Random Regret Minimization as a Choice Model: Theoretical and Empirical Comparisons with RUM-Modelling

Caspar G. Chorus *,a, Theo A. Arentze b, Harry J.P. Timmermans b

a Section of Transport Policy and Logistics Organization, TU Delft
b Urban Planning Group, TU Eindhoven

Abstract
This paper provides an in-depth comparison between choice models based on Random Regret Minimization (RRM) and Random Utility Maximization (RUM). Whereas RUM has been choice-mode ling’s workhorse for decades, RRM has been recently introduced to incorporate the role of negative emotions in decision making, while maintaining high levels of econometric tractability. First, we formally show and illustrate how RRM differs from RUM by accommodating compromise alternatives with a mediocre performance on all attributes. Then, using two datasets concerning revealed parking choices and stated route choices of travellers, we show how this conceptual difference translates into a potential advantage for RRM in terms of model fit when choice situations contain compromise alternatives. Finally, using two additional datasets, we explore the general applicability of RRM. Managerial implications are provided.

Keywords: Random Regret Minimization, Random Utility Maximization, Compromise effect

Submitted for presentation only, at the inaugural International Choice Modelling Conference.

* Tel.: +31152788546; Fax: +31152782719; E-mail address: c.g.chorus@tudelft.nl (C.G. Chorus). P.O. Box 5015, 2600 GA, Delft, The Netherlands
1. Introduction

Regret is considered an important determinant of choice-behaviour 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) and the management sciences (e.g. Savage, 1954; Bell, 1982). Simply put, regret is what you experience when a non-chosen alternative performs better than the chosen one, and regret-based theories and models are built around the notion that individuals minimize anticipated regret – rather than maximizing utility – when choosing.

Recently, the notion of minimizing anticipated regret as a determinant of choice has been translated into a generic discrete-choice formulation: 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\(^1\) and multi-attribute contexts using tractable logit-type probabilities. It allows for a straightforward estimation, based on observed choices, of parameters reflecting decision-makers’ valuation of alternatives and their attributes. This paper provides an elaborate, both theoretical and empirical, comparison of RRM with its natural counterpart: Random Utility Maximization (RUM).

Conceptually, the main difference between the two approaches is that most RUM-models (at least the conventional RUM formulations we consider, such as (Mixed) MNL-models) postulate that the utility of an alternative only depends on its own performance; in contrast, RRM assumes that regret associated with an alternative depends on its performance, at the attribute level, relative to other alternatives in the choice set. The contribution of this paper lies in studying in-depth how this conceptual difference translates into differences in terms of model properties and empirical performance of both approaches. We focus on riskless\(^2\) choices. This paper contributes to Chorus et al. (2008) in particular by a) providing an elaborate theoretical analysis of the underlying reasons for the differences between RRM and RUM in terms of predicting choice probabilities for multinomial choice sets, and b)

---

\(^1\) In fact, RRM reduces to RUM when choices are binary. Chorus et al. (In Press) exploit this property by introducing a concavity-convexity parameter which allows for a nested hypothesis-testing of RRM’s model fit versus that of RUM in the context of binary choices. However, RRM is originally designed for the analysis of multinomial choices, and the properties it displays in a multinomial context (such as accommodation of the compromise effect) are absent in a binary context.

\(^2\) For an application of RRM in the context of risky choices (not involving a comparison with RUM), see Chorus et al. (2008).
presenting an empirical comparison of RRM with RUM on four revealed and stated choice datasets.

Section 2 presents the RRM-choice model. Section 3 formally shows, and illustrates using a fictive example, how and why RRM and RUM differ in terms of predicted choice outcomes: relative to RUM-models, RRM recognizes so-called ‘compromise’ alternatives with mediocre performance on each attribute. Section 4 shows how RRM outperforms RUM, in terms of model fit, on a revealed parking choice dataset. Using a stated route choice dataset, Section 5 provides a more detailed empirical analysis of the differences in goodness-of-fit between RUM and RRM, and how they arise from the conceptual differences between the two approaches. The general applicability of RRM is discussed in Section 6. Section 7 provides conclusions and managerial implications, and discusses avenues for further research.

2. The RRM-model in brief

We here present the RRM-approach to choice modelling. A more elaborate discussion of how RRM relates to other regret-based approaches such as Regret Theory (Bell, 1982; Fishburn, 1982; Loomes & Sugden, 1982) can be found in Chorus et al. (2008). For ease of presentation, we consider a three-alternative, three-attribute choice situation; it is easily seen that formulations extend to other multinomial, multi-attribute choice situations.

Consider an individual who faces a choice between alternatives \( i, j \) and \( k \). Observed attributes are \( x, y, \) and \( z \): \( i = \{x_i, y_i, z_i\}, j = \{x_j, y_j, z_j\} \) and \( k = \{x_k, y_k, z_k\} \). Each alternative has an associated measure of random regret \( RR_i, RR_j, RR_k \) and when choosing, the individual minimizes this random regret. As a consequence, the random regret that is associated with the choice situation \( (RR) \) equals \( \min \{RR_i, RR_j, RR_k\} \). We define alternative-specific random regrets \( RR_i, RR_j, RR_k \) as follows (taking alternative \( i \) as an example):

\[
RR_i = \beta_i^n + \max \left\{ \sum \max \left\{ 0, \beta_i \cdot (v_j - v_i) \right\}, \sum \max \left\{ 0, \beta_i \cdot (v_k - v_i) \right\} \right\} + \epsilon_i \quad (1).
\]
Here, \( v_j - v_i \) (or: \( v_k - v_i \)) represents the difference between \( i \) and \( j \) (or: \( k \)) in terms of attribute \( v \). Attributes are measured in terms of numerical values (e.g. price is measured in dollars or euros). Parameters \( \beta_v \), to be estimated, represent the importance attached to the attributes and reflects – through its sign – whether or not higher attribute values are preferred over lower ones. Then, \( \max\{0, \beta_v \cdot (v_j - v_i)\} \) gives \( R^*_v \), being the regret that is associated with alternative \( i \), when comparing it with \( j \) on attribute \( v \). Regret exists only when \( \beta_v \) and \( v_j - v_i \) have the same sign (note that the sign of \( \beta_v \) is also estimated). Take for example the situation where the considered attribute \( v \) represents price: when lower prices are preferred over higher ones (i.e. the estimated price-parameter \( \beta_v \) is negative), and \( j \)'s price is lower than that of \( i \), comparing \( i \) and \( j \) in terms of price implies a strictly positive regret associated with \( i \). The larger the price-difference, the larger the associated regret. If \( j \) has a higher price than \( i \), there is no regret and – importantly – the magnitude of the difference in favour of \( i \) 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. That is, no a priori knowledge or expectations from the side of the analyst is needed. Also random parameters can be accommodated this way, acknowledging that part of the population may evaluate a certain attribute positively, and another part negatively. Furthermore, note that parameters in a RRM model represent the importance of attributes in terms of individuals’ valuation of differences in attribute-levels, whereas in most RUM models parameters refer to attribute-levels themselves. We here focus on linear-in-parameter specifications.

