EJTIR

Issue 10(2)
June 2010
pp. 181-196
ISSN: 1567-7141
www.ejtir.tbm.tudelft.nl

A New Model of Random Regret Minimization

Caspar G. Chorus[1]
Section of Transport and Logistics, Delft University of Technology

A new choice model is derived, rooted in the framework of Random Regret Minimization (RRM). The proposed model postulates that when choosing, people anticipate and aim to minimize regret. Whereas previous regret-based discrete choice-models assume that regret is experienced with respect to only the best of foregone alternatives, the proposed model assumes that regret is potentially experienced with respect to each foregone alternative that performs well. In contrast with earlier regret-based discrete-choice approaches, this model can be estimated using readily available discrete-choice software packages. The proposed model is contrasted theoretically and empirically with its natural counterpart, Random Utility Maximization's linear-additive MNL-model. Empirical comparisons on four revealed and stated travel choice datasets show a promising performance of the RRM-model.

Keywords: Discrete Choice Analysis; Random Regret Minimization; semi-compensatory choice; choice set-specific preferences

1. Introduction

Anticipated regret is considered an important determinant of choice-behavior in a variety of disciplines, including marketing (e.g. Simonson, 1992; Zeelenberg & Pieters, 2007), microeconomics (e.g. Loomes & Sugden, 1982; Sarver, 2008), psychology (e.g. Zeelenberg, 1999; Connolly, 2005), the management sciences (e.g. Savage, 1954; Bell, 1982) and transportation (e.g., Chorus et al., 2006a, b). Simply put, regret is what you experience when a foregone alternative performs better than the chosen one, and regret-based choice-theories and -models are built around the notion that individuals minimize anticipated regret – rather than maximizing anticipated utility – when choosing. Recently, the idea that minimizing anticipated regret determines choices has been translated into a generic approach for modeling discrete (travel) choice-behavior: this Random Regret Minimization-approach (RRM, Chorus et al. (2008)) is developed for the econometric analysis of risky as well as riskless choices in multinomial and multi-attribute contexts. It allows for the estimation, based on observed choices, of parameters reflecting decision-makers' valuation of alternatives and their attributes. RRM's contribution to discrete choice-modeling lies in its ability to capture semi-compensatory behavior (that is: RRM postulates that improvement of one attribute of an alternative not necessarily offsets an equally large decline in performance of another attribute) as well as choice set-specific preferences, within a model that is as tractable and parsimonious as RUM's linear-additive MNL-model.

[1] Jaffalaan 5, 2628 BX, Delft, The Netherlands, T: +31152788546, F: +31152782719, E: c.g.chorus@tudelft.nl

However, the RRM-model specification proposed in Chorus et al. (2008) faces two important limitations: first, it postulates that regret is only experienced and anticipated with respect to the best of foregone alternatives, whereas intuition would suggest that the presence of additional foregone alternatives that perform better than the considered one, though worse than the best of foregone alternatives, adds to the regret anticipated by the decision-maker. Second, the model specification's likelihood function is non-smooth, which creates difficulties with respect to the derivation of marginal effects and elasticities, and triggers a need for customized optimization routines to successfully estimate the model. This means that model estimation relies on handwritten code, which in turn hampers the model's general applicability, especially among practitioners.

This paper derives a new model form, rooted in the RRM-approach, that alleviates these two limitations: first, regret is postulated to be anticipated with respect to all foregone alternatives that perform better than a considered one in terms of one or more attributes. Second, the proposed model form features a smooth likelihood-function and can be estimated using standard discrete-choice software packages. The result is a behaviorally intuitive RRM-model, which is easy to implement and estimate.

In addition to presenting this new model, the paper provides a theoretical and empirical comparison with its natural counterpart, the linear-additive MultiNomial Logit-model (MNL) rooted in the paradigm of Random Utility Maximization (McFadden, 1974). Empirically, RRM is compared with RUM's MNL-model in terms of its model fit and predictive validity on four travel choice-datasets involving mode-route choice, destination choice, parking choice and the acquisition of Public Transport information. The focus of this paper is on riskless, multinomial choices.

The new RRM-model is derived in Section 2. Section 3 presents a theoretical comparison with RUM-based linear-additive MNL-models. Section 4 presents the empirical comparisons between the proposed RRM-model and MNL. Conclusions and directions for further research are provided in Section 5, which also presents a discussion of the RRM-model's applicability for various types of travel choice analysis.

2. A New RRM-model

Assume the following choice situation: a decision-maker faces a set of *J* travel alternatives, each being described in terms of *M* attributes x_m that are comparable across alternatives. The focus in the remainder of this Section, is on predicting the choice probability for an alternative *i* from this set. Before introducing the new RRM-model, note as a reference point that a conventional, linear-additive utilitarian specification would assign the following deterministic utility to alternative *i*: $V_i = \sum_{m=1..M} \beta_m x_{im}$. Adopting the classical Random Utility-Maximization (RUM) paradigm (that is: adding iid Extreme Value Type I-distributed errors to the deterministic utilities of all alternatives to represent heterogeneity in unobserved utility) implies the following MNL-formulation of the resulting choice probability (McFadden, 1974): $P_i = \exp\left(V_i\right) \Big/ \sum_{j=1..J} \exp\left(V_j\right)$.

2.1 A new model of Random Regret Minimization[2]

The new RRM-model postulates that when choosing between alternatives, decision-makers aim to minimize anticipated random regret, and that the level of anticipated random regret that is

[2] See Appendix 1 for a presentation of the RRM-approach as proposed in Chorus et al. (2008).

associated with the considered alternative i is composed out of an iid random error ε_i which represents unobserved heterogeneity in regret and is Extreme Value Type I-distributed, and a systematic regret R_i.

Systematic regret is in turn conceived to be sum of all so-called binary regrets that are associated with bilaterally comparing the considered alternative with each of the other alternatives in the choice set: $R_i = \sum_{j \neq i} R_{i \leftrightarrow j}$. As such, the new RRM-model postulates that each foregone alternative that performs better than the considered alternative adds to the regret anticipated by the decision-maker.[3]

The level of binary regret associated with comparing the considered alternative with another alternative j is conceived to be the sum of the regrets that are associated with comparing the two alternatives in terms of each of their M attributes[4]: $R_{i \leftrightarrow j} = \sum_{m=1..M} R^m_{i \leftrightarrow j}$. This attribute level-regret in turn either equals zero (in case the considered alternative performs better than the other alternative on that particular attribute), or it equals the weighted difference in attribute-performance. Attribute weights β_m represent tastes, and their size and sign are estimable. Two iid errors $\upsilon_{0m}, \upsilon_{xm}$ (Extreme Value Type I-distributed) are added to represent unobserved heterogeneity in attribute perceptions and weights: $R^m_{i \leftrightarrow j} = \max\left\{0 + \upsilon_{0m}, \beta_m \cdot \left(x_{jm} - x_{im}\right) + \upsilon_{xm}\right\}$.

