Specification of regret-based models of choice behaviour: formal analyses and experimental design based evidence—commentary

Caspar G. Chorus¹ · Sander van Cranenburgh¹

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract A recent paper published in this journal compares two regret based choice models, and concludes that one of them is theoretically inferior and has a worse empirical performance in the context of a particular data set [Rasouli and Timmermans, Transportation 6:1–22, 2016]. Unfortunately, those conclusions are ill-founded as they are based on a misinterpretation and misrepresentation of one of the two considered models. Furthermore, the paper overlooks highly relevant recent work on the topic, and contains insufficient empirical analyses. Together, these issues make that the paper provides a confusing addition to the literature. With the aim of lifting some of this confusion, this commentary sets out to highlight, and correct where possible, the paper’s shortcomings.

Keywords Random regret minimization · Discrete choice theory · Regret · RRM · Commentary

Introduction

In a recent contribution to this journal, Rasouli and Timmermans compare two different specifications of RRM models. One specification was introduced in a paper co-authored by Timmermans (Chorus et al. 2008), the other was presented in a paper published two years later (Chorus 2010). In the remainder of this commentary, we will refer to the former as RRM2008, and the latter as RRM2010. In their paper (from here on denoted RT2016), they conclude that 2010RRM is theoretically problematic, and that it has a worse empirical performance on a particular data set. Unfortunately, the theoretical conclusion is based on a flawed representation and interpretation of RRM2010 (and of regret based discrete choice models in general). Furthermore, RT2016 overlooks highly relevant recent work on the...
topic which makes their comparison to a considerable extent irrelevant; and the paper contains insufficient empirical analyses.

Importantly, this commentary will only focus on correcting flaws in reasoning and representation, and critical omissions of past work. Our aim is not to evaluate the outcome of the model comparison in itself (RT2016 prefer 2008RRM over 2010RRM). We are aware that, from their inception, RRM models have sparked considerable debate, (i) in terms of their potential and limitations as an alternative to linear in parameter RUM models and (ii) in terms of the ‘proper’ way to model regret minimizing behaviour. We have always kept an open mind regarding these debates, and will continue to do so; indeed, we believe academic debate is a potentially very effective route to scientific knowledge accumulation.

However, in order to be effective, a debate about the relative quality of different models must be based on a correct description of these models, a proper discussion of relevant past work, adequate interpretation of the results of these models, and a sufficiently rich empirical context. Unfortunately, RT2016 falls short on all of these aspects. This is likely to lead to considerable confusion among readers, and may hamper scientific knowledge accumulation. Therefore, this commentary aims to resolve this confusion; it certainly does not aim to stifle academic debate.

In the remainder of this commentary, we will address these issues in RT2016; we start with a rebuttal of erroneous theoretical claims in RT2016, and continue with a discussion of overlooked past work on the topic. We conclude by considering the empirical content of RT2016. We have chosen to avoid discussion of several minor mistakes and omissions, for reasons of conciseness. The Appendix provides notation for the models discussed in this commentary. Note that our notation follows that in our previous papers, and differs from that of RT2016.

Finally, note that RT2016 compare two model specifications (RRM2008 and RRM2010) which differ in terms of both the shape of the so-called attribute regret function as well as in terms of the number of competitor alternatives used for computing total regret. In this commentary (with the exception of the Section discussing empirical analyses), we focus only the former of these two dimensions, as it is this dimension which causes most issues in RT2016. In other words, this commentary (and the next two sections in particular) focuses on refuting claims made in RT2016 concerning different specifications of the attribute regret function, which computes the level of regret which is associated with comparing a considered alternative with another alternative, in terms of a particular attribute.

### Theoretical misconceptions and misinterpretations

At various places, RT2016 claim that RRM2010 is “theoretically inferior” to RRM2008, that it has “theoretical shortcomings” and is “ill-founded” under some conditions; even that it is “theoretically problematic”. The reason why RT2016 believes that RRM2010 is

---

1 Regarding the aggregation of attribute regret to total regret, RRM2008 postulates that only regret with respect to the best performing competitor counts, while RRM2010 postulates that regret with respect to every better performing competitor is relevant. There are conflicting opinions in the literature as to which variant is to be preferred for theoretical reasons (e.g., Bell 1982; Quiggin 1994); we are agnostic in this regard, and consider it a mostly empirical question which rule should be used in the context of a particular data set.