Total regret associated with alternative \( i \), when comparing it with \( j \) (denoted \( R^*_i \)) consists of the sum of regrets associated with comparisons between the two alternatives across all observed attributes. Total deterministic regret associated with alternative \( i \) (denoted \( R_i \)) equals the maximum of regrets associated with the comparisons with all other alternatives plus a possible constant\(^3\). This formulation implies that only regret with respect to the best performing alternative counts.

We assume that the analyst is unable to faultlessly assess regret for each choice situation. Resulting random regret across alternatives, situations and individuals is captured in iid Extreme Value Type I error terms \( \varepsilon_i, \varepsilon_j, \varepsilon_k \) with variances \( \pi^2/6 \). Since minimization of random regret is mathematically equivalent to maximization of minus random regret, and

\(^3\) In our three-alternative situation, only two constants are identified.
given the symmetrical nature of the iid errors\(^4\), choice probabilities take on the well-known Multi Nomial Logit-form (for example, take the probability that \(i\) is chosen):

\[
P_i = \frac{\exp(-R_i)}{\exp(-R_i) + \exp(-R_j) + \exp(-R_k)} \tag{2}
\]

Here,

\[
R_j = \beta_j^0 + \max \left\{ \sum_{v=x,y,z} \max \{0, \beta_v \cdot (v_j - v_j)\}, \sum_{v=x,y,z} \max \{0, \beta_v \cdot (v_k - v_j)\} \right\}
\]

\[
R_k = \max \left\{ \sum_{v=x,y,z} \max \{0, \beta_v \cdot (v_j - v_j)\}, \sum_{v=x,y,z} \max \{0, \beta_v \cdot (v_j - v_k)\} \right\}
\]

As has been illustrated in Chorus et al. (2008), random parameters \(\beta\) can be easily accommodated, leading to Mixed Logit-choice probabilities. As such, RRM remains tractable when accounting for panel effects, nesting effects and taste heterogeneity.

3. **How RRM differs from RUM in terms of predicted choice-behaviour**

Assume the choice situation described in Section 2 and ignore, for clarity of exposition, attribute \(z\). Make the following additional assumptions: \(\beta_x, \beta_y\) are strictly positive\(^5\) (higher attribute values are preferred over low ones – this assumption is not critical for any of the results obtained here, and is made only for ease of presentation), as are \(x_i - x_k, x_k - x_j, y_j - y_k, y_k - y_i\). In words: alternative \(i\) performs better than all other alternatives on \(x\) and worse than all others on \(y\); the same holds for \(j\) vice versa, while

\(^4\) Note that the use of Extreme Value Type I errors implies that random regret may become negative, a notion that is counterintuitive given our specification of deterministic regret as a non-negative quantity. However, it should be noted that this counterintuitivity is immaterial since choice probabilities and their interpretation are based only on the resulting logarithmically distributed differences in random regret.

\(^5\) Note that we assume here that parameters indicating the importance of attributes are the same in RRM and RUM models. As can be seen further below, empirical analyses show that this may be realistic for some datasets, and less so for others. We make this assumption to facilitate an easily interpretable comparison between the two approaches.
alternative k has a mediocre performance on both attribute x and y. Deterministic regret for each alternative is derived by applying the equations presented in Section 2. Deterministic utility is derived in the classical linear-in-parameters RUM-fashion: \( V_i = \beta_x \cdot x_i + \beta_y \cdot y_i \), \( V_j = \beta_x \cdot x_j + \beta_y \cdot y_j \) and \( V_k = \beta_x \cdot x_k + \beta_y \cdot y_k \). Errors are iid Extreme Value Type I for both RRM and RUM.

The difference between choice probabilities generated by RRM and RUM can be put in words as follows (see the Appendix for proofs): RRM predicts a higher choice probability for the alternative that has a mediocre performance on each attribute, than does RUM. The magnitude of this difference is largest when the extent to which the alternative with mediocre performance (k) is outperformed on one attribute by i, equals the extent to which it is outperformed on the other attribute by j (i.e. when alternative k’s mediocrity is most pronounced). Formally, the above can be denoted as follows: \( P_{k_{RRM}} > P_{k_{RUM}} \), and \( P_{k_{RRM}} - P_{k_{RUM}} \) is largest when \( \beta_x \cdot (x_i - x_k) = \beta_y \cdot (y_j - y_k) \). In sum, it appears that the key difference between RRM and RUM originates from RRM’s recognition of the compromise effect, whose existence has been firmly established in a series of empirical studies in the context of consumer choice (e.g. Simonson, 1989; Wernerfelt, 1995; Kivetz et al., 2004).

It is important to note here that many theoretical explanations have been suggested for the empirically well-documented notion that an alternative gains market share when ‘moving’ towards a compromise-position in a given choice set. Most studies emphasize that the compromise effect is a consequence of the fact that often, a smaller or greater amount of time elapses between making a choice and consuming the chosen alternative. For example, it has been suggested (Simonson, 1989) that the compromise effect may arise from the decision-maker’s uncertainty regarding the importance he will attach in the (near) future to different attributes: given such ‘taste-uncertainty’, choosing the ‘safe’ or compromise option that performs reasonably well on each attribute is a good strategy. As a referee noted, an alternative explanation can be found in the role of “temporal stochastic inflation” (Salisbury and Feindberg, 2008): the “uncertainty about precisely how enjoyable each available item will appear at time of consumption” may also explain the compromise-effect. Our study, instead of considering discrepancies between current and future preferences, puts forward RRM-behaviour as another possible reason underlying compromise-effects, without aiming at a comparison with the abovementioned perspectives.

