Context-dependent models versus a context-free model

A comprehensive comparison for Swiss and German SP and RP data sets

Author(s):
Belgiawan, Prawira F.; Dubernet, Ilka; Schmid, Basil; Axhausen, Kay W.

Publication date:
2017

Permanent link:
https://doi.org/10.3929/ethz-b-000228019

Originally published in:
Arbeitsberichte Verkehrs- und Raumplanung 1281

Context-dependent models versus a context-free model: A comprehensive comparisons for Swiss and German SP and RP data sets

Prawira F. Belgiawan (corresponding author)

Institute for Transport Planning and Systems, ETH Zurich, Zurich, Switzerland and School of Business and Management, Bandung Institute of Technology, Bandung, Indonesia

Address: Jl. Ganesha 10, 40132, Bandung, Indonesia. E-mail: fajar.belgiawan@sbm-itb.ac.id (present address); Phone: +62-22-2531923

Ilka Dubernet,

Institute for Transport Planning and Systems, ETH Zurich, Zurich, Switzerland

Address: Stefano-Franscini-Platz 5, 8093 Zurich, Switzerland. E-mail: ilka.dubernet@ivt.baug.ethz.ch; Phone: +41 44 633 30 92

Basil Schmid

Institute for Transport Planning and Systems, ETH Zurich, Zurich, Switzerland

Address: Stefano-Franscini-Platz 5, 8093 Zurich, Switzerland. E-mail: basil.schmid@ivt.baug.ethz.ch; Phone: +41 44 633 30 89

Kay W. Axhausen

Institute for Transport Planning and Systems, ETH Zurich, Zurich, Switzerland

Address: Stefano-Franscini-Platz 5, 8093 Zurich, Switzerland. E-mail: axhausen@ivt.baug.ethz.ch; Phone: +41 44 633 39 43

Context-dependent models versus a context-free model: A comprehensive comparisons for Swiss and German SP and RP data sets

The random regret minimization (RRM) model considers the relative performance of the alternatives and is therefore context-dependent. In RRM, an individual, when choosing between alternatives, is assumed to minimize anticipated regret as opposed to maximize his/her utility. There are three variants of RRM, the classical CRRM, the μRRM, and the P-RRM. There is also a further approach called relative advantage maximization (RAM). We compare multinomial logit with the four mentioned alternatives. We use stated choice data sets which include mode choice, location choice, parking choice, carpooling, car-sharing. We compare the performance of those five models by their model fit, values of travel time savings (VTTS), and elasticities. Looking at the model fit, RAM outperforms the other models in five cases, whereas the PRRM does so in two cases and μRRM only for one case. The VTTS and elasticities vary substantially which is relevant for cost-benefit analysis or simplified modelling approaches.

Keywords: Context-dependent models; Random Regret Minimization; RRM variants; Relative Advantage Maximization

1 Introduction

When facing multiple alternatives, it is reasonable to assume that people tend to choose an alternative which maximizes their utilities. This concept is widely known as random utility maximization (RUM), with the statistical models allowing for perception differences. In transportation research, one of the most famous model implementing this is the multinomial logit (MNL) formulation (Ben-Akiva and Lerman, 1985; McFadden, 1973). Recently there is a growing interest in implementing an alternative modeling approach called random regret minimization (RRM) (Chorus et al., 2008; Chorus, 2010). In RRM, an individual, when choosing between alternatives, is assumed to minimize anticipated regret as opposed to maximizing his/her utility. RRM is a context-dependent modeling approach since the decision to choose one alternative depends on the relative performance of the chosen alternative's attributes against the other alternatives' attributes. This modelling technique has been applied to route choice (Chorus, 2012a, 2012b; Chorus and Bierlaire, 2013; Chorus et al, 2013a; Jang et al., 2017; Li and Huang, 2017), travel information acquisition choice, parking lot, shopping location (Chorus, 2010), automobile fuel choice (Chorus et al, 2013b; Hensher et al. 2013), willingness to pay for advanced transportation services, salary and travel time trade-offs (Hess et al, 2014), activity start time (Golshani, et al., 2018), and for freight transport (Boeri and Masiero, 2014).

RRM has several variants, the classical one (Chorus, 2010), the GRRM (Chorus, 2014), the μRRM (Van Cranenburgh et al. 2015), and the PRRM (Van Cranenburgh et al. 2015). There have been many comparisons on the performance of RRMs with RUM. Chorus et al. (2014) listed 43 empirical studies comparing RUM and RRM from 2010 to 2014. Regarding model fit, in 15 cases RRM outperforms RUM and in another 15 cases, it is the other way round. The other 13 empirical studies show neither of these two modeling approaches outperforms each other. Chorus et al. (2014) also listed seven out of 43 empirical studies that measured a hit rate, which is the percentage of observation correctly predicted by the model, and shows that RRM outperforms RUM in three cases. In two cases the RUM hit rate is higher, while for other two cases both models perform equally well.

Leong and Hensher (2015) compare the value of travel time savings (VTTS) from the results of RUM, RRM, Hybrid RRM, and their new context-dependent alternative model, relative advantage maximization (RAM). They show that the difference in mean VTTS between RUM-RRM and RUM-Hybrid RRM is small but statistically significant for seven route choice data sets from Australia and New Zealand. Chorus and Bierlaire

(2013) compare RUM and RRM elasticities for the case of route choice and found that travel time elasticities of the RRM model are nearly 10% greater compared to RUM. Similarly, for a route choice case, Thiene et al. (2012) showed that for most attributes RRM model elasticities were about 10% greater than the RUM model. For the case of preference of alternative fuel car use, Hensher et al. (2013) compared RRM and RUM elasticities and found a substantial difference in the elasticities with the RRM being higher.

Other than RRM, there is another context-dependent modeling approach that recently has been introduced, RAM (Leong and Hensher, 2015). There have not been many empirical studies comparing the performance of RAM with RUM or RRMs except for route choice models comparison by Leong and Hensher (2015). They found that RAM have a potential in producing a better model fit and obtaining "more precise" model outputs such as VTTS.

It appears that most empirical studies tested the difference of RUM with context-dependent modeling approaches in terms of model fit. Few exceptions compared them in terms of prediction accuracy, VTTS, and demand elasticities. From most of the cases mentioned above, we cannot say for sure which modeling approach is better. Different data sets and contexts might produce different results and biases.

Therefore, the objective of this paper is to compare RUM, RRMs, and RAM comprehensively in term of model fit, prediction accuracy, VTTS, and demand elasticities for travel time and cost. By comparing those different approaches, we might find which model gives the best fit, which modeling approach accurately predicts the choice compared to other approaches. Hopefully, we can contribute to the greater body of RRM and RAM literature through this first comparison of Swiss data sets and one German data set.

In section 2, we discuss the alternative modeling approaches to RUM, their properties, and variants, followed by section 3 where we describe the data sets, and we present the results for the different modeling approaches including prediction accuracy. In section 4 we discuss the VTTS followed by section 5 where we discuss the demand elasticities and hit rates. We discuss another potential modelling alternative in Section 6. Finally, in section 7 the conclusions are drawn. In addition, we derive the methods to calculate the VTTS and elasticities, where they had not been derived before.

2 Alternatives to RUM

2.1 Random regret minimization

Random regret minimization (RRM) was first introduced by Chorus et al. (2008) as a model of travel choice. According to Chorus at al. (2008) in RRM, an individual chooses between alternatives wishing to avoid a situation where a non-chosen alternative turns out to be more attractive than the chosen one, causing regret. Thus, the individual when choosing between alternatives is assumed to minimize anticipated regret as opposed to maximize his/her utility. Chorus (2010) admitted that this first RRM approach has two limitations. Therefore, he improved the model to alleviate those limitations with a new RRM-approach. This new RRM approach (Chorus, 2010) is now widely known as Classical RRM (Van Cranenburgh and Prato, 2016).

In the Classical RRM (CRRM) framework, for a person q, the regret associated with an alternative i is obtained given by the following formula (Chorus, 2012a):

$$RR_{iq} = R_{iq} + \varepsilon_{iq} = \alpha_i + \sum_{j \neq i} \sum_k \ln\left(1 + \exp\left[\beta_k \cdot \left(X_{kjq} - X_{kiq}\right)\right]\right) + \varepsilon_{iq} \tag{1}$$

Where, RR_{iq} : random regret for an alternative i for person q

R_{iq} : systematic regret for alternative i for person q

ε_{iq} : unobserved regret for alternative i for person q

α_i : alternative specific constant

β_k : estimable parameter associated with generic attribute X_k

X_{kjq}, X_{kiq} : values associated with generic attribute X_k for, respectively, person q choosing alternative i over competitor alternative j.

Similar to the RUM formulation of choice probabilities (McFadden, 1973), the classical RRM framework assumes that the error term in Eq. 1 is identically and independently distributed (i.i.d) Extreme Value Type I with a variance of $\pi^2/6$. In the RRM setting, the formulation for the choice probabilities is:

$$P_{iq} = \frac{\exp(-R_{iq})}{\sum_{\substack{i \in J \\ j=1}}^{J} \exp(-R_{jq})} \tag{2}$$

The next variant of RRM idea proposed by Chorus (2014) is called Generalized-RRM (GRRM). This model generalizes the classical RRM by replacing the one inside logarithmic function with a regret-weight parameter γ. Van Cranenburgh et al. (2015) introduced a different version of RRM called µRRM. In this type of RRM, a scale parameter (µ) enters the model as an additional degree of freedom which allows for flexibility of the regret function level attribute. The µRRM generalized the CRRM by allowing to estimate the variance of the error term. The formula for µRRM is as follows (Van Cranenburgh et al. 2015):

$$RR_{iq}^{\mu\mathrm{RRM}} = \alpha_i + R_{iq}^{\mu\mathrm{RRM}} + \varepsilon_{iq} = \alpha_i + \sum_{j \neq i} \sum_k \ln\left(1 + \exp\left[\frac{\beta_k}{\mu} \cdot \left(X_{kjq} - X_{kiq}\right)\right]\right) + \varepsilon_{iq} \tag{3}$$

where $\varepsilon_{iq} \sim i.i.d.EV(0, \mu)$

The formulation for the choice probabilities is as follows:

$$P_{iq}^{\mu RRM} = \frac{\exp\left(-\mu R_{iq}^{\mu RRM}\right)}{\sum_{\substack{i\in J\\ j=1}}^{J}\exp\left(-\mu R_{jq}^{\mu RRM}\right)} \tag{4}$$

The latest version of RRM is also introduced by Van Cranenburgh et al. (2015), P-RRM. The P-RRM is a limiting case of the μRRM model. Classical RRM model and any other RRM variants postulate that both regrets and rejoices are experienced. According to Van Cranenburgh et al. (2015), the P-RRM yields the strongest regret minimization behavior possible within the RRM framework since it postulates no rejoice, which is the opposite of regret, is experienced.