As an illustration of the crucial notion of attribute level-regret, take for example the situation where the considered attribute m represents price and ignore for the moment random errors: when lower prices are preferred over higher ones (i.e. the estimated price-parameter is negative), and the price of alternative j is lower than that of alternative i, comparing them in terms of price implies a strictly positive regret associated with alternative i. That is, $\max\left\{0, \beta_m \cdot \left(x_{jm} - x_{im}\right)\right\} > 0$. The larger the price-difference, the larger the associated regret. If alternative j has a higher price than alternative i, there is no regret and the magnitude of the difference in favor of j does not matter in that case. It is crucial to note here again that the sign of the betas is estimated together with their magnitude. Also random parameters can be accommodated this way, allowing for instance for the possibility that part of the population may evaluate a particular attribute positively, and another part negatively.

To arrive at choice probabilities, errors υ need to be 'integrated out' first. Taking as an example the comparison between alternatives i and j in terms of attribute m, the integration of attribute level-regret $R^m_{i \leftrightarrow j}$ over $f\left(\upsilon_{0m}, \upsilon_{xm}\right)$ results in the following Logsum-formulation for the expected maximum (Ben-Akiva and Lerman, 1985):

[3] Of course, one may also postulate that the regret associated with the best of foregone alternatives receives a higher weight than the regret associated with other foregone alternatives, e.g. by means of adding estimable weights to the additive regret-function. However, this approach would increase the number of parameters consumed, which runs against the implicit aim of presenting a regret-based choice model that is as parsimonious as RUM's linear-additive MNL-model. Furthermore, empirical identification of these weights is likely to create difficulties, since they are confounded with the attribute weights. Finally, it should be noted that the regret associated with foregone alternatives that perform relatively well is by definition already larger than the regret associated with foregone alternatives that perform less well. In a sense, adding weights would imply a double counting of regret (hence the foreseen difficulties in terms of empirical identification).

[4] Constants can be incorporated by means of parameters associated with dummy-variables that represent the total of unobserved characteristics of an alternative. Alternatively, constants may be simply added to the regret-function.

$$\int_{\upsilon}\left[R_{i\leftrightarrow j}^{m}\cdot f\left(\upsilon\right)\right]d\upsilon=\ln\left(1+\exp\left[\beta_{m}\cdot\left(x_{jm}-x_{im}\right)\right]\right) \quad (1)$$

Figure 1 (where "A" stands for $\beta_m \cdot \left(x_{jm} - x_{im}\right)$) illustrates how the use of the Logsum smoothens binary regret. Whereas the RRM-specification without errors υ (like the one in Chorus et al. (2008)) is discontinuous at zero, the Logsum-approach results in a continuous function for every value of A.

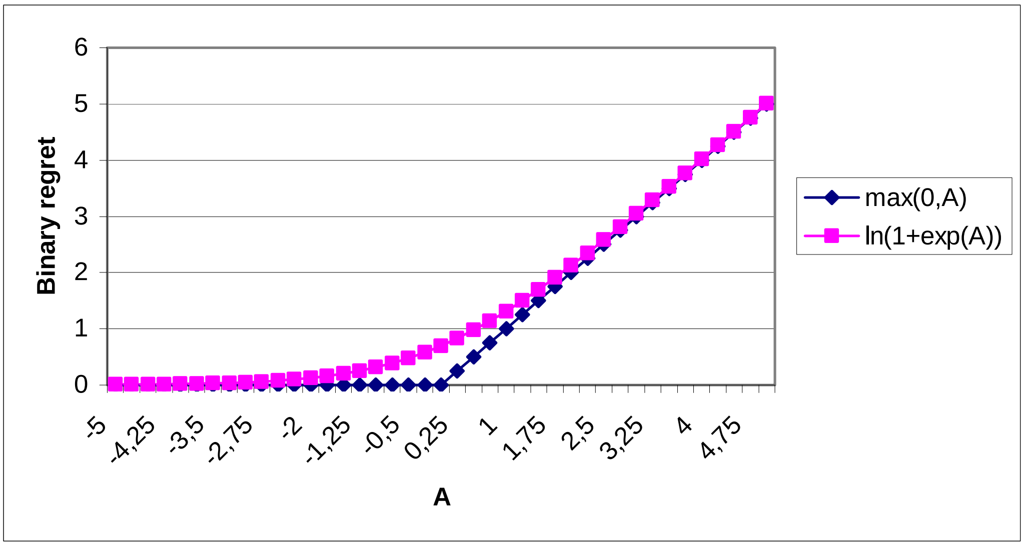


Figure 1. Using the Logsum to smoothen binary regret

Systematic regret R_i can now be rewritten into the following formulation (the ~-sign in $\tilde{R}_i$ reflects that errors υ are integrated out):

$$\tilde{R}_i = \sum_{j\neq i}\sum_{m=1..M}\ln\left(1+\exp\left[\beta_m\cdot\left(x_{jm}-x_{im}\right)\right]\right) \quad (2)$$

Acknowledging that minimization of random regret is mathematically equivalent to maximizing the negative of random regret, choice probabilities may be derived using a variant of the well known multinomial logit-formulation: the choice probability associated with alternative *i* equals:

$$P_i = \exp\left(-\tilde{R}_i\right) \Big/ \sum_{j=1..J}\exp\left(-\tilde{R}_j\right) \quad (3)$$

Note that the resulting likelihood function is smooth and that the model can be coded and estimated using standard discrete choice-software packages. The correspondence of the proposed RRM-model with the linear-additive MNL-model is striking: apart from the fact that both result in logit-choice probabilities, both models are equally parsimonious. Each parameter estimated for a RRM-model has a counterpart in a linear-additive MNL-model[5]. In this light it is interesting to

[5] Note that in this paper, the focus is on parameters associated with the attributes (*x*) of alternatives, not with characteristics of the decision-maker (sociodemographics (*z*)). However, RRM can deal with sociodemographics and associated variables γ by allowing for a combined utility-regret function, where sociodemographic variables enter the utilitarian part of the function, and the attributes the regret-part. Choice probabilities then take the

note that when choice sets are binary, the proposed RRM-model –like its predecessor (see Chorus et al. (2009))– reduces to RUM's linear-additive binary logit-model (see Appendix 2 for a derivation of this result). Triggered by this correspondence between the new RRM-model and the linear-additive MNL-model in terms of their formal model form, the following two Sections provide theoretical and empirical comparisons between the two approaches. As will become clear throughout these comparisons, the two models differ in terms of predicted choice probabilities and parameter estimates in interesting ways.

