# Does the decision rule matter for large-scale transport models?

van Cranenburgh, Sander; Chorus, Caspar G.

**DOI**
10.1016/j.tra.2018.01.035

**Publication date**
2018

**Document Version**
Final published version

**Published in**
Transportation Research Part A: Policy and Practice

Transportation Research Part A 114 (2018) 338–353

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

# Does the decision rule matter for large-scale transport models?☆

Sander van Cranenburgh*, Caspar G. Chorus

*Delft University of Technology, Jaffalaan 5, 2628 BX Delft, The Netherlands*

ARTICLE INFO

*Keywords:*
Large-scale transport models
Random Regret Minimization
Discrete choice modelling
Decision rules

ABSTRACT

This paper is the first to study to what extent decision rules, embedded in disaggregate discrete choice models, matter for large-scale aggregate level mobility forecasts. Such large-scale forecasts are a crucial underpinning for many transport infrastructure investment decisions. We show, in the particular context of (linear-additive) utility maximization (RUM) and regret minimization (RRM) rules, that the decision rule matters for aggregate level mobility forecasts. We find non-trivial differences between the RUM-based and RRM-based transport model in terms of aggregate forecasts of passenger kilometers, demand elasticities, and monetary benefits of transport policies. This opens up new opportunities for policy analysts to enrich their sensitivity analysis toolbox.

## 1. Introduction

Large-scale transport models are typically built on disaggregate discrete choice models based on linear-additive Random Utility Maximization (RUM) decision rules (e.g. de Jong et al., 2007; Hess et al., 2007). However, despite the strong foundations of such RUM-based discrete choice models in micro-economic theory and their computational tractability there has been a rapidly growing interest in the development of non-RUM discrete choice models within the travel behavior research community. There is a growing consensus now, that these 'behavior inspired' choice models form a useful addition to the toolbox of travel demand modelers, by capturing behavioral phenomena such a boundedly rational, semi-compensatory decision-making and choice set composition effects (Chorus et al., 2008; Hess et al., 2012; Leong and Hensher, 2012; Guevara and Fukushi, 2016).

However, despite this growing interest in non-RUM choice models in the travel behavior research community these models have not yet been implemented in large-scale transport models to forecast macro-level mobility patterns. Instead, the literature has predominantly focused on comparisons between RUM and non-RUM choice models at the disaggregate level, e.g. in terms of differences in model fit in the context of a given dataset. As a consequence, at present it is unknown whether large-scale transport models based on non-RUM choice models would in fact produce different aggregate level predictions than their counterparts based on RUM models. Crucially, these aggregate level forecasts, rather than micro-level analyses, form the basis for many transport policies and infrastructure investment decisions. This implies that there is currently no empirical ground for the often voiced expectation that increasing the behavioral realism of micro-level travel behavior models via the use of non-RUM choice models would lead to different (and perhaps improved?) aggregate level mobility forecasts, and, as a consequence, to different (and perhaps better informed?) transport policy making. This lack of evidence regarding the usefulness of employing non-RUM choice models for large-scale travel demand forecasting is currently considered one of the main limitations of these models (Chorus, 2014).

☆ 22nd International Symposium on Transportation and Traffic Theory.

* Corresponding author.
*E-mail address:* s.vancranenburgh@tudelft.nl (S. van Cranenburgh).

https://doi.org/10.1016/j.tra.2018.01.035

Available online 21 February 2018

Motivated by the significant scientific and societal relevance of this ‘aggregate-disaggregate gap’ in the travel behavior modelling literature, we[1] have developed what we believe is the world’s first discrete choice-based large-scale transport model built on a non-RUM decision rule. Specifically, we devised a Random Regret Minimization (RRM) based counterpart of the internationally renowned Dutch National Transport model (henceforth abbreviated as LMS, for ‘Landelijk Model Systeem’). RRM models are non-RUM discrete choice models built on the notion that regret can be an important co-determinant of choice behavior (Chorus, 2010). RRM models are designed to accommodate for semi-compensatory and reference-dependent choice behaviours, such as the compromise effect (Chorus and Bierlaire, 2013). Since their relatively recent introduction these RRM models have found their way to leading textbooks (Hensher et al., 2015) and software packages (e.g. Greene, 2012; Vermunt and Magidson, 2014), and have been used in a wide variety of travel behavior studies during the past few years (Chorus et al., 2014).

This study presents the first aggregate level comparison of mobility forecasts produced by RUM and non-RUM (i.e., RRM) based large-scale transport models. As a case study, we investigate a fictive “High Frequency Rail” (HFR) policy scenario. In this scenario train frequencies are substantially intensified as compared to the reference scenario. In the context of this policy scenario, we analyse and compare the predictions of the RUM-based and RRM-based LMS, in terms of various relevant mobility indicators such as the predicted number of tours, the predicted tour-length, and the total passenger kilometers per mode of transport; all at the national (Dutch) level.

As a secondary, methodological contribution, we introduce a technique which allows for very substantial computation time savings when estimating so-called P-RRM models (Van Cranenburgh et al., 2015a) on data sets characterized by large choice sets, as is common in transportation (e.g., destination, route choice models).

The remaining part of this paper is organized as follows. Section 2 gives a brief description of the LMS and presents the steps taken to develop the RRM-LMS. Section 3 presents baseline year forecasts and elasticities of demand, and compares the results of the RUM-LMS with those of the RRM-LMS. Section 4 presents our case study. We analyse and compare the predictions of the RUM-LMS and the RRM-LMS in the context of a HFR policy scenario.

## 2. Development of the RRM-LMS

### 2.1. The LMS in a nutshell

The LMS is a nation-wide model system for The Netherlands. It was developed in the 1980s for long-term strategic policy analysis. Nowadays, its use is obligatory (in the Netherlands) for appraisal of large transport infrastructure. Like many national transport models developed in Europe, the LMS is a tour-based model system, although in the assignment model the tours are decomposed into unconnected trips. The model operates on a national level comprising of 1406 transport analysis zones and differentiates between 9 travel purposes: Commute, Business, Education, Shopping, Other, Work-Business, Work-Other, Child-Education, and Child-Other. Furthermore, it makes use of a detailed segmentation of the population.

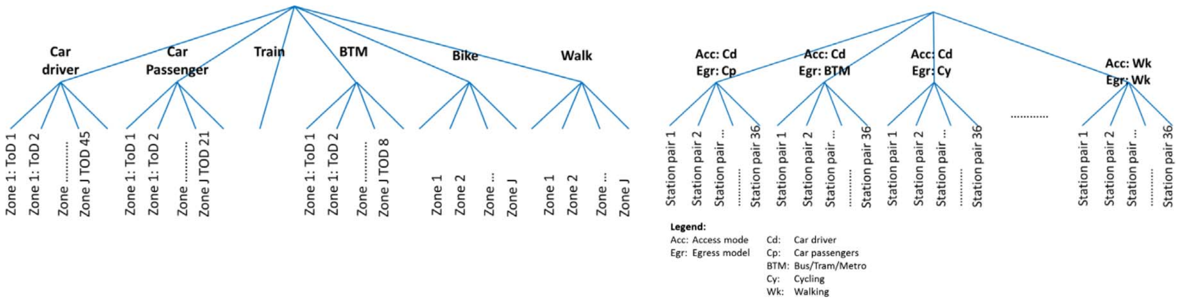
Several discrete choice models are embedded in the LMS, e.g. to model Car ownership, Tour generation, Mode and destination choice, and Departure time choice. These models operate at the household level or at the person level, and differentiate between travel purposes (except the car ownership model). The assignment model is not based on discrete choice models. The discrete choice models in the LMS are estimated in a Non-Normalized Nested Logit (NL) form (Daly and Zachery, 1978; Daly, 1987). As such, inclusive values or LogSums (LS) carry information about the decisions made on the lower levels to upper levels, in a sequential procedure, see e.g. Ortúzar and Willumsen (2011) for more details on Nested Logit specifications.

The core of the LMS is the forecasting system, called SES (Sample Enumeration System). This module samples individuals from the data, and computes choice probabilities for each individual in this sample. For each Transport Analysis Zone (TAZ) the sample is reweighted. SES consists of two primary sub modules: BASMAT and GM. Module BASMAT generates the base matrices based for each time-of-day and travel motive. Module GM determines the growth in travel demand for the future year relative to the base year. A pivot-point procedure is then used to construct the forecasted OD matrices. This pivot-point procedure enhances the accuracy of the model system’s forecasts as model systems are usually better in predicting changes relative to a base-year situation than in predicting absolute numbers (Daly et al., 2005).

The GM sub module consists of Tour generation and Mode-Destination-Time-of-Day (MD-ToD) discrete choice models. With the exception of the Time-of-Day choice, the choice models within the LMS are estimated on Revealed Preference (RP) data obtained by the Dutch National travel survey called MON.[2] These data are collected on a yearly basis in The Netherlands with the aim to provide insights on the daily mobility behavior. Each survey wave comprises of over 40,000 Dutch residents (CBS, 2010). To estimate the choice models data from 3 survey waves are used: 2008, 2009, and 2010. The Time-of-Day models are estimated on Stated Preference (SP) data, rather than on RP data. The SP data consist of more than 1 000 respondents which were recruited from an existing panel or from short recruitment interviews at Dutch train stations and at a petrol station beside a motorway (see de Jong et al., 2003 for more details on the survey).

[1] In close collaboration with Significance consultancy (Marits Pieters, Jaap Baak, Gerard de Jong) and The Netherlands road authority/Rijkswaterstaat (Frank Hofman).

[2] Recently, the Dutch Mobility survey is renamed into OViN.




Predominant Nesting structure of the MD-ToD models Nesting structure of the Train station choice models.