As said, a formal exposition of the sources of this conceptual difference in terms of (not) capturing the compromise effect between RRM and RUM can be found in the Appendix.
Intuitively, the difference is best understood by going through a concrete example. We use an example inspired by the one used in Kivetz et al. (2004): consider a choice between laptops that vary in terms of speed (GHz) and Random Access Memory (GB). Assume for simplicity of argumentation that both attributes are valued at a rate of 1 / unit. Three laptops are available: \( i = \{1\text{GHz}, 3\text{GB}\} \), \( j = \{2\text{GHz}, 2\text{GB}\} \), \( k = \{3\text{GHz}, 1\text{GB}\} \). Thus, \( j \) is a typical compromise alternative, with average speed and average memory when compared to its competitors. Given the importance of the attributes (1 / unit), deterministic utility of all laptops equals 4 and a conventional MNL-model would assign equal choice probabilities of 33% to each of them. In contrast, RRM arrives at very different choice probabilities: \( P(i) = 0.21 \), \( P(j) = 0.58 \) and \( P(k) = 0.21 \). In other words, RRM predicts a much higher choice probability for the compromise alternative.

To see the reasons behind RRM’s recognition of the compromise alternative, it is insightful to describe in detail what happens when one of \( j \)’s attributes is improved at the cost of the other attribute. For example, assume that \( j \)’s speed is improved to 2.5GHz at the cost of an equal decrease in memory towards 1.5GB. Importantly, note that in the context of an MNL-model, this change in \( j \)’s attribute values is irrelevant since it does not change deterministic utilities. However, as we will see now, the change is highly relevant in the context of RRM (see Table 1, where \( j^* \) represents laptop \( j \) after the costly increase in speed).

**Table 1: Regret-components for laptop \( j \), before and after a costly improvement in speed**

<table>
<thead>
<tr>
<th></th>
<th>Speed</th>
<th>Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>( i )</td>
<td>1GHz</td>
<td>3GB</td>
</tr>
<tr>
<td>( k )</td>
<td>3GHz</td>
<td>1GB</td>
</tr>
<tr>
<td>( j )</td>
<td>2GHz</td>
<td>2GB</td>
</tr>
<tr>
<td>( j^* )</td>
<td>2.5GHz</td>
<td>1.5GB</td>
</tr>
<tr>
<td>( R_{ij} )</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>( R_{ik} )</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>( R_{ji} )</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( R_{ik} )</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( R_{ji} )</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>( R_{kj} )</td>
<td>1.5</td>
<td>0.5</td>
</tr>
<tr>
<td>( R_{kj} )</td>
<td>2</td>
<td>0.5</td>
</tr>
</tbody>
</table>

First consider the effect of \( j \)’s increase in speed on the overall regret associated with \( j \): \( R_j \).

Due to \( j \)’s increase in speed \( R_{jk} \) drops from 1 to 0.5 (that is: \( j \)’s regret vis-à-vis \( k \) decreases).

This decrease in regret associated with the comparison of \( j \) with the fastest laptop \( k \) is fully in line with intuition. However, this decrease is irrelevant for the computation of \( j \)’s overall regret (\( R_j \)), since this overall regret is defined as the maximum, not the sum, of the regrets.
associated with comparisons with $i$ and $k$. This maximum does not decrease as a result of the decrease in $R_{jk}$; in fact it increases due to an increase in $R_{ji}$ ($j$’s decrease in memory makes that it suffers more regret from the comparison with the laptop with largest memory). In other words, whereas laptop $j$’s regret is divided evenly across $i$ and $k$, $j^*$ ’s regret is not evenly distributed anymore. It is this loss in symmetry of regrets that leads to a higher overall regret $R_j$ in the context of the max-operation.

Having discussed the impact of the costly increase in speed on $R_j$, we now investigate its impact on the regrets associated with the other two alternatives, starting with $R_i$. As expected, slowest laptop $i$’s regret associated with the comparison with $j$ increases due to $j$’s increase in speed. However, $R_j$ ’s increase from 1 to 1.5 is irrelevant for the computation of overall regret $R_i$, since there is more regret associated with $i$’s comparison with $k$: the difference in speed between $i$ (1GHz) and $k$ (3GHz) remains larger than the speed difference between $i$ (1GHz) and $j^*$ (2.5GHz). Again, the max-operation makes $j$’s speed improvement of no importance to overall regret. Regarding $R_k$, a similar mechanism is at work, but in the opposite direction: $R_{kj}$ decreases due to $j$’s decrease in memory ($j$’s increase in speed is irrelevant since $k$ remains the fastest laptop available), but this decrease is irrelevant since $R_{ki}$ remains at a level of 2 due to $k$’s relatively large difference with $j$ in terms of memory.

In sum, $j$’s increase in speed, coming at the cost of an as large increase in memory, leads to an increase in $j$’s overall regret $R_j$, while having no (increasing) effect on $R_i$ and $R_k$. As a result, RRM predicts that choice probabilities will change in the following way: $P(i)$ and $P(k)$ increase from 0.21 to 0.27 at the cost of $P(j)$ which decreases from 0.58 to 0.45. Figure 1 visualizes the effect of such costly improvements in general, and illustrates how RRM strongly penalizes $j$’s move from a compromise alternative towards an alternative with somewhat more extreme performance on relevant attributes.

In short, the high market share of compromise alternatives predicted by RRM (relative to the corresponding market share predicted by RUM) follows from the regret-symmetry exhibited by compromise alternatives. In the following two Sections, we study how this

---

6 Note that the intuition behind applying the max-operator, not the sum, to compute overall regret stems from the idea that overall regret emerges from a comparison with the best available alternative, not from a comparison with every other available alternative.
A conceptual difference between RRM and RUM translates into differences in terms of model with empirical data.

4. RRM versus RUM: Evidence from a revealed parking choice dataset

Parking policies form an increasingly important means to manage congestion in the vicinity of attractive destinations, such as shopping centres, city centres, universities, business districts and recreational facilities (e.g. Anderson & de Palma, 2007). For such policies to be effective, parking choice behaviour needs to be properly understood, which has led to a number of recent parking choice modelling and data collection efforts (e.g. Lam et al., 2006).