The formula for systematic regret of the P-RRM model (Van Cranenburgh et al., 2015) is as follows:

$$R_{iq}^{P-RRM} = \alpha_i + \sum_k \beta_k X_{kjiq}^{P-RRM} \quad \text{where } X_{kjiq}^{P-RRM} = \begin{cases} \sum_{j\neq i}\max\left(0, X_{kjq} - X_{kiq}\right) \text{ if } \beta_k > 0 \\ \sum_{j\neq i}\min\left(0, X_{kjq} - X_{kiq}\right) \text{ if } \beta_k < 0 \end{cases} \tag{5}$$

The computation of the X-vector (X_{kjiq}^{P-RRM}) is linear and can be done prior to estimation. There is the prerequisite that the signs of the taste parameters are known prior the estimation. Once the X-vectors are obtained, the estimation of the P-RRM model is similar to the estimation of a linear additive RUM model i.e. with the assumption of error term to be i.i.d Extreme Value Type I. The formulation of the choice probabilities is:

$$P_{iq}^{P-RRM} = \frac{\exp\left(-R_{iq}^{P-RRM}\right)}{\sum_{\substack{i\in J\\ j=1}}^{J}\exp\left(-R_{jq}^{P-RRM}\right)} \tag{6}$$

2.2 Relative advantage maximization

Similar to RRM, relative advantage maximization (RAM) also compares the chosen alternative with competing alternatives. However, there is a key difference in the way in which RAM explicitly takes into account the disadvantages and advantages of an alternative. The advantages of alternatives are expressed as a ratio of the sum of advantages and disadvantages.

Leong and Hensher (2015) formulate the disadvantage of person q choosing alternative i over a competing alternative j for attribute k, denoted by D_{kijq}, which is given by:

$$D_{kijq} = \ln\left(1 + \exp\left[\beta_k \cdot \left(X_{kjq} - X_{kiq}\right)\right]\right) \tag{7}$$

Leong and Hensher (2015) assume that disadvantages and advantages are symmetric: The advantage of person q choosing alternative i over j with respect to attribute k equals the corresponding disadvantage of person q choosing alternative j over i with respect to the same attribute. This is given by

$$A_{kijq} = D_{kjiq} = \ln\left(1 + \exp\left[\beta_k \cdot \left(X_{kiq} - X_{kjq}\right)\right]\right) \tag{8}$$

Now, the definition of A_{kijq} is an advantage of the person q choosing alternative i over j, and the definition of D_{kijq} is a disadvantage of the person q choosing alternative j over i. The formula of the total advantage and disadvantage over all alternatives' attributes are as follows:

$$A_{ijq} = \sum_k A_{kijq} \quad \text{and} \quad D_{ijq} = \sum_k D_{kijq} \tag{9}$$

The relative advantage of the person q choosing alternative i over j according to Leong and Hensher (2015) is as follow:

$$RA_{ijq} = \frac{A_{ijq}}{A_{ijq} + D_{ijq}} \tag{10}$$

The observed component of utility for the person q choosing an alternative i is written as linear combination with the MNL formulation. The formula for systematic utility is as follows:

$$V_{iq}^{RAM} = \alpha_i + \sum_{k'} \beta_{k'} X_{k'iq} + \sum_{\substack{i \in J \\ j \neq i}} RA_{ijq} \tag{11}$$

With $X_{k'iq}$ referring to a context-independent attribute k' for person q choosing an alternative i, the RAM model allows for a combination of context-independent preferences and context-dependent preferences. In this paper, we compare the standard RUM model (MNL) with the classical RRM (Chorus, 2010), and the μRRM as well as the P-RRM (Van Cranenburgh et al., 2015). We also compare those approaches with the new RAM approach (Leong and Hensher, 2015). Although the RAM approach allows for incorporation of context-independent attributes, in this paper, we only use context dependent generic attributes k.

3 Model Estimation

3.1 Data Description

Chorus (2010) shows that for binary choice situations, the RRM reduces to the linear-additive RUM. Therefore, in this paper, we select data sets where respondents face at least three alternatives. Note that, these three alternatives can also include an opt-out alternative (Hess et al., 2014).

Table 1 shows the information regarding the data sets used, while the description of the data sets can be found in the next subsection. The data sets are stated choice data sets, and one RP data set all collected in Switzerland. Since RRM is choice set dependent, meaning that choosing an alternative is influenced by the presence of other alternatives in term of their attribute values, we only use a parsimonious model formulation using two generic attributes: travel time (TT) and travel cost (TC).

The Swissmetro was a major innovation proposed for the Swiss transport system. Abay (1999) conducted revealed preference (RP) and stated preference (SP) survey of long-distance road and rail travelers. The details of the data sets can be found in Bierlaire et al. (2001) and Axhausen (2013). For long distance travel, there are three alternatives: Train, Swissmetro (SM), and car. For this paper, we only include data where respondents faced all three choice alternatives. Thus SP data with only two alternatives (Train and SM) are omitted. In total, 5607 observations from 1192 respondents were used for model estimation.

Weis et al. (2012) assessed the effect of parking availability on travelers' behavioral responses in Switzerland. They assumed that in addition to the trade-off between travel time and fuel or transit cost, parking search times and cost have a substantial impact on travelers' decision. Therefore, they conducted a stated choice study of parking, location, and mode choice to assess those choices. The detail of the study is explained in Weis et al. (2012; 2013). We use the data sets to run the models on three different choice sets: location choice, parking choice, and mode choice.

For location choice, there are two alternative locations and one "none of these" option as an opt-out alternative. In total 6301 observations of 631 respondents were used. For the parking choice, there are three alternative choices: parking A, parking B, and the opt-out alternative. In total 5853 observations from 585 respondents were used. We also estimate a mode choice model. There are four mode choice alternatives: walk, bicycle, car, transit. For longer distance travel, walk and bicycle might not be available. Moreover, during the experiment, none of the respondents faced all four available alternatives together. Therefore, for this paper, we only take short distance trips where respondents are facing three choices: walk or bike, car, and transit. In total, only 1666 observations from 168 respondents were used.

The next data sets come from two SP experiments that were conducted to estimate the potential of carpooling in Switzerland. In order to gain insight about user perception regarding innovative modes, the SP part was composed of two different experiments, one of them including car-sharing as an alternative. The details of the survey are available in Ciari and Axhausen (2012; 2013a; 2013b). For the experiment, which includes car-sharing, there are three alternative modes: car-sharing, car, and transit. In total, there are 4350 observations from 735 respondents. In the other experiment, there are four alternative modes: carpooling as a driver (CPD), carpooling as a passenger (CPP), transit, and car. In total, 3975 observations from 511 respondents were used. Note that car is the only alternative that available across all 3975 observations, however since all observations have three available alternatives, we include all 3975 observations in the model.

Still within Swiss context, Schmutz (2015) used data from the Swiss Microcensus 2010 for his study. The Mobility and Transport Microcensus is a survey conducted every five years that provides detailed information on mobility behavior of the Swiss residents. The official data set includes around 300,000 stages, 210,000 trips and 65,000 tours starting and ending at home. In the work of Schmutz (2015), only the travel behavior part of the main survey for travel on one allocated day per individual have been used. Schmutz (2015) presents MNL models for five levels of aggregation: stage, sub-tour, tour, trip, and day plan. In this paper, we only use the trip data set. The alternatives for mode choice are walk, bike, car, and transit. After some filtering, where we only use observations that have all four alternatives available and reasonable walk and bike travel time for all observations, we obtain 33,942 observations. The details of the data set can be found in Schmutz (2015).

The last data set is the German value of time (VOT) and value of reliability (VOR) study by Dubernet and Axhausen (2017). In 2012 Germany's Federal Ministry of Transport and Digital Infrastructure (BMVI) initiated several projects in preparation of the new Federal Transport Investment Plan (BVWP) 2030. This set includes the German VOT and VOR project which estimated VOT and VOR for personal and business travel. From May 2012 until January 2013 nationwide data of more than 3000 participants was collected in a combined two-stage revealed and stated preference survey. For our analysis in this paper, we used responses from 2058 respondents where each of them answered eight different SP mode choice scenarios. In total 15681 observations are included in our model. The alternatives for mode choice are walk, bike, transit, coach, car and airplane. The coach alternative is for long distance travel and was legallized in 2012 in Germany. During the survey, the coach option was still more a hypothetical alternative. The transit option covers local public transport and long distance train. The details of the survey and data set can be found in Dubernet and Axhausen (2017).

3.2 Estimation Results

For the Swissmetro data set, other than generic attributes, travel time and travel cost, we added alternative specific constants (ASC) for each mode to the utility function/regret function, and for identification we set the Swissmetro ASC to zero.

The location choice and parking choice are un-labelled data sets. We use a similar method as in Hess et al. (2014) for the opt-out alternative case. In the first and second utility/regret function we multiply time and cost parameters with respective attributes. Then we include the third utility/regret function where there is only one parameter "none" to be estimated. For mode choice, only car and transit are available across 1666 observations. For those who have the walk alternative, there is no bike alternative and vice versa. We added four ASCs, and we set transit ASC to zero.

For the case of car-sharing, we set the transit ASC to zero in our utility/regret function. As for the carpooling case, only the car alternative is available across 3975 observations. Therefore, we set the ASC of car to zero. For RP mode choice we set the ASC of transit to zero. Finally for the German VOT, only transit and car alternatives are available across all 15681 observations, therefore for this data set we set the ASC of car to zero.

All models are estimated using PythonBiogeme (Bierlaire, 2016). The results for MNL, CRRM, μ RRM, PRRM, and RAM for the eight data sets are presented in Table 2 below. We presented the generic attributes, time and cost, the scale parameter for μRRM, and the ASCs.

For all models, the parameters of time and cost are significant with the expected sign (negative). However, we need to be careful in interpreting these parameters. In the MNL, a parameter estimate refers to increase or decrease in the utility of an alternative caused by a one-unit or one standard deviation increase in an attribute's value. Therefore in the case of our MNL models, the increase by one unit of In the RRM context, a parameter estimate reflects the *potential* increase or decrease in regret associated with comparing a considered alternative with another alternative in term of one unit increase in an attribute's value. As discussed above, RRM is context dependent. Whereas in RUM, the attributes of other alternatives is irrelevant, in RRM attribute levels of others alternatives influence the increase/decrease the regret of the chosen alternative. For RAM context, a parameter estimate reflects the *potential* increase or decrease in relative advantage associated with comparing a considered alternative with another alternative in terms of one unit increase in an attribute's value.

We found the results of ASCs are as expected, with the sign different between utility maximization and regret minimization. For example, in the case of Swissmetro, the negative signs of ASC train and car in utility maximization context (MNL and RAM) indicate that ceteris paribus, those modes are less preferred compared with Swissmetro. Since the CRRM ASCs are insignificant, we looked at µRRM and PRRM, where the ASCs are positive and significant. In a regret minimization context these results indicate that those modes increase regret as they are less preferred compared with Swissmetro. The opt-out alternatives results for location and parking choice are as expected. Negative signs in RUM context indicate that those values reduce the utility while in the RRM context, positive signs indicate that those values increase regret. In several data sets, parking mode choice, car-sharing, carpooling, German VOT, we can observe that car is more preferred compared to other modes. However, we obtain the opposite result in the RP mode choice data set where car is less preferred compared to transit.