**Fig. 1.** Nesting structures.

### 2.2. Developing the RRM-LMS

To develop the RRM-based LMS, all RUM-based discrete choice models in the GM module are replaced by RRM-based counterparts. Estimations were conducted based on exactly same data as were being used for estimation of the RUM-models. Moreover, we used exactly the same model specifications in terms of explanatory variables, parameterizations and nesting structures as were used in the RUM-LMS. By doing so, we ensured that in case differences between the RUM-LMS and the RRM-LMS are found, these can be attributed to the differences in the underlying decision rule (rather than potentially being the result of differences in model specifications). However, note that due to this choice we expect to find higher model fits for the RUM models than for the RRM discrete choice models. After all, when developing a large-scale model for each choice model various utility specifications are tested (e.g. in terms of – interactions between – explanatory variables and nesting structures), and the specification having the highest model fit is usually chosen. Hence, the utility/regret specifications in both the RRM-LMS and the RUM-LMS are optimized for RUM.

Since the tour-generation models and the car-ownership models in the LMS are based on binary logit specifications, there is no point in replacing these models into RRM-based counterparts. After all, differences between RUM behavior and RRM behavior only manifest in the context of three or more alternative (Chorus, 2010). As such, to develop the RRM-LMS we needed to replace the RUM MD-ToD choice models with RRM counterparts.

The left hand-side plot in Fig. 1 shows the predominant nesting structure that is used in the MD-ToD models. As shown, mode choice is in the upper level, and the destination choice and the ToD choice are at the same level below the mode choice. Time-of-day choices are modelled at different levels of resolution, depending on the mode of transport. Since the LMS was originally developed with a particular focus on forecasting car travel demand, Car time-of-day choice is modelled at the highest resolution. Specifically, for mode Car 45 combinations of departure time and arrival time are modelled, while for BTM 8 combinations of departure time and arrival time are modelled. For modes Bike and Walk departure and return time combinations are not modelled at all.

A separate choice model is estimated for mode Train. This model predicts the choices for the departure and arrival train stations jointly with the egress and access modes. To identify the available departure and arrival train stations a relatively simple heuristic is used based on geometric search distances (see Significance, 2012 for more details). At urbanized places up to six candidate train (departure or arrival) stations are within reach. Therefore, up to 36 train station pairs are available, depending on the origin and destination zones. Furthermore, the station choice models encompass 5 access modes (Car driver, Car passenger, Bus/Metro/Tram, Bike and Walk), and 4 egress modes (Car passenger, Bus/Metro/Tram, Bike, and Walk), leading to up to 19 feasible mode-pairs. Accordingly, in total the station choice models consist of $36 \times 19 = 684$ alternatives.

The right-hand side plot in Fig. 1 shows the nesting structure of the Train station choice models. Access and egress mode combinations are at the top-level of the nesting structure. At the level below are the station pairs.

In the LMS the Train station model and the MD-ToD model are estimated sequentially, rather than simultaneously using Full Information Maximum Likelihood estimation. Thereby, computational time is reduced. In Nested Logit models the behavioral relationships between choices at each level of the nest are captured via the inclusive value (the LogSum) and its associated nest parameter. The inclusive value is basically an index of the expected maximum utility from the choice of alternatives at the lower-level (s) of the nesting tree.

To calibrate the MD-ToD and the Station models, first the Train station choice model is estimated. The inclusive value from the Train station choice model is then used to estimate the MD-ToD model. This procedure is repeated until the nest parameter of the MD-ToD model is sufficiently close[3] to the 'pre-set' nest parameter used in Train station choice model. For most travel purposes, 'convergence' occurred after five to ten rounds for both the RUM-LMS and RRM-LMS.

### 2.3. RRM model specifications

For reasons of space limitations and to avoid repetition, we in this paper only discuss those aspects of RRM models which are

[3] That is, the difference between the two $< 0.005$.

essential to our analyses; more detail on RRM models can be found in papers cited below. In RRM models decision makers are assumed to choose the minimum regret alternative. Regret is postulated to be experienced when a competitor alternative $j$ outperforms the considered alternative $i$ with regard to one or more attributes $m$. The overall regret of an alternative is typically conceived to be the sum of all the pairwise regrets that are associated with bilaterally comparing the considered alternative with the other alternatives in the choice set.

The predominant mathematical form of RRM models is given in Eq. (1), where $RR_{in}$ denotes the random regret for decision maker $n$ considering alternative $i$, $R_{in}$ denotes the observed part of regret, and $\varepsilon_{in}$ in denotes the unobserved part of regret. In the core of RRM models is the so-called attribute level regret function: $r_{ijmn} = f(\beta_m x_{jmn} - x_{imn})$. This function maps the difference between the levels of attributes $m$ of the competitor alternatives $j$ and the considered alternative $i$ onto regret.

$$RR_{in} = R_{in} + \varepsilon_{in} \quad where\ R_{in} = \sum_{j \neq i} \sum_{m} r_{ijmn} \tag{1}$$

After the first RRM discrete choice model was proposed in Chorus et al. (2008), various new types of RRM models have been developed, such as the G-RRM, μRRM, and P-RRM model (see Van Cranenburgh and Prato, 2016 for a recent overview). The choice models in the RRM-LMS are estimated in a P-RRM form (Van Cranenburgh et al., 2015a,b), see Eq. (2). The principal reason for using this model, rather than, for instance, the more frequently encountered RRM model proposed in Chorus (2010) is that the P-RRM model postulates the strongest degree of regret minimizing behavior within the RRM modelling framework. Therefore, by opting for the P-RRM model we maximize the chance of observing aggregate level differences between the RRM-LMS and RUM-LMS (if these exist). This, in turn, enables us to acquire insights on the extent to which aggregate mobility forecasts of RUM-based large-scale transport models (*in casu*: the RUM-LMS) are robust with regard to the specification of the underlying decision rule.

$$R_{in} = \sum_{m} \sum_{j \neq i} \max(0, \beta_m [x_{jmn} - x_{imn}]) \tag{2}$$

The RRM models are estimated in a hybrid RRM-RUM specification (Chorus et al., 2013). In this hybrid specification dummy variables and Alternative-Specific Constants (ASCs) are modelled as RUM, whereas generic attributes (such as, travel cost, travel time, travel distance, etc.) are modelled as RRM, see equation 3. Because of the binary nature of dummy variables and ASCs, treatment of these variables is mathematically equivalent under RUM and RRM (apart from a non-linear transformation which has no impact on model fit or on the behavior that is imposed by the model) (Chorus, 2012a; Hess et al., 2014). However, from a computational perspective modelling dummy variables and ASCs in a RUM way is preferred: it strongly reduces computational efforts since no pairwise comparisons need to be computed. Additionally, it eases comparison of parameter estimates associated with dummy variables and ASCs across RUM and RRM models.

$$UR_{in} = V_{in} - R_{in} \tag{3}$$

Specifically, Eq. (4) gives the hybrid Utility-Regret function that is estimated for alternatives involving all modes except Train, where $D_1$ to $D_k$ denote dummy variables (typically associated with socio-demographic characteristics, such as age and education level). For alternatives reached by mode Train, the Utility-Regret function is given by Eq. (5). Since the Train station choice model is not simultaneously estimated (see Section 2.2) the LogSum (LS) enters the utility/regret function in the MD-ToD models.

The RRM LS – which equals the expected minimum regret, see Eq.[4] (6) – is fundamentally different from its RUM counterpart (Chorus, 2012b). One key aspect in which the RRM LS differs from the RUM LS is that it is not monotonous. That is, an improvement (e.g. a travel time reduction) does not always result in a reduction of the expected minimum regret. Although the properties of the RRM LS align well with regret psychology (see Chorus, 2012a for a discussion), they seem to be less well aligned with consistency criteria usually required for policy evaluation. To be more specific, recent studies (Dekker, 2014; Dekker and Chorus, 2017) have shown that conventional approaches to obtain Value-of-Time (VoT) estimates (i.e., taking the ratio of partial derivatives of the value function with respect to time and cost) and Consumer Surplus estimates (i.e., using the difference between the Logsum before and after a policy) do not apply in the context of RRM models, and neither in the context of non-linear (in price) and context dependent RUM models. This is due to a combination of reasons, two of which in particular hamper a consistent connection between RRM and welfare economics: first, because regret is determined in terms of attribute *differences*, there is no such thing as a 'regret of income'. For example, if all alternatives in a choice set would be made $x$ euro cheaper – which implies an increase in income of $x$ euro – this does not affect regret levels. As a consequence, it is impossible to translate the impact of a policy on regret into a monetary value thereof. Second, in contrast to standard RUM models, RRM models postulate that when an attribute of one alternative is changed (e.g. a travel time decrease as a result of a policy), this will affect all alternatives' regret levels. As such, taking the partial derivative of regret with respect to a change in travel time and cost does not fully capture the impact of these changes on regret levels experiences by the traveler. Although a more in-depth discussion of these issues – and of approaches that help overcome some of them – is beyond the scope of this paper, they should be kept in mind when assessing the value of RRM as a framework for policy analyses. For more details, the reader is referred to the above cited papers; in the remainder of this paper, we will focus on generating forecasts using RRM based large scale models (and in the sub-section focusing on Cost Benefit Analysis, we will use simple rules of thumb to derive welfare effects).

[4] The constant C in Eq. (6) represents the fact that the absolute value of (expected) regret cannot be measured (just as in RUM models).

$$UR_i = ASC_{\text{mode}} + D_1 + ...+D_k - \sum_m \sum_{j\neq i} \max(0,\beta_m\cdot[x_{jm}-x_{im}]) \tag{4}$$

$$UR_i = ASC_{\text{train}} + \beta_{TrLogSum}\cdot LS_i^{RRM} \tag{5}$$

$$LS_i^{RRM} = -\ln\left(\sum_j e^{-R_j}\right) + C \tag{6}$$

Lastly, in RRM models it is necessary to account for the choice set size when that varies across observations. As explained in Van Cranenburgh et al., 2015b), in RRM models (except for the model proposed in Chorus et al. (2008), which does not include comparisons with all alternatives in the choice set) regret level differences increase with increasing choice set size. As a consequence, not accounting for variation in choice set size results in behaviorally unrealistic forecasts and inferior performance of RRM models in the context of data sets with varying choice set sizes. Therefore, we used the choice set size correction factor proposed in Van Cranenburgh et al. (2015b): $\widetilde{R}_{in} = (1/J_n)R_{in}$, where $\widetilde{R}_{in}$ denotes the choice set size corrected regret of alternative $i$ for observation $n$, and $J_n$ denotes the choice set size of observation $n$.