3. Theoretical Comparisons with RUM

Before illustrating some of the most eye-catching differences between RUM's linear-additive MNL-model and the new RRM-model using numerical examples, this section starts by formally highlighting their most important difference: in contrast with RUM's MNL-model and in spite of the use of iid errors – the new RRM-model does not exhibit the IIA-property. To see this, it suffices to inspect the ratio of choice probabilities for alternatives i and j from a choice set containing alternatives i, j and k:

$$\frac{P_i}{P_j} = \frac{\exp\left(-\tilde{R}_i\right) \Big/ \left(\exp\left(-\tilde{R}_i\right) + \exp\left(-\tilde{R}_j\right) + \exp\left(-\tilde{R}_k\right)\right)}{\exp\left(-\tilde{R}_j\right) \Big/ \left(\exp\left(-\tilde{R}_i\right) + \exp\left(-\tilde{R}_j\right) + \exp\left(-\tilde{R}_k\right)\right)} = \frac{\exp\left(-\tilde{R}_i\right)}{\exp\left(-\tilde{R}_j\right)} =$$

$$\frac{\exp\left(-\sum_{m=1..M} \ln\left(1+\exp\left[\beta_m \cdot \left(x_{jm}-x_{im}\right)\right]\right) - \sum_{m=1..M} \ln\left(1+\exp\left[\beta_m \cdot \left(x_{km}-x_{im}\right)\right]\right)\right)}{\exp\left(-\sum_{m=1..M} \ln\left(1+\exp\left[\beta_m \cdot \left(x_{im}-x_{jm}\right)\right]\right) - \sum_{m=1..M} \ln\left(1+\exp\left[\beta_m \cdot \left(x_{km}-x_{jm}\right)\right]\right)\right)} \tag{4}$$

Clearly, alternative k's attributes x_{km} enter the choice probability ratio, which therefore becomes dependent on k. As a result, the popularity of i, relative to j, depends on both their performances relative to one another and alternative k. In other words, in contrast with RUM's linear-additive MNL-model, its RRM-counterpart postulates that preferences for alternatives are choice-set specific.

The remainder of this section illustrates by means of numerical examples how this difference between the two models translates into differences in predicted choice probabilities. Assume the following arbitrary choice situation: a decision-maker faces a choice between alternatives i, j and k which are fully described in terms of two attributes, x and y. Both attributes are equally important (i.e., both are assigned unit weights) and higher values are preferred over lower ones[6]. Alternative i is defined as {1,2}, while j is defined as {2,1}. Alternative k is defined as $\{x_k, y_k\}$, which values vary between zero and three. In words: i (j) has a poor performance in terms of x (y) and a strong one in terms of y (x), while k's performance on either attribute is varied from very poor to very strong. Figure 2 (left panel) plots the choice probability for k, as a function of x_k and y_k, in the context of a linear-additive MNL-model (red) and the new RRM-model (blue). Figure 2 (right

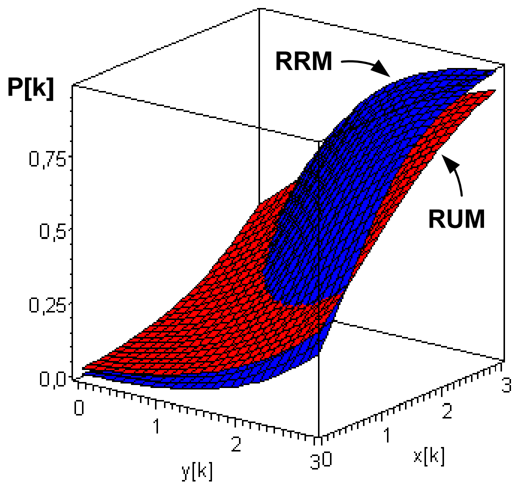
following form (e.g., for the case of one sociodemographic variable, denoting the decision-maker by n): $P_i^n = \exp\left(\gamma_i^n z^n - \tilde{R}_i\right) \Big/ \sum_{j=1..J} \exp\left(\gamma_j^n z^n - \tilde{R}_j\right)$.

[6] Note that the assumption that attribute-weights (parameters) are the same for RUM- and RRM-models turns out to be somewhat restrictive in light of the empirical evidence reported in the next Section. This should be kept in mind when interpreting the numerical example.

panel) provides the same plot, holding x_k fixed at 1.5. Four observations can be made, based on these plots.

Firstly, it appears that RRM penalizes a relatively poor performance of *k* heavier than does RUM: the lower left corner of the left panel shows lower RRM-choice probabilities than RUM-choice probabilities for alternative *k*. Secondly, it appears that RRM rewards a relatively strong performance of *k* heavier than does RUM. Thirdly, as can be most easily seen on the right panel, the difference between RRM and RUM is larger in terms of rewarding good performance of *k* than it is in terms of penalizing poor performance.

Fourthly, moving from the edges of the graphs to the center, it appears that RRM assigns higher choice probabilities to alternative *k* than does RUM, when *k* is positioned right in between the remaining two alternatives in terms of attribute-performance (i.e., when both x_k and y_k lie in the [1,2]-interval). This accommodation of *k* as an 'in between' alternative is in line with the so-called compromise effect, which is empirically well-documented in a range of choice-situations involving choice from, for example, apartments, political manifestos, mouth washes and investment portfolios (e.g. Simonson, 1989; Wernerfelt, 1995; Kivetz et al., 2004) but has not yet received attention in travel choice research. The compromise effect states that alternatives with an 'in-between' performance on all attributes, relative to the other alternatives in the choice set, are generally favored by choice-makers over alternatives with a poor performance on some attributes and a strong performance on others. See Appendix 3 for a more detailed discussion of how RRM captures the compromise effect.