We present the model fit comparison in Table 2 that consists of final log-likelihood (F-LL), Akaike information criterion (AIC) and Bayesian information criterion (BIC). Looking at AIC and BIC, we found that RAM outperforms other models in the case of Swissmetro, car-sharing, carpooling, RP mode choice, and German VOT. We also found that µRRM outperforms other models in the case of parking location, while PRRM outperforms other models in the case of a parking location and parking mode choice. Ideally, µRRM should never outperform any other model due to its special properties. It should be between PRRM, CRRM, or MNL. However, since opt-out alternatives are a special case, there is a possibility that µRRM outperforms other models.

Regarding the comparison of MNL and CRRM in terms of model fit, we found that only two times MNL outperforms CRRM, namely in the case of Swissmetro and carpooling. This result underlines the finding in the literature that neither of these two models is always clearly superior.

Another method to give confidence about the model fit is the scale parameter test. Van Cranenburgh et al. (2015) stated that μRRM model has three special cases. If we fix μ arbitrarily high, the μRRM model exhibits RUM behavior, but if we fix μ arbitrarily close to zero, the μRRM model exhibits PRRM behavior, finally, if we fix the μ equal to one, the μRRM model becomes CRRM. Therefore on each of eight data sets, we perform three μ tests. The first test is fixing μ at 10 and then compare the model fit with MNL. The second test is fixing μ at 1 and then compare the model fit with CRRM. For the third test, we fix μ at 0.01 and then compare the model fit with PRRM.

For all of eight data sets, when we fix the μ at 10, we found that almost all μRRM F-LL is similar with MNL log-likelihood except for the unlabeled data sets. When we fix μ at 1, we obtained a F-LL of the μRRM identical to CRRM in all cases. Finally, when we fix μ at 0.01 or any number close to zero, only in three datasets, we obtained similar F-LL as PRRM. Those data sets are parking mode choice, carpooling and car-sharing. Overall the test result gives us the confidence in the MNL, CRRM, and PRRM results.

For Swissmetro data set, the estimate of μ is 1.21 and significantly different from one. That means the F-LL of μRRM should be between CRRM and MNL, which in our case is as expected. For parking location and choice, the μ are arbitrarily high, which means that the F-LLs should be closer to MNL F-LLs. However, due to the presence of an opt-out alternative with only a constant, the F-LL of μRRM will not approach MNL. For parking mode choice, the μ is 1.17 and significantly different from one. That means the F-LL of µRRM should be between CRRM and MNL, which in our case closer to CRRM.

For car-sharing and car-pooling, the µ's are estimated to be closer to zero and are significantly different from one. That means that the F-LL of µRRM should be closer to PRRM F-LL. We found that for those two data sets, the F-LL of each µRRM is closer to PRRM F-LL, which is as expected. For RP mode choice, we found an unexpected case that the µRRM F-LL is not between MNL and CRRM while the µ estimation indicates that µRRM F-LL should be between MNL and CRRM. This is the only RP data set that we tested, therefore this might be a special case for this data set. Finally, for the German VOT, µ is 0.24 and significantly different from one. That means the µRRM F-LL should be closer to PRRM. We found that the µRRM F-LL is between CRRM and PRRM F-LL which is still acceptable.travel time and travel cost of an alternative decreases the utility of that alternative.

3.3 Prediction Accuracy

To test the prediction accuracy of our models, we present the in-sample hit rate, the predicted probabilities, and also the out-of-sample hit rate. The hit rate, which is the percentage of observations correctly predicted by the model, can also be one of the indicators measuring the goodness of fit of a choice model. Hit rate refers to the fit between actual choice observed in the data and the predicted choice obtained by using the model itself. A hit is observed when the probability of the chosen alternative is predicted to be the highest compared to the other alternatives. The higher the hit rate the closer we can say that our model represents reality. We also present the predicted probabilities, which is the one when the model predicts the chosen alternative as the "hit."

To do validation, we also calculate the out of sample hit rate. First, we divide the data set into two parts, the first part, two-third of the data was used for estimation. Then, we used the estimation result to simulate the probability for the other one-third of the data. For validation, we repeated this procedure five times and took the average percentage as the out-of-sample hit rate.

In Table 3, we present the prediction accuracy of five modeling approaches across seven data sets. In the first five columns, we present the hit rate, predicted probability, and out-of-sample hit rate of five models. In the following column, we present the percentage of observations where all models produce the same outcome regardless of the observed choice. In the last column, we show the percentage of observations for which all models correctly predict the chosen alternative.

We found two data sets where the hit rate is above 80%: the cases of parking choice (SP unlabeled data) and RP mode choice. In other four data sets, Swiss metro, location choice, parking mode choice, and German VOT the hit rate is approximately 60%. As for the carpooling data sets the hit rate of all models are below 50% except for the RAM model. The RAM model, in general, shows the highest hit rates except for the case of the parking data sets. Overall, we can say that RUM is outperformed by other approaches in all data sets in terms of hit rate.

The predicted probabilities show quite different results from the hit rates. For example, in the case of Swissmetro, the highest hit rate results from the RAM approach. However the highest predicted probability are shown for μRRM. Interestingly, we can also observe that for the RAM approach, especially for parking location, and parking choice, the predicted probabilities are substantially lower. However, for the other data sets, we can observe that the predicted probabilities of RAM are higher than the other approaches. For the out-of-sample hit rates, we can observe that overall, the out-of-sample hit rates are lower than the all sample hit rates. Although we can also observe that in some cases the out-of-sample hit rates percentage is higher than the all sample hit rate. This is the case in the Swissmetro CRRM, car-sharing CRRM and RAM.

It is interesting to see the distribution of the prediction rate where all models predict the same outcome. For the most of our data sets, the prediction rate is above 80%, more specifically for three data sets, the prediction rate is above 90%. The highest prediction rate can be found in the case of RP mode choice which is almost 100%. The lowest prediction rate is in the case of carpooling; this might be due to the difference in the choice set, some have no walk alternative while the rests have no bike alternative.

The rates for all models to predict the right outcome can be seen in the last column. All the percentages are slightly below the hit rates of all the five modeling approaches for each respective data set. The substantial difference between the percentages of all models predicting correctly, and the hit rate can be found in the case of parking mode choice. This might be due to some observations facing zero alternative for a particular mode (walk or bike). The same reason might be applied to the carpooling and German VOT data set where in the case of carpooling not all of the observations facing all four alternatives and for the case of German VOT none of the observations facing all six alternatives.

3.4 Non-trading behavior

Non-traders are respondents in a stated preference survey who always make the same choice decision regardless of the available alternatives' attributes. In our SP data sets, we have found the percentages of non-traders as shown in Table 4. In this section, we show how non-traders can be predicted from our five modeling approaches.

In the case of Swissmetro, the share of non-traders of the train alternative is only 2.1% (13 respondents), which for this alternative cannot be predicted by all five modeling approaches. For the Swissmetro alternative, we can see that there are 21.7% non-traders. From this number, MNL and μRRM can correctly predict non-traders in 83.7%. For the car alternative, PRRM can give a higher prediction rate than the others, but a low one only.

We could not observe a high share of non-traders for Parking mode choice and also car-pooling. Most of the observed non-trader rates are below 10%. However, in the case of car-sharing, we found high percentages of car non-traders and transit non-traders. For car non-traders, the MNL showed the highest prediction rate, while for transit case, μRRM and PRRM showed the highest prediction rate. Finally, in the case of German VOT, we can observe a quite high share of car non-traders (41.6%), and for the case of car non-traders, MNL showed the highest prediction rate.

We do not present the un-labelled data set as usually there should be much less or even no non-traders in un-labelled experiment. In general, we could say that there is no modeling approach better than the others for all contexts. There is a case where MNL outperforms others, but the RRMs and RAM performs well. However for the cases where the non-trader rates are substantially higher, MNL showed the highest prediction rate compared to others.

4 Value of Travel Time Savings

The value of travel time savings (VTTS) is an important concept for travel demand analysis. It measures how much money (e.g. CHF) a person is willing to pay for a unit reduction in travel time (e.g. an hour). The VTTS for the MNL model can be obtained from Eq.12 below.

$$VTTS_{iq}^{MNL} = 60 \times \frac{\partial V_{iq} / \partial TT_{iq}}{\partial V_{iq} / \partial TC_{iq}} = 60 \times \frac{\beta_{TT}}{\beta_{TC}} \quad (12)$$

Where V_{iq} represents systematic utility for an alternative i for person q, TT_{iq} represents travel time associated with person q choosing an alternative i, and TC_{iq} represent travel cost associated with person q choosing an alternative i. The parameters of travel time and travel cost are represented by β_{TT} and β_{TC} respectively. Since RUM is not context dependent, the VTTS for an alternative is not influenced by other alternatives as in the case of RRM. The methods to measure VTTSs for context-dependent choice models are described below.

4.1 Method to measure context-dependent choice VTTS

4.1.1 CRRM VTTS

To measure the VTTS for CRRM we need to derive the systematic regret of the person q choosing the alternative i with respect to attribute X_{kiq}. The derivation is shown in Eq. 13 below, with more details in Appendix 1.

$$\frac{\partial R_{iq}}{\partial X_{kiq}} = \sum_{\substack{i \in J \\ j \neq i}} \frac{-\beta_k \cdot \exp\left[\beta_k \cdot \left(X_{kjq} - X_{kiq}\right)\right]}{1 + \exp\left[\beta_k \cdot \left(X_{kjq} - X_{kiq}\right)\right]} = \sum_{\substack{i \in J \\ j \neq i}} \frac{-\beta_k}{\frac{1}{\exp\left[\beta_k \cdot \left(X_{kjq} - X_{kiq}\right)\right]} + 1}$$
$$= \sum_{\substack{i \in J \\ j \neq i}} \frac{-\beta_k}{\exp\left[-\beta_k \cdot \left(X_{kjq} - X_{kiq}\right)\right] + 1} \quad (13)$$

Eq. 13 enters the VTTS formula as shown in Eq. 14 below, which also presented in Chorus (2012c).

$$VTTS_{iq}^{CRRM} = 60 \times \frac{\partial R_{iq} / \partial TT_{iq}}{\partial R_{iq} / \partial TC_{iq}} = 60 \times \frac{\sum_{j \neq i} -\beta_{TT} / \left(\exp\left[-\beta_{TT} \cdot \left(TT_{jq} - TT_{iq}\right)\right] + 1\right)}{\sum_{j \neq i} -\beta_{TC} / \left(\exp\left[-\beta_{TC} \cdot \left(TC_{jq} - TC_{iq}\right)\right] + 1\right)} \quad (14)$$

Eq. 14 implies that VTTS measures will generally change when choice set changes in terms of alternatives. Changes in attributes of competing for an alternative as well as changes in attributes of the chosen alternative will influence the VTTS.