### 2.4. Computational aspects of estimating RRM models

A key aspect when developing large-scale transport models is to maintain reasonable computational efforts (Daly and Sillaparcharn, 2000). This is also, and particularly, the case for the development of the RRM-LMS. The dominant factor for the computational effort in RRM models is the choice set size. As the choice set size is typically large in large-scale transport models (it may involve thousands of alternatives), this poses a real challenge. Section 2.4.1 elaborates on computational aspects of RRM models in the context of large-scale transport models. Section 2.4.2 presents a new method that facilitates fast estimation of P-RRM models in the context of large-scale transport models.

#### 2.4.1. Computational aspects of estimating RRM models

Estimation time of RRM models in the context of large-scale applications may be exceedingly high. This is a direct consequence of the behavioral postulate in RRM models[5] that every alternative is compared with every other competitor alternative that is present in the choice set. Guevara et al. (2016) empirically illustrates that the time to estimate RRM models grows quadratically with the choice set size. In contrast, in RUM models estimation time is (roughly) linear with choice set size. While the quadratic growth in estimation time is insignificant for choice situations in which the choice set size is small to moderately large (e.g. up to 20 alternatives), in the context of large-scale transport models this causes severe computational challenges.

To illustrate the impact of choice set size in RRM models on estimation time, consider the Mode-Destination choice model of the LMS. This choice model comprises of 1406 destinations and 6 modes of transport, leading to a total of 8436 alternatives. Therefore, computing the regret levels for all 8436 alternatives involves evaluating $8436 \times 8435 \approx 72$ million pairwise comparisons per observation per attribute, at each iteration step of the estimation process. It goes without saying that this vastly exceeds current computational resources.

However, specifically for the P-RRM model estimation times can be reduced to a very substantial extent. Van Cranenburgh et al. (2015a,b) show that to estimate P-RRM models, the pairwise comparisons only need to be evaluated once, prior to the estimation, rather than multiple times (i.e., at each iteration step during the estimation). By doing so, substantial amounts of estimation time can be saved. As such, this model is much better suited for large-scale applications than most other RRM models. More specifically, Van Cranenburgh et al. (2015a,b) show that when the signs of the taste parameters $\beta_m$ are known before estimation (which is usually the case in transport contexts), the betas in Eq. (2) can be placed outside the max operator, see Eq. (7), where $\beta_m^+$ denotes positive taste parameters, and $\beta_m^-$ denotes negative taste parameters. As a result, the terms within the second summations are no longer a function of the parameters that need to be estimated, and therefore can be computed prior to the estimation. When the P-RRM attribute levels are computed prior to estimation the P-RRM model takes a linear-additive form, see Eq. (8). This makes the model relatively easy to estimate as the log-likelihood function is globally concave (under standard MNL error term assumptions).

$$R_{in} = \sum_m \beta_m^+ \sum_{j\neq i} \max(0,[x_{jmn}-x_{imn}]) + \sum_m \beta_m^- \sum_{j\neq i} \min(0,[x_{jmn}-x_{imn}]) \tag{7}$$

$$R_{in} = \sum_m \beta_m \bar{x}_{imn} \quad \textit{where } \bar{x}_{imn} = \begin{cases} \sum_{j\neq i} \max(0,x_{jmn}-x_{imn}) & if \beta_m > 0 \\ \sum_{j\neq i} \min(0,x_{jmn}-x_{imn}) & if \beta_m < 0 \end{cases} \tag{8}$$

#### 2.4.2. A computationally efficient method to compute P-RRM attribute levels

Even in case the pairwise comparisons need to be computed only once the computational efforts to do so can still be exceedingly high. Therefore, this subsection presents a computationally efficient method to compute P-RRM attribute levels. To achieve this aim,

[5] of the form of Eq. (1).

we capitalize on the linear-additive nature of the P-RRM model. This enables us to construct the P-RRM attribute levels orders of magnitudes faster than when using a naïve approach in which all pairwise comparisons are computed and summed consecutively.

Suppose that the choice set in choice observation $n$ consists of $J_n$ alternatives. Furthermore, suppose the signs of the taste parameters $\{\beta_1 ... \beta_M\}$ are known to the analyst. As the analyst knows the signs of the taste parameters, the alternatives can be sorted, from the best to the worst based on their performance on each of the attributes (e.g. highest level of comfort to lowest level of comfort). Note that sorting is, numerically speaking, a relatively 'cheap' mathematical operation.

Let $k$ denote the rank of the alternatives with respect to attribute $m$, where alternative $k = 1$ is the best performing alternative (e.g. highest level of comfort) and alternative $k = K$ is the worst performing alternative (e.g. lowest level of comfort). Then, the P-RRM attribute vector of alternative $k$ caused by attribute $m$, denoted $\bar{x}_{kmn}$, for alternatives $k = 1 \dots K$ is given in Eq. (9). Note we dropped the subscript $n$ from legibility. As can be seen, the P-RRM attribute vector equals zero for the best performing alternative. This is correct in the sense that no regret is experience by the best performing alternative in terms of attribute $m$.

$$\bar{x}_{1m} = 0 \; \bar{x}_{2m} = [x_{1m}-x_{2m}] \; \bar{x}_{3m} = [x_{1m}-x_{3m}] + [x_{2m}-x_{3m}] \; \vdots \; \bar{x}_{km} = [x_{1m}-x_{km}] + [x_{2m}-x_{km}] + \dots + [x_{k-1m}-x_{km}] \; \vdots \; \bar{x}_{Km} = [x_{1m}-x_{Km}] + [x_{2m}-x_{Km}] + \dots + [x_{K-1m}-x_{Km}] \tag{9}$$

Hence, the P-RRM attribute level of rank-ordered alternative $k$ with regard to attribute $m$ takes the following form (Eq. (10)).

$$\bar{x}_{km} = \sum_{w=1}^{k-1} [x_{wm}-x_{km}] \tag{10}$$

Next, we eliminate all redundant pairwise comparisons. Due to the linear nature, each pairwise comparison $[x_{wm}-x_{km}]$ can be constructed as a linear combination of either one or two 'principle pairwise comparisons' having the following form: $[x_{1m}-x_{wm}]$.[6] This implies that we can reduce the number of pairwise comparisons that need to be computed from $J_n \cdot (J_n-1)$ to $(J_n-1)$. Capitalising on this property, the attribute level of rank-ordered alternative $k$ with regard to attribute $m$ is given in Eq. (11).

$$\bar{x}_{km} = \left\{ (k-1)[x_{1m}-x_{km}] - \sum_{w=2}^{k-1} [x_{1m}-x_{wm}] \right\} \tag{11}$$

To test the improvement in computational performance of the proposed method a synthetic data set was created. This data set consisted of $N = 100$ choice observations. Alternatives comprised of just one attribute. Attribute levels $x_{imn}$ were randomly generated by taking draws from the unit interval. We found that the computationally efficient method is several orders of magnitude faster than a naïve approach. Using the presented method it takes about 36 s to compute the vector of P-RRM attribute levels for a data set consisting of 10,000 alternatives. In contrast, it takes about 24 h to obtain the same vector using a naïve approach. In fact, using the presented method it is still technically feasible to compute the P-RRM attribute levels even for the largest applications in current state-of-art large-scale transport models (which typically consist of about 100,000 alternatives).

### 2.5. Estimation result

This subsection presents estimation results for each of the choice models that were replaced by P-RRM counterparts in the process of developing the RRM-LMS. To enhance interpretations, estimation results for the P-RRM models are presented alongside with the original linear-additive RUM results. Section 2.5.1 presents results for the station choice model; Section 2.5.3 presents results for the MD-ToD choice models.

#### 2.5.1. Station choice models

Table 1 shows the summary statistics of the RUM and P-RRM Station choice estimation results. In the Station choice models Travel purposes Commute and Business, and Shopping and Other are lumped together. Therefore, in total three Station choice models are estimated. Table 1 reveals that for all three travel motives the linear-additive RUM models fit the data better. Taking the large number of observations into account, for travel purposes Commute & Business and Education the difference in model fit is relatively small: $\Delta LL/obs < 0.1$. However, the performance in terms of model fit of the P-RRM model for travel purposes Shopping & Other is considerably poorer than that of the RUM model. Again, note that this was expected, since – for reasons of allowing for consistent model comparisons – the model specification was optimized for RUM, and then copied into the P-RRM model.

To enhance interpretation of the outcomes of this research, we discuss the estimation results in detail for one travel purpose, namely: Commute & Business (Table 2). Full estimation results for the other travel purposes can be found in Significance (2016). Based on Table 2 a number of inferences can be made. Firstly, the signs of the parameter estimates are consistent across the RUM and P-RRM models. Secondly, the relative sizes are consistent for most parameters across the two models. However, as there is no straightforward conceptual connection between RUM and RRM parameters it is not meaningful to interpret these differences in a behavioral (or statistical) sense. Thirdly, and perhaps most noteworthy, the parameter that captures the effect of the number of connecting stations from the departing and the arrival stations: 'Connections' is, relatively speaking (i.e., compared to other parameters), substantially larger in the P-RRM model than in the RUM model. This suggests that the number of Connections is, relatively speaking, more important in the P-RRM model than in the RUM model.

[6] For instance: $[x_{3m}-x_{4m}]$ can be constructed as $[x_{1m}-x_{4m}]-[x_{1m}-x_{3m}]$.