2 But note that Prato (2014) contains analyses that refute claims in RT2016 that the additive nature of total regret in RRM2010 would be problematic in a route choice context including route overlap.
theoretically problematic, can be found in the Introduction (Page 4, below Eq. 6). Here, RT2016 refer to 2010RRM when they claim that

"the specification is theoretically no longer consistent with the concept of regret as the value of the regret function is positive even if the chosen alternative turns out to be better than the foregone choice alternative. That is, Eq. (6) does not approximate zero if attribute differences are small and/or the number of choice alternatives is large. Moreover, the new specification estimates regret to be higher than on the original specification, due to the constant 1 in the formulation."

This phrase, which is crucial to most of RT2016’s theoretical argumentation, embodies several major misunderstandings of discrete choice theory and regret theory.3

Discrete choice theory: absolute levels of regret are meaningless

The absolute level of regret (or: utility) is meaningless. As such, it is pointless to speak about ‘positive’ or ‘negative’ regret (or: utility) levels. As discussed in every textbook on discrete choice modelling (e.g., Train 2009), differences in utility (in this case: regret) levels determine choice behaviour, not their absolute levels. Adding (or, subtracting) some constant to each alternative results in exactly the same choice model.

RT2016, in subsection ‘Non-regret regime’, discusses the size of attribute regret (under RRM2010) when the considered alternative and a competitor alternative perform equally well on a particular attribute. Very straightforward mathematics gives the answer: ln(2) or 0.693147 (Eq. (9) in RT2016). RT2016 then argues that “Thus, even though theoretically regret is zero in this case, the [RRM2010] specification gives a positive value for regret”. This makes the RRM2010 model “ill-founded”, according to the authors. Unfortunately, in light of the previous paragraph, this claim itself is easily seen to be ill-founded.

In fact, Chorus (2014a) and Van Cranenburgh et al. (2015) provide an in-depth discussion of this point in the context of regret models, explaining what happens when, in the context of the RRM2010 model, a constant of size ln(2) is subtracted from the regret level which is associated with comparing a considered alternative with a competitor alternative in terms of a particular attribute (see Fig. 1, where \(x_{jm}\) and \(x_{im}\) denote the attribute levels of the competitor and considered alternative). As alluded to above, in terms of every relevant model property (e.g., model fit, parameter estimates, choice probability predictions), the shift of the regret level is—of course—irrelevant. The resulting model is exactly the same as the original model. In this ‘Modified’ formulation of RRM2010, regret becomes ‘negative’ when a considered alternative outperforms a competitor alternative. While at first sight this may seem behaviourally intuitive, we see no reason for favouring such a ‘modified’ regret function (perhaps other than for the sake of didactical purposes).

RT2016 is simply wrong to base their conclusion that RRM2010 is theoretically problematic, on an inspection of absolute regret levels. As with utility-based choice models, it is the shape of the regret function, not its position relative to the vertical axis, that determines choice behaviour imposed by the model. See Van Cranenburgh and Prato (2016) for visual illustrations of various recently proposed regret functions, which further emphasize this point.

3 It should be noted here, that the authors of RT2016 have put forward similar erroneous claims in another paper (Rasouli and Timmermans 2014). Unfortunately, claims made in that paper have been cited in a leading textbook recently (Hensher et al. 2015; p 365).
Correspondence with original Regret Theory: the role of rejoice

Here, we will explain that, in contrast to RT2016’s conclusion that the attribute regret function in RRM2010 is “theoretically no longer consistent with the concept of regret”, RRM2010 is in fact considerably closer to original Regret Theory than is RRM2008. As mentioned above, it is the shape of the regret function that determines the behaviour implied by the RRM model. Now, it can be easily seen (and it has been observed many times before) that the max operator used in RRM2008 implies that regret is insensitive to attribute changes when a considered alternative is as good as or better than a competitor on a particular attribute, before and after the change.

