we use a further level of latent classes, as illustrated in Equation 3, where we use $K_s = 2$, $\forall s$ in our example.

Within the RUM classes, the deterministic utility for alternative i ($i = 1, \dots, 3$) for respondent n in choice task t is given:

$$\begin{aligned} V_{n,t,i,k} = \; & \delta_{\mathsf{RUM},i,k} \\ & + \beta_{\mathsf{RUM},TT,k} TT_{n,t,i} \\ & + \beta_{\mathsf{RUM},LF,k} \ln\left(F_{n,t,i}\right) \\ & + \beta_{\mathsf{RUM},RD,k} RD_{n,t,i} \\ & + \beta_{\mathsf{RUM},ED,k} ED_{n,t,i} \\ & + \beta_{\mathsf{RUM},C,k} C_{n,t,i} \end{aligned} \tag{10}$$

where $k = 1, 2$ refers to the two RUM specific classes. With the notation used here, $\delta_{\mathsf{RUM},i,k}$ is an alternative specific constant which is set to zero for $i = 3$, $TT_{n,t,i}$ refers to travel time, $\ln\left(F_{n,t,i}\right)$ is the natural logarithm of the fare attribute (following evidence of strong decreasing marginal sensitivities), $RD_{n,t,i}$ is the rate of delays, $ED_{n,t,i}$ is the expected delay (rate multiplied by average delay), and $C_{n,t,i}$ is the rate of crowding. No consistent significant effects were found for the delay information attribute.

With an assumption of type I extreme value errors, the probability of respondent n choosing alternative i in choice task t, conditional on RUM class k, is now simply given by the well known Multinomial Logit (MNL) formula as:

$$P_{\mathsf{RUM},n,t,i,k} = \frac{e^{V_{n,t,i,k}}}{\sum_{i=1}^{3} e^{V_{n,t,i,k}}} \tag{11}$$

The paradigm used in the second set of classes is that of regret minimisation. The fundamental assumption in regret theory is that final utility depends not merely on the realised outcome but also on what could have been obtained by selecting a different course of action. This means that the model incorporates anticipated feelings of regret (rejoice) that would be experienced once ex-post decision outcomes are revealed to be "unfavourable" ("favourable"). The value of an alternative can thus only be assigned following a cross-wise evaluation of alternatives. The main differences compared to expected utility theory is that regret minimisation does not rely on transitivity and that choice probabilities depend on examination of the full set of alternatives.

Following Chorus (2010), the deterministic regret for alternative i ($i = 1, \dots, 3$) for respondent n in choice task t is given:

$$\begin{aligned} R_{n,t,i,k} = \; & \delta_{\mathsf{RRM},i,k} \\ & + \sum_{j \neq i} \ln\left(1 + e^{\beta_{\mathsf{RRM},TT,k}\left(TT_{n,t,j} - TT_{n,t,i}\right)}\right) \\ & + \sum_{j \neq i} \ln\left(1 + e^{\beta_{\mathsf{RRM},LF,k}\left(\ln(F_{n,t,j}) - \ln(F_{n,t,i})\right)}\right) \end{aligned}$$

$$
\begin{aligned}
&+\sum_{j\neq i}\ln\left(1+e^{\beta_{\text{RRM},RD,k}(RD_{n,t,j}-RD_{n,t,i})}\right)\\
&+\sum_{j\neq i}\ln\left(1+e^{\beta_{\text{RRM},ED,k}(ED_{n,t,j}-ED_{n,t,i})}\right)\\
&+\sum_{j\neq i}\ln\left(1+e^{\beta_{\text{RRM},C,k}(C_{n,t,j}-C_{n,t,i})}\right),
\end{aligned}
\tag{12}
$$

again with $\delta_{\text{RRM},i,k}$ being an alternative specific constant which is set to zero for $i=3$. The regret is informed by all the pairwise comparisons, where regret for alternative i increases whenever an alternative $j\neq i$ performs better than i on a given attribute. Working again under the assumption of type I extreme value errors, the probability of respondent n choosing alternative i in choice task t, conditional on RRM class k, is now simply given by the analogue of a MNL formula as:

$$
P_{\text{RRM},n,t,i,k}=\frac{e^{-R_{n,t,i,k}}}{\sum_{i=1}^{3}e^{-R_{n,t,i,k}}},
\tag{13}
$$

where the negative signs relate to minimising rather than maximising regret.