**Table 1**
Estimation results station choice models.

| | Commute & Business | Education | Shopping & Other |
|---|---|---|---|
| No. observations | 791 | 415 | 165 |
| No. parameters | 21 | 19 | 17 |
| Null LL | −4725 | −2434 | −985 |
| LL RUM | −2182 | −955 | −391 |
| LL RRM | −2229 | −973 | −448 |
| $LL_{RUM} - LL_{RRM}$ | 48 | 19 | 57 |
| ΔLL/obs | 0.060 | 0.045 | 0.346 |

**Table 2**
Estimation results train station choice models (Purpose: Commute).

| MODEL | | RUM Nested Logit | | | | P-RRM Nested Logit | | | |
|---|---|---|---|---|---|---|---|---|---|
| No. observations | | 791 | | | | 791 | | | |
| No. parameters | | 21 | | | | 21 | | | |
| Null Log-likelihood | | −4725 | | | | −4725 | | | |
| Final Log-likelihood | | −2182 | | | | −2229 | | | |
| $\rho^2$ | | 0.54 | | | | 0.53 | | | |
| | | Est | Std err | *t*-val | p-val | Est | Std err | *t*-val | p-val |
| *Generic parameters* | | | | | | | | | |
| | AccessTime_Car | −0.14 | 0.013 | −11.0 | .00 | −0.10 | 0.011 | −8.5 | .00 |
| | AccessTime_BTM | −0.06 | 0.007 | −9.3 | .00 | −0.07 | 0.009 | −7.6 | .00 |
| | AccessTime_Cy | −0.19 | 0.010 | −18.6 | .00 | −0.24 | 0.015 | −16.4 | .00 |
| | AccessTime_Wk | −0.16 | 0.012 | −13.6 | .00 | −0.22 | 0.018 | −12.2 | .00 |
| | EgressTime_BTM | −0.06 | 0.007 | −8.2 | .00 | −0.06 | 0.009 | −6.9 | .00 |
| | EgressTime_Cy | −0.18 | 0.018 | −10.0 | .00 | −0.21 | 0.028 | −7.8 | .00 |
| | EgressTime_Wk | −0.14 | 0.009 | −16.8 | .00 | −0.19 | 0.012 | −15.5 | .00 |
| | Connections | 2.48 | 0.180 | 13.8 | .00 | 5.07 | 0.436 | 11.6 | .00 |
| *ASCs and Dummy variables* | | | | | | | | | |
| | Cd_Access | −3.30 | 0.324 | −10.2 | .00 | −4.45 | 0.425 | −10.5 | .00 |
| | Cp_Access | −4.80 | 0.533 | −9.0 | .00 | −6.58 | 0.738 | −8.9 | .00 |
| | BTM_Access | −3.42 | 0.403 | −8.5 | .00 | −4.38 | 0.520 | −8.4 | .00 |
| | Cy_Access | −0.54 | 0.193 | −2.8 | .01 | −0.30 | 0.282 | −1.1 | .28 |
| | Cp_Egress | −6.39 | 0.645 | −9.9 | .00 | −9.30 | 0.997 | −9.3 | .00 |
| | BTM_Egress | −3.49 | 0.327 | −10.7 | .00 | −4.36 | 0.412 | −10.6 | .00 |
| | Cy_Egress | −2.84 | 0.265 | −10.7 | .00 | −4.07 | 0.378 | −10.8 | .00 |
| | AET_Cp | −0.18 | 0.022 | −8.0 | .00 | −0.22 | 0.033 | −6.6 | .00 |
| | BTM_Access_Urb4 | 1.17 | 0.307 | 3.8 | .00 | 1.62 | 0.477 | 3.4 | .00 |
| | BTM_Access_Urb5 | 1.40 | 0.341 | 4.1 | .00 | 1.94 | 0.525 | 3.7 | .00 |
| | BTM_Egress_Urb4 | 0.19 | 0.313 | 0.6 | .55 | 0.19 | 0.490 | 0.4 | .70 |
| | BTM_Egress_Urb5 | 1.11 | 0.259 | 4.3 | .00 | 1.31 | 0.393 | 3.3 | .00 |
| *Nest parameter* | | | | | | | | | |
| | theta | 0.88 | 0.064 | 13.9 | .00 | 0.55 | 0.043 | 12.7 | .00 |

AET: Generalised Travel time; Cy: Cycling; BTM: Bus/Tram/Metro; Urb: Urbanisation Level; Cd: Car driver; Wk: Walking; Cp: Car passenger.

### *2.5.2. Further investigation of the RRM LogSum*

As explained in Section 2.2, the MD-ToD model and the Station choice model are sequentially estimated. To do so, the LS of the station choice model is fed into the MD-ToD model. For this reason, and before proceeding with the estimation results of the MD-ToD models, this section analyses the RUM and RRM LogSums. We analyse the distribution of the RUM LS and of RRM LS implied by the estimated station choice models, and investigate the relation between the RUM LS and the RRM LS. For these analyses, we use a typical, representative origin,[7] being Almere Centrum (Almere is a medium sized city located centrally in the Netherlands). From this origin 1406 destinations can be reached by Train via several departure and arrival train stations, and using different access and egress modes. Each destination is associated with a LS. In the RUM-context the LS represents the expected maximum utility of the set of Train alternatives to reach that destination (e.g. using different station pairs, or egress and access modes). Likewise, in the RRM context the LS represents the expected minimum regret of the set of Train alternatives to reach that destination.

Fig. 2 shows two subplots. The left-hand side plot depicts the density of RUM LS and RRM LS distributions. The shapes of the empirical density functions differ in terms of the means and dispersions. The right-hand side plot of Fig. 2 shows a scatter plot. It reveals that the RUM and RRM LogSums (hence: the expected maximum utility and the expected minimum regret, respectively) are

[7] We tested the robustness of these analyses using observations with other origins. These gave similar results.

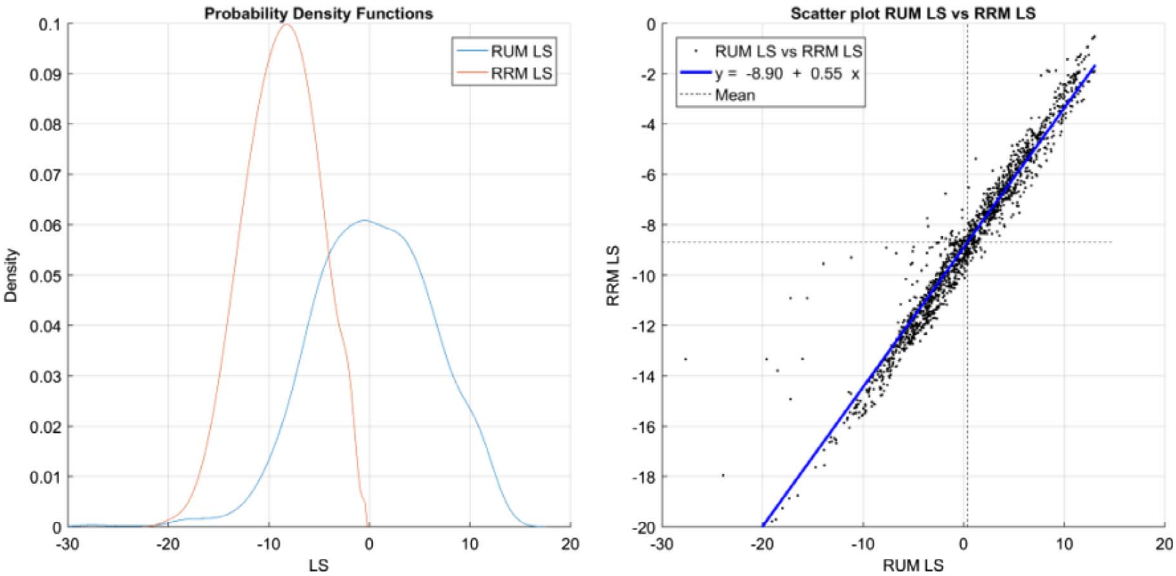



**Fig. 2.** Empirical relation between RUM LS and RRM LM.

strongly correlated. In line with expectations, we see that a high expected maximum utility (hence: destinations having attractive Train alternatives) associates with a high (i.e. close to zero) expected minimum regret. The relation appears to be mostly linear. The regressed blue line shows that an increase of the RUM LS of 1 is 'equivalent' to an on average increase of 0.55 of the RRM LS. Furthermore, Fig. 2 shows that the mean RUM LS is about 0.4, whereas the mean RRM LS is about −8.7. However, this difference is irrelevant; as the absolute level of (expected) utility (or regret) cannot be measured it has no behavioral meaning.

### *2.5.3. Mode-Destination-Time-of-Day choice models*

Table 3 shows the summary statistics of P-RRM and RUM MD-ToD model estimation results, for all travel purposes. Similar as for the Train station models, it shows that for most travel purposes the RUM model outperforms the RRM model in terms of model fit. Only for travel purpose Child-Education the RRM model performs slightly better in terms of model fit than the RUM model. However, taking into account the number of observations, the model fit differences are small for all travel purposes.