For example, if a considered route (A) is already 10 min faster than a competing route (B), RRM2008 postulates that in- or decreasing A’s travel time by, e.g., 5 min does not change route A’s regret. In other words, in the RRM2008 model, an alternative’s regret is invariant with respect to attribute changes in the domain of rejoice (which is the name economists have given to the behavioural opposite of regret). In sum, the RRM2008 model postulates that as long as a considered alternative performs better than a competing alternative on a given attribute, it does not matter—in terms of the considered alternative’s regret—how much better it performs.

In contrast, the RRM2010 model postulates that also in the domain of rejoice, the regret/rejoice function is sensitive to attribute changes. That is, the RRM2010 model postulates

![Attribute level regret function](image)

**Fig. 1** Attribute level regret functions of RRM2010 and ‘Modified RRM2010’

4 But note that we do not wish to state that RRM2010 is therefore theoretically superior; we wish to refrain from such normative evaluations.
that when a considered alternative performs better than a competing alternative on a given attribute, it does matter—in terms of the considered alternative’s regret—how much better it performs. But—and this is a crucial aspect of any model inspired by the original Regret Theory (Loomes and Sugden 1982)—the shape of the regret function in RRM2010 postulates that sensitivity to changes in attributes is greater in the domain of regret than in the domain of rejoice. This is clearly illustrated in Fig. 1, where the right hand side represents the domain of regret, and the left hand side the domain of rejoice. For example, the RRM2010 model postulates that, if a considered route (A) is 10 min faster than a competing route (B), in-or decreasing A’s travel time by, e.g., 5 min does change A’s regret/rejoice. Moreover, in line with Regret Theory, the RRM2010 model postulates that the effect of A’s travel time change would be even larger if A was the slowest route already.

In sum, RRM2008 can be considered an extreme case of Regret Theory, as it gives no weight to attribute changes in the domain of rejoice. RRM2010, in line with Regret Theory, postulates that regret levels are more sensitive to attribute changes in the domain of regret than in the domain of rejoice, but that also the domain of rejoice is relevant. RT2016’s claim that RRM2010 is “theoretically no longer consistent with the concept of regret” is false and misleading: if correspondence to the original Regret Theory (and, for that matter, to Prospect Theory/Loss Aversion) is a criterion, then the opposite statement would be considerably much closer to the truth.

New model developments which have been overlooked in RT2016.

In their abstract, RT2016 state that “Two different [regret-based] model specifications have been introduced in the transportation literature”. This may have been true in 2010, but no longer so; over the past six years, several developments in RRM modelling have been reported in the literature. In fact, much of the theoretical confusion that emerges in RT2016 could have been avoided if this more recent work on regret-based choice models would have been considered by RT. Specifically, consideration of two recent developments, published in other leading journals in our field (Chorus 2014a; Van Cranenburgh et al. 2015), is helpful to address misconceptions in RT2016. Furthermore, these recent developments make the comparisons presented in RT2016 largely redundant. See the Appendix for notation of these and previously discussed RRM models.

First, RT2016 incorrectly state that, in 2010RRM, a “constant 1 is added to avoid that the logarithmic function is undefined”. Note that the constant was not just “added” in RRM2010; but resulted from mathematical derivation (see Chorus 2010). And note that removing the constant does not lead to an undefined logarithmic function; on the contrary, it is easily seen that dropping the constant makes that the logarithm and the exponential function cancel out against each other. In fact, as Chorus (2014b) has shown, the result would be (an alternative mathematical specification of) a linear in parameters RUM model. Figure 1 in Chorus (2014a) clearly illustrates how the constant should be interpreted as a

5 Choice modellers familiar with Prospect Theory models (Kahneman and Tversky 1979) and models based on the notion of Loss Aversion (Tversky and Kahneman 1991), will notice the conceptual similarities between those models and RRM models: loss aversion postulates that losses with respect to a reference point loom larger than gains of equal magnitude. In RRM models, the reference point consists of the attribute level of a competing alternative.