For our empirical assessment of RRM versus RUM, we make use of a revealed parking choice dataset 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 choice set of 14 available parking lots at the university campus. Please see van der Waerden et al. (2008) for an elaborate discussion of sample characteristics and the interview process, and for an interpretation of parameter estimates in terms of their parking policy-implications. Table 2

![Figure 1: The effect of an increase in the speed of compromise laptop j, at the cost of an equal decrease in memory, on its market share as predicted by RRM. As 'Speed' moves from 1 to 3 GHz, 'Memory' moves from 3 to 1 (and vice-versa) by definition. The two black dots refer to the values for alternative j as presented in Table 1.](image-url)
presents the best parking choice model for the RRM and RUM approaches (we tested a number of specifications with varying numbers of constants – both the RRM and RUM model fitted best when constants were estimated for parking lots 1, 2, 3, 5, 6, 7, 9, 10, 12 and 13)\(^7\).

Table 2: RUM- and RRM-estimation on revealed parking choices

<table>
<thead>
<tr>
<th>Variables</th>
<th>RUM Parameter</th>
<th>RUM t-Statistic</th>
<th>RRM Parameter</th>
<th>RRM t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NR_SPACES</td>
<td>0.7161</td>
<td>0.866</td>
<td>2.4415</td>
<td>5.165</td>
</tr>
<tr>
<td>ROOM_MANEUV</td>
<td>0.3840</td>
<td>0.722</td>
<td>0.1503</td>
<td>0.090</td>
</tr>
<tr>
<td>RIGHT_OF_WAY</td>
<td>-0.6005</td>
<td>-1.273</td>
<td>2.0857</td>
<td>1.784</td>
</tr>
<tr>
<td>DISTANCE</td>
<td>-3.6336</td>
<td>-17.049</td>
<td>-6.6249</td>
<td>-8.688</td>
</tr>
<tr>
<td>CONSTANT_1</td>
<td>0.4790</td>
<td>0.592</td>
<td>1.2430</td>
<td>2.228</td>
</tr>
<tr>
<td>CONSTANT_2</td>
<td>1.9217</td>
<td>3.923</td>
<td>0.9796</td>
<td>-3.820</td>
</tr>
<tr>
<td>CONSTANT_3</td>
<td>0.8233</td>
<td>0.629</td>
<td>1.4743</td>
<td>2.308</td>
</tr>
<tr>
<td>CONSTANT_5</td>
<td>1.3567</td>
<td>3.018</td>
<td>-1.3402</td>
<td>-3.795</td>
</tr>
<tr>
<td>CONSTANT_6</td>
<td>0.1033</td>
<td>0.286</td>
<td>-0.7359</td>
<td>-2.413</td>
</tr>
<tr>
<td>CONSTANT_7</td>
<td>0.5477</td>
<td>-1.192</td>
<td>1.2401</td>
<td>3.943</td>
</tr>
<tr>
<td>CONSTANT_9</td>
<td>-1.2605</td>
<td>-0.706</td>
<td>2.8255</td>
<td>2.170</td>
</tr>
<tr>
<td>CONSTANT_10</td>
<td>1.6934</td>
<td>2.220</td>
<td>-0.5556</td>
<td>-1.193</td>
</tr>
<tr>
<td>CONSTANT_12</td>
<td>0.8529</td>
<td>2.077</td>
<td>-0.5777</td>
<td>-2.350</td>
</tr>
<tr>
<td>CONSTANT_13</td>
<td>0.0151</td>
<td>0.035</td>
<td>-1.1002</td>
<td>-3.312</td>
</tr>
</tbody>
</table>

NR_SPACES refers to the actual number of spaces available at a particular parking lot (ranging from 30 to 298) divided by 100, ROOM_MANEUV refers to the availability of extra space for manoeuvres (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). Finally, CONSTANT_X refers to the constant for parking lot X (note that the best models for both RUM and RRM contained 10 constants, which makes the other parking lots serve as base options in terms of unobserved attributes).

The difference in goodness of fit in favour of RRM appears to be non-negligible, consisting of a 1.5 percentage points difference in rho square with respect to RUM. Since both

---

\(^7\) Note that all estimated models presented in this paper were coded in GAUSS 7.0, using the MaxLik-routine.

\(^8\) “Parameter” refers to the \( \beta \)’s presented in section 2. “Variables” refers to the \( \nu \)’s presented in section 2.
models use the same number of parameters, RRM’s better fit purely stems from its model structure, as highlighted earlier – the main difference with RUM being that RRM’s structure accommodates for the compromise effect. Concerning specific parameter estimates, both models generate expected signs for all significant parameters. Note that RRM has more significant parameters than RUM. Where the RUM model exhibits a substantial and significant parameter for DISTANCE, RRM features a significant parameter for RIGHT_OF_WAY as well, and has more significant constants than the best RUM-model.

Practically, RRM-estimates – in line with intuition – show that distance to and from the workplace is the most important variable explaining parking choice, followed by the number of spaces available. Whether or not one has right of way when leaving the parking lot, and whether or not there is room for manoeuvring, are of lesser importance. The presence of many significant (and substantial) constants signal that many other factors are important for explaining parking choice behaviour. Note the difference in sign between significant RRM- and RUM-constants: contrary to RUM-constants, positive RRM-constants refer to higher regret, and lower choice probabilities.

Although the empirical estimation presented here signals RRM’s potential in terms of model fit, the size of the choice set (14 alternatives) makes it difficult to relate RRM’s performance difference with RUM to their conceptual differences: any choice set of this size will contain many alternatives with mediocre performance in terms of one or more relevant attributes, making it difficult to identify specific compromise alternatives and their associated differences in terms of RUM- and RRM-based choice probabilities. To provide a more detailed account of how the empirical differences between RRM and RUM can be related to their conceptual differences, the next Section presents RRM- and RUM-estimations on an additional dataset.

5.  RRM versus RUM: Further evidence from a stated route choice dataset

Whereas parking policies aim at combating congestion within cities, road pricing is generally considered the most promising policy instrument to reduce congestion on highways (e.g. Rouwendal & Verhoef, 2006). In the process of designing effective road pricing schemes, detailed knowledge is needed about how travellers makes trade-offs between money, travel time and travel time variability. This consideration has led to a large and growing body of

\[9\] See van der Waerden et al. (2008) for estimation of more complex models on the available data, incorporating a number of interaction effects.
Stated Preference-based modelling efforts to estimate the monetary value of time and reliability (e.g. Small et al., 2005).