Our combined model structure now makes use of two layers of latent classes, i.e. replacing Equation 3 by:

$$
\begin{aligned}
LC_n=&\,\pi_{n,\text{RUM}}\left[\varpi_{n,\text{RUM},1}LC_{n,\text{RUM}}\left(\beta_{\text{RUM},1}\right)+\varpi_{n,\text{RUM},2}LC_{n,\text{RUM}}\left(\beta_{\text{RUM},2}\right)\right]\\
&+\pi_{n,\text{RRM}}\left[\varpi_{n,\text{RRM},1}LC_{n,\text{RRM}}\left(\beta_{\text{RRM},1}\right)+\varpi_{n,\text{RRM},2}LC_{n,\text{RRM}}\left(\beta_{\text{RRM},2}\right)\right],
\end{aligned}
\tag{14}
$$

where $LC_{n,\text{RUM}}\left(\beta_{\text{RUM},k}\right)$ is a product of MNL probabilities for the sequence of alternatives chosen by respondent n, using parameters $\beta_{\text{RUM},k}$, with $LC_{n,\text{RRM}}\left(\beta_{\text{RRM},k}\right)$ being a corresponding product of RRM probabilities. In our work, we did not parameterise the class allocation probabilities with socio-demographics, such that we simply have:

$$
\pi_{n,\text{RRM}}=\frac{e^{\delta_{\pi,\text{RRM}}}}{1+e^{\delta_{\pi,\text{RRM}}}}
\tag{15}
$$

$$
\varpi_{n,\text{RUM},1}=\frac{e^{\delta_{\varpi,\text{RUM},1}}}{1+e^{\delta_{\varpi,\text{RUM},1}}}
\tag{16}
$$

$$
\varpi_{n,\text{RRM},1}=\frac{e^{\delta_{\varpi,\text{RRM},1}}}{1+e^{\delta_{\varpi,\text{RRM},1}}}
\tag{17}
$$

with $\delta_{\pi,\text{RRM}}$, $\delta_{\varpi,\text{RUM},1}$ and $\delta_{\varpi,\text{RRM},1}$ being estimated parameters, and with $\pi_{n,\text{RUM}}=1-\pi_{n,\text{RRM}}$, $\varpi_{n,\text{RUM},2}=1-\varpi_{n,\text{RUM},1}$ and $\varpi_{n,\text{RRM},2}=1-\varpi_{n,\text{RRM},1}$.

This completes the specification for the simple latent class model. In the hybrid model, we additionally make use of a latent variable α_n, where we did not include a deterministic component within Equation 6 owing to a desire to not confound the drivers of decision rule heterogeneity with heterogeneity caused by socio-demographic factors. In the presence of the latent variable α_n, the class allocation probabilities at the paradigm level now become:

$$
\pi_{n,\text{RRM}}=\frac{e^{\delta_{\pi,\text{RRM}}+\tau_{RRM}\alpha_n}}{1+e^{\delta_{\pi,\text{RRM}+\tau_{RRM}\alpha_n}}},
\tag{18}
$$

while still having that $\pi_{n,\text{RUM}} = 1 - \pi_{n,\text{RRM}}$.

As outlined in the theory section, this latent variable is now also used to explain the values for a number of indicators in the measurement model component of the hybrid structure, where the indicators used in our work relate to a respondent's level of satisfaction with their real world commute, in particular how well a respondent's current commute journey lines up with their aspirations. The stated disparities between ideal/acceptable and current values are measurable indicators of underlying discontent or regret feelings. Our specific hypothesis is that a respondent who is more likely to be driven by regret minimisation is less likely to have settled on a current commute journey which performs poorly against their desired values on one or more key characteristics. In doing this, we hypothesise that travellers have a certain amount of influence over their commute journey and over time align it with their aspirations. The approach used in this paper is to test whether the size of the gap between ideal or acceptable and current values is related to the predisposition to use a regret-minimising decision rule. This is based on an underlying assumption that the stated gap between ideal/aceptable outcomes and current experiences gives a quantifiable measure of satisfaction with the respondent's real world commute journey.