6 Introductions to these more recent models, as well as to more conventional RRM models, can be found at www.advancedrrmmodels.com. In addition, that website contains code for different softwares, as well as several data sets, which may be freely used to try out different RRM models.
measure of regret aversion: as it moves away from one and starts to approach zero, the convex regret function starts to ‘fan out’ and approximate a linear function where regret is no longer overweighted compared to rejoice (this eventually implies classical linear in parameters RUM behaviour when the constant equals zero). Furthermore, that paper explains and shows how this regret aversion parameter can be estimated from choice data together with taste parameters, in a so-called G-RRM model (where ‘G’ stands for Generalized). When RT in their Conclusions and Discussion section claim that RRM2010’s “problems would be less if a smaller value [of the constant] had been chosen”, they are clearly misunderstanding the role of this constant, and they are overlooking recent studies (Chorus 2014b; Mai et al. 2015) which show that the constant is in fact an estimable regret aversion parameter, governing the relative weight of regret and rejoice.

Second, it has been convincingly shown recently (Van Cranenburgh et al. 2015), that the attribute regret functions embedded in RRM2008 and RRM2010 are in fact two very special cases of a more generic attribute regret function, whose shape can be estimated. That is, the μRRM model proposed in Van Cranenburgh et al. (2015) contains an estimable parameter (μ). If it approaches zero, the attribute regret function embedded in RRM2008 is obtained⁷; if it equals one, the attribute regret function embedded in RRM2010 is obtained (and if it becomes arbitrarily large, a linear in parameters RUM model is approximated). In other words, μ governs the shape of the attribute regret function, and as such, the degree to which regret is more important than rejoice: if it approaches zero, rejoice becomes irrelevant, as in RRM2008. If it equals one, regret is more important than rejoice but also rejoice is relevant. If it becomes large and positive, regret and rejoice become equally important (giving rise to linear in parameters RUM behaviour). Beyond these special cases, estimable parameter μ can take on every value in the domain [0, +∞], governing the shape of the regret function—i.e., the extent to which regret is more important than rejoice.

In sum, where RT2016 make erroneous claims about theoretical inferiority/superiority of the attribute regret functions embedded in two particular RRM models, they fail to acknowledge that these regret functions are special cases of a more recently proposed, much richer and more generic model (μRRM). In fact, this more generic model allows for the empirical estimation—as opposed to ‘theoretical’ assertion—of the degree to which regret weighs more heavily than rejoice (which is conceptually similar to estimating a Loss Aversion parameter in Prospect Theory inspired models). This makes the theoretical discussion in RT2016, as far as the shape of the regret function is concerned, largely redundant (in addition to being flawed, as explained further above).

Empirical comparisons based on only one dataset and two model specifications

This brings us to the topic of empirical analyses. Although in their early days, RRM models were often compared to linear in parameter RUM models based on one or a couple of datasets (also by the first author of this commentary), it has since been well established (e.g. Hess et al. 2012; Boeri et al. 2014; Hess et al. 2014; Chorus et al. 2014; Hess and Chorus 2015) that the relative performance of—different specifications of—RRM models varies widely across data sets, and even across different classes of individuals within the

⁷ But note that the total regret (i.e., aggregated over all attributes) as computed in μ RRM model presupposes that comparisons with all competing alternatives are relevant, whereas RRM2008 only focuses on the comparison with the best competing alternative. For the present discussion, this distinction is irrelevant.
same data set. Several dozens of empirical comparisons between RRM (in various specifications) and RUM which have been reported in the literature over the past few years, further illustrate this point. Regarding the empirical performance of RRM models with different shape of the attribute regret function, Van Cranenburgh et al. (2015) use ten different data sets to confirm that also this aspect is highly data set specific.

