Towards more behaviourally robust travel demand forecasts: Catering to utility maximisers and regret minimisers

Thijs Dekker
Sander van Cranenburgh
Caspar Chorus

UNITE Conference,
Copenhagen, September 2013

Outline

1. Motivation and research objective

2. The RRM model

3. Empirical performance of the RRM model

4. Forecasting using alternative decision rules

5. Forecasting using RRM: Dutch National Model

----

6. Conclusions and discussion

1. Motivation and research objective

Different stages in forecasting travel demand:

1. Model estimation / calibration :
   - Discrete choice models based on RUM premises (e.g. four stage models – trip generation, mode choice, route choice) .
2. Forecasting / sample enumeration:
   - Choice probabilities for specific alternatives and specific types of agents.
   - Weighting choice probabilities to obtain market shares.
3. Sensitivity analyses:
   - Testing different model parameters
   - Evaluating alternative demographic patterns
   - No test on behavioural decision rules other than RUM

- Significant evidence of heterogeneity in decision rules (Leong and Hensher, 2012; Chorus, 2013)
- Decision rules affect choice probabilities → different forecasts

1. Motivation and research objective

Primary research question:

- How to accommodate forecasts from alternative behavioural decision rules in travel demand forecasting exercises?
  - Random Regret Minimization framework an obvious counterpart to RUM

Secondary research objective:

- Evaluate the implementability of the RRM model for aggregate forecasting
  - To what extent will RRM predictions differ from RUM?
  - How to deal with these differences (see main question)?
  - What challenges can be foreseen?

2. The RRM model

Random Regret Minimization as an alternative decision rule:

- Considerable support for regret minimization in psychological literature
- RRM translates this notion into a tractable discrete choice model
- Introduced as a counterpart of RUM (Chorus 2008, 2010)
- Growing body of empirical studies on differences between (linear additive) RUM and RRM
  - Environmental, Health and Transport Economics

Current view:

- RRM has proven itself relative to RUM
  - Econometrics work out; comparable model fit
  - Intuitive explanations for deviations from RUM
- Waiting for the 'next level' of the RRM model:
  - Implementation in forecasting exercises (current paper)
  - Development of a consistent welfare framework (work in progress)

2. The RRM model

- Regret minimization postulates that a decision-maker chooses the alternative with minimum random regret (*RR*$_i$)
- Observed regret *R*$_i$ is a function of the performance of the alternative, <u>relative</u> to the performance of all other alternatives in the choice set.
- Performance contrasted at the attribute level