The fact that the answers to the questions on acceptable and ideal values were captured after the stated choice questions has two key implications. Firstly, there is no way that us asking questions about these aspirations can have any impact on the behaviour in the stated choice survey, i.e. drawing a respondent's attention to their current situation will not have any influence on their stated choices, as the choices have already been made. Secondly, we are not arguing that someone adopts a specific type of behaviour in the context of our stated choice survey, but rather, that someone is or is not a regret-minimizer as a character trait - i.e. this is constant for a given person. This is then the motivation for linking their responses as to the satisfaction with their current commute to this underlying character trait, where our argument is that someone who is more likely to be a regret minimiser is less likely to have chosen in real life a commute journey which performs poorly compared to aspiration levels. Small disparities between aspirations and reality are thus a manifestation of a regret minimising personality rather than a trigger of regret minimising behaviour in the stated choice task.

It should again be acknowledged that the actual emotions which underlie the answers to the follow-up questions could just as likely be manifestations of disappointment or frustration, rather than regret per se. So while the indicators do not necessarily refer directly to regret, all these emotions have a negativity about them which is likely to manifest itself more in a regret than a utility concept, and thus respondents with smaller disparities are arguably more likely to exhibit regret minimising traits than utility maximising traits.

We use eight separate indicators measuring the difference between the acceptable (respectively ideal) values for the four core attributes and the corresponding value for the reference commute. For fare, we again worked with the logarithm of fare, and all eight indicators were zero centred so as to avoid the need to estimate a constant in the measurement equation. This thus gives us:

$$I_{n,1} = TT_{n,1} - TT_{n,ideal} - \sum_{n=1}^{N} \frac{TT_{n,1} - TT_{n,ideal}}{N}$$

$$I_{n,2} = TT_{n,1} - TT_{n,acc} - \sum_{n=1}^{N} \frac{TT_{n,1} - TT_{n,acc}}{N}$$

$$
\begin{aligned}
I_{n,3} &= \ln\left(F_{n,1}\right) - \ln\left(F_{n,ideal}\right) - \sum_{n=1}^{N} \frac{\ln\left(F_{n,1}\right) - \ln\left(F_{n,ideal}\right)}{N} \\
I_{n,4} &= \ln\left(F_{n,1}\right) - \ln\left(F_{n,acc}\right) - \sum_{n=1}^{N} \frac{\ln\left(F_{n,1}\right) - \ln\left(F_{n,acc}\right)}{N} \\
I_{n,5} &= RD_{n,1} - RD_{n,ideal} - \sum_{n=1}^{N} \frac{RD_{n,1} - RD_{n,ideal}}{N} \\
I_{n,6} &= RD_{n,1} - RD_{n,acc} - \sum_{n=1}^{N} \frac{RD_{n,1} - RD_{n,acc}}{N} \\
I_{n,7} &= C_{n,1} - C_{n,ideal} - \sum_{n=1}^{N} \frac{C_{n,1} - C_{n,ideal}}{N} \\
I_{n,8} &= C_{n,1} - C_{n,acc} - \sum_{n=1}^{N} \frac{C_{n,1} - C_{n,acc}}{N}
\end{aligned} \tag{19}
$$

These eight indicators were then used in the model using a continuous measurement equation, as in Equation 8, without the estimation of the now redundant constant. An illustration of the combined model structure is presented in Figure 2.