For further illustration, we here present new results, in the context of ten different data sets (i.e., the same ones that were considered in Van Cranenburgh et al. 2015). We provide model fit (final LogLikelihood) for a variety of RRM models, including the 2008RRM and 2010RRM models. Table 1 lists, from left to right, the choice context, and model fit for: a linear in parameters RUM model, RRM2008, RRM2010, the G-RRM model, and the $\mu$RRM model (all in Multinomial Logit form). Results can be summarized as follows:

- Concerning the comparison between RRM2008 and RRM2010, which is the topic of RT2016, it turns out that RRM2008 (the model preferred by RT2016) has a better model fit on only two out of ten data sets. RRM2010 fits the other eight data sets better. If one would like to base a generic behavioural conclusion on these analyses (something we would advise against), that conclusion would be that RRM2010 is to be preferred over RRM2008 as a model of decision making.

- Taking also the other models into account, it is directly seen that model fit varies greatly. Unsurprisingly (as it is the most generic model at the cost of one extra parameter), the $\mu$RRM has the best fit on all but one data sets. One some data sets, large values for $\mu$ are found, implying no regret aversion (and linear additive RUM behaviour). On other data sets, values close to zero are obtained, implying extreme levels of regret aversion. The special case of 2010RRM (i.e., $\mu \approx 1$) is obtained for two data sets. In short, results indicate that the shape of the regret function, estimated by means of a regret aversion parameter, varies greatly across data sets.

In light of these new results, as well as results reported in the literature over the past few years, it may be said that the empirical analysis of RT2016 certainly provides another data point to the discussion, but certainly does not warrant drawing conclusions as to the relative performance of different RRM models.

In sum, if one wants to empirically compare different specifications of RRM, our advice would be to not do so, based on a single dataset. Furthermore, we argue that in light of recent developments (see previous section) a comparison between only the special cases embodied by RRM2008 and RRM2010 is not particularly insightful. Rather, we advise to estimate the generic $\mu$RRM model, to consider each data set in separation, and to refrain from drawing generic conclusions as to the (estimated) shape of the regret function—it will generally differ a lot across data sets. In their Conclusions and Discussion section, RT2016 do mention that replication of their results on other data sets is needed, but we believe that this should have been done in the paper itself, and also that other regret functions than just the two special cases looked at in RT2016 should have been considered.

---

8 We use the exact same specifications as used in RT2016 (Eqs. 1, 2 and 6). That is, for 2008RRM we use the max-operator in the attribute regret function, and to compute total regret we only consider the best competitor alternative. For 2010RRM we use the so-called LogSum-operator in the attribute regret function, and to compute total regret we consider all competitor alternatives.
Conclusion

As said in the Introduction, we most welcome theoretical and empirical comparisons between different decision rules embedded in choice models, including but not limited to comparisons between regret-based models and utility-based models, and between different types of regret-based models (as in RT2016). We believe such studies to have a lot of potential to enrich our knowledge of choice models and ultimately, of human decision making behaviour; leading to better forecasting and more informed policy making.

However, such comparisons need to be done carefully, with proper acknowledgement of the basics of discrete choice theory and the state-of-the-art in regret based modelling, and based on a sufficiently rich empirical context. This commentary explains where RT2016 falls short in these regards and it tries to set the record straight, with the aim to avoid confusion caused by erroneous claims.

To summarize

- RT2016 claims that the RRM model proposed in Chorus (2010) has theoretical problems, is ill-founded and is theoretically inferior to another RRM model proposed in Chorus et al. (2008). We show that these claims are unjustified, as they are based on a misconception of Discrete Choice Theory in general and Regret Theory in particular.
- RT2016 compares two RRM models which were proposed more than 6 years ago, while more recent work has convincingly shown that the attribute regret functions embedded in the two considered models are special (extreme) cases of a more generic RRM model; what’s more, it is now widely acknowledged that the shape of the regret function can and should be estimated from choice data, rather than fixed by the analyst a priori as in RT2016.
- The empirical performance of models which embed different shapes of the regret function, varies a lot across data sets. Therefore, RT2016’s conclusion that one particular model form (RRM2008) performs better than another one (RRM2010) in the context of just one data set, is not of much value. Our own analyses based on ten data sets and five model types, including RRM2008 and RRM2010, show that these claims are unjustified, as they are based on a misconception of Discrete Choice Theory in general and Regret Theory in particular.