Summation over alternatives *j*

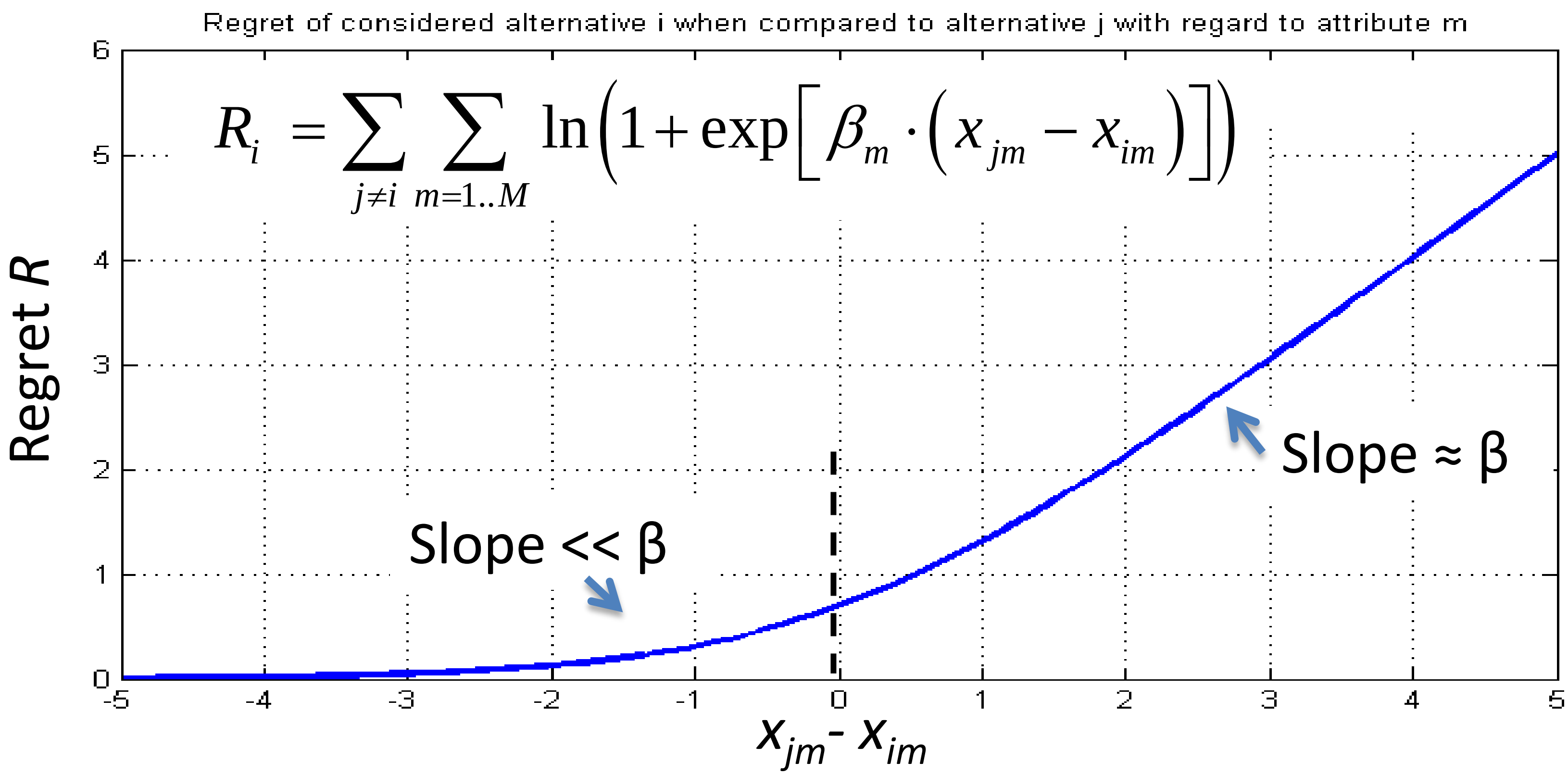
Summation over attributes *m*

$$RR_i = \sum_{j \neq i} \sum_{m=1..M} \ln\left(1 + \exp\left[\beta_m \cdot \underline{\left(x_{jm} - x_{im}\right)}\right]\right) + \varepsilon_i$$

Random Regret associated with alternative *i*

Pairwise comparison of attribute *m* of alternative *i* with attribute *m* of alternative *j*

2. The RRM model



⇒ RRM predicts that having an increasingly <u>poor</u> performance on one attribute causes <u>much</u> additional regret, while having an increasingly <u>strong</u> performance on another attribute does <u>not</u> necessarily <u>compensate</u> for this.

⇒ Therefore, RRM predicts that it is relatively effective (in terms of avoiding regret and gaining market share) to select a compromise alternative.

2. The RRM model

- When -ε is *iid* type I EV, then the well-known and convenient MNL closed-form expression for choice probabilities is obtained:

$$RR_i = R_i + \varepsilon_i \quad \overset{\varepsilon \sim iid\ EV}{\Rightarrow} \quad P_i = \frac{\exp\left(-R_i\right)}{\sum_{j=1..J} \exp\left(-R_j\right)}$$

- Using flexible specifications of the error term, correlation structures can be captured. This translates into well-known model forms like the Nested Logit model, Mixed Logit models (Error Components or Random Parameters), Probit model, ...
- RRM choice probabilities can directly replace RUM choice probabilities in forecasting exercises

2. RRM model vs. linear-additive RUM (MNL-form)

	RRM	RUM
Differences		
Satisfies IIA		X
Strong foundation in neo-classical welfare economics		X
Captures semi-compensatory behaviour: such as the compromise effect	X	
In line with observations in behavioural / psychological economics (i.e. context effects)	X	

Features of the RRM model:

- Preferences are context-dependent in a predictable fashion
- Composition of the choice set matters
- Produces relevant output for forecasting:
  - Choice probabilities
  - Elasticities
  - Willingness-to-pay / Value of Time
  - Logsums

3. Empirical performance of the RRM model

Chorus et al. (2013a)

- Overview of 19 peer-reviewed articles
- 33 empirical comparisons between RUM and RRM model
- Comparisons on:
  - Model fit and external validity
  - Choice probabilities, market shares and elasticities
  - Willingness-to-pay and related welfare measures

3. Empirical performance: Model fit

Model fit:

- RRM or Hybrid RRM-RUM model on average outperforms RUM
- Differences are (very) small, but significant

External validity:

- Predictive performance on hold-out data again comparable (e.g. hit-rate)
- Results not necessarily consistent with differences in model fit

Similarities not surprising due to close connection between RUM and RRM, and the aggregate nature of these two measures:

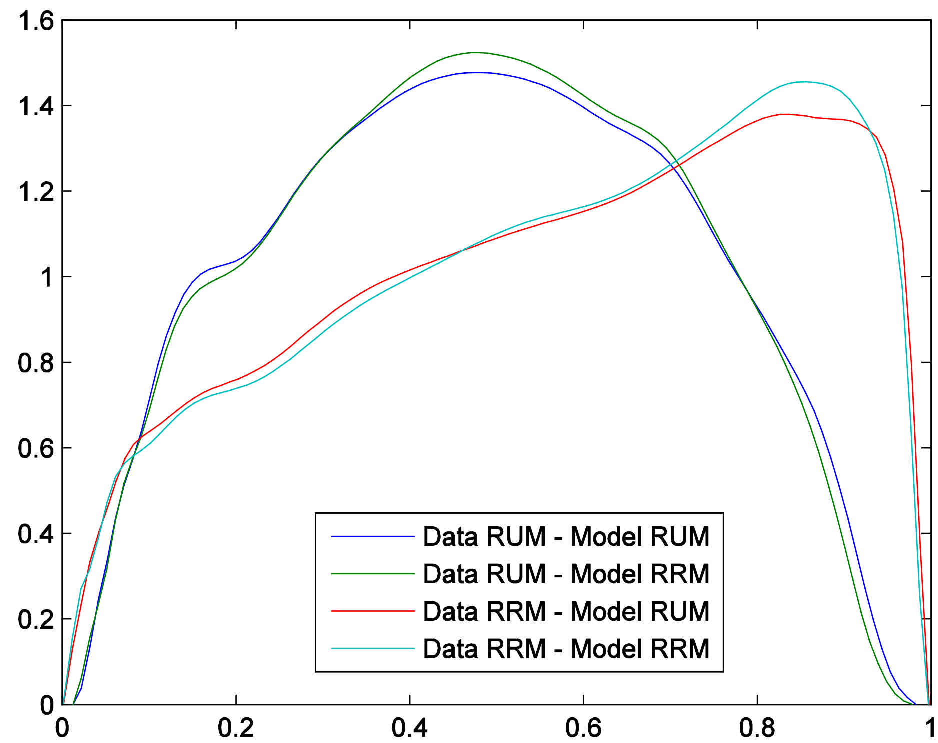
1. Logit type choice probabilities
2. Same # of parameters (d.o.f.) to describe the same data
3. Binary RUM = Binary RRM
4. Differences at observation level are likely to be averaged out

3. Empirical performance: Model fit

Model fit – Synthetic data

- 3 alternatives, 3 attributes, 5,000 observations
- 2 datasets: RUM and RRM based decisions
- On average, similar performance

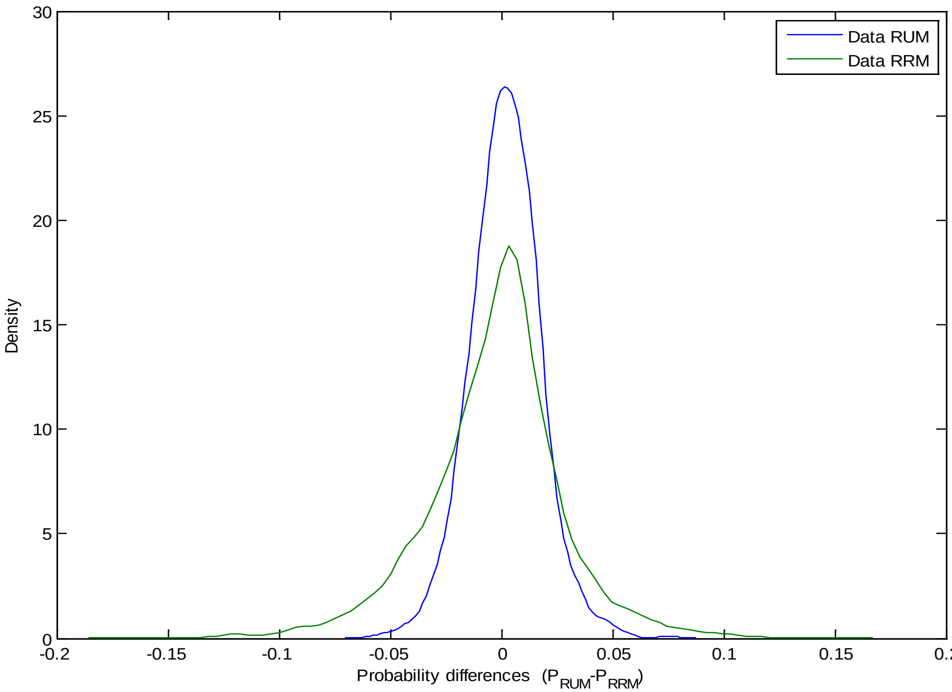
Model Fit	Data	
Model	RUM	RRM
RUM	-4.345,9	-3.651,5
RRM	-4.351,2	-3.625,4