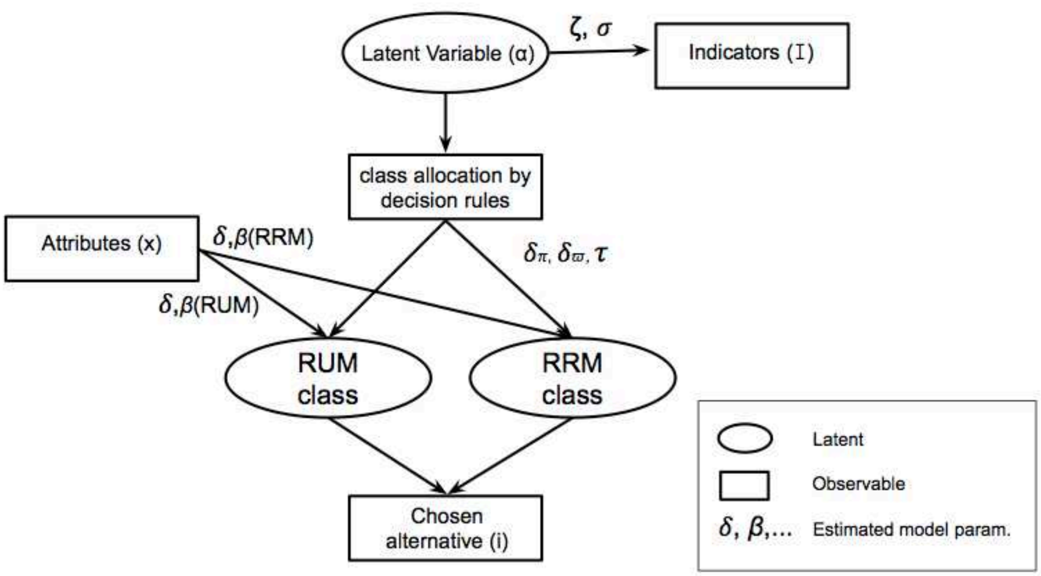

Figure 2: Full structure

4 Results

Table 2 gives an overview of the different models estimated, including the base models for which detailed results are not reported here. We can see that, in line with the majority of past work, RUM and RRM produce very similar model fit when estimated on their own, while a significant improvement is obtained in the simple latent class mixture incorporating the two paradigms (LC with $K_s = 1, \forall s$). Further significant improvements in fit are obtained when allowing for additional heterogeneity at the paradigm level (LC with $K_s = 2, \forall s$). Finally, the model fit from the hybrid structure cannot be directly compared to that of the other models given that it relates to the likelihood of both the choices and the indicator values. Furthermore, as discussed by Vij and Walker (2012), such model fit comparisons, even when factoring out the component of the likelihood relating to choices alone, are not insightful.

We next turn to the estimation results for the choice model components of the simple latent class structure (with $K_s = 2, \forall s$) and the hybrid equivalent in Table 3, where for ease of presentation, the subclasses for each paradigm are ordered such that the *larger* class goes first. We focus first on the results in terms of class allocation probabilities. Both models suggest a relatively even split between RUM and RRM, with a slightly larger probability for RRM in the simple mixture (51.13%) than in the hybrid model (47.18%). These results are not surprising given the very similar fit for the two base models. The RRM probabilities are higher than reported in Hess et al. (2012) on the same data, possibly due to the use of the 2008 variant of the RRM model in that work, which only uses comparisons with the best foregone alternative. In the paradigm specific classes, we see a split into two rather evenly sized groups for RRM in both models, while the split for RUM is slightly less balanced, especially in the base model without the latent variable component. Finally, in the hybrid model, the presence of α_n in the class allocation probabilities in Equation 18 leads to variation in the probabilities for classes across respondents as a function of the latent variable. Specifically, we see that we obtain a wide 95% confidence interval for the probability for the RRM class, going from 16.77% to 79.85%, where this is then also reflected in the probabilities for the RUM class.

In addition to estimates and t-ratios, Table 3 also gives ratios against the log-fare coefficients for parameters in each class, multiplied by 10 to make them applicable to a base fare of $\pounds 10$. While these equate to willingness-to-pay measures, it should be noted that this interpretation does not apply in the RRM models. Turning to the detailed estimation results, we first focus on the two RUM classes. We note that for $k = 1$, i.e. the first class, the relative sensitivities to travel time, the log of fare and the rate of delays is similar in both models, where $\beta_{\text{RUM},RD,1}$ is not statistically significant in either model (and positive in the base model). The relative importance of crowding is very similar in both models, where the significance in the hybrid model is however lower for $\beta_{\text{RUM},C,1}$. For $\beta_{\text{RUM},ED,1}$, we see a big drop in significance in the hybrid model, along with a smaller relative sensitivity (by a factor of about three). The differences are much more marked in the second class, i.e. for $k = 2$. We observed a much higher fare sensitivity in this class in the base model, to the point where the results would suggest that this class primarily captures respondents with very high cost sensitivities and little importance for the remaining attributes, also reflected in the insignificant positive estimate for $\beta_{\text{RUM},ED,2}$ in the base model. When comparing the results across the two classes, there is in fact a suggestion that both classes in the base model primarily capture respondents with high fare sensitivity. This is an initial observation of potentially higher risk of confounding between decision rule heterogeneity and taste heterogeneity in the base model.