For our analysis, we use a three alternative, five attribute stated route choice dataset (Hensher et al., 2007). Data consisted of 300 car-drivers that made 16 choices each, between their current or reference route and two alternative routes with varying trip attributes: free flow travel time (TT_FREEFLOW), slowed down travel time (TT_SLOW), travel time variability (TT_VAR), vehicle running cost (RUNN_COSTS) and the costs associated with road pricing (TOLL_COSTS). Time-related attributes are in minutes, costs-related attributes in Australian dollars. Variability in travel time was conceptualized as the difference between the longest and shortest trip time. Besides parameters for these attributes, constants were estimated for alternative 1 (CONSTANT_1, for the reference alternative) and alternative 2 (CONSTANT_2). A statistically efficient experimental design was used to generate choice sets pivoted around the traveller’s current trip.

First, RRM and RUM-models were estimated on the complete dataset. Table 3 clearly shows that RUM fits the data better than does RRM (the difference in rho-square being 2.6 percentage points). Note that obtained parameter estimates and $t$-values are comparable across the two models: travel time in congested conditions is valued more negatively than travel time in free flow conditions, the negative effect of travel time variability is minor, and toll costs are valued more negatively than petrol costs.

As a second analysis, all choice situations were selected where one of the non-reference alternatives is a compromise alternative – a compromise alternative being defined as an alternative that features at least 3 (out of 5) compromise-attributes. A compromise-attribute is in turn defined as an attribute whose value is in between the corresponding attribute values of the other two alternatives. By definition, the remaining two non-compromise alternatives each feature zero, one or two compromise-attributes, totalling up to no more than two across alternatives. This selection resulted in a set of 618 observations. As shown in Table 3, RRM slightly outperforms RUM in terms of fitting this subset of the data (the difference in rho-square being 0.7 percentage points). Apparently, RRM does a little better than RUM in terms of explaining choice-behaviour in the context of choice situations containing a compromise alternative. However, recognizing that RUM fits better on the

---

10 Note that making this selection implies losing the statistical efficiency of the overall design. Given the scope of this paper, we accept this to enable a more detailed comparison between RRM and RUM.

11 We excluded reference alternatives from this selection process to avoid confounding reference effects with compromise effects. See Hess et al. (2008) for a study into reference effects, using the same data.
dataset as a whole, it seems that RUM does a considerably better job in explaining choice-behaviour in choice situations that do not contain a compromise alternative. Note that again, parameter estimates and t-values are comparable across models. Also note that the reduction in magnitude of t-values (compared with those obtained from the full dataset) results from the loss in design efficiency in combination with the drop in number of observations.

Table 3: RUM- and RRM-estimation on stated route choices

<table>
<thead>
<tr>
<th>Variables</th>
<th>Full dataset</th>
<th>Presence of a compromise alternative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RUM</td>
<td>RRM</td>
</tr>
<tr>
<td>TT_FREEFLOW</td>
<td>-0.0572</td>
<td>-17.590</td>
</tr>
<tr>
<td>TT_SLOW</td>
<td>-0.0745</td>
<td>-25.451</td>
</tr>
<tr>
<td>TT_VAR</td>
<td>-0.0037</td>
<td>-1.758</td>
</tr>
<tr>
<td>RUNN_COSTS</td>
<td>-0.2340</td>
<td>-12.668</td>
</tr>
<tr>
<td>TOLL_COSTS</td>
<td>-0.2845</td>
<td>-26.706</td>
</tr>
<tr>
<td>CONSTANT_1</td>
<td>0.0521</td>
<td>0.765</td>
</tr>
<tr>
<td>CONSTANT_2</td>
<td>0.1092</td>
<td>2.614</td>
</tr>
<tr>
<td>0-LL</td>
<td>-5232</td>
<td></td>
</tr>
<tr>
<td>Final LL</td>
<td>-4021</td>
<td></td>
</tr>
<tr>
<td>rho-square</td>
<td>0.232</td>
<td></td>
</tr>
<tr>
<td>Number of cases</td>
<td>4800</td>
<td></td>
</tr>
</tbody>
</table>

Although the above analyses provide support for the notion that RRM’s potential to outperform RUM in terms of model fit is related to its treatment of compromise alternatives, we perform a brief additional analysis at the aggregate level to further investigate the relation between the models’ empirical performance differences and their conceptual differences. Specifically, we compare actual and predicted market shares of compromise alternatives in the final subset of 318 stated choices. The compromise alternative is chosen 39 times, implying a market share of 12.26%. The estimated RUM model slightly underestimates this share by predicting an average choice probability for the compromise alternative of 11.90%. RRM predicts an average choice probability of 12.38%, coming closer to the actual share.

In sum: although the differences found in this Section’s analyses are of minor or at best moderate magnitude, they do provide preliminary support for the proposition that RRM’s potential to outperform RUM results from RRM’s recognition of the compromise effect.

---

12 Reported t-values are based on robust standard errors, to account for the panel structure of the data.
6. General applicability of RRM

We here argue and illustrate that RRM, due to its behavioural premises, is expected to be less suitable than RUM for the analysis of choice from either i) choice or consideration sets that are imputed\(^\text{13}\) by the analyst and ii) choice or consideration sets that contain a no-choice option.

6.1. General applicability of RRM: Theory

Suboptimal performance of RRM in the context of imputed choice sets is expected because the RRM-approach, in contrast to RUM, predicts the popularity of an alternative (partly) in terms of its position relative to other alternatives in the choice set. More specifically, the notion of compromise alternatives (favoured by RRM) is only meaningful within the context of a particular choice set. As a result, RRM is expected to be less suitable for analyzing choices from sets that are imputed by the analyst: a compromise alternative in this imputed choice set may not necessarily be the compromise effect in the set actually considered by the individual. In contrast, RUM-models that exhibit the IIA-property have been well-known to be rather robust vis-à-vis choice set misspecifications.\(^\text{14}\) Note that choice (consideration) set imputation is by definition non-existent in Stated Preference data-collections, but it may occur in Revealed Preference settings, especially when the number of available alternatives – the universal set – is large (e.g. Shocker et al., 1991).

In the context of choice sets containing a no-choice option, suboptimal performance of RRM is expected since it postulates that choice is determined on an attribute-by-attribute comparison. Since the no-choice option has, by definition, no attribute in common with any of the other alternatives, no meaningful measure of regret can be assigned in those situations. In a RUM context, the no-choice option is represented in the estimation process by means of a constant (or: all other alternatives are assigned a constant relative to the no-choice option that then is assigned a utility of zero). It is easily seen that such an approach is meaningless in a

---

\(^{13}\) With “imputed” we mean to say that, in the presence of a very large set of objectively available choice options, the analyst makes an inference about what constitutes the set of alternatives that the decision-maker has actually considered when making his choice (all other alternatives play no role in the process of estimating the choice model).

