

## Spijtminimalisatie als discrete-keuzemodel: Een overzicht van empirische vergelijkingen met het nutsmaximalisatie-model

**Caspar G. Chorus**  
Technische Universiteit Delft<sup>1</sup>

---

### Samenvatting

Dit artikel presenteert een overzicht van empirische studies naar discrete-keuzemodellen die gebaseerd zijn op het idee van spijtminimalisatie. Deze zogenaamde Random Regret Minimization-benadering is een kleine twee jaar geleden gepresenteerd. Dit artikel geeft een overzicht van 21 recente empirische studies waarin de vergelijking tussen RRM en RUM gemaakt wordt. Er wordt langs drie dimensies onderzocht op welke wijze, en in welke mate, de twee modeltypen verschillen: modelfit, externe validiteit, en beleidsrelevante output (zoals elasticiteiten en reistijdwaarderingen). Op hoofdlijnen zijn de resultaten van deze vergelijking van empirische studies als volgt: in termen van modelfit zijn beide modelvormen aan elkaar gewaagd. De verschillen in modelfit tussen de twee modeltypen zijn in het algemeen statistisch significant, maar bijna altijd zeer klein. Met betrekking tot de externe validiteit geldt ook dat het verschil tussen RRM en RUM klein is. Niettegenstaande deze slechts kleine verschillen in modelprestaties, kunnen de verschillen tussen de twee modeltypen in termen van beleidsrelevante output substantieel zijn. De implicaties van deze resultaten voor de mogelijke rol die de RRM-benadering in de (transport-)beleidspraktijk kan spelen, worden uitgebreid besproken.

*Trefwoorden:* Random regret; Random utility; spijtminimalisatie; nutsmaximalisatie; literatuuroverzicht

---

### Inleiding

Sinds de introductie van het Random Utility Maximization-paradigma halverwege de jaren '70 (McFadden, 1974), is de populariteit van op nutsmaximalisatie gebaseerde discrete-

---

<sup>1</sup> Sectie Transport en Logistiek; Faculteit Techniek, Bestuur en Management; Technische Universiteit Delft. Jaffalaan 5, 2628 BX, Delft. T: 015-2788546, F: 015-2782719; E: [c.g.chorus@tudelft.nl](mailto:c.g.chorus@tudelft.nl)

keuzemodellen sterk gegroeid, zowel binnen als buiten de transportwetenschappen (Ben-Akiva & Lerman, 1985; Train, 2003). Deze populariteit is onder meer te danken aan het feit dat deze RUM-modellen op eenduidige en elegante wijze gefundeerd zijn in axioma's uit de neoklassieke welvaartseconomie (Small & Rosen, 1981). Een minstens even belangrijke verklaring voor het succes van deze modellen is te vinden in het feit dat ze relatief makkelijk hanteerbaar zijn: de meeste modelspecificaties hebben 'gesloten vorm'-formuleringen voor keuzekansen en zijn (mede daardoor) makkelijk schatbaar en interpreteerbaar. Dit laatste geldt in het bijzonder voor de meest bekende van alle RUM-modellen, het multinomiale logit (MNL) model met lineair-additieve nutsspecificatie<sup>2</sup>.

Niettegenstaande het succes van deze lineair-additieve, op nutsmaximalisatie gebaseerde, modellen, is er al geruime tijd veel aandacht voor alternatieve gedragsregels en specificaties. Denk hierbij bijvoorbeeld aan de groeiende populariteit van Prospect Theory (Kahneman & Tversky, 1979) en gerelateerde modelvormen (Van de Kaa, 2010), of aan modelvormen gebaseerd op behoeftebevrediging (Arentze & Timmermans, 2009). Andere voorbeelden zijn discrete-keuze modellen die gebaseerd zijn op de door Tversky ontwikkelde theorie van Elimination-by-Aspects (Tversky, 1972; Hess *et al.*, 2012) of modelvormen die veronderstellen dat keuzemakers 'IF-THEN'-beslisregels hanteren (Arentze & Timmermans, 2007).

Het op spijtminimalisatie gebaseerde discrete-keuzemodel (Random Regret Minimization, of vanaf hier: RRM) dat de focus is van dit artikel, valt binnen deze traditie van alternatieve modelvormen. Het RRM-model (Chorus, 2010)<sup>3</sup> veronderstelt dat mensen bij het maken van keuzes niet zozeer uit zijn op het kiezen van het alternatief met het hoogste nut, maar dat zij vooral de situatie willen vermijden dat een niet-gekozen alternatief het beter doet dan het gekozen alternatief op één of meerdere kenmerken (wat spijt zou opleveren). Zoals in een eerdere publicatie in dit tijdschrift is uiteengezet (Chorus & de Jong, 2010), genereert het RRM-model logit-keuzekansen en is het als zodanig schatbaar met behulp van de meeste standaard software pakketten zoals Biogeme (versie 1.3 en 1.8) en NLOGIT (versie 5).<sup>4</sup> De door het RRM-model voorspelde gedragspatronen verschillen in diverse aspecten van de door lineair-additieve RUM-modellen voorspelde patronen. Met name voorspelt het RRM-model dat de mate waarin veranderingen in prestaties van een alternatief op verschillende attributen elkaar kunnen compenseren<sup>5</sup> afhangt van de relatieve positie van het alternatief in de keuzeset. Volgens het

---

<sup>2</sup> Met de term 'lineair-additieve nutsspecificatie' wordt het volgende bedoeld: het systematische nut van een alternatief is een gewogen sommatie van attribuutwaarden (die aangeven hoe een keuzealternatief scoort op de verschillende kenmerken) en te schatten parameters (die het gewicht van de verschillende attributen aangeven). De keuzemaker wordt verondersteld op deze lineair-additieve wijze de multi-dimensionele prestaties van de verschillende alternatieven in kaart te brengen en vervolgens het alternatief met het hoogste nut te kiezen. Wanneer er in het vervolg van dit artikel gesproken wordt van RUM-modellen, wordt deze specifieke modelvorm bedoeld.

<sup>3</sup> Reeds in 2007 werd een eerste versie van het RRM-model gepresenteerd in het proefschrift van de auteur van dit artikel (zie ook Chorus *et al.* (2008) voor een tijdschrift-versie van het desbetreffende hoofdstuk). Die versie van het RRM-model had als nadeel dat er knikken in de spijtfunctie zaten door het gebruik van max-operators. Dat zorgde voor problemen op het vlak van modelschatting en de afleiding van elasticiteiten en marginale effecten. De RRM-versie die gepresenteerd werd in Chorus (2010) heeft deze beperkingen niet, en is mede daardoor in termen van populariteit de eerdere versie voorbijgestreefd. Dit artikel richt zich op deze 2010-versie van het RRM-model.

<sup>4</sup> Schatting met het Alogit-pakket is op dit moment nog niet mogelijk, gegeven dat die software maximaal één logaritmische term in de nutsfunctie toelaat. Hierover is contact geweest met de auteur van dit softwarepakket (Andrew Daly). Deze ziet mogelijkheden om RRM-schatting mogelijk te maken in Alogit, maar het is nog onduidelijk op welke termijn dit concreet kan resulteren in compatibiliteit tussen RRM en Alogit.

<sup>5</sup> Denk bijvoorbeeld een daling in reistijd die een stijging in reiskosten kan compenseren.