<table>
<thead>
<tr>
<th>Choice content</th>
<th>Model fits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Linear-additives</td>
</tr>
<tr>
<td>Parking lots</td>
<td>-406</td>
</tr>
<tr>
<td>Shopping dest</td>
<td>-2305</td>
</tr>
<tr>
<td>Departure time</td>
<td>-795</td>
</tr>
<tr>
<td>Policies</td>
<td>-240</td>
</tr>
<tr>
<td>Routes</td>
<td>-1092</td>
</tr>
<tr>
<td>Routes</td>
<td>-1272</td>
</tr>
<tr>
<td>Car types</td>
<td>-4193</td>
</tr>
<tr>
<td>Routes</td>
<td>-4024</td>
</tr>
<tr>
<td>Routes</td>
<td>-2613</td>
</tr>
<tr>
<td>Dating</td>
<td>-3688</td>
</tr>
</tbody>
</table>

Table 1 Model fit results for ten data sets and five model types, including RRM2008 and RRM2010.
different data sets suggest that while RRM2008 performs better than RRM2010 on two data sets, it performs worse on the other eight. More generally speaking, we find—as expected—that the estimated shapes of attribute regret functions differ substantially across data sets, implying that one should be extremely cautious to generalize results based on just one data set.

The aim of this commentary was to resolve confusion that emerges from the misconceptions and misinterpretations presented in RT2016. It is certainly not meant to discredit the authors of that paper. On the contrary, we note that Soora and Harry at the end of their paper indicate that they plan to continue their work on regret modelling, and we look forward to collaborate with them on this fascinating topic, following up on previously successful collaborations like the one which resulted in the RRM2008 model (Chorus et al. 2008).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix: Notation of different RRM models discussed in this commentary

The total (systematic) regret associated with a considered alternative is given as follows, for the different RRM models discussed in this commentary:

2008RRM (Chorus et al. 2008)

\[ R_i = \max_{j \neq i} \left\{ \sum_m \max \{ 0, \beta \cdot (x_{jm} - x_{im}) \} \right\} \]

2010RRM (Chorus 2010)

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

G-RRM (Chorus 2014a)

\[ R_i = \sum_{j \neq i} \sum_m \ln \left( e^m + \exp \left[ \beta_m \cdot (x_{jm} - x_{im}) \right] \right) \]

\( \mu \)RRM (van Cranenburgh et al. 2015)

\[ R_i = \sum_{j \neq i} \sum_m \mu_m \cdot \ln \left( 1 + \exp \left( \frac{\beta_m \cdot x_{jm} - x_{im}}{\mu_m} \right) \right) \]

Here, \( R \) denotes total (systematic) regret, \( i \) denotes the considered alternative, \( j \) a competitor alternative, \( m \) an attribute, \( x \) an attribute level, and \( \beta \) a taste parameter. In the G-RRM model, \( \gamma \) represents an estimable regret aversion parameter which may take on values between zero and one. In the \( \mu \)RRM model, \( \mu \) represents an estimable regret aversion parameter which may take on values between zero and \( +\infty \). See main text, as well as cited papers, for in-depth discussions of models arising from different values for these regret aversion parameters.
Note that the above notation only refers to systematic or ‘observed’ regret; depending on the chosen error term specification, Multinomial or Mixed Logit choice models are obtained.

References

Prato, C.G.: Expanding the applicability of random regret minimization for route choice analysis. Transportation 41(2), 351–375 (2014)
Rasouli, S., Timmermans, H.: Applications of theories and models of choice and decision-making under conditions of uncertainty in travel behavior research. Travel Behav. Soci. 1(3), 79–90 (2014)

Caspar G. Chorus is Professor of choice behavior modeling, and head of the Transport and Logistics group, at TU Delft. His main research aim is to increase the behavioral realism of (travel) choice models, and by implication to provide a stronger foundation for (transport) policy development.

Sander van Cranenburgh is Assistant Professor at the Transport and Logistics Group, Delft University of Technology. His main research areas are Discrete choice modelling and Large-scale travel demand models. He is specifically interested in exploring new methodologies and developing new estimation techniques, ultimately improving insights on travel behaviour.