For reasons of word limitations, we do not discuss the estimation details in full detail. A full description of the estimation results for the all purposes can be found in Significance (2016). In general, the estimation results of the MD-ToD models provide the same picture as for the Train stations choice models. Signs and relative sizes are consistent across the two models for almost all parameters. Typically, we see two noteworthy differences. Firstly, the ASCs for model Train are highly negative in the RUM models, whereas they are positive in the RRM models. However, this has no substantive meaning as the LS is only identified up to a constant (which represents the fact that the absolute value of utility or regret cannot be measured); the large difference between the Train ASCs rather stems from the fact that the mean of the RRM LS is unequal to the mean of the RUM LS (see Section 2.5.2). Secondly, close inspection of the part-worth utilities and regrets show that – on average – the difference in part-worth utility between the most attractive and least attractive Train alternative in a choice set for the RUM-LMS is larger than for the RRM-LMS. Given that the model fit and the estimates for the constants and dummy variables are roughly the same, this signals that the destination choices predicted by the MD-ToD model in the RUM-LMS are likely to be considerably more deterministic (i.e., less subject to random noise) than those predicted

**Table 3**
Estimation results Mode-Destination-Time-of-Day models.

| | Commute | Business | Education | Shopping | Other | Work-Business | Work-Other | Child-Education | Child-Other |
|---|---|---|---|---|---|---|---|---|---|
| No. observations | 33,803 | 3100 | 8614 | 24,039 | 36,206 | 1086 | 652 | 9150 | 8095 |
| No. parameters | 84 | 52 | 74 | 87 | 90 | 19 | 20 | 13 | 23 |
| Null Log-likelihood | −2,70,972 | −20,032 | −63,033 | −2,00,518 | −3,07,793 | −9651 | −5642 | −76,910 | −70,117 |
| *RUM* | | | | | | | | | |
| Final Log-likelihood | −1,36,570 | −12,260 | −24,583 | −61,960 | −1,19,424 | −4777 | −1467 | −16,074 | −21,596 |
| McFadden's ρ2 | 0.50 | 0.39 | 0.61 | 0.69 | 0.61 | 0.50 | 0.74 | 0.79 | 0.69 |
| *RRM* | | | | | | | | | |
| Final Log-likelihood | −1,36,903 | −12,329 | −24,689 | −62,042 | −1,19,880 | −4808 | −1467 | −16,071 | −21,654 |
| McFadden's ρ2 | 0.49 | 0.38 | 0.61 | 0.69 | 0.61 | 0.50 | 0.74 | 0.79 | 0.69 |
| ΔLL/obs | 0.010 | 0.022 | 0.012 | 0.003 | 0.013 | 0.029 | 0.000 | 0.000 | 0.007 |

**Table 4**
Indices passenger kilometers by mode RUM2010/MON.

| RUM2010/MON | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Train | Car driver | Car passenger | BTM | Bike | Walk | Total |
| Education | 0.96 | 0.85 | 1.13 | 0.94 | 1.05 | 1.41 | 1.02 |
| Commute | 0.97 | 1.04 | 1.55 | 0.95 | 1.04 | 2.07 | 1.09 |
| Business | 0.98 | 1.43 | 1.49 | 0.85 | 1.20 | 2.52 | 1.41 |
| Shopping | 0.65 | 0.90 | 1.15 | 1.02 | 1.23 | 1.73 | 0.96 |
| Other | 0.74 | 0.94 | 0.92 | 0.79 | 0.91 | 1.20 | 0.97 |
| Total | 0.89 | 1.05 | 1.13 | 0.94 | 1.07 | 1.37 | 1.09 |
| RMSE | 0.19 | 0.21 | 0.34 | 0.12 | 0.15 | 0.92 | 0.19 |

in the RRM-LMS.

## 3. Base year forecasts and aggregate demand elasticities

This section explores the aggregate level differences in mobility forecasts. Section 3.1 discusses the results for the base year. The degree to which the base year can be accurately forecasted is a first sign of how well the model performs. Section 3.2 discusses implied aggregate demand elasticities.

### 3.1. Base year forecasts

To explore the differences in the base year forecasts we primarily use indices: ratios of outcomes, such as the number of passenger kilometers forecasted by the RRM-LMS divided by the observed number of passenger kilometers in the data. The reason for reporting indices, e.g. rather than absolute numbers, is that we believe these are more insightful given the objective of this study: to acquire insights on whether changing the decision rule of the disaggregate discrete choice models embedded in large-scale transport models would lead to different aggregate level mobility forecasts. In order to make the differences in forecasts more tangible we also report Root-Mean-Square-Errors[8] (RMSEs). These give a good measure for the aggregate accuracy of the forecasts. Finally, for reasons of space limitations we report only the results for the number of passenger kilometers. Somewhat similar results are obtained of the number of tours (see Significance, 2016 for more details).

#### 3.1.1. Passenger kilometers by mode

Tables 4 and 5 show the forecasted number of passenger kilometers for the base year divided by the observed number of passenger kilometers in the MON data, for respectively the RUM-LMS and the RRM-LMS, by mode and by travel purpose. They reveal that several predictions are quite off, both for the RUM-LMS as well as the RRM-LMS predictions. Both model systems systematically underestimate the number of passenger kilometers by Train. Furthermore, when we more closely inspect and compare these two tables we see that cases of 'misprediction' are usually consistent across the two model systems. This suggests that the inaccuracy in predicting the number of passenger kilometers is unrelated to the decision rule that is imposed by the discrete choice models embedded in the model systems. Finally, looking at the RMSEs, we see that the RRM-LMS performs somewhat worse than the RUM-LMS in terms of prediction accuracy of the number of passenger kilometers, in particular with regard to modes Car driver and BTM.

Table 6 presents the ratio of the number of passenger kilometers predicted by the RRM-LMS and the number of passenger kilometers predicted by the RUM-LMS, by mode and by travel purpose. Two things catch the eye. Firstly, Table 6 shows that the RRM-LMS systematically predicts almost 10% less passenger kilometers by Train. Secondly, there are two instances in which the predictions of the RUM-LMS and the RRM-LMS differ considerably. The RRM-LMS predicts substantially more Car driver kilometers for purpose Business than the RUM-LMS, while it predicts substantially less BTM passenger kilometers for purpose Other than the RUM-LMS. It is unclear where these differences precisely stem from, or how they possibly relate to the differences in the underlying decision rules between the two model systems.

### 3.2. Aggregate demand elasticities

Demand elasticities derived from large-scale transport models are generally considered to be of considerable policy relevance. Therefore, we conducted an extensive analysis of the differences in the implied elasticities between the RUM-LMS and the RRM-LMS. Demand elasticities measure the percentage change in the aggregate demand in response to a percentage change in some Level-of-Service attribute $X_j$, see equation 12 where $Q_i$ denotes the demand for alternative $i$ and $E_i$ denotes the aggregate demand elasticity.

[8] $RMSE = \sqrt{\frac{\sum_n (\hat{\theta}_n - \theta_n)^2}{N}}$.

**Table 5**
Indices passenger kilometers by mode RRM2010/MON.

| RRM2010/MON | Train | Car driver | Car passenger | BTM | Bike | Walk | Total |
|---|---|---|---|---|---|---|---|
| Education | 0.85 | 0.86 | 1.05 | 0.96 | 1.05 | 1.38 | 0.97 |
| Commute | 0.90 | 1.04 | 1.48 | 0.95 | 1.04 | 2.04 | 1.07 |
| Business | 0.90 | 1.84 | 1.53 | 0.85 | 1.14 | 2.40 | 1.75 |
| Shopping | 0.66 | 0.90 | 1.11 | 1.07 | 1.23 | 1.72 | 0.95 |
| Other | 0.68 | 0.93 | 0.91 | 0.49 | 0.91 | 1.21 | 0.95 |
| Total | 0.81 | 1.07 | 1.10 | 0.89 | 1.07 | 1.37 | 1.08 |
| RMSE | 0.23 | 0.39 | 0.33 | 0.24 | 0.13 | 0.87 | 0.34 |

**Table 6**
Indices passenger kilometers by mode RRM2010/RUM2010.

| RRM2010/RUM2010 | Train | Car driver | Car passenger | BTM | Bike | Walk | Total |
|---|---|---|---|---|---|---|---|
| Education | 0.89 | 1.01 | 0.93 | 1.03 | 1.00 | 0.97 | 0.95 |
| Commute | 0.93 | 0.99 | 0.96 | 1.00 | 1.01 | 0.98 | 0.98 |
| Business | 0.92 | 1.29 | 1.03 | 1.00 | 0.95 | 0.96 | 1.24 |
| Shopping | 1.01 | 1.00 | 0.97 | 1.05 | 1.00 | 1.00 | 0.99 |
| Other | 0.91 | 0.99 | 0.99 | 0.62 | 1.00 | 1.00 | 0.98 |
| Total | 0.91 | 1.02 | 0.97 | 0.95 | 1.00 | 1.00 | 1.00 |

$$E_i = \frac{\%\Delta Q_i}{\%\Delta X_j} \tag{12}$$

To compute the percentage change in demand $\%\Delta Q_i$ two simulations are conducted using both model systems. In the first run $X_j$ (e.g. the car driver cost) is increased by 10%, in the second run the car driver costs are set to their original value. The percentage change in aggregate demand is then computed using: $\Delta\%Q_i = (Q_{+10\%} - Q_{\text{Ref}})/Q_{\text{Ref}}$, where $Q_{Ref}$ denotes the demand in the reference scenario and $Q_{+10\%}$ denotes the demand in the increased cost scenario.

Table 7 presents the aggregate demand elasticities for the passenger number of kilometers. Elasticities are derived for Car driver cost, Car driver time, Car passenger time, BTM cost, BTM in-vehicle time, Train cost, Train time and Train frequency. As can be seen, with the exception for Train frequency, the demand elasticities implied by the RUM-LMS and the RRM-LMS are quite close to one another. That is, the differences are less than $|0.2|$. These results suggest that the aggregate demand elasticities are rather robust towards the underlying decision rule. However, we see considerable differences between the RUM and RRM demand elasticities for Train frequency. In fact, the elasticities implied by the RUM-LMS are a factor 2.5–5 larger than the elasticities implied by the RRM-LMS, depending on the travel purpose.