Turning next to the RRM classes, we note that the relative importance of the fare attribute is more similar across the two classes in the base model, while the hybrid model clearly points

Table 2: Model fit summary

	Observations			
	choices	indicators	par	log-likelihood
RUM	3,680	0	7	-3,401.68
RRM	3,680	0	7	-3,402.59
LC ($K_s = 1, \forall s$)	3,680	0	15	-3,171.61
LC ($K_s = 2, \forall s$)	3,680	0	31	-3,025.64
Hybrid ($K_s = 2, \forall s$)	3,680	2,944$^{\$}$	48	-9,533.25

$^{\$}$: 8 indicators per respondent

towards one class ($k = 2$) capturing respondents with higher fare sensitivity. The fact that such a clear segmentation is absent in the base model, together with the earlier observations of both RUM classes in the base model capturing high fare sensitivity suggests that the base model may indeed be subject to more confounding. Looking in more detail at the two RRM classes, we see very similar estimates in both models for $k = 1$. The second class on the other hand is less comparable, even after accounting for the higher overall importance of fare in the hybrid model with $k = 2$. Finally, it is worth paying some attention to the constant for the first alternative which always uses the attribute levels for the reference journey. We see that overall, these constants have more importance in the RRM classes than in the RUM classes. A significant negative value for the constants for the first alternative would imply a higher rate of choosing that alternative (given the minimisation of regret) and this would be consistent with individuals having aligned their real world commute with underlying preferences, in line with our hypothesis. This is the case with $k = 1$ for both models. With $k = 2$, the constants in the hybrid model are no longer significant, where this is in line with this class being driven mainly by fare sensitivity. The positive estimate for $\delta_{RRM,1,2}$ in the base model is more difficult to explain and further suggests that the patterns of heterogeneity retrieved by the base model, which does not include the latent variable, are driven by other factors.

We can also see from Table 3 that the latent variable α_n has a significant influence on the class allocation probabilities, with a more positive value for α_n leading to a higher probability for the RRM class (positive sign of τ_{RRM} in Equation 18). The effect is strong enough to lead to a wide 95% confidence interval for the RRM share, as already discussed above. At the same time, the latent variable α_n explains the values of the eight indicator variables. Remembering that for each indicator, a higher value equates to a greater positive difference between the concerned attribute's performance for the reference commute trip and the ideal or acceptable trip, the negative signs we see for each ζ parameter in Table 4 suggests that a respondent with a more positive latent variable is less likely to have a current commute journey which is substantially worse than the ideal or acceptable trip on any of the four main characteristics. With the exception of reliability, the effect of the latent variable is stronger for the difference to acceptable than ideal. It should be acknowledged that the model is somewhat more successful in making a link with the stated satisfaction with current crowding and reliability conditions than with travel time and fare, and this is in part a reflection of more variability across reference trips in these measures. Overall however, the results give clear empirical support to our hypothesis that a respondent who has an

Table 3: Estimation results for choice model components and implied class allocation probabilities

RUM: choice model component

	base model			hybrid		
	est	rob. t-rat	$\frac{10\beta.}{\beta_{LF}}$	est	rob. t-rat	$\frac{10\beta.}{\beta_{LF}}$
$\delta_{\text{RUM},1,1}$	0.5345	3.32	-0.47	0.3462	1.91	-0.41
$\delta_{\text{RUM},2,1}$	-0.1014	-0.67	0.09	-0.0264	-0.18	0.03
$\beta_{\text{RUM},TT,1}$	-0.1475	-6.44	0.13	-0.1256	-7.30	0.15
$\beta_{\text{RUM},LF,1}$	-11.4640	-11.82	-	-8.5000	-5.71	-
$\beta_{\text{RUM},RD,1}$	0.0981	0.94	-0.09	-0.0907	-1.07	0.11
$\beta_{\text{RUM},ED,1}$	-0.3699	-4.53	0.32	-0.0852	-1.64	0.10
$\beta_{\text{RUM},C,1}$	-0.2653	-2.85	0.23	-0.1692	-1.91	0.20
$\delta_{\text{RUM},1,2}$	0.4063	1.90	-0.12	-0.3114	-1.30	2.38
$\delta_{\text{RUM},2,2}$	0.5492	1.83	-0.17	0.3820	2.83	-2.92
$\beta_{\text{RUM},TT,2}$	-0.0917	-4.12	0.03	-0.0299	-2.69	0.23
$\beta_{\text{RUM},LF,2}$	-32.9400	-4.05	-	-1.3103	-2.84	-
$\beta_{\text{RUM},RD,2}$	-0.5718	-2.40	0.17	-0.3550	-4.29	2.71
$\beta_{\text{RUM},ED,2}$	0.0253	0.51	-0.01	-0.1736	-3.11	1.32
$\beta_{\text{RUM},C,2}$	-0.5787	-2.05	0.18	-0.3365	-4.07	2.57
$\delta_{\pi,\text{RUM}}$	0	-	-	0	-	-
$\delta_{\varpi,\text{RUM},1}$	0.6274	1.97	-	0.3980	1.30	-
τ_{RUM}	-		-	0	-	-