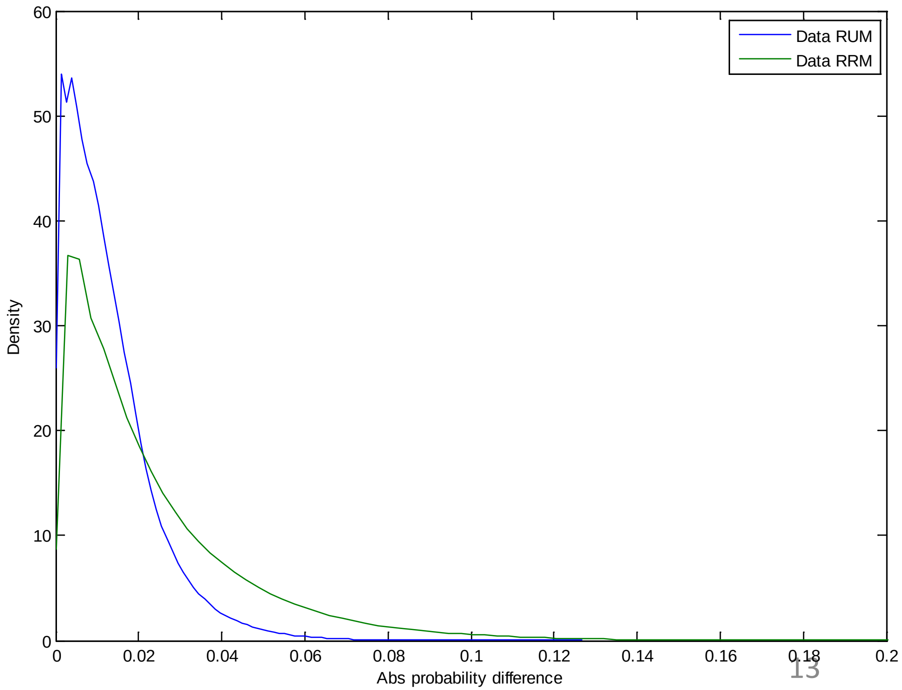


3. Empirical performance: Choice probabilities

Disaggregate comparison – Synthetic data

- Substantial differences in choice probabilities at the choice task level
- Differences in choice probabilities up to 17.2 pct points (empirical 19%)
- Different ‘winners’ in 3% of cases (7% in Chorus et al. 2013b - AFVs)





3. Empirical performance: Choice probabilities

Disaggregate comparison

- Differences in choice probabilities related to choice task composition
- Most prominent: ‘compromise effect’
  - Average performance on all attribute levels (middle alternative)
  - Chorus and Bierlaire (2013): 27% RRM vs. 23% RUM; aggregate level
  - Similar observations in Chorus et al. (2013b), de Bekker-Grob and Chorus (2013)

Implications for market shares / aggregate demand forecasting:

- Less clear, since differences are weighted across different types of agents and choice situations:
  1. Clearly specified forecasting scenarios more likely to generate differences in predicted markets shares etc. (limited possibilities for averaging out)
  2. Differences across types of agents still relevant for policy makers

3. Empirical performance: Elasticities

Elasticities

- Directly comparable between RUM and RRM
- ..vary by observation in *both* the RUM and RRM model
- Overall, average elasticities are reported without confidence intervals
  - Thiene et al. (2012) – RRM higher elasticities for 6/8 attributes, but lower cost elast
  - Greene et al. (2012) – RRM lower cost elasticity
  - Hensher et al. (2011) – RRM consistently higher elasticities
  - Chorus and Bierlaire (2013) – No significant differences