4.1.2 μRRM VTTS

The derivative of the systematic regret of the μRRM model is shown in Appendix 2. The formula for deriving the μRRM VTTS is shown in Eq. 15 below.

$$VTTS_{iq}^{\mu RRM} = 60 \times \frac{\partial R_{iq}^{\mu RRM} / \partial TT_{iq}}{\partial R_{iq}^{\mu RRM} / \partial TC_{iq}} = 60 \times \frac{\sum_{j \neq i} -\frac{\beta_{TT}}{\mu} \Big/ \left(\exp\left(-\frac{\beta_{TT}}{\mu} \left[TT_{jq} - TT_{iq} \right] \right) + 1 \right)}{\sum_{j \neq i} -\frac{\beta_{TC}}{\mu} \Big/ \left(\exp\left(-\frac{\beta_{TC}}{\mu} \left[TC_{jq} - TC_{iq} \right] \right) + 1 \right)} \quad (15)$$

4.1.3 PRRM VTTS

Van Cranenburgh and Prato (2016) show the derivative of the systematic regret for PRRM model with respect to attribute X_{kiq} as shown in Eq.16 below:

$$\frac{\partial R_{iq}}{\partial x_{kiq}} = \begin{cases} -\beta_k \sum_{j \neq i} 1 & \text{if } \beta_k < 0 \text{ and } x_{kjq} < x_{kiq} \text{ or } \beta_k > 0 \text{ and } x_{kjq} > x_{kiq} \\ 0 & \text{if } \beta_k < 0 \text{ and } x_{kjq} > x_{kiq} \text{ or } \beta_k > 0 \text{ and } x_{kjq} < x_{kiq} \end{cases} \quad (16)$$

Since we only have two generic attributes and we already know in advanced that our parameter estimates are both negative, then we use the upper left part of Eq.16. Thus, part of Eq.16 enter Eq.17 for deriving the PRRM VTTS.

$$VTTS_{iq}^{PRRM} = 60 \times \frac{\partial R_{iq} / \partial TT_{iq}}{\partial R_{iq} / \partial TC_{iq}} = 60 \times \frac{-\beta_{TT} \sum_{\substack{j \neq i \\ TT_{jq} < TT_{iq}}} 1}{-\beta_{TC} \sum_{\substack{j \neq i \\ TC_{jq} < TC_{iq}}} 1} \quad (17)$$

Let us recall the properties of PRRM as shown in Eq.5, X_{kjiq}^{PRRM} is obtained by the summation of $\min(0, X_{kiq} - X_{kjq})$ in the case of a negative parameter. Therefore in the condition where the chosen alternative is outperformed by the competing for an alternative, the derivative of systematic regret with respect to travel time or travel cost will become zero. If that is the case, we will have an infinite VTTS for the this person and alternative.

4.1.4 RAM VTTS

Leong and Hensher (2015) have already derived an equation for measure RAM VTTS, as shown in Eq. 18:

$$VTTS_{iq}^{RAM} = 60 \times \frac{\partial V_{iq}^{RAM} / \partial TT_{iq}}{\partial V_{iq}^{RAM} / \partial TC_{iq}} = 60 \times \frac{\sum_{j \neq i} \frac{D_{ijq} \frac{\partial A_{ijq}}{\partial TT_{iq}} - A_{ijq} \frac{\partial D_{ijq}}{\partial TT_{iq}}}{\left[A_{ijq} + D_{ijq}\right]^2}}{\sum_{j \neq i} \frac{D_{ijq} \frac{\partial A_{ijq}}{\partial TC_{iq}} - A_{ijq} \frac{\partial D_{ijq}}{\partial TC_{iq}}}{\left[A_{ijq} + D_{ijq}\right]^2}} \tag{18}$$

The derivation of advantage and disadvantage of the person q choosing alternative i over j is in Eq. 19 below:

$$\frac{\partial A_{ijq}}{\partial X_{kiq}} = \frac{\beta_k}{\exp\left[\beta_k \cdot \left(X_{kjq} - X_{kiq}\right)\right] + 1} \text{ and } \frac{\partial D_{ijq}}{\partial X_{kiq}} = \frac{-\beta_k}{\exp\left[-\beta_k \cdot \left(X_{kjq} - X_{kiq}\right)\right] + 1} \tag{19}$$

4.2 VTTS result and discussion

In this subsection, we present the result of VTTS, weighted sample mean value and standard deviation for each alternative for the five models in Table 5.

For un-labelled SPs it does not make sense to present values for the two alternatives as the order in the choice experiment (left or right alternative) does not matter and is quite random. Therefore we only present the mean VTTS from one alternative. Normally, for whichever alternative chosen, the VTTS for MNL is the same (as travel time and cost are generic), however with the weight, the VTTS for each alternatives is different. As for the other context-dependent modelling approaches, we obtain a value for each alternative.

Looking at the results for all data sets, it appears that the values are as expected. However, we found some strange cases where the values are quite low, as in the case for RP mode choice. The value of car and transit VTTS are very low, below 10 CHF/hour which is unexpected. We found also some strange cases where the values are very large, especially for µRRM and RAM model. For RAM model, we found a strange result for the parking choice data set. As for µRRM model we found four high results. One in car-sharing data set for transit alternative and two in German VOT data set for PT and car alternative. Another high result can be found in car-pooling data sets for CPP alternative. In these four data sets, during the survey, the choice games do not only include time and cost trade-offs but also other alternative specific attributes. The extreme outliers result from a choice situation where the trade-off was not only between time and cost but also between one of the other attributes (e.g. number of transfers). This effect especially occurs in the µRRM framework and has to be further investigated in the future.

For a better depiction of the VTTSs distribution, especially for those five strange results discussed above, we plot the VTTS by choice situation for each alternative with a box plot. We only present four out of five strange results. The boxplot for the parking choice data set can be seen in Figure 1. On the x-axis, we present the five context-dependent models. At the y-axis, we present the VTTS in CHF per hour. Due to some outliers we receive a substantially higher mean value for the RAM case. However if we look at Figure 1, we can see that the median value of RAM model is actually quite low, below 10 CHF/hour. As for the other RRM models, their median value are close to MNL median.

In Figure 2, we present the box-plot for car-sharing data sets, specifically the transit alternative. Due to outliers we obtained very high mean value for µRRM VTTS. We can see in the boxplot that the median value of µRRM is far above the MNL median.

We also found that the CRRM median value is above MNL while other two modeling approaches median values are below MNL. We present the boxplot of carpooling data set, specifically carpooling as passenger in Figure 3. We can see in the table that similar to car-sharing, the median VTTS of CRRM, µRRM, and also RAM are above MNL while the PRRM median VTTS is below MNL. Finally, we present the box plot of German VOT data set, specifically the car alternative in Figure 4. As seen in the figure, the median value of all alternatives are below MNL VTTS.

Overall, we have presented VTTS for all data sets and all modelling approaches. Note that for context-dependent modelling approaches, difference choice-set and different modelling approaches produce different VTTS. For example, in the case of Swissmetro, we can see that the context-dependent VTTS for the newly introduced mode are below the MNL VTTS. A similar case can be found in car-sharing where the VTTS of car-sharing alternative overall below other competing alternatives. In the German VOT data we can see that the context-dependent VTTS for newly introduced mode, the coach, are below the MNL (except in the µRRM case).

5 Travel Time and Cost Elasticities

Direct elasticities are another way to compare the behavioral implications of RUM and RRM. The direct elasticities derive from the RUM and RRM model shows the relationship between a percentage change in the magnitude of the attribute and the percentage change in the probability of choosing an alternative based on respected attribute. To measure the disaggregate direct point elasticities of RUM model, we can use the formula from Ben-Akiva and Lerman (1985) shown in Eq. 20. In the following subsection, we show how to measure direct elasticities for the three variant RRM models.

$$E_{iqX_{kiq}} = \frac{\partial P_{iq}}{\partial X_{kiq}} \cdot \frac{X_{kiq}}{P_{iq}} = (1 - P_{iq}) \cdot \beta_k \cdot X_{kiq} \tag{20}$$

5.1 Context-dependent elasticities measurement

Hensher et al. (2013) derive for the first time an equation to measure RRM (CRRM) elasticities as shown in Eq. 21. This equation according to Van Cranenburgh and Prato (2016) can also be used to measure the PRRM and µRRM elasticities.

$$E_{iqX_{kiq}} = \left(-\frac{\partial R_{iq}}{\partial X_{kiq}} + \sum_{\substack{i \in J \\ j \neq i \\ j=1}}^{J} P_{jq} \cdot \frac{\partial R_{jq}}{\partial X_{kiq}} \right) \cdot X_{kiq} \tag{21}$$

5.1.1 CRRM elasticities

The formula to measure the disaggregate direct point elasticities for CRRM (as discussed in Hensher et al. 2013) is as follow.

$$E^{CRRM}_{iqX_{kiq}} = \left(-\frac{\partial R^{CRRM}_{iq}}{\partial X_{kiq}} + \sum_{\substack{i \in J \\ j \neq i \\ j=1}}^{J} P_{jq} \cdot \frac{\partial R^{CRRM}_{jq}}{\partial X_{kiq}} \right) \cdot X_{kiq} =$$

$$\left(\left(-\sum_{\substack{i \in J \\ j \neq i \\ j=1}}^{J} \frac{-\beta_k}{\exp\left[-\beta_k \cdot \left(X_{kjq} - X_{kiq}\right)\right] + 1} \right) + \left[\sum_{\substack{i \in J \\ j \neq i \\ j=1}}^{J} P_{iq} \frac{-\beta_k}{\exp\left[-\beta_k \cdot \left(X_{kjq} - X_{kiq}\right)\right] + 1} + \left(\sum_{\substack{i \in J \\ j \neq i \\ j=1}}^{J} P_{jq} \frac{\beta_k}{\exp\left[\beta_k \cdot \left(X_{kjq} - X_{kiq}\right)\right] + 1} \right) \right] \right) \cdot X_{kiq} \tag{22}$$

5.1.2 µRRM elasticities

From Appendix 1, we find that the derivative of µRRM with respect to attribute X_k as shown

$$\frac{\partial R_{iq}^{\mu RRM}}{\partial X_{kiq}} = \sum_{\substack{i \in J \\ j \neq i \\ j=1}}^{J} \frac{-\frac{\beta_k}{\mu}}{\exp\left(-\frac{\beta_k}{\mu}\left[x_{kjq} - x_{kiq}\right]\right) + 1} \tag{23}$$

In order to measure µRRM disaggregate direct point elasticities, we derive the formula from Eq. 21 and Appendix 2 to get Eq. 24 below.

$$E_{iqX_{kiq}}^{\mu RRM} = \left(-\frac{\partial R_{iq}^{\mu RRM}}{\partial X_{kiq}} + \sum_{\substack{i \in J \\ j \neq i \\ j=1}}^{J} P_{jq} \cdot \frac{\partial R_{jq}}{\partial X_{kiq}} \right) \cdot X_{kiq}$$

$$= \left(\left(-\sum_{\substack{i \in J \\ j \neq i \\ j=1}}^{J} \frac{-\frac{\beta_k}{\mu}}{\exp\left(-\frac{\beta_k}{\mu}\left[x_{kjq} - x_{kiq}\right]\right) + 1} \right) + \left[\left(\sum_{\substack{i \in J \\ j \neq i \\ j=1}}^{J} P_{iq} \frac{-\frac{\beta_k}{\mu}}{\exp\left(-\frac{\beta_k}{\mu}\left[x_{kjq} - x_{kiq}\right]\right) + 1} \right) + \left(\sum_{\substack{i \in J \\ j \neq i \\ j=1}}^{J} P_{jq} \frac{\frac{\beta_k}{\mu}}{\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq} - x_{kiq}\right]\right) + 1} \right) \right] \right) \cdot X_{kiq} \tag{24}$$