RRM: choice model component

	base model			hybrid		
	est	rob. t-rat	$\frac{10\beta.}{\beta_{LF}}$	est	rob. t-rat	$\frac{10\beta.}{\beta_{LF}}$
$\delta_{RRM,1,1}$	-1.5751	-8.16	5.80	-1.6072	-8.30	5.78
$\delta_{RRM,2,1}$	-0.4866	-2.55	1.79	-0.4481	-2.42	1.61
$\beta_{RRM,TT,1}$	-0.0280	-3.49	0.10	-0.0276	-3.22	0.10
$\beta_{RRM,LF,1}$	-2.7138	-2.76	-	-2.7822	-2.97	-
$\beta_{RRM,RD,1}$	-0.2062	-3.05	0.76	-0.1872	-2.68	0.67
$\beta_{RRM,ED,1}$	-0.0577	-2.49	0.21	-0.0891	-1.91	0.32
$\beta_{RRM,C,1}$	-0.2238	-2.98	0.82	-0.2207	-3.24	0.79
$\delta_{RRM,1,2}$	0.3702	1.09	-3.57	-0.1386	-0.75	0.07
$\delta_{RRM,2,2}$	-0.3633	-2.87	3.50	-0.4404	-1.52	0.22
$\beta_{RRM,TT,2}$	-0.0201	-3.06	0.19	-0.0562	-4.89	0.03
$\beta_{RRM,LF,2}$	-1.0381	-3.47	-	-20.0590	-5.83	-
$\beta_{RRM,RD,2}$	-0.2367	-4.31	2.28	-0.1938	-1.60	0.10
$\beta_{RRM,ED,2}$	-0.0384	-1.40	0.37	-0.0682	-0.46	0.03
$\beta_{RRM,C,2}$	-0.1559	-3.15	1.50	-0.2747	-3.36	0.14
$\delta_{\pi,RRM}$	0.0450	0.27	-	-0.1127	-0.37	-
$\delta_{\varpi,RRM,1}$	0.0703	0.16	-	0.1361	0.49	-
τ_{RRM}	-		-	0.7600	2.57	-

	class allocation probabilities within rules				overall					
	RUM-A	RUM-B	RRM-A	RRM-B	RUM-A	RUM-B	RRM-A	RRM-B	RUM	RRM
Base model	65.19%	34.81%	51.76%	48.24%	31.86%	17.01%	26.46%	24.66%	48.87%	51.13%
Hybrid model: 5^{th} percentile point	-	-	-	-	12.05%	8.10%	8.95%	7.81%	20.15%	16.77%
Hybrid model: mean	59.82%	40.18%	53.40%	46.60%	31.59%	21.22%	25.20%	21.99%	52.82%	47.18%
Hybrid model: 95^{th} percentile point	-	-	-	-	49.79%	33.44%	42.64%	37.21%	83.23%	79.85%

Table 4: Estimation results for measurement model component of hybrid structure

par	est	rob. t-rat
$\zeta_{\text{time ideal}}$	-3.0804	-2.41
$\sigma_{\text{time ideal}}$	15.176	8.71
$\zeta_{\text{time acceptable}}$	-3.7674	-3.15
$\sigma_{\text{time acceptable}}$	13.134	8.48
$\zeta_{\text{log-fare ideal}}$	-0.05814	-2.12
$\sigma_{\text{log-fare ideal}}$	0.39272	10.56
$\zeta_{\text{log-fare acceptable}}$	-0.07216	-2.00
$\sigma_{\text{log-fare acceptable}}$	0.4282	6.70
$\zeta_{\text{crowding ideal}}$	-0.94657	-5.33
$\sigma_{\text{crowding ideal}}$	2.4254	19.96
$\zeta_{\text{crowding acceptable}}$	-0.94933	-5.03
$\sigma_{\text{crowding acceptable}}$	2.2044	19.51
$\zeta_{\text{reliability ideal}}$	-2.2996	-14.80
$\sigma_{\text{reliability ideal}}$	1.0781	12.49
$\zeta_{\text{reliability acceptable}}$	-2.2985	-13.64
$\sigma_{\text{reliability acceptable}}$	0.85885	10.70

underlying character trait which makes him/her more likely to make choices that can be better explained by regret minimisation than utility maximisation is also more likely to have chosen in real life a commute journey that is more closely aligned with his/her aspirations.