Elasticities: synthetic data

- Again, at observation level sometimes higher elasticities for RRM
- Differences directly related to the shape of regret function
  - RUM more responsive when alternative performs well on that attribute

3. Empirical performance: WTP

Willingness-to-pay / Value of Time

- Marginal rate of substitution well-defined in RUM literature
- Chorus et al. (2012a, 2013c) develop an RRM alternative
  - Definition: MRS to keep regret of alternative *i* constant
  - Neglects impact of x_{im} on R_j
  - Redefinition of indifference concept required (work in progress)
- Observed differences in trade-offs directly related to semi-compensatory behaviour (up to 20%, de Bekker-Grob and Chorus 2013)

Logsum

- Describes the expected minimum regret from a choice set
- In contrast to RUM, does not necessarily improve when an alternatives performance is improved on an attribute (context dependency!)
- Extension to measure of Consumer Surplus only possible in hybrid RUM-RRM (linear treatment of cost attribute)

3. Empirical performance: Implications for forecasting

The RRM model:

- Can provide the necessary inputs for the forecasting exercise
- ….but welfare measure such as VoT and Consumer Surplus work-in-progr.
- Differences between RUM and RRM often averaged out,
- ….but possibly substantial for specific scenarios
- Direction of these averages easily explained by the context dependency of the RRM model

Preliminary conclusion:

- Forecasting is possible with the RRM model
- Implications for future behaviour can be different from RUM in *different* directions
- RRM is conceptually ready for the ‘next level’

4. Uncertainty in decision rules

- Hard to decide about the ‘true’ model based on model fit and external validity, but relatively easy to estimate both types of models

Options:

1. Arbitrary selection criterion to select either RUM or RRM
   - Best model fit
   - RUM proven track record, needs no introduction
   - Neglects uncertainty about the decision rule
2. Implement both the RUM and RRM model
   - Conduct similar sensitivity analysis and establish confidence intervals
   - Forecast are ‘robust from a behavioural perspective’
3. Use model averaging approaches

4. Uncertainty in decision rules: model weights

Models selection based on model fit:

- Classical estimation: Akaike (or Bayesian) Information Criterion
- Bayesian estimation: Marginal likelihood
  - Both approaches include a penalty for fit and additional parameters
  - AIC and marginal likelihood contain no information on relative importance

Transform into model weights:

- 'Probability that *Mi* is the best model after observing the data'
- Bayes Rule:

$$p(M_i \mid y) = \frac{p(M_i)\,p(y \mid M_i)}{\sum_j p(M_j)\,p(y \mid M_j)}$$

- Akaike weights: (logit equivalent)

$$w_i(AIC) = \frac{\exp\left\{-\frac{1}{2}\Delta_i(AIC)\right\}}{\sum_j \exp\left\{-\frac{1}{2}\Delta_j(AIC)\right\}}$$

4. Uncertainty in decision rules: averaging

Model averaging:

- For details see Wagemakers and Farrell (2004), or Hoeting et al. (2007)
- Model specific prediction of concept of interest *g(β)*, conditional on *β*
- Average across models:

$$E\left[g(\beta)\,|\,y\right]=\sum_{j}^{J}E\left[g\left(\beta_j\right)|\,y,M_j\right]p\left(M_j\,|\,y\right)$$

$$E\left[g(\beta)\,|\,y\right]=\sum_{j}^{J}E\left[g\left(\beta_j\right)|\,y,M_j\right]w_j\left(AIC\right)$$

- Intuitive approach to provide a single measure of interest to policy makers whilst taking into account uncertainty about the underlying behavioural decision rule
- Empirical applications extend beyond RUM-RRM comparison

5. RRM-based Dutch National Model

Dutch national model:

- Tool for policy evaluation of large transport projects
- Medium to long-term forecasts on national scale
- Underlying choice models based on RUM premises and model choices at the individual level