5.1.3 PRRM elasticities

We can use the same formula as Eq. 16 to derive PRRM disaggregate direct point elasticities. Van Cranenburgh and Prato (2016) already derived the formula as shown in Eq. 25 below.

$$E_{iqX_{kiq}}^{PRRM} = \left(-\frac{\partial R_{iq}}{\partial X_{kiq}} + \sum_{\substack{i \in J \\ j \neq i \\ j=1}}^{J} P_{jq} \cdot \frac{\partial R_{jq}}{\partial X_{kiq}} \right) \cdot X_{kiq}$$

$$\text{where } \frac{\partial R_{iq}}{\partial X_{kiq}} = \begin{cases} -\beta_k \sum_{j \neq i} 1 \text{ if } X_{kjq} < X_{kiq} \\ 0 \text{ if } X_{kjq} > X_{kiq} \end{cases} \text{ and } \frac{\partial R_{jq}}{\partial X_{kiq}} = \begin{cases} \beta_k \text{ if } X_{kjq} > X_{kiq} \\ 0 \text{ if } X_{kjq} < X_{kiq} \end{cases} \tag{25}$$

5.1.4 RAM elasticities

We derive the formula to measure RAM elasticities in Appendix 3, as follows:

$$E^{RAM}_{iqX_{kiq}} = \frac{\partial P_{iq}}{\partial X_{kiq}} \cdot \frac{X_{kiq}}{P_{iq}} = \frac{\partial \ln P_{iq}}{\partial X_{kiq}} \cdot X_{kiq} = \left(\frac{\partial RA_{ijq}}{\partial X_{kiq}} - \sum_{\substack{i \in J \\ j \neq i \\ j=1}}^{J} P_{jq} \frac{\partial RA_{jiq}}{\partial X_{kiq}} \right) \cdot X_{kiq} \tag{26}$$

The derivation for $\frac{\partial RA_{ijq}}{\partial X_{kiq}}$ can be seen in Leong and Hensher (2015) to measure RAM VTTS as shown in Eq. 18. While the derivation of $\frac{\partial RA_{jiq}}{\partial X_{kiq}}$ is as follows:

$$\frac{\partial RA_{jiq}}{\partial X_{kiq}} = \frac{D_{jiq} \cdot \frac{\partial A_{jiq}}{\partial X_{kiq}} - A_{jiq} \cdot \frac{\partial D_{jiq}}{\partial X_{kiq}}}{\left[A_{jiq} + D_{jiq} \right]^2} \tag{27}$$

The derivation of advantage and disadvantage of choosing alternative *j* over alternative *i* is similar to Eq. 19 as shown below:

$$\frac{\partial A_{jiq}}{\partial X_{kiq}} = \frac{-\beta_k}{\exp\left[-\beta_k \cdot \left(X_{kjq} - X_{kiq} \right) \right] + 1} \quad \text{and} \quad \frac{\partial D_{jiq}}{\partial X_{kiq}} = \frac{\beta_k}{\exp\left[\beta_k \cdot \left(X_{kjq} - X_{kiq} \right) \right] + 1} \tag{28}$$

Substituting Eq. 28 to Eq. 26, the formula to derive disaggregate direct point elasticities of RAM is as follows:

$$\begin{aligned} E^{RAM}_{iqX_{kiq}} &= \left(\frac{\partial RA_{ijq}}{\partial X_{kiq}} - \sum_{\substack{i \in J \\ j \neq i \\ j=1}}^{J} P_{jq} \frac{\partial RA_{jiq}}{\partial X_{kiq}} \right) \cdot X_{kiq} \\ &= \left(\begin{array}{l} \left(\sum_{\substack{i \in J \\ j \neq i \\ j=1}}^{J} \frac{D_{ijq} \cdot \frac{\partial A_{ijq}}{\partial X_{kiq}} - A_{ijq} \cdot \frac{\partial D_{ijq}}{\partial X_{kiq}}}{\left[A_{ijq} + D_{ijq} \right]^2} \right) + \left(-\sum_{\substack{i \in J \\ j \neq i \\ j=1}}^{J} P_{iq} \frac{D_{ijq} \cdot \frac{\partial A_{ijq}}{\partial X_{kiq}} - A_{ijq} \cdot \frac{\partial D_{ijq}}{\partial X_{kiq}}}{\left[A_{ijq} + D_{ijq} \right]^2} \right) \\ + \left(-\sum_{\substack{i \in J \\ j \neq i \\ j=1}}^{J} P_{jq} \frac{D_{jiq} \cdot \frac{\partial A_{jiq}}{\partial X_{kiq}} - A_{jiq} \cdot \frac{\partial D_{jiq}}{\partial X_{kiq}}}{\left[A_{jiq} + D_{jiq} \right]^2} \right) \end{array} \right) \cdot X_{kiq} \end{aligned} \tag{29}$$

5.2 *Travel time elasticities*

To compare direct elasticities between models, we need to calculate the aggregate direct point elasticities for each model. The formula, presented in Atasoy et al (2013), is shown below:

$$E^{W_i}_{iqx_{kiq}} = \sum_{q=1}^{Q_s} E_{iqX_{kiq}} \frac{w_q P_{iq}}{\sum_{q=1}^{Q_s} w_q P_{iq}} \quad (30)$$

where w_q represents the sample weight for individual q from sample Q_s from population Q and $E_{iqX_{kiq}}$ is the disaggregate elasticity of demand of individual q for variations in attribute x_{kiq}. We weighted each observation in our data sets according to the representation of its age and gender category in Swiss population (Swiss Federal Statistical Office, 2011) and German population (Statistisches Bundesamt, 2014).

We present the travel time elasticities across the five models and seven data sets in Table 6. In the first column, we present the alternatives followed by the MNL elasticities and RRMs elasticities in the next column. As the RRM is the alternative model to the well-established RUM, therefore in the next column we present the percentage difference between RUM and the four other context-dependent models.

The sign of all the time elasticities measurement are as expected which means that a percentage increase in travel time will on an average reduces the probability of choosing an alternative. From the Swissmetro data set, we can see that travel time for the Swissmetro alternative is nearly inelastic across all models. That means 10% increase in travel time for Swissmetro on average will not have a substantial impact on the reduction of Swissmetro probability. But that is not the case for the train alternative with the MNL model, where a 10% increase of traintravel time on average will give a 17% reduction in the respective choice probability. In the context of RRM, a 10% increase of train travel time results in a 26% reduction of the respective choice probability, takes into account the level of travel time associated with car travel time and Swissmetro travel time. The percentage difference between RRM and MNL, with RRM being substantially higher, might suggest the idea that when the wrong choice may have been taken, this amplifies the response away from normal RUM based elasticity.

In general, across all data sets and alternatives, we can see a substantial difference between MNL and CRRM with CRRM values being higher, except for some cases, for example, three cases in RP mode choice and four cases in German VOT. From the behavioral perspective, this might suggest that accounting for possibility of making the wrong choices might amplifies the behavioral responses that one would normally attribute to a RUM based elasticity (as also discussed in Hensher et al., 2013)

In the next column, we can see that the percentage difference between MNL and μRRM are substantially high especially for car-sharing, car-pooling, and German VOT. Note that these are cases where the scale parameter (μ) are estimated closer to zero. The next column shows the percentage difference between MNL and PRRM. PRRM is a special case of μRRM where it postulates no rejoices, just pure regret. Therefore, it makes sense that across all data sets and alternatives we can see a substantially higher difference compared to the difference between RUM and CRRM (except where there are strange results of μRRM). Finally, for the RAM model; it is interesting that across all alternatives and data sets, the RAM models are inelastic.

5.3 Cost elasticities

We present the average of cost elasticities across four models and six data sets in Table 7. For walk and bike alternative the travel cost is zero. This cases are therefore removed from the table.

Overall the travel cost attributes are less elastic compared to the travel times. For MNL and PRRM, we can observe that all cases are inelastic except the plane option in German VOT case. For CRRM approach, all cases are inelastic while for the RAM case, only train alternative in Swissmetro and coach alternative in German VOT are elastic. Regarding µRRM, we can observe that travel cost are mostly elastic across all alternatives in the case of car-sharing, car-pooling, and German VOT. These are cases where the scale parameters were estimated near zero, closer to PRRM.

Regarding the percentage difference between MNL and other modelling approaches, we can observe that the percentage differences are quite high in the case of CRRM and PRRM with PRRM differences being substantially higher compared to CRRM difference. For RAM case, we can observe that the difference are substantially higher than the case of car-sharing, car-pooling and also German VOT. Similar to the travel time elasticities, cost elasticities difference between MNL and µRRM are substantially higher in those cases where the scale parameter are closer to zero.

6 Discussion and Another Modelling Alternative

Having investigated all of the modelling approaches, the model fit comparison, the prediction accuracy, VTTS and elasticities, the next question would be, what modelling approach should we use? We present a summary assessment of the models in Table 8 below.

The positive sign on the left hand side of the slash symbol represents a model that outperforms other models in term of model fit, while an "x" sign represents a model that outperforms other models in term of hit rate. The dollar sign represents the highest probability and "o" sign represents the highest out-of-sample hit rate. From the Table 8, we can see that MNL is not the best in any data set (except for parking mode choice hit rate and German VOT out-of-sample hit rate), neither does CRRM (except for location hit rate and out-of-sample hit rate). It appears that there is no clear model which can be deemed best of the MNL and CRRM. Therefore we might say that when choosing modelling technique, it is up to the modeler, if he/she is a utility maximizer then MNL is the necessary approach, but if he/she is a regret minimizer, then choosing CRRM, which is already an established modelling technique might be the alternative.

Regarding the hit rate, CRRM shares the same hit rate with the other two random regret models in the case of location choice. In the case of location choice, the best model fit is obtained by µRRM. The µRRM also holds the best hit rate in the case of parking choice, where in the same data set, PRRM obtains the best model fit. PRRM also has the best model fit for the case of parking mode choice. Overall it appears that regret minimization models are powerful in the context of unlabeled data sets with an opt-out alternative.

Finally, we can observe that for most of the labeled data sets RAM outperforms other models in term of model fit and prediction rate. This result is quite surprising but also show us that there is a stronger alternative modelling approach in terms of model fit and prediction accuracy. The notion behind the RAM technique is utility maximization with the relative advantage approach. Looking from a regret minimization perspective, we can also turn relative advantage maximization into relative regret minimization (RERM). Using the formula of total advantage and disadvantage shown in Eq. 9, we can propose relative regret formula as shown in Eq. 31 below.

$$RER_{ijq} = \frac{D_{ijq}}{A_{ijq} + D_{ijq}} \tag{31}$$

The formula for systematic regret is as follows:

$$R_{iq}^{RERM} = \alpha_i + \sum_{\substack{i \in J \\ j \neq i}} RER_{ijq} \tag{32}$$

Unlike the RAM formula, in RERM, which based on the regret minimization framework, we do not incorporate context-independent attribute k' for person q choosing an alternative i. The formulation of the choice probabilities is:

$$P_{iq}^{RERM} = \frac{\exp\left(-R_{iq}^{RERM}\right)}{\sum_{\substack{i \in J \\ j=1}}^{J} \exp\left(-R_{jq}^{RERM}\right)} \tag{33}$$

The result of the RERM can be seen in Table 9.