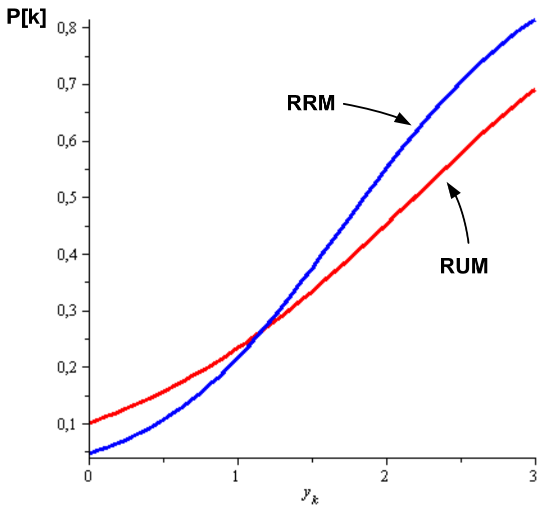


Figure 2. Comparing choice probabilities generated by RRM and RUM

Whereas this Section focused on conceptual differences between RRM and RUM, the next Section studies empirical differences in terms of estimation results and model performance in the context of four stated- and revealed-choice datasets. It is expected that to some extent, conceptual differences highlighted here will vanish due to the process of fitting the two models on the same data. On the other hand, some differences are expected to result in different empirical outcomes, such as the asymmetry between RRM's penalization of inferiority and its rewarding of superiority, and RRM's accommodation of compromise alternatives.

However, note that obtained empirical results will not be analyzed explicitly in terms of how they might relate to the conceptual differences between RRM and RUM that have been discussed here. Firstly, as shown in the previous section's formal analyses, the most striking conceptual difference between the two models is that RRM's choice probability predictions are highly choice set-specific. This makes it difficult to formulate empirical expectations for other than the most trivial choice situations (trinomial choice sets with two attributes per alternative). Related to this, it is intrinsically difficult to trace back empirical differences between RRM and RUM to conceptual ones, in the absence of data specifically collected with the aim of empirically identifying these rather subtle conceptual differences. Finally, and this is likely to be related to the absence of data collected with the aim of discriminating between RRM and RUM's MNL, found empirical differences between RRM and RUM are small on the used datasets. This hampers a reliable statistical assessment of where these differences stem from in terms of the conceptual differences between the two modeling approaches. The conclusions Section discusses what type of data, to be collected in future research efforts, may facilitate a reliable and meaningful interpretation of empirical differences between RRM and RUM in light of their conceptual differences.

Before moving on to the empirical analyses, it is worth mentioning that (and: why) RRM is not well suited for dealing with choice sets containing a 'no choice' option. First, note that since RRM relies on attribute-by-attribute comparisons, it has little to say about how products are being evaluated that have no, or only very few, attributes in common. As a result, RRM is likely to encounter conceptual difficulties when choice behavior is modeled in the presence of a no-choice option (which has by definition no attributes in common with any of the other alternatives). In a RUM-setting, a traveler's choice for a no-choice option can be conceived as meaning to say that the best of the choice options was perceived as not being good enough. Obviously, this reasoning does not hold in a RRM-context, since a choice set consisting of high-performing alternatives can generate the same level of regret as a choice set with poor performing alternatives: it is the alternatives' performance *relative* to their competitors that generates regret. As a result, the meaning of a no-choice option and the associated estimated constant in an RRM model is ambiguous. Throughout the next Section, it has been taken care of that the datasets used for empirical analysis do not contain a 'no-choice' option.

4. Empirical Comparisons with RUM

In this Section, estimation results on four datasets are reported. It is worth mentioning at this point that these were the only four datasets on which the new RRM-model has been estimated and compared with the linear-additive MNL-model. Whenever multiple observations were available per individual, robust standard errors are reported. For reasons of space limitations, and in line with the scope of this paper, details about the data collection efforts, as well as about the interpretation of parameter estimates, are left out whenever possible.

4.1 Travel mode-/route-choices

Choice-data were collected using a multimodal travel simulator (Chorus et al., 2007). Thirty-one participants to the simulator-based choice experiment were asked to perform trips with one of a range of car- and train-options. They had the opportunity to acquire all sorts of travel information before and during the trip. It may be assumed that participants, when faced with a choice situation, went through a two-step process, involving a decision whether or not to acquire information, and involving a choice between information options (in case it was decided to acquire information) or a choice between travel options (in case it was decided not to acquire information). The analyses that are presented below focus on those 416 cases where a) more than two travel alternatives were available[7], and b) a travel alternative was chosen (instead of an

[7] Since, in a binary context, RRM reduces to RUM (see Appendix 2).

information option). In line with the above described two-step decision process, it is assumed that the choice set faced by the individual in these cases contained travel alternatives only (and no 'no choice' option in the form of travel information acquisition). Two thirds of the data (280 cases) were used for estimation, the remaining 136 cases served as hold out sets.

Travel alternatives are characterized in terms of the following attributes: expected in-vehicle travel time (TRAVEL_TIME_EXP, minutes), travel time uncertainty[8] (TRAVEL_TIME_UNC, minutes), expected travel costs (TRAVEL_COSTS_EXP, Euros), travel costs uncertainty (TRAVEL_COSTS_UNC, Euros), expected waiting time (WAITING_TIME_EXP, minutes), waiting time uncertainty (WAITING_TIME_UNC, minutes), expected seat availability (SEAT_EXP, 0 = no seat available, 0.5 = 50% chance that a seat is available, 1 = seat available), seat availability uncertainty (SEAT_UNC, 0 = no uncertainty about availability, 1 = uncertainty about availability). In line with intuition, car-alternatives were assumed to have zero expected waiting time (and zero waiting time uncertainty), and that a seat was always available for the car option (SEAT_EXP = 1; SEAT_UNC = 0). Finally, a dummy-attribute reflected the nature of the travel mode (CAR_DUMMY, 0 = train, 1 = car). Initially, this dummy was allowed to vary randomly across individuals, allowing for the possibility of panel and nesting effects. However, the parameter's standard deviation appeared insignificant at any reasonable level, so the model was re-estimated with a non-random dummy. Table 1 presents estimation results[9].

Table 1. Travel mode-/route-choices (estimation results)

	RUM		RRM	
	Parameter	*t*-Statistic	Parameter	*t*-Statistic
CAR_DUMMY	-0.6626	-1.603	-0.1865	-0.901
TRAVEL_TIME_EXP	-0.1265	-7.140	-0.0725	-7.216
TRAVEL_TIME_UNC	-0.0348	-3.250	-0.0218	-3.367
TRAVEL_COSTS_EXP	-0.5670	-7.607	-0.2888	-7.988
TRAVEL_COSTS_UNC	-0.1479	-2.505	-0.0760	-2.485
WAITING_TIME_EXP	-0.1345	-2.794	-0.0596	-3.055
WAITING_TIME_UNC	0.0405	1.260	0.0177	0.962
SEAT_EXP	1.2333	3.055	0.6045	3.356
SEAT_UNC	0.0977	0.215	0.1622	0.595
0-Log-Likelihood	-372		-372	
Log-likelihood at convergence	-211		-209	
rho-square	0.433		0.440	
AIC	441		435	
BIC	474		468	
Number of cases	280		280	

[8] Uncertainty regarding in-vehicle and waiting times and travel costs was communicated to participants in terms of certainty intervals, i.e., ranges of time and costs within which the participants were told (correctly) that actual values would almost certainly fall. The lengths of these certainty intervals are taken to be proxies of the uncertainty associated with travel times, waiting times and travel costs, respectively.