\(^{14}\) One could, of course, also argue that the empirically well-established compromise effect provides a good reason to be very careful with choice set imputations, irrespective of the choice-paradigm used for model estimation.
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’ relative performance that generates regret, rather than their overall or average performance. Clearly, an individual is more likely to choose the no-choice option when all alternatives perform poorly than when all alternatives perform very well. However, since both choice sets might generate equal amounts of regret there is a mismatch between observed and predicted behaviour and as a result, the estimated constant for the no-choice option has no clear behavioural meaning. The no-choice constant can therefore never be as effective as in a RUM-setting, where the absolute value of its magnitude does reflect the overall performance of available choice-options (to be more precise: the performance of the best of available choice options).15

Before we move to an empirical illustration of the argument made above, it is important to note that the expected limited applicability of RRM when choice sets are imputed, or contain no-choice options, does not follow from specific model artefacts; rather, it follows from the basic behavioural premises underlying the use of regret and attribute-level comparisons in decision-making.

6.2. General applicability of RRM: Empirical illustration

In the remainder of this Section, we investigate whether the theoretically expected relatively poor performance of RRM in choice situations of Types I or II results in an actual difference in terms of goodness of fit, when compared with RUM.

Concerning Type I choice sets (choice or consideration sets that are imputed by the analyst), we use a dataset containing shopping centre-choices (Arentze et al., 2005). Without going into much detail concerning the background of the data-collection and response group characteristics, we note that 1503 revealed choices were observed from consumers that went shopping with the aim of buying groceries. Variables included were FLOOR_GROCERIES (m² floorspace concerning groceries in a particular centre), FLOOR_OTHER (m² 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 (please refer to Arentze et al. (2005) for details). Table 4 shows the estimation results for both the RUM- and RRM-models.

15 See Dhar (1997) for a detailed discussion of the (dis-)advantages of adding a no-choice option to a choice task.
Table 4: RRM-performance when choice sets are imputed

<table>
<thead>
<tr>
<th>Variables</th>
<th>RUM</th>
<th>RRM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parameter</td>
<td>t-Statistic</td>
</tr>
<tr>
<td>FLOOR_GROCERIES</td>
<td>0.1060</td>
<td>6.689</td>
</tr>
<tr>
<td>FLOOR_OTHER</td>
<td>0.0110</td>
<td>4.977</td>
</tr>
<tr>
<td>DISTANCE</td>
<td>-0.0448</td>
<td>-8.965</td>
</tr>
<tr>
<td>0-Log-Likelihood</td>
<td>-2419</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood at convergence</td>
<td>-2305</td>
<td></td>
</tr>
<tr>
<td>rho-square</td>
<td>0.047</td>
<td></td>
</tr>
<tr>
<td>Number of cases</td>
<td>1503</td>
<td></td>
</tr>
</tbody>
</table>

The difference in goodness of fit, as expected, is in favour of the RUM-model, which is a sign that RRM indeed has some difficulty with handling imputed choice sets. Of course, this difference in rho-square is a rather aggregate measure that provides little information on the exact source RRM’s inferior performance. However, because we lack data about the true choice sets of individual shoppers, we cannot point at specific examples in the dataset where RRM makes a wrong prediction due to a choice-set misspecification. Note that parameter estimates do not differ much between the two models and suggest the importance of both floorspace – for groceries as well as other categories – and distance as drivers of shopping destination choice; a finding that is in line with intuition and earlier literature on this topic (e.g. Oppewal et al., 1997). Rather low rho-squares suggest that besides these variables a number of other variables play an important role as well.16

Concerning the empirical testing of RRM in the context of choice sets of Type II (choice or consideration sets that contain a no-choice option), we use stated choice data relating to travellers’ willingness to pay for travel information (Molin et al., 2007). Again, we will not go into much detail here concerning the specifics of the data-collection effort. 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). Three tasks referred to the situation where no transfer was needed during a train trip; three tasks referred to the situation where a transfer was needed towards a high-frequency train service (TRANSFER_HIGH); the other three tasks referred to the

16 Please see Arentze et al. (2005) for estimation of more complex and elaborate joint models of shopping purpose and shopping destination, in the context of a number of shopping purposes other than groceries.
situation where, during the trip, an interchange towards a low-frequency service was to be made (TRANSFER_LOW). The order of these contexts was varied systematically.

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 traveller 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 3 shows estimation results of both RUM- and RRM-models.

Note that the RRM-model dealt with the no-choice option as follows: for each of the three information options, regret was computed using the RRM-model presented in Section 2, taking the set of three information options as choice set. The three obtained measures of regret where compared to a context-specific constant (that may be interpreted as a regret-threshold): when the regret associated with the lowest regret-alternative was higher than that of the constant (threshold), the no-choice option was selected. Otherwise, the minimum regret alternative was chosen. Note also that robust t-statistics were used to account for the dataset’s panel structure. NO_INFO, TRANSFER_HIGH and TRANSFER_LOW refer to constants being estimated for the no-choice option under the different contexts: the two transfer-constants need to be added to the NO_INFO constant to get the constant of the no-choice option for a given transfer context.

Clearly, and as expected, RUM outperforms RRM in terms of Goodness of Fit, although again, parameter estimates and their significance levels do not vary that much between the two models. An additional analysis of predicted choices is performed as a further illustration of how the difference in performance is indeed related to RRM’s expected inability to provide correct choice probability-predictions for the no-choice option in particular. Using the final parameter estimates of RRM and RUM, we computed choice probabilities for the no-choice option in the context of the top forty observations of the dataset. It appears that RRM structurally underestimates the likelihood of the no-choice option, whereas RUM does a much better job in correctly assessing this likelihood: for the
seven times the no-choice option is chosen RRM achieves an average likelihood of .51, which is – as expected – substantially lower than RUM’s likelihood of .57. Apparently, application of the estimated RUM-model leads to the correct conclusion that in these seven choice situations the respondent perceived the presented choice-alternatives as unsatisfactory. In contrast, the estimated RRM model focuses on the performance of the three choice-options relative to each other, and apparently finds that this comparison between choice-options generates relatively little regret across these alternatives. As a result, application of the RRM-model underestimates the probability associated with the no-choice option.