Overall the results of RERM are similar to RAM, especially for the generic attributes time and cost. The magnitude of ASCs are also exactly the same except that the sign is different. For example in the Swissmetro case, a positive car ASC can be interpreted as, ceteris paribus, car giving more regret compared to Swissmetro. Therefore, people prefer to choose Swissmetro compared to car. The only difference in magnitude can be found in the opt-out alternative. For parking location and parking choice, the opt-out for RAM are -0.70 and -0.86 respectively, while for RERM the opt-out are 2.70 and 2.85 respectively.

The model fit of RAM and RERM is also similar. We also simulate the model to find the hit rate, out-of-sample hit rate and also predicted probabilities. Their results are the same as for the RAM so we do not present them here. We also calculate the VTTS and aggregate direct point elasticities. We found that the aggregate direct point elasticities of RERM are also the same with RAM. However, for VTTS, we found some cases where there are significant difference between the results of RAM and RERM as shown in Table 10 below.

We present the weighted sample mean and standard deviation of the VTTS and also the difference between the mean (t-test) in the last column. We can observe that for the case of Swissmetro, the mean value for three modes are significantly different. The mean VTTS of Train and SM VTTS for RERM are substantially higher than RAM. However, for car, the mean VTTS of RAM is substantially higher than RERM. For the unlabeled case with an opt-out alternative, we can observe significant differences where the RERM VTTS for both data sets are substantially higher than RAM.

In the case of car-sharing, we can observe significant difference between RAM and RERM where for all modes RAM VTTS are substantially higher than RERM. In RP mode choice, we found that there is a small but significant difference between RAM and RERM VTTS for car modes. For parking mode choice, carpooling, RP mode choice (except car), and German VOT, there is no significant difference between the RAM VTTS and RERM VTTS.

Overall, the significant difference between RAM VTTS and RERM VTTS for some data sets might indicate that the modeler might use the RERM as another alternative modelling approach to measure VTTS for transportation infrastructure projects. The significant difference of VTTS for two unlabeled data sets might indicate that if this model is applied to other un-labeled data sets such as route choice experiments, the RAM and RERM VTTS result might be significantly different, however further testing is necessary to confirm that.

To summarize, it appears there is no clear winner between RUM and classic RRM, although our results show that random regret minimization appears powerful for the unlabeled case and several labeled data sets. It is clear that RAM results outperform other modeling approaches, but mostly for labeled data sets. We might also say that these five alternative modelling approaches also belong to the family of similarity measures such as the C-logit model (see for example Rieser-Schüssler and Axhausen, 2007). This approach obviously has no behavioral interpretation, but is a possibility to capture the correlations between the alternatives better.

7 Conclusions

There have been many empirical studies which compare the performance of RUM with context-dependent alternative modeling approaches such as RRM. In many cases, empirical studies reported a better goodness of fit of RRM compared to the RUM. However, there are also similar numbers of other empirical studies which reported that RUM model fit is better than RRM. Apparently, there has not been a consistent result for which modeling approach is better than other. While most previous empirical studies reported model fit, few of them presented the prediction accuracy, VTTS and also demand elasticity to compare those modeling approaches.

Our goal is to comprehensively compare five different modeling approaches, RUM, CRRM, µRRM, PRRM, and RAM using Swiss and German data sets. We presented model fit, prediction accuracy, non-traders, VTTS and demand elasticity across five modeling approaches and eight data sets. With only two generic attributes, time and cost, we found in our model that the parameters of those attributes are significant with expected sign.

Our comparison of MNL and CRRM underlines the literature results that neither of the two approaches, RUM and RRM is superior in all cases. The addition of µRRM in the model comparison can give us more confidence in the modelling result due to the properties of µRRM. We can test µRRM by fixing the scale parameter to arbitrarily high and compare the result with MNL. When we fix the scale parameter to one, we obtain CRRM result, and finally, when we fix the scale parameter arbitrarily closer to zero, we obtain a result closer to PRRM. We have tested these three properties for our eight data sets, and we obtain results as expected. Looking at our scale parameter estimate, we have confidence that our model is correct since the model fit close to the model fit MNL when it is high, close to CRRM when the estimation is close to one, and closer to PRRM when the estimation is closer to zero.

In term of prediction rate, our results show that in many cases the RUM and RRM hit rate is almost the same. Surprisingly the hit rate of the new RAM model is slightly higher than for the others especially for labeled data and RP data. We also present predicted probabilities and out-of-sample hit rate where the results are slightly lower than the all sample hit rate. We found an interesting result that more than 80% of almost all models predict the same outcome. This indicates that whatever model we use, we might end up obtaining the same outcome. If we want to use the model which has higher model fit, then we might be able to use these comparisons, but in term of correct prediction, it might be a different case.

For the VTTS, we find that for MNL the result is the same for all alternatives in the case of Swissmetro, car-sharing, carpooling, and RP mode choice. But for the case of the other four data sets, we can obtain different mean VTTS. We found some under-estimated value in the case of RP mode choice, where the mean VTTS is too low. We also found some cases that show the mean VTTS is ludicrously high, for example, $7.10*10^{10}$ CHF/hour for the case of transit alternative in car-sharing. These strange results require further examination. We presented boxplots for some alternative where we found strange, mean values. The VTTS results open up new approaches for the policy analyst to introduce new alternative modes, as can be seen from the case of new modes such as Swissmetro and long distance bus.

In terms of time and cost aggregate direct point elasticities, the signs are as expected. But we found that in many cases the difference between MNL and other models is substantially high. For regret case, this might be due to the potential regret that will be faced by the person choosing that alternative. A similar explanation might be applied to RAM.

Overall, we can observe that from model fit and prediction accuracy results, the RAM approaches is superior compared to others especially for the case of labeled data sets. These results inspire us to propose a new approach which still similar with RAM but from the regret minimization perspective, which we named Relative Regret Minimization (RERM). Since we tested the model with only two generic attributes, time and cost, we obtain similar results with RAM except for the sign of the ASCs. The model fit, prediction accuracy and aggregate direct point elasticities of RERM is also similar with RAM, the only significant difference we found is in the VTTS measurement.

There are some limitations of this study. First, we only use two generic attributes for all our models. While this makes comparison easier, we cannot capture other relevant factors that might influence the decision, especially in the multi-attribute choice context. There are some consequences of leaving out attributes in the estimation, one example are the high values for the VTTS for four data sets, since during the survey, the choice games do not only include time and cost trade-offs but also other alternative specific attributes. Second, we have tried to be as comprehensive as possible as we include both labeled and unlabeled data sets in our presentation. However, the unlabeled data sets that we have are the ones where we have an opt-out alternative. We do not have an unlabeled data set where there are three alternatives with three generic attributes. For future study, it would be better to add more RP data so that we can better compare and draw safer conclusion. Since the modeling approaches that we presented here include context-dependent models, different choice sets and different contexts might produce different results. Therefore more empirical results are necessary especially for the RAM and RERM approaches.

Acknowledgement

We would like to thank Caspar Chorus, Sander Van Cranenburgh, and Michiel Bliemer for their suggestions regarding the RRM modeling. We are grateful for having a discussion with Michel Bierlaire, Claude Weiss, and Francesco Ciari regarding the seven data sets.

Disclosure Statement

No potential conflict of interest was reported by the authors.

References

Abay, G. (1999) Nachfrageabschätzung Swissmetro: Eine stated-preference Analyse, *Berichte des Nationalen Forschungsprogrammes 41 "Verkehr und Umwelt"*, **F1**, EDMZ, Bern.

Atasoy, B., A. Glerum and M. Bierlaire (2013) Attitudes towards mode choice in Switzerland, *disP – The Planning Review*, **49**, 101-117.

Axhausen, K.W. (2013) SwissMetro, *Travel Survey Metadata Series*, **42**, Institute for Transport Planning and System (IVT), ETH Zürich, Zürich.

Ben-Akiva, M. and S.R. Lerman (1985) *Discrete Choice Analysis*: *Theory and Application to Travel Demand*, MIT Press, Cambridge, MA.

Bierlaire, M (2016) PythonBiogeme: a short introduction. *Report TRANSP-OR 160706 Series on Biogeme*, Transport and Mobility Laboratory, School of Architecture, Civil and Environmental Engineering, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

Bierlaire M., K.W. Axhausen and G. Abay (2001) The acceptance of modal innovation: The case of Swissmetro, paper presented at the *1st Swiss Transport Research Conference*, Ascona, March 2001.

Boeri, M., and L. Masiero (2014) Regret minimisation and utility maximisation in a freight transport context, *Transportmetrica A: Transport Science*, **10** (6), 548-560.

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

Chorus, C. G. (2012a) Logsums for utility-maximizers and regret-minimizers, and their relation with desirability and satisfaction, *Transportation Research Part A: Policy and Practice*, **46** (7) 1003-1012.

Chorus, C. G. (2012b) Regret theory-based route choices and traffic equilibria, *Transportmetrica*, **8** (4), 291-305.

Chorus, C.G. (2012c) Random Regret Minimization: An Overview of Model Properties and Empirical Evidence, *Transport Reviews*, **32** (1) 75-92.

Chorus, C. G. (2014) A generalized random regret minimization model, *Transportation Research Part B: Methodological*, **68**, 224-238.

Chorus, C. G. and M. Bierlaire (2013) An empirical comparison of travel choice models that capture preferences for compromise alternative, *Transportation*, **40** (3) 549-562.

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

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

Chorus, C.G., M. J. Koetse and A. Hoen (2013b) Consumer preferences for alternative fuel vehicles: Comparing a utility maximization and a regret minimization model, *Energy Policy*, **61**, 901-908.

Chorus, C.G., S. Van Cranenburgh and T. Dekker (2014) Random regret minimization for consumer choice modelling: Assesment of empirical evidence, *Journal of Business Research*, **67** (11) 2428-2436.

Ciari, F. and K.W. Axhausen (2012) Choosing carpooling or carsharing as a mode: Swiss stated choice experiments, paper presented at the *91st Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2012.

Ciari, F. and K.W. Axhausen (2013a) Carpooling Stated Preference (SP), *Travel Survey Metadata Series*, **46**, IVT, ETH Zürich, Zürich.

Ciari, F. and K.W. Axhausen (2013b) Carsharing Stated Preference (SP), *Travel Survey Metadata Series*, **47**, IVT, ETH Zürich, Zürich.

Dubernet, I. and K.W. Axhausen (2017) The German value of time (VOT) and value of reliability (VOR) study: The survey work, *Arbeitsberichte Verkehrs- und Raumplanung, IVT, ETH Zurich, Zurich.*

Golshani, N., R. Shabanpour, J. Auld and A. Mohammadian (2018) Activity start time and duration: incorporating regret theory into joint discrete–continuous models. *Transportmetrica A: Transport Science*, 1-19.

Hensher, D.A., W.H. Greene and C.G. Chorus (2013) Random regret minimization or random utility maximization: An exploratory analysis in the context of automobile fuel choice, *Journal of Advanced Transportation*, **47** (7) 667-678.

Hess, S., M.J. Beck and C.G. Chorus (2014) Contrast between utility maximisation and regret minimisation in the presence of opt out alternatives, *Transportation Research Part A: Policy and Practice*, **66**, 1-12.