[9] When inspecting estimation results it should be kept in mind that the interpretation of parameters differs between RUM- and RRM-models, although in both paradigms, parameter signs reflect whether or not higher values of an attribute are preferred over lower ones, and absolute magnitude reflects the importance of the attributes. Where a RUM-parameter gives the effect of a unit increase in attribute value on utility, a RRM-parameter gives the effect of a unit increase in attribute value- *differences* (across alternatives) to regret.

All significant parameters are of the expected sign. In terms of the comparison between RRM and RUM's linear-additive MNL-model, it appears that the RRM-model fits the data slightly better than its utilitarian counterpart. Significance levels of parameters are fairly similar, although estimates themselves differ: on the average, RRM-estimates of significant parameters are about twice as close to zero as their RUM-counterparts. Relative importance of attributes differ modestly across models. RRM also slightly outperforms RUM in terms of out-of-sample predictions as measured in terms of mean predicted choice probabilities for chosen alternatives (0.564 versus 0.558).

4.2 Travel information acquisition choices

Stated Choice data was collected relating to travelers' willingness to pay for advanced Public Transport information (Molin et al., 2009). In total, 204 individuals were interviewed while riding a train. They were faced with nine choice tasks containing three alternative travel information types and a 'none of these' option (resulting in 1836 observed choices). In line with the expectation that the RRM-model would experience conceptual difficulties when dealing with the no-choice option (see Section 3 for a discussion of this issue), all 1463 cases were selected where one of the three information alternatives was chosen. Like was the case in 4.1, it was assumed that respondents go through a two-step process, involving a decision whether or not to acquire information, and involving a choice between information options. In line with this reasoning, it is assumed here that the choice set faced by the individual in the 1463 cases contained the three information alternatives only. Again, two thirds of the data (1000 cases) were used for estimation, the remaining 463 cases served as hold out sets.

Table 2. Travel information acquisition choices (estimation results)

Variables	RUM		RRM	
	Parameter	*t*-Statistic	Parameter	*t*-Statistic
TIMES+SEARCH	0.2821	2.537	0.3706	2.502
TIMES+ADVICE	0.6487	5.733	0.8452	5.997
INFO_INIT	0.1465	1.407	0.2350	1.849
BOTH_INIT	0.2376	1.930	0.3721	2.500
UNRELIABILITY	-0.0824	-3.294	-0.0738	-2.839
PRICE	-0.8806	-18.663	-0.9670	-19.392
0-Log-Likelihood	-1109		-1109	
Log-likelihood at convergence	-818		-822	
rho-square	0.262		0.259	
AIC	1650		1657	
BIC	1685		1692	
Number of cases	1000		1000	

Four attributes were varied in the Stated Choice task. First, type of provided information could take on the levels TIMES (only travel time information is provided – this level serves as a base level), TIMES+SEARCH (in addition to travel time information, the service also provides an option to search for alternative routes), TIMES+ADVICE (in addition to travel times, advice on the best route is also provided). Second, it was varied who takes the initiative to get information: TRAVELER_INIT (only the traveler can take the initiative to acquire information – this level serves as a base level), INFO_INIT (only the service can take the initiative to provide information) and BOTH_INIT (both can take the initiative). Third, the UNRELIABILITY represented the number of minutes maximum deviation between informed travel times and true travel times (levels: 0, 2.5 and 5). Fourth, PRICE per message could take on the levels 0, 0.15 and 0.30 euro. Table 2 shows estimation results of both RUM- and RRM-models.

All significant parameters have a plausible sign. In terms of the comparison between RRM and RUM's linear-additive MNL, it appears that the RUM-model fits the data marginally better than its regret-based counterpart. Significance levels of parameters are fairly similar, although estimates themselves differ: RRM-estimates of significant parameters are, on the average, about 50% larger than their RUM-counterparts. An exception is the valuation of information unreliability: measured in terms of Euros, the RRM-model suggests a valuation of unreliability that is 40% more negatively than its utilitarian counterpart. Note that in terms of out-of-sample predictions, RRM marginally outperforms RUM in terms of mean predicted choice probabilities for chosen alternatives (0.537 versus 0.535).

4.3 Parking lot choices

A revealed parking choice dataset has been collected by van der Waerden et al. (2008) at the campus of Eindhoven University of Technology. In total, 517 useable interviews from car-drivers were obtained relating to, among other things, their most recent parking choice from the set of 14 available parking lots at the university campus. Table 3 presents the best parking choice model for the RRM and RUM models (note that while a full-constant model was estimated as well, it delivered no significant increases in model fit for both RUM as well as RRM). Again, two thirds of the data (350 cases) were used for estimation, the remaining 167 cases served as hold out sets.

NR_SPACES refers to the actual number of spaces available at a particular parking lot (ranging from 30 to 298), ROOM_MANEUV refers to the availability of extra space for making maneuvers (dummy coded, 1 = yes). RIGHT_OF_WAY refers to whether or not one has right-of-way when leaving the parking lot (dummy coded, 1 = yes), and DISTANCE refers to the distance to and from one's workplace from the parking lot (1 = approximately 100 meters, 2 = approximately 300 meters and 3 = approximately 500 meters).

Table 3. Parking lot choices (estimation results)

Variables	RUM		RRM	
	Parameter	*t*-Statistic	Parameter	*t*-Statistic
NR_SPACES	0.5612	6.137	0.0867	6.076
ROOM_MANEUV	0.6527	3.515	0.0907	3.298
RIGHT_OF_WAY	0.2028	1.060	0.0339	1.228
DISTANCE	-5.8073	-5.854	-1.4439	-3.201
0-Log-Likelihood	-924		-924	
Log-likelihood at convergence	-406		-405	
rho-square	0.561		0.562	
AIC	819		817	
BIC	835		833	
Number of cases	350		350	

Again, significant parameters have the expected sign, significance levels being roughly equal across models. RRM-estimates are about seven times smaller than their RUM-counterparts, with the exception of the parameter relating to distance: this attribute plays a much more important role in the RRM-model than in RUM's linear-additive MNL-model. A very small goodness-of-fit difference is observed, in favor of RRM. Both models score equally well in terms of out-of-sample predictions (mean predicted choice probabilities for chosen alternatives equal 0.232).