RRM-model is een verslechtering van een attribuut waarop het alternatief relatief *goed* scoort makkelijk te compenseren door een verbetering van een ander attribuut. Daarentegen voorspelt RRM dat een verslechtering van een attribuut waarop het alternatief al relatief *slecht* scoort veel minder makkelijk te compenseren is door een verbetering van een ander attribuut. Deze door het RRM-model voorspelde specifieke vorm van semi-compensatorisch gedrag resulteert in een aantal systematische verschillen tussen RRM en RUM in termen van voorspelde keuzekansen<sup>6</sup>.

Dit artikel zet de op dit moment beschikbare empirische vergelijkingen tussen RRM en RUM op een rij, en probeert zodoende een beeld te schetsen van de verschillen tussen de twee genoemde modelvormen in termen van (i) modelfit (sectie 2), (ii) voorspelkracht op 'hold-out samples' (sectie 3), en (iii) verschillen in beleidsimplicaties (zoals in termen van parameterschattingen, betalingsbereidheid, elasticiteiten, Logsums, en de effecten van beleid op marktaandeelen van alternatieven (sectie 4)).<sup>7</sup> Een bredere discussie van de gerapporteerde verschillen tussen RRM en RUM wordt gepresenteerd in sectie 5, waarbij uitgebreid stil wordt gestaan bij de implicaties van de empirische verschillen tussen RRM en RUM voor de mogelijke rol die het RRM-model in de beleidspraktijk kan spelen.

Gegeven dat het RRM-model en zijn eigenschappen eerder al uitgebreid zijn besproken in dit tijdschrift en elders (zie de referenties hierboven), is er voor gekozen om het model slechts zeer summier en in een appendix te bespreken. Een recent gepubliceerd (Engelstalig) dictaat (Chorus, 2012b) geeft een uitgebreide bespreking van het RRM-model.

Tot slot van deze inleiding nog twee kanttekeningen, waarbij ik uit het oogpunt van leesbaarheid de ik-vorm hanteer: ten eerste zal het de lezer opvallen dat ik bij het merendeel van de besproken studies (co-)auteur ben. Gegeven de oorsprong van het RRM-model is dit misschien niet vreemd, maar men kan zich afvragen (en één van de referenten deed dit ook) of een dergelijk overzichtsartikel niet nuttiger zou zijn wanneer er een groter aantal studies is verschenen van de hand van andere auteurs. Mijn persoonlijke mening hierover is als volgt: als lezer vind ik het altijd prettig om in een relatief vroegtijdig stadium een overzicht te krijgen van een nieuwe methode, model of theorie. Vanuit die gedachte ben ik dit artikel gaan schrijven, slechts twee jaar na introductie van het RRM-model en op basis van deels (grotendeels?) eigen werk. Hierbij speelde mee dat er – zoals in tabel 1 valt te zien – zeker niet sprake is van een eenzijdige literatuurlijst: naast mijn eigen naam is er sprake van 21 andere (co-)auteurs. Dit geeft aan dat de besproken artikelen zeker niet puur als eigen werk zijn te bestempelen. Een tweede kanttekening die kan worden gemaakt (en ook deze is door een referent gemaakt) is dat het vreemd kan overkomen wanneer een 'uitvinder' van een methode of theorie zijn of haar eigen werk bespreekt in een overzichtsartikel ('de slager die zijn eigen vlees keurt'). Ook hier geldt dat ik dit als lezer van dergelijke artikelen persoonlijk nooit storend heb gevonden, en dat ik met die gedachte in het achterhoofd dit artikel heb geschreven. Wel heb ik, mede met hulp van de desbetreffende referent, zeer mijn best gedaan om zo objectief en voorzichtig mogelijk de verschillen tussen de twee modelbenaderingen (RRM en RUM) te bespreken, zeker wanneer het waardeoordeel betreft. Of ik daar in geslaagd ben of niet, is het best door de lezer te bepalen.

---

<sup>6</sup> Recente overzichten van de eigenschappen van het RRM-model zijn te vinden in Chorus (2012a, b).

<sup>7</sup> Een voorzet tot een dergelijke inventarisatie is gedaan in sectie 4 (twee pagina's lang) van Chorus (2012a). Die studie bevatte echter maar de helft van de hier besproken studies, en was veel beperkter in scope en mate van detail.

## 1. De 'data': artikelen met empirische vergelijkingen tussen het RRM- en het RUM-model

Vanuit het oogpunt van transparantie (in de zin dat de lezer de hier gepresenteerde observaties en resultaten moet kunnen verifiëren) richt dit artikel zich op empirische vergelijkingen tussen RRM en RUM die ofwel verschenen zijn in reeds gepubliceerde (of in druk zijnde) artikelen, ofwel worden beschreven in artikelen die momenteel beoordeeld worden voor publicatie (en beschikbaar zijn bij de auteur van dit artikel). De 21 bij de auteur van dit artikel bekende empirische studies die voldoen aan deze voorwaarden, zijn gerapporteerd in tabel 1. De verschillende kolommen betreffen: het nummer van de studie, waarnaar gerefereerd wordt in de rest van dit artikel; de namen van de auteurs; de status van het artikel (jaartal indien gepubliceerd, anders 'in druk' of 'in review'); de verschillende aspecten van de vergelijking met RUM die in het artikel aan bod komen en relevant zijn voor dit overzicht; het type keuze; en het type dataverzameling (Stated Preference versus Revealed Preference<sup>8</sup>). Elke rij betreft een andere dataset, al zijn de keuzetypes soms gelijk aan die in andere rijen.<sup>9</sup> Een 'x' in een cel geeft aan dat de desbetreffende studie (rij) aandacht besteedt aan de desbetreffende vergelijking tussen RUM en RRM (kolom).

Tabel 1: de voor dit overzicht gebruikte artikelen

ID	Auteurs	Status	Model-fit	Voorspelkracht	Beleidsrelevante output	Type keuze	Type data
1	Chorus	2010	x	x		route-/vervoerswijzekeuzes	SP
2	Chorus	2010	x	x		reisinformatie-acquisitie	SP
3	Chorus	2010	x	x		parkeerkeuzes	RP
4	Chorus	2010	x	x		bestemmingskeuzes	RP
5	Chorus de Jong	2011	x		x	vertrektijdstipkeuzes	SP
6	Chorus Annema Mouter van Wee	2011	x		x	beleidskeuzes politici	SP

<sup>8</sup> Zoals een referent aangaf, is het opvallend en mogelijk wat verontrustend dat het merendeel van de gepresenteerde studies (16 van de 21) gebaseerd is op SP-data (in plaats van op RP-data). Zoals ook aangegeven in Chorus (2012c) verdient het aanbeveling om bij het vergelijken van verschillende modeltypen te proberen om zoveel mogelijk gebruik te maken van RP-data, aangezien dit voor een hogere externe validiteit van empirische vergelijkingen kan zorgen.

<sup>9</sup> Empirische studies die de 2008-variant (Chorus *et al.*, 2008) van het RRM-model schatten, zoals Hess *et al.* (2012) worden in dit artikel buiten beschouwing gelaten. Met uitzondering van studies 5, 17 en 18 zijn de RRM- en RUM-modellen in MNL-vorm geschat. De genoemde uitzonderingen betreffen zogenaamde error-component-Mixed Logit modellen.