Jang, S., S. Rasouli and H. Timmermans (2017) Bias in random regret models due to measurement error: formal and empirical comparison with random utility model. *Transportmetrica A: Transport Science*, **13** (5), 405-434.

Leong, W. And D.A. Hensher (2015) Contrast of Relative Advantage Maximisation with Random Utility Maximisation and Regret Minimisation, *Journal of Transport Economics and Policy*, **49** (1) 167-186.

Li, M. and H. J. Huang (2017) A regret theory-based route choice model, *Transportmetrica A: Transport Science*, **13** (3), 250-272.

McFadden, D. (1973) Conditional Logit Analysis of Qualitative Choice Behavior, in P. Zarembka (Ed.) *Frontiers in Econometrics*, 105-142, Academic Press, New York.

Rieser-Schüssler, Nadine, and Kay W. Axhausen (2007) Recent developments regarding similarities in transport modelling, paper presented at the *7th Swiss Transport Research Conference*, Ascona, September 2007.

Schmutz, S. A. (2015) Effect of analytical units and aggregation rules on mode choice models, Master Thesis, IVT, ETH Zurich, Zurich.

Statistisches Bundesamt (2014) Bevölkerung nach Alter in Jahren und Geschlecht für Gemeinden, Ergebnisse des Zensus am 9. Mai 2011, https://www.destatis.de/DE/Methoden/Zensus_/Downloads/2F_BevoelkerungAlterGeschlecht.html, accessed July 13th, 2017.

Swiss Federal Statistical Office (2011) Permanent and non permanent resident population by canton, sex, citizenship, country of birth and age, 2010-2011, https://www.pxweb.bfs.admin.ch/, accessed July 03rd, 2017.

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

Van Cranenburgh, S. and C.G. Prato (2016) On the robustness of random regret minimization modelling outcomes towards omitted attributes, *Journal of Choice Modelling*, **18**, 51-70.

Van Cranenburgh, S., C.A. Guevara and C.G. Chorus (2015) New insights on random regret minimisation model, *Transportation Research Part A: Policy and Practice*, **74**, 91-109.

Weis, C., M. Vrtic, P. Widmer and K.W. Axhausen (2012) Influence of parking on location and mode choice: A stated choice survey, paper presented at the *91st Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2012.

Weis, C., M. Vrtic, P. Widmer and K.W. Axhausen (2013) Influence of Parking on Location and Mode Choice: A Stated Choice Survey, *Travel Survey Metadata Series*, **41**, IVT, ETH Zürich, Zürich.

Appendix. 1. Derivation of CRRM

In this appendix, we present comprehensively how to derive the systematic regret of CRRM in order to measure VTTS and elasticities. The derivation of RRM elasticities has previously shown in Hensher et al. (2013) as well as in Van Cranenburgh and Prato (2016). The derivation of systematic regret for an alternative i for person q with respect to attribute X_{kiq} is shown below:

$$\begin{aligned}
\frac{\partial R_{iq}^{CRRM}}{\partial x_{iq}} &= \frac{\partial}{\partial x_{kiq}} \sum_{\substack{i \in J \\ j \neq i}} \ln\left(1 + \exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right)\right) = \sum_{\substack{i \in J \\ j \neq i}} \frac{\partial}{\partial x_{kiq}} \ln\left(1 + \exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right)\right) \\
&= \sum_{\substack{i \in J \\ j \neq i}} \frac{1}{1 + \exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right)} \cdot \frac{\partial}{\partial x_{kiq}} \left(1 + \exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right)\right) \\
&= \sum_{\substack{i \in J \\ j \neq i}} \frac{\frac{\partial}{\partial x_{kiq}}(1) + \frac{\partial}{\partial x_{kiq}} \left(\exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right)\right)}{1 + \exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right)} \\
&= \sum_{\substack{i \in J \\ j \neq i}} \frac{0 + \exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right) \cdot \frac{\partial}{\partial x_{kiq}} \left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right)}{1 + \exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right)} \\
&= \sum_{\substack{i \in J \\ j \neq i}} \frac{\exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right) \cdot \beta_k \left(\frac{\partial}{\partial x_{kiq}}\left(x_{kjq}\right) - \frac{\partial}{\partial x_{kiq}}\left(x_{kiq}\right)\right)}{1 + \exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right)} \\
&= \sum_{\substack{i \in J \\ j \neq i}} \frac{\exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right) \cdot \beta_k (0 - 1)}{1 + \exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right)} \\
&= \sum_{\substack{i \in J \\ j \neq i}} \frac{-\beta_k \cdot \exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right)}{1 + \exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right)} = \sum_{\substack{i \in J \\ j \neq i}} \frac{\left(-\beta_k \cdot \exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right)\right) / \exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right)}{\left(1 + \exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right)\right) / \exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right)} \\
&= \sum_{\substack{i \in J \\ j \neq i}} \frac{-\beta_k}{\frac{1}{\exp\left(\beta_k \left[x_{kjq} - x_{kiq}\right]\right)} + 1} = \sum_{\substack{i \in J \\ j \neq i}} \frac{-\beta_k}{\exp\left(-\beta_k \left[x_{kjq} - x_{kiq}\right]\right) + 1}
\end{aligned} \tag{34}$$

The formula to measure direct elasticities for RRM is as follows:

$$
\begin{aligned}
E_{iqX_{kiq}} &= \frac{\partial P_{iq}}{\partial X_{kiq}} \cdot \frac{X_{kiq}}{P_{iq}} = P_{iq} \cdot \frac{\partial \ln P_{iq}}{\partial X_{kiq}} \cdot \frac{X_{kiq}}{P_{iq}} = \frac{\partial \ln P_{iq}}{\partial X_{kiq}} \cdot X_{kiq} \\
&= \left(\frac{\partial \ln\left(\frac{\exp\left(-R_{iq}\right)}{\sum_j \exp\left(-R_{jq}\right)} \right)}{\partial X_{kiq}} \right) \cdot X_{kiq} = \left(-\frac{\partial R_{iq}}{\partial X_{kiq}} - \frac{\partial \ln\left(\sum_j \exp\left(-R_{jq}\right)\right)}{\partial X_{kiq}} \right) \cdot X_{kiq} \\
&= \left(-\frac{\partial R_{iq}}{\partial X_{kiq}} - \frac{\partial \ln\left(\sum_j \exp\left(-R_{jq}\right)\right)}{\partial X_{kiq}} \right) \cdot X_{kiq} \\
&= \left(-\frac{\partial R_{iq}}{\partial X_{kiq}} - \frac{1}{\sum_j \exp\left(-R_{jq}\right)} \cdot \frac{\partial\left(\sum_j \exp\left(-R_{jq}\right)\right)}{\partial X_{kiq}} \right) \cdot X_{kiq} \\
&= \left(-\frac{\partial R_{iq}}{\partial X_{kiq}} - \frac{1}{\sum_j \exp\left(-R_{jq}\right)} \cdot \sum_J \frac{\partial\left(\exp\left(-R_{jq}\right)\right)}{\partial X_{kiq}} \right) \cdot X_{kiq} \\
&= \left(-\frac{\partial R_{iq}}{\partial X_{kiq}} + \frac{1}{\sum_j \exp\left(-R_{jq}\right)} \cdot \sum_J \exp\left(-R_{jq}\right) \frac{\partial R_{jq}}{\partial X_{kiq}} \right) \cdot X_{kiq} \\
&= \left(-\frac{\partial R_{iq}}{\partial X_{kiq}} + \sum_J P_{jq} \frac{\partial R_{jq}}{\partial X_{kiq}} \right) \cdot X_{kiq}
\end{aligned}
\qquad (35)
$$

As mentioned by Van Cranenburgh and Prato, the formula in Eq. 31 can be used to measure elasticities for all RRMs. The formula to measure systematic regret for an alternative j for person q with respect to attribute X_{kiq} is shown below:

$$\frac{\partial R_{jq}^{CRRM}}{\partial x_{kiq}} = \frac{\partial}{\partial x_{kiq}} \sum_{\substack{i\in J\\ j\neq i}} \ln\left(1+\exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)\right) = \sum_{\substack{i\in J\\ j\neq i}} \frac{\partial}{\partial x_{kiq}} \ln\left(1+\exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)\right)$$

$$= \sum_{\substack{i\in J\\ j\neq i}} \frac{1}{1+\exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)} \cdot \frac{\partial}{\partial x_{kiq}} \left(1+\exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)\right)$$

$$= \sum_{\substack{i\in J\\ j\neq i}} \frac{\frac{\partial}{\partial x_{kiq}}(1) + \frac{\partial}{\partial x_{kiq}}\left(\exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)\right)}{1+\exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)}$$

$$= \sum_{\substack{i\in J\\ j\neq i}} \frac{0+\exp\left(\beta_k\left[x_{kjq}-x_{kiq}\right]\right)\cdot \frac{\partial}{\partial x_{kiq}}\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)}{1+\exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)} \tag{36}$$

$$= \sum_{\substack{i\in J\\ j\neq i}} \frac{\exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)\cdot \beta_k\left(\frac{\partial}{\partial x_{kiq}}(x_{kiq}) - \frac{\partial}{\partial x_{kiq}}(x_{kjq})\right)}{1+\exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)}$$

$$= \sum_{\substack{i\in J\\ j\neq i}} \frac{\exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)\cdot \beta_k(1-0)}{1+\exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)} = \sum_{\substack{i\in J\\ j\neq i}} \frac{\beta_k \cdot \exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)}{1+\exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)}$$

$$= \sum_{\substack{i\in J\\ j\neq i}} \frac{\left(\beta_k \cdot \exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)\right)/\exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)}{\left(1+\exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)\right)/\exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)}$$

$$= \sum_{\substack{i\in J\\ j\neq i}} \frac{\beta_k}{\frac{1}{\exp\left(\beta_k\left[x_{kiq}-x_{kjq}\right]\right)}+1} = \sum_{\substack{i\in J\\ j\neq i}} \frac{\beta_k}{\exp\left(\beta_k\left[x_{kjq}-x_{kiq}\right]\right)+1}$$