4.4 Shopping location choices

A dataset was collected concerning shopping center-choices (Arentze et al., 2005), containing 1503 revealed choices from consumers that went shopping with the aim of buying groceries. Variables included were FLOOR_GROCERIES (m^2 floorspace concerning groceries in a particular center), FLOOR_OTHER (m^2 floorspace concerning other items) and DISTANCE (seconds, taking into account the used travel mode). Given the very large number of available stores in the study-area (Noord-Brabant province, The Netherlands), choice sets were imputed: next to the chosen alternative, four other stores were selected based on shortest travel time calculations. Again, two thirds of the data (1000 cases) were used for estimation, the remaining 503 cases serving as hold out cases. Table 4 shows estimation results for RUM- and RRM-models.

Table 4. Shopping location choices (estimation results)

Variables	RUM		RRM	
	Parameter	*t*-Statistic	Parameter	*t*-Statistic
FLOOR_GROCERIES	0.1153	5.578	0.0752	5.717
FLOOR_OTHER	0.0148	5.328	0.0045	3.280
DISTANCE	-0.0492	-7.646	-0.0169	-7.053
0-Log-Likelihood	-1609		-1609	
Log-likelihood at convergence	-1514		-1510	
rho-square	0.060		0.062	
AIC	3033		3027	
BIC	3048		3042	
Number of cases	1000		1000	

Parameter estimates have the expected sign, significance levels again being roughly equal across models. While RRM marginally outperforms RUM's linear-additive MNL-model, rather low rho-squares suggest that besides the three variables used for estimation, a number of other variables play an important role as well. RRM-estimates are closer to zero than those provided by the RUM-model, and relative importance of variables differs quite substantially between RRM and RUM. For example, m^2 floorspace concerning groceries plays a twice as important role, relative to the other attributes, in the context of RRM than it does in RUM. RRM also slightly outperforms RUM's linear-additive MNL-model in terms of mean predicted choice probabilities for chosen alternatives in the hold out sample (0.227 versus 0.224).

4.5 Summarizing empirical results

In terms of model fit, RRM outperforms RUM on three out of four datasets (RUM outperforming RRM only on the travel information acquisition dataset). Differences, however, are marginal. In terms of out-of-sample predictions, measured in terms of mean predicted choice probabilities for chosen alternatives, RRM also outperforms RUM on three out of four datasets, including the travel information acquisition dataset (both models score equally well on the parking lot choice dataset). Again, found differences are small.[10]

In sum, it seems fair to draw two preliminary conclusions from these empirical analyses: first, it appears that the differences between RUM's linear-additive MNL-model and the new RRM-model in terms of model fit and predictive performance are small indeed. In light of the fact that both models are equally parsimonious, this is no big surprise. However, note that relative importance of attributes (and hence, practical implications derived from estimation results) can differ more substantially across models. As a second preliminary conclusion, it appears that, out of eight performance measurements (four regarding model fit and four regarding predictive validity), the new RRM-model (slightly) outperforms RUM's linear-additive MNL-model on six occasions, while the latter (slightly) outperforms RRM on only one occasion – the remaining measurement resulting in a draw. All in all, this provides an indication of the potential of the new RRM-model for discrete choice analysis.

5. Conclusions and directions for further research

This paper proposed a new discrete choice model, rooted in the Random Regret Minimization-approach. Like the initially proposed RRM-specification (Chorus et al., 2008), the proposed model postulates that when choosing, people anticipate and aim to minimize regret with respect to foregone alternatives. However, whereas the initially proposed RRM-model assumed that regret is only experienced with respect to the best of foregone alternatives, the proposed model assumes that regret is experienced with respect to all alternatives that perform better than a considered alternative in terms of one or more attributes. Furthermore, unlike the initially proposed RRM-specification, the proposed model features a smooth likelihood-function and allows for being estimated using standard discrete-choice software packages. In this paper, this new RRM-model is contrasted conceptually with RUM's workhorse, the linear-additive MNL-model. Empirical comparisons between the two approaches on four revealed- and stated-choice datasets show a promising performance of the new RRM-model.

Given that RRM is a new kid on the block, the most obvious direction for further research lies in further empirical comparisons between RRM and established models like RUM's linear (Mixed) MNL. In general, these comparisons should preferably be based on data that allow for an unambiguous interpretation of estimation results in light of conceptual differences between the two approaches. In this light, three sorts of data-collection efforts seem particularly promising when it comes to comparing RRM's potential with that of its utilitarian counterparts.

First, whereas this paper has focused on riskless choice, the notion of regret is of course quite compatible with risky choice as well (see for example Chorus et al., 2008). It would be interesting to see how the new RRM-model performs, relative to other (RUM-based) models, in such risky choice contexts.

Second, a host of literature has identified choice situations and contexts that induce feelings of (anticipated) regret. For example, Zeelenberg and Pieters (2007) find that "regret is experienced when decisions are difficult and important and when the decision maker expects to learn the outcomes of both the chosen and rejected options quickly", and that "regret is anticipated when significant others in the decision maker's social network view the decision as important". In these situations, one would expect the RRM-approach to perform particularly strong, a hypothesis that can be tested empirically. Projected on a travel choice environment, it might be hypothesized that, for example, lane-changing choices ("decision maker expects to learn the outcomes of both

[10] Note that the RRM-specification presented in Chorus et al. (2008) was outperformed by both the new formulation as well as RUM's linear-additive MNL-model on almost all occasions.

the chosen and rejected options quickly"), travel choices under emergency evacuation-circumstances ("decisions are difficult and important") and car-type choices ("significant others in the decision maker's social network view the decision as important") are likely to trigger choice behavior governed by anticipated regret-minimization.

Third, given that one of the more salient characteristics of RRM lies in its potential to capture compromise effects, a fruitful research avenue would be to study RRM's performance in the context of choice situations where compromise effects have been found to play an important role. Simonson (1989) identifies a number of situations where compromise effects are potentially important, including situations where subjects expect to justify their decisions to others, and situations where elaborate, difficult decisions are to be made. Interestingly, but in line with the theoretical findings derived in Section 3, these conditions appear to be similar to the ones identified by Zeelenberg and Pieters (2007) as triggers of anticipated regret. Of course, in order to identify compromise effects and as a result optimally discriminate between the RRM-model and competitors, choice situations like described directly above should preferably contain (outspoken) compromise alternatives. One type of choice-situation in the transportation domain where compromise alternatives exist almost by definition is that of multiattribute ordered choices such as involving car-ownership levels: owning one car is more (less) expensive than owning no car at all (two cars), but delivers more (less) convenience. In this respect, owning one car is a compromise alternative positioned in between owning no car at all, or owning two cars. However, also in many other travel choice contexts, compromise alternatives are likely to exist (and they can of course be created artificially using Stated Choice-experimental design techniques).