7	Thiene Boeri Chorus	2012	x		x	keuzes voor recreatieve activiteiten in natuurpark	SP
8	Boeri Longo Doherty Hynes	2012	x		x	recreatieve bestemmings-keuzes kajakkers	RP
9	Hensher Greene Chorus	in druk	x		x	autotypekeuzes	SP
10	Chorus Rose Hensher	in druk	x	x	x	routekeuzes	SP
11	Chorus Rose	2012	x	x		keuzes voor datingprofielen	SP
12	Chorus	2012	x			routekeuzes forenzen	SP
13	Kaplan Prato	in druk	x	x	x	uitwijkacties in de auto	RP
14	Chorus Arentze Hess	in review	x	x	x	keuzes voor recreatieve activiteiten van senioren	SP
15	Chorus Bierlaire	in review	x	x	x	routekeuzes forenzen	SP
16	Beck Chorus Rose Hensher	in druk	x			autotypekeuzes van huishoudens	SP
17	de Bekker- Grob Chorus	in review	x		x	vaccinatiekeuzes baarmoederhalskanker	SP
18	de Bekker- Grob Chorus	in review	x		x	keuzes voor medische behandeling osteoporose	SP
19	Chorus Koetse Hoen	in review	x	x	x	autotypekeuzes leaserijders	SP
20	Prato	in review	x			routekeuzes	RP
21	Boeri & Masiero	in review	x		x	vervoerswijzekeuzes vrachtvervoer	SP

## 2. RRM versus RUM: verschillen in modelfit

De directe vergelijking van de eind-log-likelihood van een RUM- en een RRM-model is zinnig in die zin dat beide modellen evenveel parameters consumeren, en er dus geen correctie gemaakt

hoeft te worden voor het aantal opgegeven vrijheidsgraden. Maar, gegeven dat geen van beide modellen gezien kan worden als een speciaal geval van het andere model is een test voor niet-geneste modellen nodig om te onderzoeken of verschillen in modelfit statistisch gezien significant zijn. De Ben-Akiva & Swait test (1986) is zo'n test, en deze wordt in de verschillende publicaties uit tabel 1 (en in dit overzichtsartikel) gebruikt. Deze test genereert een bovengrens voor de kans dat, gegeven dat de eind-log-likelihood van model A hoger is (dichter bij nul is) dan die van model B, model B toch de beste weergave is van het data-genererend proces. Als zodanig is deze bovengrens te beschouwen als een conservatieve schatting van de p-waarde die de significantie van het verschil in modelfit aangeeft. Alle 21 in tabel 1 opgenomen studies rapporteren een vergelijking in modelfit. De meeste studies rapporteren zelf deze statistische vergelijking, en waar dat niet het geval is (bijvoorbeeld studies 1 t/m 4) is deze test *ex post* uit te voeren op basis van de gerapporteerde schattingsresultaten.

De conclusies zijn als volgt: voor 12 van de 21 studies fit het RRM-model de data significant beter dan het RUM-model. In vier van die 12 studies (studies 3, 5, 6 en 19) is het verschil slechts significant op 10%-niveau; in twee gevallen (studies 1 en 16) op 5%-niveau en in de resterende gevallen (studies 4, 9, 11, 13, 15 en 21) is er sprake van significantie bij een 1%-niveau. Voor twee studies (studies 18 en 20) geldt dat het verschil in modelfit wel in het voordeel van RRM uitvalt, maar niet significant is. Voor zeven van de 21 studies (studies 2, 7, 8, 10, 12, 14, en 17) geldt dat het RUM-model beter fit dan het RRM-model, en in dat geval zijn de verschillen telkens significant bij 1%-niveau. Belangrijker wellicht dan de statistische significantie van de verschillen in modelfit, is de vraag hoe groot (en dus: relevant) die verschillen zijn. Het blijkt dat in alle gevallen de verschillen klein zijn, en meestal pas tot uiting komen in het tweede of derde cijfer achter de komma van McFadden's rho-square (die de stijging in modelfit aangeeft ten opzichte van het nulmodel waarin alle parameters de waarde nul aannemen).

Het ligt voor de hand om verwachtingen te testen aangaande mogelijke determinanten van de verschillen in modelfit tussen RRM en RUM: zijn er sociaal-demografische segmenten die eerder geneigd zijn volgens RRM-of RUM-premissen te kiezen, of zijn er contextuele factoren die zorgen voor een structureel betere fit van het ene of het andere modeltype? Voor zover het sociaal-demografische determinanten betreft, bieden de studies uit tabel 1 (en aanvullende niet gepubliceerde analyses uitgevoerd door de auteur van dit artikel) weinig aanknopingspunten: er zijn meestal wel verschillen in modelfit te vinden tussen verschillende segmenten, maar deze lijken in eerste instantie niet structureel te zijn (generiek over datasets). Bovendien zijn ook deze verschillen in modelfit eerder *significant* dan *relevant* te noemen, dus het lijkt niet direct aanbevelenswaardig om uitgebreider te verkennen welke mogelijk structurele verschillen tussen sociaal-demografische segmenten er opduiken in termen van de relatieve modelfit van RRM en RUM.

Voor wat betreft de rol van contextuele factoren is er meer duidelijkheid – ook al geldt hier dat het relatief kleine aantal empirische studies dat tot nog toe beschikbaar is, reden is voor voorzichtigheid met het trekken van algemene conclusies. De sociaalpsychologische literatuur biedt aanknopingspunten voor een analyse van de rol van contextuele factoren als determinant van de verschillen in modelfit tussen RRM en RUM: in een reeks aan empirische studies (zie Zeelenberg & Pieters (2007) voor een overzicht) is aangetoond dat de minimalisatie van geanticipeerde spijt met name een rol speelt wanneer keuzemakers de desbetreffende keuze moeilijk en/of belangrijk vinden, en/of wanneer ze inschatten dat ze de gemaakte keuze mogelijk tegenover anderen moeten (kunnen) verantwoorden. Met dit gegeven in het achterhoofd, is het niet verrassend dat studies naar autotypekeuzes (9 en 16), uitwijkmanoeuvres (13), beleidskeuzes van politici (studie 6), keuzes op datingwebsites (studie 11) en keuzes van vrachtvervoerders (21) een betere fit voor RRM laten zien, terwijl keuzes met betrekking tot, bijvoorbeeld, vrijetijdsbesteding (studies 7, 8 en 14) een betere fit voor RUM tonen. Nogmaals: de verschillen in modelfit zijn klein en daardoor niet noodzakelijkerwijs relevant. Tot slot: niet

gepubliceerde analyses op de datasets van studies 11 en 19 laten zien dat het RRM-model met name goed fit op dat deel van de steekproef dat relatief lang deed over het maken van keuzes in Stated Choice-experimenten. Ook dit resultaat is in lijn met de hierboven genoemde bevindingen uit de sociaalpsychologische literatuur.

Zoals een referent aangaf, zou het interessant zijn om – wanneer er meer studies beschikbaar zijn gekomen (minimaal dertig, maar liefst meer) – een statistische meta-analyse te doen waarin de effecten van socio-demografische segmenten, keuzecontexten en type data op de relatieve modelfit van RUM en RRM op een meer systematische wijze kunnen worden gemeten.