## 4. Case study: High Frequency Rail scenario

In this section we investigate a fictive "High Frequency Rail" (HFR) policy scenario. We specifically chose this case study (e.g. rather than a pricing-policy scenario) for three reasons. The first and most important reason was that the results from the base year analyses showed that most substantial differences between the RUM-LMS and the RRM-LMS were found for mode Train (see Tables 6

**Table 7**
Passenger kilometer demand elasticities.

| | Commute | | Business | | Education | | Shopping | | Other | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | RUM | RRM | RUM | RRM | RUM | RRM | RUM | RRM | RUM | RRM | RUM | RRM |
| Car driver cost | −0.51 | −0.53 | −0.17 | −0.10 | – | – | −0.46 | −0.52 | −0.29 | −0.37 | −0.41 | −0.44 |
| Car driver time | −0.88 | −0.82 | −0.88 | −0.75 | −1.54 | −1.54 | −1.35 | −1.26 | −1.32 | −1.18 | −1.02 | −0.93 |
| Car passenger time | −1.50 | −1.48 | −1.05 | −0.91 | −1.68 | −1.82 | −1.93 | −1.90 | −1.55 | −1.49 | −1.59 | −1.54 |
| BTM cost | −0.55 | −0.65 | −0.12 | −0.08 | – | – | −0.54 | −0.67 | −0.24 | −0.29 | −0.30 | −0.36 |
| BTM in-vehicle time | −0.81 | −0.80 | −0.88 | −0.94 | −1.03 | −1.10 | −0.96 | −0.86 | −0.81 | −0.80 | −0.90 | −0.92 |
| Train cost | −0.91 | −0.97 | −0.07 | −0.09 | – | – | −0.99 | −0.96 | −0.69 | −0.70 | −0.59 | −0.63 |
| Train time | −0.74 | −0.81 | −0.41 | −0.51 | −1.02 | −1.13 | −0.50 | −0.52 | −0.45 | −0.47 | −0.77 | −0.84 |
| Train frequency | 0.70 | 0.16 | 0.51 | 0.11 | 0.47 | 0.16 | 0.29 | 0.09 | 0.25 | 0.10 | 0.57 | 0.15 |

**Table 8**
Differences in the percentage increase in the number of tours and the number of passenger kilometers for the RUM-LMS.

| **Δ%Tours RUM** | Train | Car driver | Car passenger | BTM | Bike | Walk | Total |
|---|---|---|---|---|---|---|---|
| Education | 0.11 | -0.03 | -0.02 | -0.03 | -0.01 | -0.01 | 0.00 |
| Commute | 0.21 | -0.01 | -0.01 | -0.03 | -0.02 | -0.02 | 0.00 |
| Business | 0.15 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | 0.00 |
| Shopping | 0.08 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Other | 0.05 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| **Total** | 0.16 | -0.01 | 0.00 | -0.02 | -0.01 | 0.00 | 0.00 |
| **Δ%Passenger kilometres RUM** | Train | Car driver | Car passenger | BTM | Bike | Walk | Total |
| Education | 0.15 | -0.02 | -0.01 | -0.03 | -0.01 | -0.01 | 0.05 |
| Commute | 0.26 | -0.01 | -0.01 | -0.03 | -0.02 | -0.02 | 0.03 |
| Business | 0.16 | 0.00 | -0.01 | -0.01 | -0.01 | -0.01 | 0.01 |
| Shopping | 0.12 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 |
| Other | 0.08 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 |
| Total | 0.20 | -0.01 | 0.00 | -0.03 | -0.01 | 0.00 | 0.02 |

and 7). The second and more pragmatic reason was that several comprehensive HFR scenario studies have been conducted before using the Dutch National model. Therefore, Level-of-Service matrices were readily available for several HFR scenarios. The third reason was that we had agreed with The Netherlands Road Authorities (the owner of the LMS) that we would only use the LMS (and make RUM-LMS and RRM-LMS comparisons) for research which would not relate to any ongoing political discourse or policy development. In the Netherlands, HFR is currently not a hot topic, politically speaking, and it hence provided a 'safe' environment to experiment with different versions of the LMS.

### 4.1. The scenario

The future year we investigate is 2030. In the HFR scenario train frequencies are substantially intensified in the year 2030 as compared to the reference scenario in which the 2010 train tables are maintained in the year 2030. In the HFR scenario the train frequency is increased by 50% on the main train lines in the corridors "Utrecht – Den Bosch", "Utrecht – Arnhem" and "Den Haag – Rotterdam". For instance, in the HFR scenario there are 6 trains per hour connecting Utrecht Central Station (the largest Train station in the Netherlands) and Schiphol International Airport, while in the baseline scenario there are 4 hourly connections.

### 4.2. Results

To investigate the difference in the forecasts between the RUM-LMS and the RRM-LMS we assess the difference in the percentage increase in the number of tours and the number of passenger kilometers between the HFR scenario and the reference scenario, relative to the base year 2010 (Eq. (13)). This value is typically of key policy relevance when assessing the benefits of a new transport policy.

Tables 8 and 9 show the differences in the percentage increase in the number of tours and the number of passenger kilometers between the HFR scenario and the reference scenario, relative to the base year 2010. Table 8 shows the results for the RUM-LMS; Table 9 shows the results for the RRM-LMS.

**Table 9**
Differences in the percentage increase in the number of tours and the number of passenger kilometers for the RRM-LMS.

| **Δ%Tours RRM** | Train | Car driver | Car passenger | BTM | Bike | Walk | Total |
|---|---|---|---|---|---|---|---|
| Education | 0.07 | -0.02 | -0.01 | -0.02 | -0.01 | -0.01 | 0.00 |
| Commute | 0.11 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | 0.00 |
| Business | 0.07 | 0.00 | 0.00 | -0.01 | 0.00 | 0.00 | 0.00 |
| Shopping | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Other | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| **Total** | 0.09 | 0.00 | 0.00 | -0.01 | 0.00 | 0.00 | 0.00 |
| **Δ%Passenger kilometres RRM** | Train | Car driver | Car passenger | BTM | Bike | Walk | Total |
| Education | 0.10 | -0.01 | -0.01 | -0.02 | -0.01 | -0.01 | 0.03 |
| Commute | 0.15 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | 0.01 |
| Business | 0.08 | 0.00 | 0.00 | -0.01 | 0.00 | 0.00 | 0.00 |
| Shopping | 0.08 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Other | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Total | 0.12 | 0.00 | 0.00 | -0.01 | 0.00 | 0.00 | 0.01 |

**Table 10**
Aggregate level output of the RUM-LMS and RRM-LMS for mode train.

| *Scenario* | *GE2030 without HFR* | | *GE2030 with HFR* | |
|---|---|---|---|---|
| | RUM-LMS | RRM-LMS | RUM-LMS | RRM-LMS |
| In-vehicle time [min/day] | 4,89,51,881 | 4,64,38,115 | 5,44,85,395 | 4,85,70,999 |
| Out of vehicle time [min/day] | 28,04,909 | 37,24,502 | 38,11,909 | 45,13,007 |
| No. trips per day [-] | 14,95,904 | 14,26,944 | 17,12,130 | 15,34,596 |
| Avg travel time [min/trip] | 34.60 | 35.15 | 34.05 | 34.59 |

$$\%\Delta Q = \frac{Q_{HFR}^{2030} - Q_{NO\,HFR}^{2030}}{Q_{Base}^{2010}} \tag{13}$$

Based on Tables 8 and 9 a number of inferences can be made. First, in line with intuition, both transport model system predict a strong increase in demand for the train mode due to the HFR policy, both in terms of the number of train tours and the number of train passenger kilometers. Second, the predictions of both models are qualitatively consistent in terms of differences across purpose: both transport models predict the strongest increase in train demand for purpose Commute, followed by Education and Business. Third, and most interestingly, the RUM-LMS is found to be significantly more sensitive towards the improvement in train frequencies than the RRM-based LMS. Depending on the travel purpose, the increase in the number of passenger kilometers predicted by the RUM-LMS is between 1.5 and 2.1 times higher than the increase predicted by the RRM-LMS. Clearly, there are non-trivial differences in the train travel demand forecasts between the RUM-LMS and the RRM-LMS. Finally, it goes without saying that based on these results alone, it is not possible to say which forecasts are more accurate.

To further highlight the relevance of these results and make them more tangible, we illustrate their implication for transport policy appraisal. In many Westerns countries Cost Benefit Analysis (CBA) is used to evaluate the costs and benefits of transport policy alternatives (Mouter et al., 2013). A large share of the total benefits in CBAs of transport policies stem from accrued travel time savings. These travel time savings, in turn, are computed using large-scale transport models, and are subsequently monetized using the so-called Value-of-Time (VoT). Ideally, the accrued benefits are computed based on disaggregate model outputs, i.e. at the level of OD relations, differentiated to travel purpose and mode. However, these disaggregate model output data are not readily available from the LMS. Therefore, as a first proxy, we compute the benefits using relatively high (i.e., aggregate) level outputs.

Table 10 shows the relevant output of the transport models to compute the accrued total benefits of the HFR policy[9]. In line with earlier results, Table 10 shows that the RRM-LMS predicts a smaller increase in train demand due to the HFR scenario, both in terms of total travel time and the total number of trips made by train. Furthermore, in line with expectations both transport models predict that the average train travel times decrease due to the HFR policy.

Table 11 shows the total travel time gains for existing trips and for new train trips. These are computed using the output of Table 10. The total travel time gains for new train trips are computed using the rule-of-half (De Jong et al., 2007). Note that, as discussed in Section 2.3, there is no straightforward connection between RRM models and the welfare economic axioms underlying classical Cost Benefit Analysis. Therefore, we here use typical RUM-techniques to translate the differences in aggregate travel times (between models and between the 'before the policy' and 'after the policy' situation) into differences in monetary benefits. Subsequently, the total benefits are computed by converting the total travel time gains into hours per year and multiplying these with the VoT. In practice, the VoT varies across travel purposes. However, this is inessential in this context, as we are predominantly interested in comparing the outcomes of the RUM and RRM based transport models. For convenience, we set the VoT to €10 per hour. The total predicted benefits are 38.2 and 36.1 million euros per year, respectively for the RUM-LMS and the RRM-LMS. Hence, the yearly benefits predicted by the RRM-LMS are about 6% lower than those predicted by the RUM-LMS. This finding lends support for the view that the presumed decision rule of the choice models embedded in large-scale travel demand models does matter for transport policy appraisal.

### 4.3. Explaining the differences in forecasts

There are at least two, potentially interrelated, explanations for the considerable difference in forecasts between the RUM-LMS and the RRM-LMS in terms of train travel demand. Below, we discuss these two explanations. However, it should be noted upfront that due to the multi-module nature of the LMS it is practically impossible to formally proof how differences in forecasts relate exactly to these explanations.