5 Discussion

This paper has added to a growing body of work that recognises that within a sample population, different decision paradigms may be better suited for one individual than for another. In contrast with other existing work, we have moved away from a purely random treatment of such decision rule heterogeneity or a treatment linking it to observed respondent characteristics. Rather, we have focussed on attempts to link the behaviour in the stated choice survey to an underlying character trait which will make some respondents more likely to make choices that can be explained by a specific decision rule. In order to identify the role of such latent components, the proposed hybrid model makes use of additional information from outside a stated choice context.

Our empirical application has focussed in particular on the case of contrasting random utility maximisation with random regret minimisation. Our results show that a link can be made between the likely decision rule for a given respondent in the stated choice scenarios and that respondent's stated satisfaction with the real world performance of their current commute journey with regard to their declared aspirational outcomes. The hypothesis is that both outcomes (stated choices and stated satisfaction with real world choices) are influenced by deep rooted character traits. In particular, findings point towards a link between the tendency for regret minimisation and the effective minimisation of disparity with desired trip features for a respondent's real world commute journey. In line with the observation that regret is linked to the motive of moving away from undesired outcomes (c.f. Crowe and Higgins, 1997), the most regret-prone respondents in our sample have, to a larger extent, aligned their reference trip performance to their aspirational values.

The use of such additional indicator variables is a key characteristic of hybrid choice models and specifically the fact that these are treated as dependent rather than explanatory variables. In the present context, this means that the latent variable is used to jointly explain the heterogeneity in class allocation probabilities and the values of the additional indicator measures. This is in contrast with simply using the indicators as explanatory variables within the class allocation probabilities, i.e. replacing α_n by $f(I_n)$ in Equation 18. As with all hybrid structures, this has the advantage of avoiding the risk of endogeneity bias, making the model suitable for forecasting and also accommodating measurement error in the indicators (cf. Abou-Zeid and Ben-Akiva, 2014). In addition however, the causality link is very clear in our specific context. Indeed, the measures relating to satisfaction with the real life commute are meant to relate to the outcome of real world choice processes that are driven by the same character traits that also influence the choice processes in the stated choice component. This is different from an assumption that the satisfaction with real life commute journeys influences the choice of decision rule in the hypothetical choice scenarios.

The approach presented in this paper permits analysts to gain further insights into behavioural patterns and the process by which decision rules may be adopted. Much work remains to be done, including testing the framework on other data or on other decision rules. Other than the deeper insights into decision processes, a potential advantage suggested for the hybrid model in the empirical work is a reduced risk of confounding between decision rule heterogeneity and simple heterogeneity in sensitivities. This would arise as any implied heterogeneity in decision rule also needs to be consistent with the measurement model component of the hybrid structure.

There is ample scope for future work in this area. A key issue remains the choice of appropriate indicators for the measurement component of the model, and here the onus is on analysts to make appropriate decisions at the survey design stage. The use of a richer set of indicators also opens up possibilities of using multiple latent variables that relate to different character traits. Even with the data used here, other possibilities would have arisen, such as using additional indicators that focus on the differences between ideal and acceptable values. There of course also remain possibilities of linking the choice of decision rules to the values of presented alternatives, but this moves us away from the notion that the likely decision rule is influenced in particular by underlying character traits.

We have also paid limited attention to the actual implications of the results for the different model components, focussing instead on testing our underlying hypothesis. The field is only just starting to explore the actual benefits of allowing for alternative decision rules, for example looking at implications in terms of forecasting performance. Furthermore, stated choice data is not well suited to computing elasticities, while there is still no clarity in terms of how monetary valuations, which would provide another means of comparison, might be obtained from RRM models.

Acknowledgements

The authors are grateful for the comments of four anonymous referees which have helped to significantly improve the paper.

References

Abou-Zeid, M., Ben-Akiva, M., 2014. Hybrid choice models. In: Hess, S., Daly, A. (Eds.), Handbook of Choice Modelling. Edwar.

Abou-Zeid, M., Ben-Akiva, M., Bierlaire, M., Choudhury, C., Hess, S., 2010. Attitudes and Value of Time Heterogeneity. In: Van de Voorde, E., Vanelslander, T. (Eds.), Applied Transport Economics - A Management and Policy Perspective. De Boeck Publishing, pp. 523–548.