Tot slot van deze sectie een wat meer anekdotische bespreking van het in de context van dit artikel relevante begrip ‘publication bias’: gegeven dat de RRM-benadering nieuw is (zeker in vergelijking met het in het vakgebied zeer bekende RUM-paradigma), ligt het voor de hand dat wetenschappelijke tijdschriften geneigd zijn om minder positief te staan tegenover ingediende artikelen die empirische resultaten laten zien die in het nadeel van het RRM-model zijn en in het voordeel van het RUM-model. De vraag die zou kunnen opkomen bij redactie en referenten is: “waarom ruimte vrijmaken voor een model dat een alternatief wil bieden voor RUM, maar slechter fit bij de data”.<sup>10</sup> Dit fenomeen zou de resultaten over modelfit zoals besproken in deze sectie potentieel vertekenen. Het is echter de ervaring van de auteur dat referenten en redacteuren van wetenschappelijke tijdschriften in het algemeen niet reageren zoals hierboven: het kwam bij de auteur van dit artikel slechts eenmaal voor dat een referent vroeg om het verwijderen uit het manuscript van die casestudies waar RRM het minder goed deed dan RUM. Deze verwijdering van casestudies heeft toen overigens niet plaatsgevonden, en het desbetreffende artikel is uiteindelijk elders ingediend en geaccepteerd. Ook het feit dat verschillende gepubliceerde (of voor publicatie geaccepteerde) artikelen RRM en RUM vergelijken op één dataset, waarop het RUM-model beter fit (studies 7, 8 en 12), is een teken dat er geen sprake lijkt te zijn van een al te sterke publicatiebias. Toch kan publicatiebias niet uitgesloten worden, en lijkt het als zodanig verstandig om te concluderen dat de twee modeltypen gemiddeld genomen een even goede modelfit lijken te genereren, niettegenstaande de kleine ‘voorsprong’ voor RRM betreffende de 21 in dit artikel besproken artikelen. De belangrijkste conclusie blijft overigens toch dat de verschillen in modelfit tussen RRM en RUM – hoewel meestal statistisch significant – zeer klein zijn.

### **3. RRM versus RUM: verschillen in externe validiteit**

Van de 21 studies rapporteren er negen één of meerdere maten voor de voorspelkracht van de geschatte RUM- en RRM-modellen op zogenaamde hold-out data die niet gebruikt zijn voor modelschatting<sup>11</sup>. Acht studies rapporteren een maat die gebaseerd is op de likelihood van de hold out data gegeven het geschatte model. Het betreft hier de zogenaamde mean-likelihood (studies 1, 2, 3, 4, 10, 14, 15, 19), die per waarneming (geobserveerde keuze) uit de hold-out data de waarschijnlijkheid van die observatie gegeven het geschatte model berekent, en daar

---

<sup>10</sup> Hier anticipeert de desbetreffende auteur mogelijkerwijs weer op, door minder geneigd zijn om resultaten in te dienen bij tijdschriften wanneer die ‘negatief’ zijn over het RRM-model.

<sup>11</sup> Deze voorspelkracht op hold out-data wordt door velen beschouwd als maat voor de ‘externe validiteit’ van het geschatte model, hoewel er strikt genomen geen sprake is van een werkelijk ‘externe’ validatie met behulp van een wezenlijk andere dataset.

vervolgens het gemiddelde van neemt<sup>12</sup>. Drie studies rapporteren de zogenaamde hitrate (13, 14 en 19). Deze geeft op een andere wijze aan hoe goed het geschatte model is in het identificeren van de gekozen alternatieven in de hold-out data: indien het gekozen alternatief ook het alternatief is dat door het geschatte model als meest aantrekkelijk is geïdentificeerd, wordt dit als een 'hit' beschouwd. De hitrate wordt vervolgens verkregen door het aantal hits te delen door het aantal observaties in de hold out data. Voor vijf andere studies is de hitrate wel berekend, maar niet gerapporteerd (studies 1, 2, 3, 4, 15). Hieronder wordt een overzicht gegeven van de verschillen tussen RRM en RUM in termen van alle gepubliceerde en niet-gepubliceerde resultaten.

Voor wat betreft de likelihood-gebaseerde maten, zijn de prestaties voor de genoemde studies van het RRM-model iets beter dan die van het RUM-model: in vijf gevallen (studies 1, 2, 4, 14, 19) zijn de data van de hold-out set waarschijnlijker gegeven het RRM-model dan gegeven het RUM-model. In twee gevallen (studies 3 en 15) presteren beide modellen even goed, en in één geval genereert RUM een hogere gemiddelde 'likelihood' voor de hold-out data (studie 14). Ook in termen van de hitrate doet het RRM-model het iets beter dan het RUM-model: in vier van de acht gevallen (studies 4, 13, 14, 15) genereert het RRM-model een hogere hitrate; in twee gevallen is de hitrate van het geschatte RUM-model hoger (studies 1 en 19), en in de overgebleven gevallen scoren beide modellen even goed (studies 2 en 3). Voor wat betreft de likelihood-gebaseerde maten zijn de verschillen tussen RUM en RRM zonder uitzondering zeer klein. De verschillen in hitrates zijn meer substantieel, maar nooit groter dan enkele procentpunten. Als zodanig zijn de resultaten op het vlak van externe validiteit vergelijkbaar met die op het vlak van modelfit, met uitzondering wellicht van de hitrate-resultaten.

Het blijkt dat de verschillen tussen RUM en RRM in externe validiteit niet gelijkaardig zijn tussen de verschillende maten: wanneer zowel de gemiddelde likelihood als een hitrate wordt berekend (en dit is het geval voor zeven studies) zijn er slechts twee gevallen (studies 4 en 14) waar de twee maten hetzelfde model als beste aanwijzen; in beide gevallen is dat het RRM-model. Ook blijkt dat de verschillen in modelfit zoals gerapporteerd in de vorige sectie geen goede voorspeller zijn voor de in deze sectie besproken maten voor externe validiteit. Er is maar één studie (nummer 4) waar zowel de vergelijking in modelfit als in gemiddelde likelihood als in hitrate hetzelfde model (in dat geval: RRM) als beste model aanwijzen.

Tot slot van deze sectie kan nogmaals gewezen worden op potentiële publicatie-bias (zoals in de vorige sectie uitgebreid aan de orde gekomen): het is mogelijk dat studies die gunstige resultaten voor het 'nieuwe' model (in dit geval RRM) rapporteren met betrekking tot de externe validiteit een grotere kans hebben te worden ingediend bij en/of geaccepteerd door wetenschappelijke tijdschriften. In dat licht lijkt het reëel om te concluderen dat de beide modeltypen het gemiddeld genomen even goed doen in termen van externe validiteit, ook al wijzen de in tabel 1 genoemde studies in het algemeen op wat betere prestaties van het RRM-model. Hoe dan ook: de verschillen tussen het RUM- en RRM-model, met name wanneer gemeten in termen van gemiddelde likelihood van de hold-out sample, zijn klein.