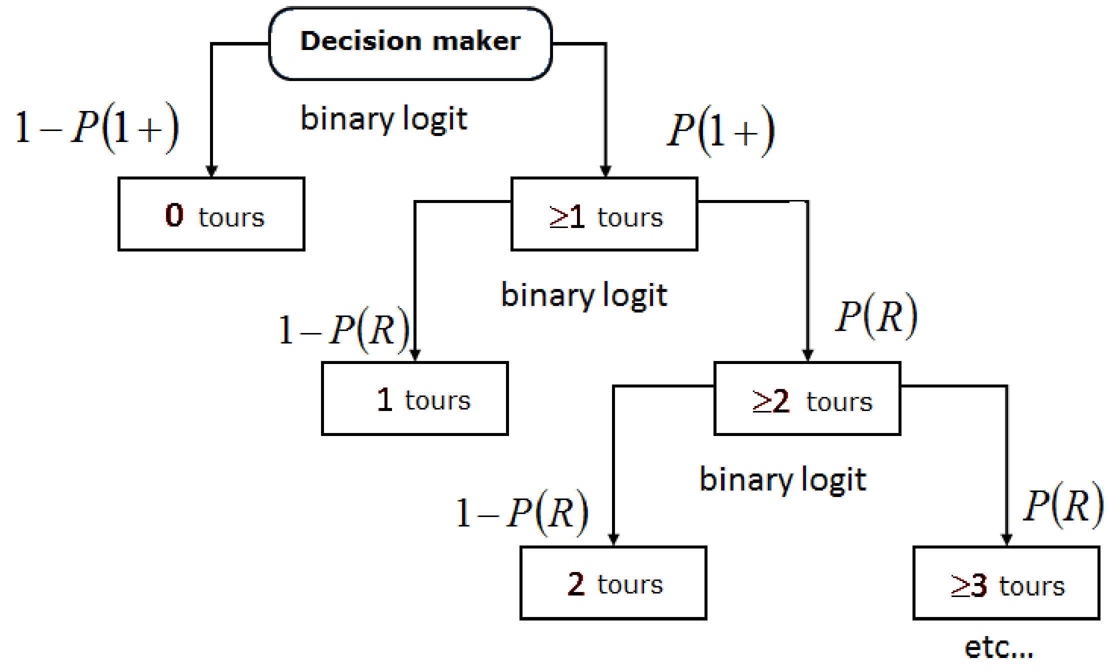
Current research project (S. van Cranenburgh):

- Develop an RRM based alternative of the underlying choice models
  - Tour-frequency models
  - Destination/mode/time-of-day models
  - Route assignment model

5. Dutch National Model: Tour frequency

Tour frequency model:

- Series of binary choices
- Binary RUM = Binary RRM → no differences expected
- However, differences may arise due to the RRM logsums directly imputed from underlying mode-destination-time-of-day model.



5. Dutch National Model: Destination/mode/…

Destination/mode/time-of-day model:

- Joint decision about destination, mode and time-of-day
  - 1380 zones, 6 modes of transport, 9 time periods (45 possible depart-return time combinations)
- Nested logit model structure
  - Mode above Time-of-day above Destinations

RRM based alternative:

1. Estimation of RRM-nested logit not an issue, no examples yet
   - Possibly different nesting structures may turn out to be optimal
2. Large number of alternatives: too many binary comparisons in RRM model
   - Sampling of alternatives in RRM + GEV models (Guevara and Ben-Akiva 2013; Guevara et al. 2013)
   - Needs empirical testing, including alternative sampling strategies

5. Dutch National Model:

Weighting the forecasts of the RUM and RRM models:

- Overall model outcome is the combination of different models
- Proposal: Aggregate fit of the individual models and number of parameters to calculate the overall AIC of the model
- Potential issue: fit cannot be used after sampling of alternatives has been applied
- Work in progress!

6. Conclusions

- Different behavioural decision rules can result in different choice probabilities = different travel demand forecasts
- Such uncertainties are not (yet) taken into account
- Model averaging approaches appear suitable to develop behaviourally robust forecasts when a clear `winner' cannot be identified

- RRM is a suitable candidate as an alternative decision rule
- Has proven itself relative to the RUM model
- Intuitive deviations from RUM due to introducing context-effects
- Differences are expected to arise mainly for specific choice situations, but may be averaged out when large number of different scenarios is evaluated
- RRM ready for the `next level': forecasting travel demand
  - Identifying welfare effects remains an issue

6. Conclusions (II)

Replacing RUM models by RRM models in the Dutch National Model:

- Theoretically, each separate model can be replaced
- Limited differences expected for tour-frequency models, only due to inclusion of RRM based log-sum
- Practical issues arise due to large number of alternatives in the destination choice and route-assignment models
  - RRM computational very intensive due to binary comparisons
  - Sampling of alternatives may offer a solution
- Impact on forecasts and model structure needs to be evaluated empirically

7. Questions and discussion

t.dekker@tudelft.nl