Anable, J., Jan. 2005. 'complacent car addicts' or 'aspiring environmentalists'? identifying travel behaviour segments using attitude theory. Transport Policy 12 (1), 65–78.

Bekker-Grob, E., Chorus, C., 10.1007/s40273-013-0059-0 2013. Random regret-based discrete-choice modelling: An application to healthcare. PharmacoEconomics 31 (7), 1–12.

Ben-Akiva, M., McFadden, D., Gärling, T., Gopinath, D., Walker, J., Bolduc, D., Börsch-Supan, A., Delquié, P., Larichev, O., Morikawa, T., Polydoropoulou, A., Rao, V., 1999. Extended framework for modeling choice behavior. Marketing Letters 10 (3), 187–203.

Boeri, M., Longo, A., Doherty, E., Hynes, S., 2012. Site choices in recreational demand: a matter of utility maximization or regret minimization? Journal of Environmental Economics and Policy 1 (1), 32–47.

Bolduc, D., Boucher, N., Alvarez-Daziano, R., Dec. 2008. Hybrid choice modeling of new technologies for car choice in canada. Transportation Research Record: Journal of the Transportation Research Board 2082, 63–71.

ChoiceMetrics, 2012. Ngene 1.1.1 ngene 1.1.1. user manual and reference guide.

Chorus, C., 2010. A new model of random regret minimization. European Journal of Transport and Infrastructure Research 10 (2), 181–196.

Chorus, C., Arentze, T., Timmermans, H., 2008. A random regret minimization model of travel choice. Transportation Research Part B 42 (1), 1–18.

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

Chorus, C., Rose, J., 2011. Selecting a date: a matter of regret and compromises. In: Hess, S., Daly, A. (Eds.), Choice modelling: The state of the art and the state of practice,. Edward Elgar, Ch. 11.

Chorus, C. G., 2012. Random regret minimization: An overview of model properties and empirical evidence. Transport Reviews: A Transnational Transdisciplinary Journal 32 (1), 75–92.

Chorus, C. G., Arentze, T. A., Timmermans, H. J. P., 2009. Spatial choice: a matter of utility or regret? Environment and Planning B: Planning and Design 36 (3), 538 – 551.

Chorus, C. G., Rose, J. M. J. M., Hensher, D. A., 2013. Regret minimization or utility maximization: it depends on the attribute. Environment and Planning B: Planning and Design 40 (1), 154 – 169.

Choudhury, C., Ben-Akiva, M., Abou-Zeid, M., 2010. Dynamic latent plan models. Journal of Choice Modelling 3 (2), 50–70.

Crowe, E., Higgins, E., 1997. Regulatory focus and strategic inclinations: Promotion and prevention in decision-making. Organizational Behavior and Human Decision Processes 69 (2), 117 – 132.

Daly, A. J., Hess, S., Patruni, B., Potoglou, D., Rohr, C., 2012. Using ordered attitudinal indicators in a latent variable choice model: A study of the impact of security on rail travel behaviour. Transportation 39 (2), 267–297.

Doornik, J. A., 2001. Ox: An Object-Oriented Matrix Language. Timberlake Consultants Press, London.

Hensher, D. A., Greene, W. H., Chorus, C. G., 2011. Random regret minimization or random utility maximization: an exploratory analysis in the context of automobile fuel choice. Journal of Advanced Transportation available online, 10.1002/atr.188.

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

Redmond, L., Mokhtarian, P., 2001. The positive utility of the commute: Modeling ideal commute time and relative desired commute amount. Transportation 28 (2), 179–205.

Stathopoulos, A., Hess, S., 2012. Revisiting reference point formation, gains–losses asymmetry and non-linear sensitivities with an emphasis on attribute specific treatment. Transportation Research Part A: Policy and Practice 46 (10), 1673 – 1689.

Swait, J., 2001. A non-compensatory choice model incorporating attribute cutoffs. Transportation Research Part B: Methodological 35 (10), 903–928.

Thiene, M., Boeri, M., Chorus, C., 2012. Random regret minimization: Exploration of a new choice model for environmental and resource economics. Environmental and Resource Economics 51 (3), 413–429.

Vij, A., Walker, J., 2012. Hybrid choice models: Holy grail... or not? paper presented at the 13^{th} International Conference on Travel Behaviour Research, Toronto.