#### **4. RRM versus RUM: verschillen in beleidsrelevante output**

Al zijn veel keuze-modelleers met name geïnteresseerd in de prestaties van modellen in termen van modelfit en externe validiteit, het is evident dat deze keuzemodellen uiteindelijk een 'hoger'

---

<sup>12</sup> Twee studies (14 en 15) rapporteren naast deze gemiddelde likelihood ook een gerelateerde maat, de rho-bar squared. Voor studie 14 valt deze maat uit in het voordeel van het RUM-model, en voor studie 15 is de rho-bar squared gelijk voor de twee modellen.



doel moeten dienen: de kwantitatieve ondersteuning van (transport-)beleid. In dat kader is het belangrijk om de verschillen tussen RUM en RRM ook te verkennen in termen van de door de twee modeltypen gegenereerde beleidsrelevante output (zoals inschattingen van elasticiteiten of marktaandeelen). Bij deze vergelijking gaat het dus niet om te bepalen welke van de twee modeltypen het 'beter' doet, maar eerder om te bepalen op welke wijze en in welke mate de twee modellen verschillen in de gegenereerde beleidsrelevante output. Deze sectie geeft zo een overzicht van de verschillen tussen RRM en RUM in termen van beleidsrelevante output. Van de 21 gerapporteerde studies zijn er 14 die op enigerlei wijze dergelijke output presenteren. Deze kan bijvoorbeeld de vorm aannemen van reistijdwaarderingen of gerelateerde maten (studies 6, 10, 14, 17), elasticiteiten (studies 7, 9, 13, 15, 21), welvaartsbaten van beleidsmaatregelen (studies 5, 8), of de voorspelde effecten van beleid op marktaandeelen/ keuzekansen van producten of diensten (studies 7, 11, 15, 17, 19, 21). Gegeven de relatief grote variëteit aan typen beleidsrelevante output die aan bod komen in de in tabel 1 gepresenteerde studies, worden deze studies stuk voor stuk doorgenomen (in plaats van de verschillende typen beleidsrelevante output thematisch te bespreken).

Studie 5 onderzoekt de effecten van een fictief spits-spreiden ('peak spreading') experiment, waarbij specifiek gekeken wordt naar de bereikbaarheidseffecten<sup>13</sup> van beleid dat de reistijd in de spits verlaagt ten koste van de reistijd aan de randen van en buiten de spits. Geschatte vertrektijdstipkeuzemodellen (RRM en RUM) worden gebruikt om de effecten van deze maatregel te berekenen in termen van daling of stijging van de verwachte spijt / het verwachte nut voor een representatieve reiziger. Analyses tonen aan dat de verschillen in bereikbaarheidseffecten kunnen oplopen tot 45%, afhankelijk van de gekozen variant van het spits-spreiden beleid. Voor de meeste varianten zijn de verschillen minder groot, met name wanneer de effecten van de beleidsmaatregel groter zijn (in termen van een grotere reductie in reistijden tijdens de spits). In die gevallen zijn de verschillen maar enkele procenten groot.

Studie 6 onderzoekt de voorkeuren van politici op lokaal (gemeentelijk) niveau voor verschillende varianten van prijsbeleid op de weg binnen de bebouwde kom. Op basis van de schattingsresultaten van RRM- en RUM-modellen wordt de populariteit van verschillende varianten voorspeld. Meer specifiek wordt er, bij wijze van voorbeeld, gekeken naar de stijging in emissiereductie die nodig is om een stijging in operationele kosten van 10% te compenseren (in termen van de keuzekans voor die variant). Het RUM model voorspelt dat hiervoor een stijging in emissiereductie van 6.6% nodig is. Het RRM model voorspelt dat de benodigde compensatie (in termen van stijging van emissiereductie) afhangt van de relatieve prestaties van de desbetreffende variant in termen van emissiereductie en operationele kosten. Scoort een variant relatief goed in termen van operationele kosten, dan is de verslechtering in dit attribuut goed te maken door een 5.5% stijging in emissiereductie. Maar: scoort een variant al relatief slecht in operationele kosten, dan wordt een verdere verslechtering pas gecompenseerd door een stijging in emissiereductie van 6%. Deze verschillen (binnen RRM en tussen RRM en RUM) zijn beperkt, wat overigens mede verklaard kan worden door de relatief slechte modelfit van beide modellen (er is relatief veel 'random' keuzegedrag).

Studie 7 richt zich op de wandelroutekeuzes van bezoekers van een natuurgebied in de Dolomieten. Elasticiteiten worden berekend (voor RRM- en RUM-modellen) van attributen zoals de aanwezigheid van bewegwijzering, de drukte op de route, en eventuele toegangskosten. Voor de meeste attributen zijn de RRM-elasticiteiten groter dan de RUM-tegenhangers - in de meeste gevallen zijn de verschillen zo'n 10% groot. Ook wordt onderzocht in welke mate het

---

<sup>13</sup> In studie 5 wordt bereikbaarheid geconceptualiseerd als het verwachte nut (spijt) van een keuzesituatie. Deels wordt gebruik gemaakt van conventionele Logsums, en deels van gerelateerde maten die volatiliteit in preferenties proberen te modelleren.

marktaandeel van een route wordt beïnvloed door de heffing van toegangskosten. Voor de gekozen heffings-variant is het door het RUM model voorspelde effect (3.10 procentpunt daling in marktaandeel) groter dan het door het RRM model voorspelde effect (2.06 procentpunt daling in marktaandeel). Hier zijn de verschillen, zeker in absolute zin, klein.

Studie 8 onderzoekt de bestemmingskeuze van kajakkers in natuurgebieden, en richt zich op de voorspelling van de welvaartseffecten van (i) de heffing van parkeerkosten en (ii) een sterk toenemende drukte op de rivier. Zogenaamde Logsums gebaseerd op RUM modellen (zie Ben-Akiva & Lerman, 1985) en op RRM modellen (zoals afgeleid in Chorus, 2012d) worden berekend om deze welvaartseffecten te kwantificeren. Verschillen tussen de gemiddelde effecten zoals voorspeld met behulp van de RUM-Logsum en de RRM-Logsum zijn, zeker in relatieve zin, groot en lopen op tot bijna 300%.

De focus van studie 9 is op preferenties van consumenten voor auto's die rijden op alternatieve brandstoffen. De auteurs berekenen elasticiteiten op basis van geschatte RUM- en RRM-modellen, en komen tot de conclusie dat de verschillen tussen de twee modeltypen substantieel zijn - ze lopen op tot 19%. Op RRM gebaseerde prijselasticiteiten blijken groter dan hun op RUM gebaseerde tegenhangers. De auteurs concluderen dat "the mean differences are such that the RUM model is not a good approximation to the RRM model if random regret is a preferred representation of behavioral response, as is the case in this empirical study" (Hensher *et al.*, in druk).

Studie 10 onderzoekt routekeuzes van Australische forenzen. RRM-reistijdwaarderingen worden gegenereerd door de partiële afgeleide van de spijtfunctie naar de reistijd te delen door de partiële afgeleide van de spijtfunctie naar de reiskosten. In lijn met de korte uitleg zoals gegeven in de introductie van dit artikel zijn de op deze wijze berekende RRM-reistijdwaarderingen afhankelijk van de relatieve prestatie van het alternatief in termen van reistijd en -kosten. Meer specifiek: in de context van een route die snel maar duur is, is de RRM reistijdwaardering (dat is: de betalingsbereidheid voor een reistijdverkortung) relatief laag, en in de context van een route die langzaam maar goedkoop is, is de RRM reistijdwaardering relatief hoog. De verschillen tussen de twee modeltypen zijn substantieel te noemen: het RUM model voorspelt een reistijdwaardering van 0.2 Australische dollars per minuut, terwijl de RRM reistijdwaardering varieert tussen 0.03 Australische dollars per minuut en 1.34 Australische dollars per minuut (gemiddelde waarde is 0.23 Australische dollars per minuut). NB: ten tijde van de desbetreffende dataverzameling (anno 2007) was de wisselkoers van de Australische dollar ten opzichte van de Euro gelijk aan 0.61 Euro / AUD.