Table 5: RRM-performance when choice sets contain a no-choice option

<table>
<thead>
<tr>
<th>Variables</th>
<th>RUM Parameter</th>
<th>t-Statistic</th>
<th>RRM Parameter</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIMES+SEARCH</td>
<td>0.2085</td>
<td>2.337</td>
<td>0.2623</td>
<td>2.866</td>
</tr>
<tr>
<td>TIMES+ADVICE</td>
<td>0.5872</td>
<td>6.578</td>
<td>0.6174</td>
<td>6.806</td>
</tr>
<tr>
<td>INFO_INIT</td>
<td>-0.0364</td>
<td>-0.433</td>
<td>-0.1519</td>
<td>-1.992</td>
</tr>
<tr>
<td>BOTH_INIT</td>
<td>0.1600</td>
<td>1.773</td>
<td>0.1687</td>
<td>1.780</td>
</tr>
<tr>
<td>UNRELIABILITY</td>
<td>-0.0756</td>
<td>-3.742</td>
<td>-0.0802</td>
<td>-3.933</td>
</tr>
<tr>
<td>PRICE</td>
<td>-0.9290</td>
<td>-26.311</td>
<td>-0.9660</td>
<td>-21.125</td>
</tr>
<tr>
<td>NO_INFO</td>
<td>-1.0613</td>
<td>-7.138</td>
<td>-1.1536</td>
<td>-8.748</td>
</tr>
<tr>
<td>TRANSFER_HIGH</td>
<td>-0.2199</td>
<td>-1.532</td>
<td>-0.0947</td>
<td>-0.675</td>
</tr>
<tr>
<td>TRANSFER_LOW</td>
<td>-0.7159</td>
<td>-4.499</td>
<td>-0.6723</td>
<td>-4.392</td>
</tr>
<tr>
<td>0-Log-Likelihood</td>
<td>-2534</td>
<td></td>
<td>-2534</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood at convergence</td>
<td>-1992</td>
<td></td>
<td>-2113</td>
<td></td>
</tr>
<tr>
<td>rho-square</td>
<td>0.214</td>
<td></td>
<td>0.166</td>
<td></td>
</tr>
<tr>
<td>Number of cases</td>
<td>1836</td>
<td></td>
<td>1836</td>
<td></td>
</tr>
</tbody>
</table>

In the context of the remaining 33 observations (where the no-choice option is not chosen), RRM predicts an average choice probability for the no-choice option of .26, which is somewhat higher than RUM’s average prediction of .24. For these observations, RRM apparently finds that comparisons between the three choice-options generate relatively high amounts of regret, which would warrant a choice for the no-choice option. In contrast, RUM focuses on whether the best of the choice-options is good enough to be chosen, and arrives at somewhat lower and more realistic probabilities associated with the no-choice option. Concerning specific parameter estimates, it appears that price is by far the most important attribute, suggesting a low willingness to pay for the information; this finding is in line with earlier empirical studies into the use of travel information services (see Chorus et al. (2006) for a review). Furthermore, it appears that a search possibility, and especially the provision of advice, is preferred over basic travel time information. Travellers are indifferent with respect
to who should take the initiative for providing / acquiring information, but attach great importance to reliability. As expected, context-specific constants for the no-choice option reflect that travellers are more inclined to acquire information when making a transfer during their trip, especially when the transfer is towards a low-frequency transit service\textsuperscript{17}.

6. Conclusions and managerial implications

This paper provides theoretical and empirical comparisons between the Random Utility Maximization (RUM)-approach to choice modelling and the recently introduced Random Regret Minimization-approach (RRM). RRM is based on the postulate that choice is determined by the decision-maker’s inclination to avoid the situation where non-chosen alternatives outperform the chosen one on one or more of its attributes.

The main theoretical contribution of this paper is the identification of the key difference between RRM and RUM in terms of choice predictions: RRM, in contrast with RUM, recognizes alternatives with mediocre performance on each attribute, rather than performing well on some attributes and poor on others. As such, RRM appears to capture the compromise-effect, an effect well documented in a number of empirical studies of consumer choice. The main empirical contribution of this paper is the illustration, based on revealed and stated choice data, of the notion that RRM has the potential to outperform RUM when choice situations contain compromise alternatives. When no compromise alternatives are present or when choice sets are imputed or contain a ‘no choice’ option, RUM appears to outperform RRM. Although found differences range from small to moderate in terms of magnitude, they do provide preliminary support for the proposition that RRM’s potential to outperform RUM is directly related to RRM’s recognition of the compromise effect in multinomial choice situations.

In terms of managerial implications, RRM offers a number of strategies for improving an alternative’s market share that are based on the compromise effect. First, to the extent that choice-behaviour is governed by RRM-premises, it pays off to position new alternatives as compromise alternatives, rather than establishing a very good performance on one attribute (relative to the performance on that attribute of the other available alternatives) at the cost of a very poor performance on another. Importantly, performance is defined here in terms of utils, not attribute values, and as such incorporates attribute importance. A second and related

\textsuperscript{17} See Molin et al. (2007) for estimation of more complex model forms, involving estimation of a variety of random parameters and context-specific interaction effects.
managerial implication suggests that it is more fruitful, *ceteris paribus*, to improve an attribute on which an alternative is performing poorly, than improving an attribute on which the alternative already has a strong performance. The underlying reason for this implication is that improvement of a poor performing attribute increases the compromise-nature of the alternative. Finally, to the extent that choice-behaviour is governed by RRM-premises, improving an attribute on which an alternative is performing poorly, at the cost of deteriorating an attribute on which the alternative already has a strong performance, may substantially improve an alternative’s market share, also when the two attributes are equally important to the consumer.