Although empirical comparisons like presented above are likely to present interesting results, another fruitful avenue for further research would be to embed regret-minimization in a latent class model where decision strategies (regret minimization, utility maximization and others) vary across classes and where class membership is a function of socio-demographics and contextual variables. Such a model would enable the identification of different segments in the population, and different choice contexts, in terms of associated types of choice behavior. See Hensher (2009) for a recent example of analyses along these lines.

Finally, note that the fact that the data used in this paper was not collected with the specific aim of identifying differences between RRM and RUM's linear-additive MNL-model, and did not refer to choice situations known to trigger anticipated regret or compromise effects, can be considered a limitation and a strength at the same time. A limitation, because the data did not allow for the *ex ante* derivation of specific hypotheses, nor for a meaningful *ex post* interpretation of empirical results in light of identified conceptual differences between RRM and RUM's linear-additive MNL-model. A strength, because the promising empirical performance of RRM on these data suggests that it may in fact be a relatively robust model of choice-behavior, rather than being a model that performs well only within particular, strictly defined, perimeters.

Acknowledgements

We would like to thank Theo Arentze, Aloys Borgers, Eric Molin, Harmen Oppewal, Ruud van Sloten, Harry Timmermans and Peter van der Waerden for providing data used for empirical analyses.

References

Arentze, T.A., Oppewal, H. and Timmermans, H.J.P. (2005). A multipurpose shopping trip model to assess retail agglomeration effects. *Journal of Marketing Research*, Vol. 42, No. 1, pp. 109-115.

Bell, D.E. (1982). Regret in decision making under uncertainty. *Operations Research*, Vol. 30, No. 5, pp. 961-981.

Ben-Akiva, M. and Lerman, S.R. (1985). *Discrete choice analysis: theory and application to travel demand.* The MIT Press, Cambridge, Mass.

Chorus, C.G., Arentze, T.A., Molin, E.J.E., Timmermans, H.J.P. and van Wee, G.P. (2006a). The value of travel information: Decision-strategy specific conceptualizations and numerical examples. *Transportation Research Part B*, Vol. 40, No. 6, pp. 504-519.

Chorus, C.G., Molin, E.J.E., van Wee, G.P., Arentze, T.A. and Timmermans, H.J.P. (2006b). Responses to transit information among car-drivers: Regret-based models and simulations. *Transportation Planning and Technology,* Vol. 29, No. 4, pp. 249-271.

Chorus, C.G., Molin, E.J.E., Arentze, T.A., Hoogendoorn, S.P., Timmermans, H.J.P. and van Wee, G.P. (2007). Validation of a multimodal travel simulator with travel information provision. *Transportation Research Part C*, Vol. 15, No. 3, pp. 191-207.

Chorus, C.G., Arentze, T.A. and Timmermans, H.J.P. (2008). A Random Regret Minimization model of travel choice. *Transportation Research Part B*, 42(1), pp. 1-18.

Chorus, C.G., Arentze, T.A. and Timmermans, H.J.P. (2009). Spatial choice: A matter of utility or regret? *Environment and Planning Part B*, Vol. 36, No. 3, pp. 538-551.

Connolly, T. and Reb, J. (2005). Regret in cancer-related decisions. *Health Psychology*, Vol. 24, No. 4, pp. 29-34.

Hensher, D.A. (2009). Attribute processing, heuristics, and preference construction in choice analysis. *Invited plenary paper for the First International Conference on Choice Analysis, Harrogate,* UK, March 29 - April 3, 2009.

Kivetz, R., Netzer, O. and Srinivasan, V. (2004). Alternative models for capturing the compromise effect. *Journal of Marketing Research*, Vol. 41, pp. 237-257.

Loomes, G. and Sugden, R. (1982). Regret-Theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, Vol. 92, No. 368, pp. 805-824.

Molin, E.J.E., Chorus, C.G. and van Sloten, R. (2009). The need for advanced Public Transport information services when making transfers. *European Journal of Transport and Infrastructure Research*, Vol. 9, No. 4, pp. 397-410.

Savage, Leonard J. (1954). *The Foundations of Statistics.* New York, Wiley.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, Vol. 19, pp. 158-174.

Simonson, I. (1992). The influence of anticipating regret and responsibility on purchasing decisions. *Journal of Consumer Research*, Vol. 19, No. 1, pp. 105-119.

Starver, T. (2008). Anticipating regret: Why fewer options may be better. *Econometrica*, Vol. 76, No. 2, pp. 263-305.

Van der Waerden, P., Borgers, A. and Timmermans, H.J.P. (2008). Modeling parking choice behavior in business areas. Paper presented at the *87th annual meeting of the Transportation Research Board*, Washington D.C.

Wernerfelt, B. (1995). A rational reconstruction of the compromise effect: Using market data to infer utilities. *Journal of Consumer Research*, Vol. 21, No. 4, pp. 627-633.

Zeelenberg, M. (1999). The use of crying over spilled milk: A note on the rationality and functionality of regret. *Philosophical Psychology*, Vol. 12, No. 3, pp. 325-340.

Zeelenberg, M. and Pieters, R. (2007). A theory of regret regulation 1.0. *Journal of Consumer Psychology*, Vol. 17, No. 1, pp. 3-18.

Appendix 1: The Chorus et al. (2008) RRM-model specification

This RRM-specification postulates that deterministic regret equals the *maximum* of all so-called binary regrets that are associated with bilaterally comparing the considered alternative with each of the remaining alternatives. The level of binary regret associated with comparing the considered alternative with another alternative is conceived to be the sum of the regrets that are associated with comparing the two alternatives in terms of each of their attributes. This attribute level-regret either equals zero (in case the considered alternative performs better than the other alternative on that particular attribute), or it equals the weighed difference in attribute-performance. The weights are represented by parameters that are estimated in the process of model calibration. Note that in contrast with the model proposed in this paper, the (2008) approach does not contain random errors at the level of attribute-regret (nor the associated Logsum-formulation). In notation, deterministic regret associated with alternative i is written as $R_i = \max_{j \neq i} \left\{ \sum_{m=1..M} \max\left\{0, \beta_m \cdot \left(x_{jm} - x_{im}\right)\right\} \right\}$. Note that the fact that only the largest of binary regrets enters overall deterministic regret R implies that only regret with respect to the best performing of the foregone alternatives counts.