Studie 11 richt zich op het verklaren en voorspellen van de keuzes tussen datingprofielen, gemaakt door bezoekers aan datingwebsites. Verschillen in voorspelde keuzekansen (tussen RRM en RUM) zijn relatief klein, en lopen op tot maximaal 4 procentpunt. Ook in deze studie zijn de verschillen herleidbaar tot (en: interpreteerbaar in termen van) de conceptuele verschillen tussen de twee modeltypen. Meer specifiek voorspelt het RRM-model dat zogenaamde compromis-alternatieven - dit zijn opties die niet extreem goed of slecht scoren op verschillende kenmerken, maar juist op alle kenmerken een gemiddelde waarde aannemen ten opzichte van de andere opties uit de keuzeset - een relatief hoog marktaandeel veroveren. In de context van de datingwebsite data komt dit compromis effect in beperkte mate tot uiting in de zin dat een kandidaat die gemiddeld scoort op belangrijke kenmerken een marktaandeel van 45% krijgt toegewezen door het RRM model en een marktaandeel van 43% door het RUM model.

Studie 13 onderzoekt uitwijkmanoeuvres van automobilisten, vlak voor een (bijna-)ongeluk. Gebruik wordt gemaakt van een dataset met daarin werkelijk geobserveerde uitwijkmanoeuvres op Amerikaanse snelwegen. Elasticiteiten worden berekend voor geschatte RRM- en RUM-modellen. Aangezien er sprake is van indicator-variabelen (zoals: "gebruik van veiligheidsgordel" (0/1)) kan er geen gebruik worden gemaakt van de conventionele

elasticiteiten. In plaats daarvan wordt gebruik gemaakt van pseudo-elasticiteiten die het effect op de kans op een bepaald type uitwijkmanoeuvre geeft indien een indicatorvariabele van '0' naar '1' verandert. Voor de overgrote meerderheid van variabelen is de RRM pseudo-elasticiteit 5% tot 10% groter dan de RUM pseudo-elasticiteit. Deze bevinding is in lijn met de resultaten zoals gerapporteerd in studies 7 en 9 (zie hierboven).

In studie 14 wordt onderzocht hoe senioren bepaalde factoren afwegen bij hun keuze tussen verschillende recreatieve activiteiten (zoals concertbezoek of museumbezoek). Op basis van geschatte RRM- en RUM-modellen worden eerst parameter ratio's berekend. Zowel in een RUM-context als in een RRM context geven deze ratio's een indicatie van het relatieve belang van attributen (niet gecorrigeerd voor schaalverschillen). Verschillen van bijna 10% worden gevonden tussen de ratio's van RUM-parameters en die van RRM-parameters. Daarnaast worden de twee geschatte modellen ingezet om voor elk alternatief in de hold-out dataset (beslaande 1/3 deel van de totale dataset) de keuzekans te berekenen. Het blijkt dat de op deze wijze voorspelde keuzekansen aanzienlijk - tot meer dan 10 procentpunt - kunnen verschillen tussen de twee modeltypen (hoewel voor de meeste alternatieven de verschillen kleiner zijn dan 5 procentpunten).

Studie 15 maakt gebruik van een stated choice dataset om de routekeuzes van Nederlandse forenzen te onderzoeken. Elasticiteiten voor geschatte RUM- en RRM-modellen verschillen nauwelijks van elkaar, met uitzondering van de reistijdelasticiteit: deze is bijna 10% groter voor het RRM model dan voor het RUM model (zie enigszins vergelijkbare resultaten gerapporteerd in studies 7, 9 en 13). De geschatte modellen zijn ook gebruikt om marktaandelen te voorspellen voor alternatieve routes. Specifiek werd er gekeken naar de mogelijke aanwezigheid van een compromis-effect. In lijn met theoretische verwachtingen (Chorus, 2010, 2012a) en eerdere empirische bevindingen (studie 11) blijkt inderdaad dat het RRM model (27%) een substantieel hogere keuzekans dan het RUM model (23%) toekent aan een route die in termen van de hierboven genoemde kenmerken gemiddeld presteert ten opzichte van concurrerende routes.

Studies 17 en 18 betreffen keuzes van patiënten tussen verschillende medische behandelingen. Geschatte RUM- en RRM-modellen worden in eerste instantie gebruikt om parameter-ratio's te berekenen (om zo het relatieve belang van parameters te achterhalen). Verschillen tussen modellen, in termen van het relatieve belang van parameters, worden gerapporteerd van maximaal 33% (studie 17), respectievelijk 16% (studie 18). Wanneer de geschatte modellen worden gebruikt om keuzekansen te voorspellen voor alternatieve behandelingen, worden verschillen van slechts enkele procentpunten gevonden; ook hier blijkt het RRM model een hogere keuzekans te voorspellen voor zogenaamde compromis-alternatieven. Het verschil voor deze compromis-alternatieven is ruim 2 procentpunt voor baarmoederhalskanker-vaccinatiekeuzes, en ruim 1.5 procentpunt voor osteoporose-behandelingen.

In studie 19 worden de preferenties van leaserijders geanalyseerd met betrekking tot 'alternatieve brandstof'-auto's. Geschatte modellen worden gebruikt om marktaandelen te voorspellen van verschillende alternatieven in het hold-out sample (op precies dezelfde wijze als in studie 14). Verschillen tussen RUM en RRM in termen van voorspelde marktaandelen zijn vaak groot: in 26% van de gevallen is het verschil tussen het door RRM en het door RUM voorspelde marktaandeel groter dan 5 procentpunt, en in 13% van de gevallen is er zelfs sprake van een verschil van meer dan 10 procentpunt. Bovendien wijzen de twee modellen in 7% van de gevallen een ander alternatief aan als het meest populaire van een bepaalde keuzeset. Het zijn, net als in eerder genoemde studies, met name compromis-alternatieven waarvoor geldt dat ze door het RRM model als winnaar worden aangeduid en door het RUM model niet.

Tot slot worden in studie 21 de vervoerswijzekeuzes van vrachtvervoerders geanalyseerd met behulp van een Stated Choice-experiment waarin kenmerken als reistijd, reiskosten en punctualiteit zijn opgenomen. Elasticiteiten worden berekend voor geschatte RUM- en RRM-

modellen. Hieruit blijkt dat voor het merendeel van de attributen de RRM-elasticiteiten groter zijn, wat in lijn is met eerdere bevindingen. Ook wordt het effect op marktaandeelen van modaliteiten voorspeld van een stijging van de reistijd met 15%. Het blijkt dat – en dit is enigszins onverwacht gegeven de richting van het verschil in de grootte van elasticiteiten – de door het RUM model voorspelde impact van deze toename in reistijd groter is dan de door het RRM model voorspelde toename.

Samenvattend kan het volgende gezegd worden over de verschillen tussen RRM en RUM in termen van beleidsrelevante output zoals elasticiteiten, reistijdwaarderingen, bereikbaarheid/welvaart, en voorspelde marktaandeelen: verschillen tussen de twee modelvormen kunnen substantieel zijn. In andere gevallen is er sprake van kleinere verschillen, en soms zijn de verschillen tussen de modellen verwaarloosbaar vanuit praktisch en/of beleidsmatig oogpunt. Twee min of meer stabiele trends lijken waarneembaar, al moet ook hierbij worden opgemerkt dat het aantal bekeken studies (veel) te klein is om harde conclusies op te baseren: het RRM model lijkt, in termen van elasticiteiten en voorspelde marktaandeelen, wat gevoeliger te zijn dan het RUM model. Meer specifiek: elasticiteiten zijn meestal groter voor het RRM model dan voor het RUM model, en de voorspelde marktaandeelen voor zeer slecht (zeer goed) presterende alternatieven zijn kleiner (groter) voor het RRM model dan voor het RUM model<sup>14</sup>.