At this point, we want to stress that the four datasets used in this paper for empirical analyses only provide a first step in exploring the performance of RRM vis-à-vis RUM. As a result, a first and most important conclusion is that more empirical testing of RRM (vis-à-vis RUM) needs to be done before a more reliable picture of its potential will arise. Clearly, we see a strong need for such more elaborate empirical testing in the near future. With respect to RRM’s potential in terms of capturing the compromise effect, an empirical comparison with other models that have been used to capture that effect (see Rieskamp (2006) for an overview of modelling approaches), such as the contextual concavity model (Kivetz et al., 2004), the relative advantage model (Tversky & Simonson, 1993; Kivetz et al., 2004) and the loss-aversion model (Tversky & Kahneman, 1991; Kivetz et al., 2004), seems an obvious and interesting avenue for further research. In addition, it seems fruitful to compare the RRM-approach with other models of choice set-specific preferences such as, for example, the sequential best-worst choice model (Marley & Louvière, 2005) – see Marley et al. (2008) for a recent overview. Finally it may be noted that, as a referee pointed out, the relative empirical performance of two different decision-models (in our case: regret minimization versus utility maximization) has been shown recently (Blavatskyy & Pogrebna, In Press) to depend on the model used for error term specification (in our case: iid Extreme Value Type I errors). This is something that deserves our attention in future research efforts.

**Acknowledgements**

We thank Aloys Borgers, Eric Molin, Harmen Oppewal, Ruud van Sloten, Peter van der Waerden, John Rose and David Hensher for providing the data needed for model estimation. Furthermore, we thank participants at the 19th Advanced Research Techniques forum of the American Marketing Association, and two anonymous referees for the International Choice Modelling Conference, for making a host of constructive suggestions.
References


**Appendix: Difference between choice probabilities generated by RRM and RUM**

First, note that application of Equations (1) and (3), given that $\beta_i, \beta_j, x_i - x_k, x_k - x_j, y_j - y_k, y_k - y_i$ are all strictly positive, leads to the following regrets for alternatives $i, j, k$:

$$R_i = \beta_y \cdot (y_j - y_i), \quad R_j = \beta_x \cdot (x_i - x_j) \quad \text{and} \quad R_k = \max \{\beta_x \cdot (x_i - x_k), \beta_y \cdot (y_j - y_k)\}.$$  

We first show that $P_{k}^{RRM} > P_{k}^{RUM}$ by showing that $[(R_i - R_k) - (V_k - V_i)] > 0$ and $[(R_j - R_k) - (V_k - V_j)] > 0$. Consider first the situation where $R_k = \beta_x \cdot (x_i - x_k)$. Then, $(R_i - R_k) - (V_k - V_i)$ can be rewritten as: $\beta_y \cdot (y_j - y_i) \cdot \beta_x \cdot (x_i - x_k) - \beta_x \cdot x_k - \beta_y \cdot y_k + \beta_x \cdot x_i + \beta_y \cdot y_i$. Rearranging gives: $(R_i - R_k) - (V_k - V_i) = \beta_y \cdot (y_j - y_k)$. Since $\beta_y$ and $y_j - y_k$ are strictly positive, their product is strictly positive as well. As a result, we obtain that $(R_i - R_k) - (V_k - V_i) > 0$ when $R_k = \max \{\beta_x \cdot (x_i - x_k), \beta_y \cdot (y_j - y_k)\} = \beta_x \cdot (x_i - x_k)$. Now,
consider the situation where \( R_k = \beta_y \cdot (y_j - y_k) \). Then, \( (R_i - R_k) - (V_k - V_i) \) can be rewritten as

\[
\beta_y \cdot (y_j - y_i) - \beta_y \cdot (y_j - y_k) - \beta_x \cdot x_i - \beta_x \cdot y_k + \beta_x \cdot x_k + \beta_y \cdot y_i.
\]

Rearranging gives:

\[
(R_i - R_k) - (V_k - V_i) = \beta_x \cdot (x_i - x_k).
\]

Since \( \beta_x \) and \( x_i - x_k \) are strictly positive, their product is strictly positive as well. As a result, we obtain that \( (R_i - R_k) - (V_k - V_i) > 0 \) when \( R_k = \beta_y \cdot (y_j - y_k) \). Follows:

\[
(\beta_y \cdot (y_j - y_k) > 0 \text{ when } R_k = \beta_y \cdot (y_j - y_k)).
\]

Similarly, it can be seen that \( (R_j - R_k) - (V_k - V_j) = \beta_y \cdot (y_j - y_k) > 0 \) when \( R_k = \beta_x \cdot (x_i - x_k) \) and that \( (R_j - R_k) - (V_k - V_j) = \beta_x \cdot (x_i - x_k) > 0 \) when \( R_k = \beta_y \cdot (y_j - y_k) \). Follows:

\[
(\beta_y \cdot (y_j - y_k) > 0 \text{ when } R_k = \beta_y \cdot (y_j - y_k));
\]

As a result,

\[
[(R_i - R_k) - (V_k - V_i) > 0] \land [(R_j - R_k) - (V_k - V_j) > 0].
\]

We go on to show that \( P_k^{BRM} - P_k^{RUM} \) is largest when \( \beta_x \cdot (x_i - x_k) = \beta_y \cdot (y_j - y_k) \), by showing that both \( (R_i - R_k) - (V_k - V_i) \) and \( (R_j - R_k) - (V_k - V_j) \) are largest when \( \beta_x \cdot (x_i - x_k) = \beta_y \cdot (y_j - y_k) \). As derived above, \( (R_i - R_k) - (V_k - V_i) = (R_j - R_k) - (V_k - V_j) = \beta_y \cdot (y_j - y_k) \) when \( R_k = \beta_x \cdot (x_i - x_k) \). Reformulating this condition gives \( \beta_y \cdot (y_j - y_k) \leq \beta_x \cdot (x_i - x_k) \). Now, it is easily seen that the largest possible values for \( (R_i - R_k) - (V_k - V_i) \) and \( (R_j - R_k) - (V_k - V_j) \), under the condition that \( R_k = \beta_y \cdot (y_j - y_k) \) (or: \( \beta_x \cdot (x_i - x_k) \leq \beta_y \cdot (y_j - y_k) \)), are obtained when \( \beta_x \cdot (x_i - x_k) = \beta_y \cdot (y_j - y_k) \). Since \( (R_i - R_k) - (V_k - V_i) = (R_j - R_k) - (V_k - V_j) = \beta_x \cdot (x_i - x_k) \) when \( \beta_x \cdot (x_i - x_k) \leq \beta_y \cdot (y_j - y_k) \), it is again easily seen that the largest possible values for \( (R_i - R_k) - (V_k - V_i) \) and \( (R_j - R_k) - (V_k - V_j) \), under the condition that \( \beta_x \cdot (x_i - x_k) \leq \beta_y \cdot (y_j - y_k) \), are obtained when \( \beta_x \cdot (x_i - x_k) = \beta_y \cdot (y_j - y_k) \).■