#### 4.3.1. The effect of general ('across the board') improvements in RRM models

The first explanation touches upon a fundamental behavioral property of RRM models, and relates to the fact that in the HFR policy scenario the performance of many train alternatives improve in roughly the same way. RRM models postulate that regret is experienced when comparing the performance of a considered alternative with the performance of its competitor alternatives. Thus,

[9] Note that we only consider Train trips here.

**Table 11**
Total benefits.

| *Benefits* | RUM-LMS | RRM-LMS |
|---|---|---|
| Travel time gains existing trips [min/day] | 8,21,890 | 8,02,457 |
| Travel time gains new trips [min/day] | 59,400 | 30,270 |
| Total travel time gains [min/day] | 8,81,290 | 8,32,727 |
| *Total travel time gains [h/yr]* | 38,18,924 | 36,08,483 |
| *Avg. VoT* | € 10 | € 10 |
| *Total monetary benefits* | € 3,81,89,244 | € 3,60,84,827 |

in RRM models only relative performance matters to determine regret levels (in contrast to RUM models, where utility is a function on an alternative's own performance only). This means, for instance, that in the hypothetical case in which the travel times to all destinations would be shortened by five minutes, all else being equal, regret levels would remain exactly the same as before the policy measure. As in the HFR scenario many train alternatives are improved in the same way, the regret levels in the Train station model are to some extent unaffected. In turn, choice probabilities of alternatives in the Train nest as well as the LS of the Train station model change only relatively mildly. As a consequence, the choice probability forecasts in the MD model, which uses the LS from the train station model, do not change much. So, the observation that the RRM-LMS is relatively insensitive for the HFR policy scenario can be explained by the behavioral premises underlying the RRM model.

To further investigate this potential explanation Fig. 3 shows two scatter plots. The *x*-axis depicts the LS without the HFR policy; the *y*-axis depicts the LS with the HFR policy. The left scatter plot shows the RUM LS; the right scatter plot shows the RRM LS, both for the same observation as was used in Section 2.5.2. Furthermore, the left-hand side plot in Fig. 4 shows the corresponding Probability Density Functions (PDFs) and the right hand-side plot shows the corresponding Cumulative Distribution Functions (CDFs) of the difference in the LogSums: $\Delta LS = LS_{HFR} - LS_{NO\,HFR}$.

First, we inspect scatter plots in Fig. 3. The RUM scatter plot (left) shows that the vast majority of the dots lie above the blue $y = x$ line. This is in line with intuition, as it means that the expected maximum utility to reach a destination by train improves due to the HFR policy. In fact, looking at Fig. 4, we see that in 95% of the cases (i.e., destinations) the RUM LS improves (i.e., $\Delta LS > 0$) due to the HFR policy. For a small number of destinations (more specifically, 5% of them) Train alternatives become less attractive due to the HFR policy. However, in the RRM scatter plot (right) a different picture emerges: although most dots lie above the blue $y = x$ line (as one would expect), also quite a few dots are located below the blue line. Fig. 4 shows that for 34% per cent of the destinations the RRM LS deteriorates, implying that the model postulates that the train alternatives to these destinations become less attractive, rather than more attractive. Moreover, the (statistical) mode in the PDF (left-hand side plot) shows that for most destinations the RRM LS does not change due to the HFR policy measure. This again signals the fundamental behavioral difference between the RUM LS and its RRM counterpart. Furthermore, it underpins the view that the obtained differences in mobility forecasts may for a considerable part be attributed to the differences in the underlying decision rule.

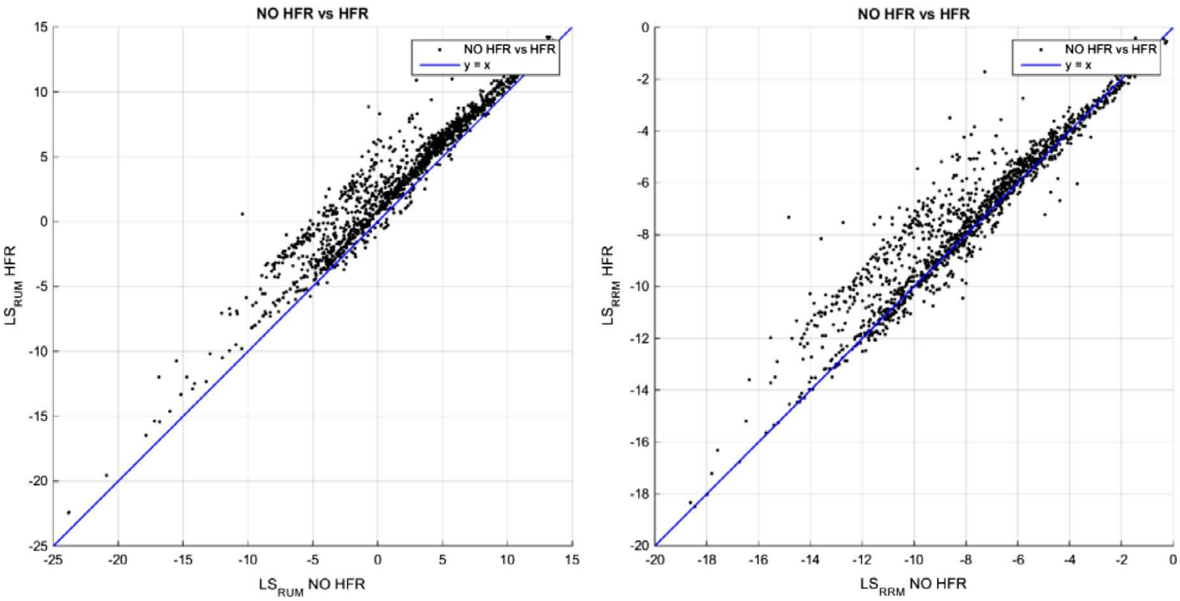



**Fig. 3.** Scatter plots RUM and RRM LogSums.

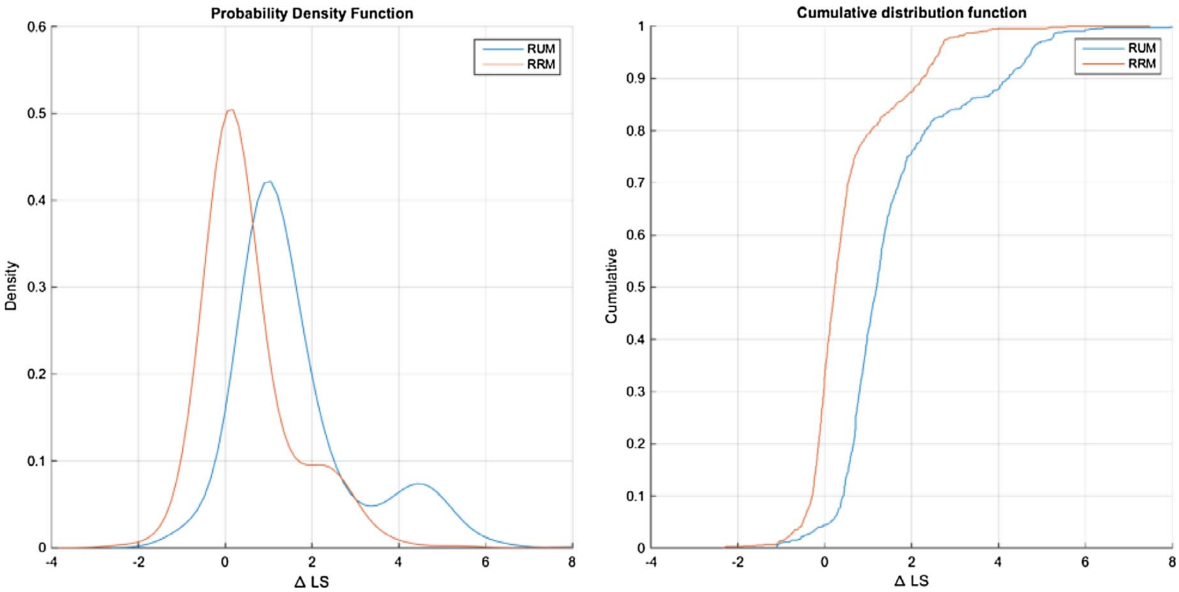



**Fig. 4.** PDF and CDF of ΔLS.

#### 4.3.2. Sequential estimation of RRM models

The second explanation is methodological rather than behavioral; it relates to the sequential estimation procedure that is used to estimate the MD-ToD and Train station choice model. As discussed in Section 2.2, in order to be able to compare the results from the RRM-LMS with those of the RUM-LMS, we used the exact same structure and estimation approach to develop the RRM-LMS. Therefore, the MD-ToD model and the Train station choice model are sequentially estimated. In this sequential estimation approach, the LS is used to transfer information from the lower level (i.e. the Train station model) towards the upper level (i.e. the MD-ToD model). To calibrate the two models, the inclusive value is determined using an iterative procedure.

While the sequential estimation approach results in consistency between the RUM and RRM transport models, it is uncertain whether a sequential estimation approach makes sense in a regret modelling context. It is important to note that in both model systems separate cost and travel time parameters are estimated in the MD-ToD and Train station choice models. Since in RUM models the utilities are only a function of the performance of the alternative itself, this is not so elegant (it does, for instance, imply that the implicit VoT can be different in the MD-ToD model than in the Train station model), but it does not fundamentally affect the properties of the model systems.

In contrast, in RRM models the impact of the sequential estimation and the use of separate cost and travel time parameters is probably more severe. To see this, as before, suppose that the travel times to all Train alternatives are shortened by five minutes, and that the MD-ToD model and the Train station model is sequentially estimated. All else being equal, this would not affect regret levels or the logsums in the Train station model. As a result, the MD-ToD model Train would not predict any new travelers to be attracted from other modes. However, when the MD-ToD model and the Train station model would have been estimated jointly using a joint, rather than a separate travel time parameter, then the five minute reduction in the travel times of all Train alternatives would imply that all non-train alternatives would experience an increase in the level of regret. As a result, that model would predict that the Train alternatives would become relatively more attractive as compared to the non-train modes. This shows that sequential estimation of Nested Logit RRM models in combination with separate parameters for generic attributes may severely jeopardize the behavioral realism of the model. As such, it seems conceivable that the obtained differences in mobility forecasts can partly be attributed to this sequential estimation procedure.