Given iid Extreme Value Type I errors added to deterministic regret, choice probabilities take on the following functional form (for example, referring to the probability that alternative 1 is chosen): $P_i = \dfrac{\exp\left(-R_i\right)}{\sum_{j=1..J} \exp\left(-R_j\right)}$.

Appendix 2: RRM reduces to linear-additive RUM in binary choice-situations

Consider a binary choice set $\{i, j\}$. Choice-probabilities, for example for alternative j, generated by the binary logit-type RUM- and RRM-models can then be written as follows: $P_{RUM}(j) = \dfrac{1}{1+\exp\left[V_i - V_j\right]}$, $P_{RRM}(j) = \dfrac{1}{1+\exp\left[R_j - R_i\right]}$. As a result, to show that RRM reduces to RUM in the context of binary choice sets, it suffices to show that $V_i - V_j = R_j - R_i$. When assuming, without loss of general applicability, that the alternatives are evaluated in terms of only two observed attributes x and y, it is easily seen that $V_i - V_j$ = $\beta_x \cdot x_i + \beta_y \cdot y_i - \beta_x \cdot x_j + \beta_y \cdot y_j = \beta_x \cdot \left(x_i - x_j\right) + \beta_y \cdot \left(y_i - y_j\right)$. Denoting, for reasons of simplicity of notation, $\beta_x \cdot \left(x_i - x_j\right)$ by ΔX and $\beta_y \cdot \left(y_i - y_j\right)$ by ΔY, the following holds:

$$R_j - R_i = \ln\left(1+\exp\left[\Delta X\right]\right) + \ln\left(1+\exp\left[\Delta Y\right]\right) - \ln\left(1+\exp\left[-\Delta X\right]\right) - \ln\left(1+\exp\left[-\Delta Y\right]\right)$$

$$= \ln\left(\frac{\left(1+\exp\left[\Delta X\right]\right)\cdot\left(1+\exp\left[\Delta Y\right]\right)}{\left(1+\exp\left[-\Delta X\right]\right)\cdot\left(1+\exp\left[-\Delta Y\right]\right)}\right) = \ln\left(\frac{\left(1+\exp\left[\Delta X\right]\right)\cdot\left(1+\exp\left[\Delta Y\right]\right)}{\frac{1+\exp\left[\Delta X\right]}{\exp\left[\Delta X\right]}\cdot\frac{1+\exp\left[\Delta Y\right]}{\exp\left[\Delta Y\right]}}\right)$$

$$= \ln\left(\frac{\left(1+\exp\left[\Delta X\right]\right)\cdot\left(1+\exp\left[\Delta Y\right]\right)}{\frac{\left(1+\exp\left[\Delta X\right]\right)\cdot\left(1+\exp\left[\Delta Y\right]\right)}{\exp\left[\Delta X\right]\cdot\exp\left[\Delta Y\right]}}\right) = \ln\left(\exp\left[\Delta X\right]\cdot\exp\left[\Delta Y\right]\right) = \ln\left(\exp\left[\Delta X\right]\right) + \ln\left(\exp\left[\Delta Y\right]\right)$$

$$= \Delta X + \Delta Y = V_i - V_j$$

Appendix 3: RRM and the compromise effect

Some additional analysis can shed more light on RRM's potential to accommodate compromise alternatives. Consider the same situation as depicted in Section 3 (Figure 2). However, consider now that y_k ranges from one to two, and that x_k is defined as follows: $x_k = 3 - y_k$. As a result, x_k ranges from one to two as well (more precisely: from two to one). For example, a decision-maker faces a choice between a congested freeway and a non-congested toll way, where tolls are dynamically adjusted to the prevailing traffic situation. Figure A3 shows the difference in choice probabilities generated for alternative *k* by RUM and RRM as a function of y_k, for varying levels of the relative importance of attributes *x* and *y*.

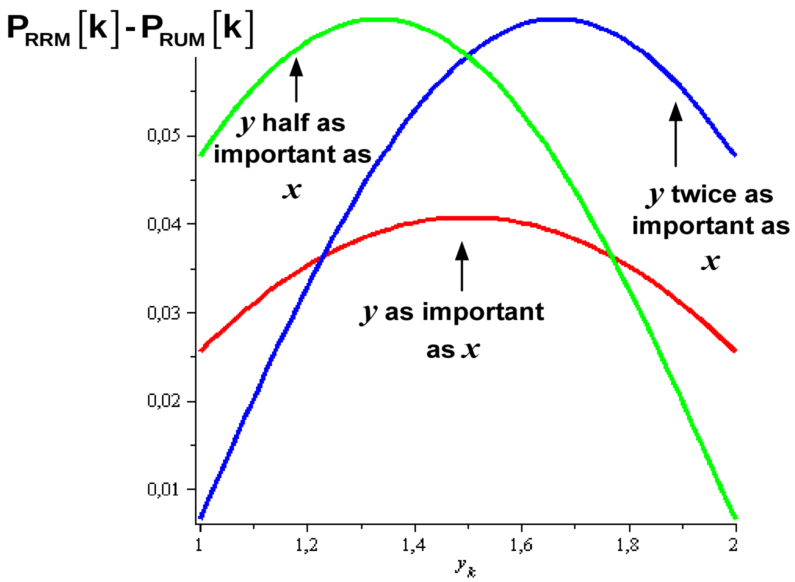


Figure A3. How RRM captures compromise effects

Consider first the situation where both attributes are equally important (red line). In that case, the difference between RRM and RUM in terms of the choice probability associated with *k* is largest when $y_k = x_k = 1.5$ or, in other words, when *k* is positioned exactly in the middle of *i* and *j* in terms of its attribute values. At that point, RRM predicts a choice probability for *k* that is 4 percentage points higher than the choice probability generated by RUM (37% versus 33%). When *y* is twice as important as *x* (blue line), the difference between the two models is largest (more than 6 percentage points) when $y_k = 1.67$, and when *y* is only half as important as *x* (green line), the difference between the two models is largest when $y_k = 1.33$.

In sum, the magnitude of the difference between choice probabilities generated for the compromise alternative by RRM and RUM, while always in favor of the compromise alternative, is largest when the extent to which the compromise alternative is outperformed on one attribute by *i*, equals the extent to which it is outperformed on the other attribute by *j* (i.e. when the compromise alternative's mediocrity is most pronounced). Importantly, as suggested by the three differently located peaks in Figure A3, performance is measured in terms of attribute levels *in combination with* their importance as perceived by the decision-maker. As such, it appears that RRM provides an intuitive and parsimonious (it consumes no more parameters than a linear-additive MNL-model) approach to capture the compromise effect.