Feitelijk ontstaat hierdoor een wat paradoxale situatie: de verschillen tussen de twee modelvormen in termen van modelfit en externe validiteit zijn vaak zeer klein (wat het moeilijk maakt één van de twee modellen tot het meest geschikte te verklaren in de context van een bepaalde dataset), maar toch kunnen ze tot wezenlijk andere beleidsimplicaties leiden. In de volgende sectie wordt dieper ingegaan op de vraag welke implicaties deze wat paradoxale situatie heeft, voor de mogelijke rol die het RRM-model kan spelen in de (transport-)beleidspraktijk.

## 5. Discussie

Het is op dit moment nog niet helder in welke situaties, in welke mate en op welke manier het recent geïntroduceerde Random Regret Minimization (RRM) model het meest effectief kan worden ingezet ter ondersteuning van de ontwikkeling van transportbeleid, in plaats van of in combinatie met het klassieke Random Utility Maximization (RUM) model. De in dit artikel gepresenteerde resultaten op het gebied van modelfit en externe validiteit maken de keuze tussen de twee modellen ook niet makkelijker: in termen van modelfit is in veel gevallen ofwel het RRM ofwel het RUM model strikt (statistisch) genomen te prefereren boven het andere model, maar de verschillen zijn bijna altijd klein. Bovendien komt het regelmatig voor dat een model dat een betere modelfit behaalt dan de alternatieve modelvorm op een bepaalde dataset, weer minder goed scoort in termen externe validiteit dan die alternatieve modelvorm. Als zodanig komt het niet vaak voor dat één van de twee paradigma's (RRM of RUM) zich op overtuigende wijze superieur toont in de context van een bepaalde dataset. Ook al maken deze marginale en soms

---

<sup>14</sup> Deze laatste deelconclusie blijkt ook te gelden in een andere studie die hier niet gerapporteerd wordt, omdat er geen sprake is van empirische modelschatting (Bekhor *et al.*, in druk). In deze studie wordt een RRM-gebaseerd gebruikersevenwicht (*Stochastic User Equilibrium*) ontwikkeld voor de evenwichtstoedeling van verkeer over netwerken. Door RUM-SUE's en RRM-SUE's te bepalen voor het zogenaamde Winnipeg-netwerk kunnen vergelijkingen tussen de twee modeltypen worden gemaakt in termen van de toedeling van verkeer aan verschillende links. Het blijkt dat, in lijn met verwachting en hierboven gepresenteerde resultaten, het RRM-SUE minder verkeer toedeelt aan slecht presterende links (links met hoge reistijden / veel congestie) dan het RUM-SUE, en dat de verschillen in flow aanzienlijk kunnen zijn.

tegenstrijdige verschillen tussen RRM en RUM in termen van modelfit en externe validiteit het vaak lastig om te kiezen voor één van de twee benaderingen, het is – zoals eerder aangegeven – wel zo dat de twee modellen wezenlijk andere beleidsrelevante output kunnen genereren. Hierdoor geldt dat de keuze voor één van de twee modelvormen vanuit beleids- en/of praktisch oogpunt wel degelijk kan uitmaken.

Dit betekent dat er in wezen drie opties overblijven voor de rol van RRM als model voor de kwantitatieve ondersteuning van (transport-)beleid. Ten eerste kan het RRM model worden toegepast in die situaties waarin het aantoonbaar beter presteert dan het conventionele RUM model, bijvoorbeeld in termen van de model fit op relevante datasets. Zoals gezegd kan gedebatteerd worden over de redelijkheid van deze aanpak gegeven (onder andere) de kleine verschillen van de twee modellen in termen van modelfit, maar strikt statistisch bezien is het verdedigbaar om op basis van een statistisch significant verschil in fit op relevante datasets voor één van de twee modellen te kiezen.

Een tweede optie is om de keuze tussen de modelvormen niet (alleen) te baseren op de vaak kleine verschillen in empirische prestaties, maar om (ook) andere criteria te hanteren bij het maken van die keuze. In dat opzicht zijn er verschillende argumenten voor het gebruik van RRM te verzinnen, maar hetzelfde geldt uiteraard ook voor het gebruik van het RUM-model. Voordeel van het conventionele RUM model is dat het de standaard (zij het niet onbetwist) is bij beleidsmakers en wetenschappers, en dat het een bewezen ‘past performance’ heeft. Het RUM model behoeft geen introductie bij beleidsondersteunende diensten, en de eigenschappen van het model zijn bekend en daardoor goed interpreteerbaar door transportwetenschappers en kwantitatief ingestelde beleidsondersteuners. Het RUM-model heeft een sterke fundering in axiomatische welvaartseconomie, wat het gebruik van RUM-gebaseerde welvaartsmaten zoals de betalingsbereidheid (om reistijdwaardering te meten) of de Logsum (om consumentensurplus te meten) vergemakkelijkt en ook rechtvaardigt. Het RRM is nog maar twee jaar oud en dus ‘the new kid in town’: het model moet zijn reputatie nog grotendeels verdienen, en zeker voor wat betreft afgeleide maten zoals de reistijdwaardering en de op RRM-gebaseerde Logsum geldt dat hun interpretatie en fundering in (of: consistentie met) welvaarts-economische principes nog werk in uitvoering zijn (Chorus & de Jong, 2011; Chorus, 2012a; Chorus, 2012d).

Aan de andere kant geldt dat het RRM model relatief goed past in een stroming die in toenemende mate populair is bij beleidsmakers – die van de zogenaamde gedragseconomie. Het RRM model is niet gebouwd op axiomatische, op instrumentele rationaliteit gebaseerde principes als dat van non-transitiviteit of het principe van de ‘Independence from Irrelevant Alternatives’. Integendeel, het RRM-model is gebouwd op basis van een bepaald gedragsidee: het niet noodzakelijkerwijs rationele concept van spijtminimalisatie<sup>15</sup>. De eigenschappen van het RRM-model (die regelmatig strijdig zijn met axioma’s uit de neoklassieke economie) blijken bovendien in het algemeen rechtstreeks terug te voeren op deze gedragsmatige premissen (Chorus 2012a, b). Wanneer bepaalde eigenschappen van het RRM-model strijdig zijn met neoklassieke axioma’s blijkt ook regelmatig dat er in de sociale psychologie overvloedig ondersteunend bewijs voor die

---

<sup>15</sup>De zinsnede ‘niet noodzakelijkerwijs’ wordt hier gebruikt omdat het zeer lastig is om een door iedereen geaccepteerde definitie of interpretatie van het begrip rationaliteit te geven. Als zodanig is het lastig om het etiket (ir-)rationeel aan het RRM-model te plakken. Aan de ene kant kan gesteld worden dat het nastreven van minimale spijt achteraf een rationeel doel is. Aan de andere kant zijn de meeste economen het er over eens dat modellen die niet voldoen aan de IIA-eigenschap (Independence of Irrelevant Alternatives), zoals het RRM model, niet rationeel kunnen worden genoemd (zie bijvoorbeeld Arrow (1950)). Zonder hier al te diep op in te gaan, lijkt het RRM-model beter bij de in opkomst zijnde stroming van ‘behavioral economics’ te passen, dan bij de meer strikte definities van rationaliteit zoals die gehanteerd worden bij de dominante en conventionele stroming van ‘neo-classical economics’.