## 5. Conclusion and discussion

This paper studies to what extent decision rules (embedded in disaggregate discrete choice models) matter for large-scale aggregate transport demand forecasts. To do so, we developed a RRM-based counterpart of the RUM-based Dutch National Transport model, compared its forecasts with those of the conventional RUM-based transport model, and illustrated the impact for transport appraisal in the context of a case study. As such, this paper steps beyond the existing body of literature on RUM-RRM comparisons, which are based on disaggregate level applications only. As a secondary contribution, we presented a technique to speed-up estimation of P-RRM models by a factor thousand or more for data sets that are characterized by large choice sets.

Our results show that the aggregate mobility forecasts made by the utility-based and regret-based transport models can differ substantially. Although at a disaggregate level model fit differences between the two choice models are rather small, their implementations in the National Transport Model result in non-trivial differences in aggregate forecasts of the number of tours, the passenger number of kilometers and demand elasticities. Furthermore, by analyzing a policy case study, we find that the impact of intensifying the train frequencies on travel demand (i.e. passenger kilometers), differs strongly between the utility- and regret-based National Transport Model. The differences in predicted travel demand are found to translate into non-negligible differences for transport appraisal.

We believe it is safe to say that the obtained differences in mobility forecasts can – for a considerable part – be understood and explained considering the fundamental differences in behavioral premises underlying the regret- and utility-based models. However, despite our efforts to isolate the 'behavioral' effect of the decision rule as best as possible, it should be noted that a large-scale and multi module model such as the LMS incorporates a wide range of sometimes rather pragmatic methodological assumptions. Although we went through a lot of effort in our analyses, to limit any such potential methodological differences between RUM-LMS and RRM-LMS (and note that we have tried to be fully transparent about remaining ones), we cannot guarantee that the obtained differences in mobility forecasts are not also partially caused by potential methodological artefacts such as relating to the sequential estimation procedure, or even simply programming errors.

Our study also new provides opportunities for policy analysts to enrich their sensitivity analysis toolbox. In current practice, sensitivity analyses of mobility forecasts are based on induced changes in contextual or population related variables (e.g. economic or population growth), or in parameters (e.g. travel time and cost penalties). This study shows that it is also possible to conduct sensitivity analysis based on the presumed underlying decision rule. Testing the sensitivity of a transport policy towards different presumed decision rules may help policy-makers to select robust transport policies, i.e. policies that perform well under a range of different assumptions regarding human decision making, and help policy-makers avoiding selecting those policies that perform well only under one particular decision rule assumption. In light of the still relatively modest understanding of travel choice behavior, we believe this type of sensitivity analysis is a welcome complement to more conventional sensitivity analysis approaches.

Finally, is important to note that although we have shown that large-scale transport models based on non-RUM can produce different aggregate level predictions than those based on RUM, we can – at this point – not draw any conclusions regarding whether the non-RUM forecasts are more accurate, or would result in better informed policy decisions. These questions are relevant avenues for future research. In particular, it would be interesting to conduct a backcasting analysis (see e.g. Rich and Hansen, 2016), or to evaluate the accuracy of the RUM- and RRM-based forecasts in hindsight, e.g. one or two decades from now, inspired by the work of Annema and De Jong (2011).

## Acknowledgements

We would like to thank Marits Pieters, Jaap Baak and Gerard de Jong of Significance Consultancy and Frank Hofman of The Netherlands Road Authority for their valuable contributions to the modelling efforts presented in this paper. Furthermore, we would like to thank two anonymous reviewers for their thoughtful comments. The authors gratefully acknowledge support from the Netherlands Organisation for Scientific Research (NWO), in the form of VIDI Grant 016-125-305.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.tra.2018.01.035.

## References

Annema, J.A., De Jong, M., 2011. The history of the transport future—evaluating Dutch transport scenarios of the past. Trans. Rev. 31 (3), 341–356.
Chorus, C.G., 2012a. Random regret minimization: an overview of model properties and empirical evidence. Trans. Rev. 32 (1), 75–92.
Chorus, C.G., 2010. A new model of random regret minimization. Eur. J. Trans. Infrastruct. Res. 10 (2), 181–196.
Chorus, C.G., 2012b. Logsums for utility-maximizers and regret-minimizers, and their relation with desirability and satisfaction. Transport. Res. Part A: Policy Practice 46 (7), 1003–1012.
Chorus, C.G., 2014. Capturing alternative decision rules in travel choice models: a critical discussion. In: Hess, S., Daly, A. (Eds.), "Handbook of Choice Modelling". Edward Elgar, pp. 290.
Chorus, C.G., Arentze, T.A., Timmermans, H.J.P., 2008. A random regret-minimization model of travel choice. Transport. Res. Part B-Methodol. 42 (1), 1–18.
Chorus, C.G., Bierlaire, M., 2013. An empirical comparison of travel choice models that capture preferences for compromise alternatives. Transportation 40 (3), 549–562.
Chorus, C.G., Rose, J.M., Hensher, D.A., 2013. Regret minimization or utility maximization: it depends on the attribute. Environ. Plann. B: Plann. Des. 40 (1), 154–169.
Chorus, C., Van Cranenburgh, S., Dekker, T., 2014. Random regret minimization for consumer choice modeling: assessment of empirical evidence. J. Bus. Res. 67 (11), 2428–2436.
Daly, A., 1987. Estimating "tree" logit models. Transport. Res. Part B: Methodol. 21 (4), 251–267.
Daly, A., Fox, J., Tuinenga, J.G., 2005. Pivot-point procedures in practical travel demand forecasting. European Regional Science Association meeting 2005, Amsterdam.
Daly, A., Sillaparcharn, P., 2000. National models. Handbook of transport modelling, pp. 421–432.
Daly, A., Zachery, S., 1978. Improved multiple choice models. In: Hensher, D. (Ed.), Determinants of Travel Choice. Saxon House, Sussex.
Dekker, T., 2014. Indifference based value of time measures for Random Regret Minimisation models. J. Choice Modell. 12, 10–20.
Dekker, T., Chorus, C.G., 2017. Consumer surplus for Random Regret Minimisation models. EURO Journal on Transport and Logistics (under second review).
de Jong, G., Daly, A., Pieters, M., van der Hoorn, T., 2007. The logsum as an evaluation measure: review of the literature and new results. Transport. Res. Part A: Policy Practice 41 (9), 874–889.
de Jong, G., Daly, A., Pieters, M., Vellay, C., Bradley, M., Hofman, F., 2003. A model for time of day and mode choice using error components logit. Transport. Res. Part E: Logistics Transport. Rev. 39 (3), 245–268.
Greene, W.H., 2012. NLOGIT Version 5.0 Reference Guide. Plainview, NY, USA.
Guevara, C.A., Chorus, C.G., Ben-Akiva, M.E., 2016. Sampling of alternatives in random regret minimization models. Transport. Sci. 50 (1), 306–321.
Guevara, C.A., Fukushi, M., 2016. Modeling the decoy effect with context-RUM Models: Diagrammatic analysis and empirical evidence from route choice SP and mode

choice RP case studies. Transport. Res. Part B: Methodol. 93 (Part A), pp. 318–337.
Hensher, D.A., Greene, W.H., Rose, J.M., 2015. Applied Choice Analysis: A Primer. Cambridge University Press, Cambridge.
Hess, S., Beck, M.J., Chorus, C.G., 2014. Contrasts between utility maximisation and regret minimisation in the presence of opt out alternatives. Transport. Res. Part A: Policy Pract. 66, 1–12.
Hess, S., Daly, A., Rohr, C., Hyman, G., 2007. On the development of time period and mode choice models for use in large-scale modelling forecasting systems. Transport. Res. Part A: Policy Pract. 41 (9), 802–826.
Hess, S., Stathopoulos, A., Daly, A., 2012. Allowing for heterogeneous decision rules in discrete choice models: an approach and four case studies. Transportation 39 (3), 565–591.
Leong, W., Hensher, D.A., 2012. Embedding decision heuristics in discrete choice models: a review. Trans. Rev. 32 (3), 313–331.
Mouter, N., Annema, J.A., van Wee, B., 2013. Attitudes towards the role of Cost-Benefit Analysis in the decision-making process for spatial-infrastructure projects: A Dutch case study. Transport. Res. Part A: Policy Pract. 58, 1–14.
Ortúzar, S.J.d.D., Willumsen, L.G., 2011. Modelling Transport. John Wiley & Sons, Chichester, West Sussex, United Kingdom.
Rich, J., Hansen, C.O., 2016. The Danish National Passenger Model–model specification and results. EJTIR 16 (4), 573–599.
Significance, 2012. Documentatie van GM2011.
Significance, 2016. Spijtminimalisatie-gebaseerd Landelijk Model Systeem: Beschrijving runs RRM-RUM.
Van Cranenburgh, S., Guevara, C.A., Chorus, C.G., 2015a. New insights on random regret minimization models. Transport. Res. Part A: Policy Pract. 74, 91–109.
Van Cranenburgh, S., Prato, C.G., 2016. On the robustness of Random Regret Minimization modelling outcomes towards omitted attributes. J. Choice Modell. 18, 51–70.
Van Cranenburgh, S., Prato, C.G., Chorus, C.G., 2015b. Accounting for variation in choice set size in Random Regret Minimization models (http://resolver.tudelft.nl/uuid:1de1446b-1feb-4213-bef4-12031fa61579).
Vermunt, J.K., Magidson, J., 2014. Upgrade manual for Latent GOLD Choice 5.0. (Belmont, MA, USA).