Appendix. 2. Derivation of μRRM

The derivation of systematic regret for an alternative i for person q with respect to attribute X_{kiq} is shown below:

$$
\begin{aligned}
\frac{\partial R_{iq}^{\mu RRM}}{\partial x_{kiq}} &= \frac{\partial}{\partial x_{kiq}} \sum_{\substack{i\in J\\ j\neq i}} \ln\left(1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)\right) = \sum_{\substack{i\in J\\ j\neq i}} \frac{\partial}{\partial x_{kiq}} \ln\left(1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)\right) \\
&= \sum_{\substack{i\in J\\ j\neq i}} \frac{1}{1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)} \cdot \frac{\partial}{\partial x_{kiq}} \left(1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)\right) \\
&= \sum_{\substack{i\in J\\ j\neq i}} \frac{\frac{\partial}{\partial x_{kiq}}[1] + \frac{\partial}{\partial x_{kiq}} \exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)}{1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)} \\
&= \sum_{\substack{i\in J\\ j\neq i}} \frac{0+\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right) \cdot \frac{\partial}{\partial x_{kiq}} \left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)}{1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)} \\
&= \sum_{\substack{i\in J\\ j\neq i}} \frac{\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right) \cdot \frac{\beta_k}{\mu} \left(\frac{\partial}{\partial x_{kiq}}\left[x_{kjq}\right] - \frac{\partial}{\partial x_{kiq}}\left[x_{kiq}\right]\right)}{1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)} \qquad (37) \\
&= \sum_{\substack{i\in J\\ j\neq i}} \frac{\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right) \cdot \frac{\beta_k}{\mu}(0-1)}{1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)} = \sum_{\substack{i\in J\\ j\neq i}} \frac{-\frac{\beta_k}{\mu} . \left(\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)\right)}{1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)} \\
&= \sum_{\substack{i\in J\\ j\neq i}} \frac{\left(-\frac{\beta_k}{\mu} . \left(\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)\right)\right) \Big/ \left(\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)\right)}{\left(1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)\right) \Big/ \left(\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)\right)} \\
&= \sum_{\substack{i\in J\\ j\neq i}} \frac{-\frac{\beta_k}{\mu}}{\frac{1}{\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)}+1} = \sum_{\substack{i\in J\\ j\neq i}} \frac{-\frac{\beta_k}{\mu}}{\exp\left(-\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)+1}
\end{aligned}
$$

The formula in Eq. 31 can be used to measure µRRM elasticities. The derivation of systematic regret for an alternative j for person q with respect to attribute X_{kiq} is shown below:

$$
\begin{aligned}
&\frac{\partial R_{jq}^{\mu RRM}}{\partial x_{kiq}} = \frac{\partial}{\partial x_{kiq}} \sum_{\substack{i\in J \\ j\neq i}} \ln\left(1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right)\right) \\
&= \sum_{\substack{i\in J \\ j\neq i}} \frac{\partial}{\partial x_{kiq}} \ln\left(1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right)\right) \\
&= \sum_{\substack{i\in J \\ j\neq i}} \frac{1}{1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right)} \cdot \frac{\partial}{\partial x_{kiq}} \left(1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right)\right) \\
&= \sum_{\substack{i\in J \\ j\neq i}} \frac{\frac{\partial}{\partial x_{kiq}}[1] + \frac{\partial}{\partial x_{kiq}} \exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right)}{1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right)} \\
&= \sum_{\substack{i\in J \\ j\neq i}} \frac{0+\exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right) \cdot \frac{\partial}{\partial x_{kiq}} \left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right)}{1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right)} \\
&= \sum_{\substack{i\in J \\ j\neq i}} \frac{\exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right) \cdot \frac{\beta_k}{\mu} \left(\frac{\partial}{\partial x_{kiq}}\left[x_{kiq}\right] - \frac{\partial}{\partial x_{kiq}}\left[x_{kjq}\right]\right)}{1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right)} \qquad (38) \\
&= \sum_{\substack{i\in J \\ j\neq i}} \frac{\exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right) \cdot \frac{\beta_k}{\mu}(1-0)}{1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right)} = \sum_{\substack{i\in J \\ j\neq i}} \frac{\frac{\beta_k}{\mu} . \left(\exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right)\right)}{1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right)} \\
&= \sum_{\substack{i\in J \\ j\neq i}} \frac{\left(\frac{\beta_k}{\mu} . \left(\exp\left(\frac{\beta_k}{\mu}\left[xx_{kiq}-x_{kjq}\right]\right)\right)\right) \Big/ \left(\exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right)\right)}{\left(1+\exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right)\right) \Big/ \left(\exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right)\right)} \\
&= \sum_{\substack{i\in J \\ j\neq i}} \frac{\frac{\beta_k}{\mu}}{\frac{1}{\exp\left(\frac{\beta_k}{\mu}\left[x_{kiq}-x_{kjq}\right]\right)}+1} = \sum_{\substack{i\in J \\ j\neq i}} \frac{\frac{\beta_k}{\mu}}{\exp\left(\frac{\beta_k}{\mu}\left[x_{kjq}-x_{kiq}\right]\right)+1}
\end{aligned}
$$

Van Cranenburgh and Prato (2016) have shown a formula to measure PRRM elasticities therefore in this paper we do not present the measurement for PRRM elasticities.

Appendix. 3. Derivation of RAM

Leong and Hensher (2015) have shown the derivatives of the systematic utility of RAM in order to measure VTTS. The derivation of systematic utility for an alternative i for person q with respect to attribute X_{kiq} is shown below:

$$\frac{\partial V_{iq}^{RAM}}{\partial x_{kiq}} = \sum_{\substack{i \in J \\ j \neq i}} \frac{\partial\left(\frac{A_{ijq}}{A_{ijq}+D_{ijq}}\right)}{\partial X_{kiq}} = \sum_{\substack{i \in J \\ j \neq i}} \frac{\frac{\partial A_{ijq}}{\partial X_{kiq}} \cdot \left(A_{ijq}+D_{ijq}\right) - A_{ijq} \cdot \frac{\partial\left(A_{ijq}+D_{ijq}\right)}{\partial X_{kiq}}}{\left(A_{ijq}+D_{ijq}\right)^2}$$

$$= \sum_{\substack{i \in J \\ j \neq i}} \frac{A_{ijq} \cdot \frac{\partial A_{ijq}}{\partial X_{kiq}} + D_{ijq} \cdot \frac{\partial A_{ijq}}{\partial X_{kiq}} - A_{ijq} \cdot \frac{\partial A_{ijq}}{\partial X_{kiq}} - A_{ijq} \cdot \frac{\partial D_{ijq}}{\partial X_{kiq}}}{\left(A_{ijq}+D_{ijq}\right)^2}$$

$$= \sum_{\substack{i \in J \\ j \neq i}} \frac{D_{ijq} \cdot \frac{\partial A_{ijq}}{\partial X_{kiq}} - A_{ijq} \cdot \frac{\partial D_{ijq}}{\partial X_{kiq}}}{\left(A_{ijq}+D_{ijq}\right)^2} \qquad (39)$$

where $\frac{\partial A_{ijq}}{\partial X_{kiq}} = \frac{\partial R_{jq}^{CRRM}}{\partial x_{kiq}} = \frac{\beta_k}{\exp\left(\beta_k\left[x_{kjq}-x_{kiq}\right]\right)+1}$

and $\frac{\partial D_{ijq}}{\partial X_{kiq}} = \frac{\partial R_{iq}^{CRRM}}{\partial x_{kiq}} = \frac{-\beta_k}{\exp\left(-\beta_k\left[x_{kjq}-x_{kiq}\right]\right)+1}$

The formula to measure direct elasticities for RAM is as follows:

$$E_{iqX_{kiq}}^{RAM} = \frac{\partial P_{iq}}{\partial X_{kiq}} \cdot \frac{X_{kiq}}{P_{iq}} = P_{iq} \cdot \frac{\partial \ln P_{iq}}{\partial X_{kiq}} \cdot \frac{X_{kiq}}{P_{iq}} = \frac{\partial \ln P_{iq}}{\partial X_{kiq}} \cdot X_{kiq}$$

$$= \left(\frac{\partial \ln\left(\frac{\exp\left(V_{iq}\right)}{\sum_j \exp\left(V_{jq}\right)}\right)}{\partial X_{kiq}} \right) \cdot X_{kiq} = \left(\frac{\partial V_{iq}}{\partial X_{kiq}} - \frac{\partial \ln\left(\sum_j \exp\left(V_{jq}\right)\right)}{\partial X_{kiq}}\right) \cdot X_{kiq}$$

$$= \left(\frac{\partial V_{iq}}{\partial X_{kiq}} - \frac{\partial \ln\left(\sum_j \exp\left(V_{jq}\right)\right)}{\partial X_{kiq}}\right) \cdot X_{kiq} = \left(\frac{\partial V_{iq}}{\partial X_{kiq}} - \frac{1}{\sum_j \exp\left(V_{jq}\right)} \cdot \frac{\partial\left(\sum_j \exp\left(V_{jq}\right)\right)}{\partial X_{kiq}}\right) \cdot X_{kiq} \qquad (40)$$

$$= \left(\frac{\partial V_{iq}}{\partial X_{kiq}} - \frac{1}{\sum_j \exp\left(V_{jq}\right)} \cdot \sum_J \frac{\partial\left(\exp\left(V_{jq}\right)\right)}{\partial X_{kiq}}\right) \cdot X_{kiq}$$

$$= \left(\frac{\partial V_{iq}}{\partial X_{kiq}} - \frac{1}{\sum_j \exp\left(V_{jq}\right)} \cdot \sum_J \exp\left(V_{jq}\right) \frac{\partial V_{jq}}{\partial X_{kiq}}\right) \cdot X_{kiq}$$

$$= \left(\frac{\partial V_{iq}}{\partial X_{kiq}} - \sum_J P_{jq} \frac{\partial V_{jq}}{\partial X_{kiq}}\right) \cdot X_{kiq}$$

The derivation of systematic utility for an alternative j for person q with respect to attribute X_{kiq} is shown below:

$$\frac{\partial V_{jq}^{RAM}}{\partial x_{kiq}} = \sum_{\substack{i\in J\\ j\neq i}} \frac{\partial\left(\frac{A_{jiq}}{A_{jiq}+D_{jiq}}\right)}{\partial X_{kiq}} = \sum_{\substack{i\in J\\ j\neq i}} \frac{\frac{\partial A_{jiq}}{\partial X_{kiq}}\cdot\left(A_{jiq}+D_{jiq}\right)-A_{jiq}\cdot\frac{\partial\left(A_{jiq}+D_{jiq}\right)}{\partial X_{kiq}}}{\left(A_{jiq}+D_{jiq}\right)^2}$$

$$= \sum_{\substack{i\in J\\ j\neq i}} \frac{A_{jiq}\cdot\frac{\partial A_{jiq}}{\partial X_{kiq}}+D_{jiq}\cdot\frac{\partial A_{jiq}}{\partial X_{kiq}}-A_{jiq}\cdot\frac{\partial A_{jiq}}{\partial X_{kiq}}-A_{jiq}\cdot\frac{\partial D_{jiq}}{\partial X_{kiq}}}{\left(A_{jiq}+D_{jiq}\right)^2}$$

$$= \sum_{\substack{i\in J\\ j\neq i}} \frac{D_{jiq}\cdot\frac{\partial A_{jiq}}{\partial X_{kiq}}-A_{jiq}\cdot\frac{\partial D_{jiq}}{\partial X_{kiq}}}{\left(A_{jiq}+D_{jiq}\right)^2} \quad (41)$$

where $\frac{\partial A_{jiq}}{\partial X_{kiq}} = \frac{\partial R_{iq}^{CRRM}}{\partial x_{kiq}} = \frac{-\beta_k}{\exp\left(-\beta_k\left[x_{kjq}-x_{kiq}\right]\right)+1}$

and $\frac{\partial D_{jiq}}{\partial X_{kiq}} = \frac{\partial R_{jq}^{CRRM}}{\partial x_{kiq}} = \frac{\beta_k}{\exp\left(\beta_k\left[x_{kjq}-x_{kiq}\right]\right)+1}$