eigenschappen is. Het meest voor de hand liggende voorbeeld hiervan is het feit dat het RRM-model het compromis-effect voorspelt. Dit effect is strijdig met het hierboven genoemde IIA-principe, maar het ondersteunend *empirisch* bewijs uit de consumentenpsychologie is zeer overtuigend; zie bijvoorbeeld Kivetz *et al.* (2004) en Müller *et al.* (2010). Meer in algemene zin geldt dat er empirisch bewijs is voor de stelling dat de minimalisatie van geanticiperde spijt een belangrijke rol speelt bij het maken van keuzes, met name wanneer deze als moeilijk en belangrijk worden ervaren – zie Zeelenberg en Pieters (2007) voor een overzichtartikel. Bovendien lijken de studies uit tabel 1 te suggereren (zie de conclusies aan het eind van de vorige sectie) dat het RRM model wat gevoeliger is – in termen van effecten op keuzekansen/marktaandeelen – voor beleidsinstrumenten die aangrijpen op de attributen van keuze-opties. Dit kan een voordeel zijn bij de ontwikkeling en *ex ante* evaluatie van beleidsmaatregelen. Afhankelijk van de smaak van de analist of beleidsmaker (dat is: afhankelijk van de gewichten die deze aan de verschillende hierboven genoemde eigenschappen hangt, mogelijk in combinatie met het criterium van modelfit) kan deze in de context van een bepaalde model-exercitie kiezen voor een RRM- of RUM-gebaseerde benadering.

Uiteraard is er ook nog de (derde) optie om niet te kiezen voor één van de twee modellen, maar ze simultaan in te zetten: op deze wijze kan er getracht worden om beleid te ontwikkelen dat zowel vanuit het spijtminimalisatie-perspectief als vanuit nutsmaximalisatie-perspectief relatief goed scoort (bijvoorbeeld in termen van verwachte mobiliteitseffecten en/of welvaartswinst). In andere woorden: op deze wijze kan geprobeerd worden om beleidsmaatregelen te ontwikkelen en/of selecteren die *robuust zijn vanuit gedragsmatig oogpunt*: de geselecteerde maatregelen hebben laten zien relatief goed te scoren, ongeacht de veronderstelde gedragsmatige premissen.

De derde optie lijkt van deze drie het meest vruchtbaar en is ook in lijn met de klassieke notie van een these-antithese-synthese. Op deze wijze kan op een relatief veilige manier geëxperimenteerd worden met het RRM model, en kan het model (verder) onder de aandacht worden gebracht van beleidsmakers en beleidsondersteuners die nu vooral gewend zijn te werken met (de uitkomsten van) RUM modellen. In de tussentijd kan worden afgewacht of het wetenschappelijke veld er in slaagt een solide welvaarts-economisch fundament te ontwikkelen voor RRM modellen, en kan worden bekeken in welke mate de groei in internationaal-wetenschappelijke interesse in het model doorzet. Tegelijkertijd kan er in de tussentijd al wel geprofiteerd worden van de additionele inzichten die het RRM kan bieden in termen van de (mobiliteits-)gedragseffecten van (transport-)beleid.

In dat kader zou het met name interessant zijn het RRM model in te zetten om zogenaamde keuzeset-effecten te bestuderen, zoals het hierboven herhaaldelijk genoemde compromis-effect. Deze effecten houden in dat de aantrekkelijkheid van een alternatief op subtiele wijze afhankelijk is van de compositie van de keuzeset waarin dat alternatief zich bevindt. Voor veel van dergelijke effecten is overtuigend empirisch bewijs voorhanden in vakgebieden als de consumentenpsychologie, en het ligt voor de hand dat dergelijke effecten ook een belangrijke rol kunnen spelen in de context van mobiliteitskeuzes. Zoals eerder in dit artikel naar voren kwam voorspelt het RRM model dat de spijt van een alternatief beïnvloed kan worden door de compositie van de keuzeset aan te passen (en zonder het alternatief zelf aan te passen); spijt bestaat immers bij de gratie van vergelijkingen met concurrerende opties. Zo voorspelt het RRM model het compromis-effect, maar ook een aantal andere potentieel relevante keuzeset effecten (Chorus, 2010, 2012a; Bekhor *et al.*, in druk). Als zodanig biedt het RRM model de mogelijkheid om de potentie van beleidsinstrumenten voor *choice set engineering* te verkennen; deze hebben tot doel om de keuzeset optimaal in te richten in het licht van bepaalde beleidsdoelen. Denk aan het optimaal positioneren van een elektrische auto, gegeven de kenmerken van concurrerende autotypes. Op deze wijze kan het RRM-model beleidsrelevante kennis aandragen die werkelijk een toevoeging vormt ten opzichte van de kennis die lineair-additieve RUM-modellen (kunnen) genereren.

## Dankwoord

Sponsoring door NWO (VENI-beurs 451-10-001) heeft bijgedragen aan de totstandkoming van dit artikel. Commentaar van twee anonieme referenten was zeer behulpzaam bij het reviseren van een eerdere versie van het artikel.

## Bibliografie

- Arentze, T.A., en H.J.P. Timmermans (2007). Parametric action trees: incorporating continuous attribute variables into rule-based models of discrete choice. *Transportation Research Part B*, 41(7), 772-783.
- Arentze, T.A., en H.J.P. Timmermans (2009). A need-based model of multi-day, multi-person activity generation. *Transportation Research Part B*, 43(2), 251-265.
- Arrow, K.J. (1950). A difficulty in the concept of social welfare. *The Journal of Political Economy*, 58(4), 328-346.
- Beck, M.J., C.G. Chorus, J.M. Rose, en D.A. Hensher (in druk). Vehicle purchasing behavior of individuals and groups: Regret or reward? *Journal of Transport Economics and Policy*.
- Bekhor, S., C.G. Chorus, en T. Toledo (in druk). A Stochastic User Equilibrium formulation for the Random Regret Minimization-based route choice model. *Transportation Research Record*.
- De Bekker-Grob, E.W., en C.G. Chorus (in review). Random regret-based discrete choice modeling: An application to health care.
- Ben-Akiva, M., en S.R. Lerman (1985). *Discrete choice analysis: theory and application to travel demand*. The MIT Press, Cambridge, Mass.
- Ben-Akiva, M., en J. Swait (1986). The Akaike likelihood ratio index. *Transportation Science*, 20(2), 133-136.
- Boeri, M., A. Longo, E. Doherty, en S. Hynes (2012). Site choices in recreational demand: A matter of utility maximization or regret minimization? *Journal of Environmental Economics and Policy*, 1(1), 32-47.
- Boeri, M., en L. Masiero (in review). Regret minimization and utility maximization in a freight transport context: An application from two stated choice experiments.
- Chorus, C.G., T.A. Arentze, en H.J.P. Timmermans (2008). A Random Regret Minimization model of travel choice. *Transportation Research Part B*, 42(1), 1-18.
- 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., en C.G. de Jong (2010). Spijtminimalisatie: Een nieuwe modelvorm en een toepassing op vertrektijdstopkeuze-data. *Tijdschrift Vervoerswetenschap*, 46(2), 70-75.
- Chorus, C.G., en C.G. de Jong (2011). Modeling experienced accessibility for utility-maximizers and regret-minimizers. *Journal of Transport Geography*, 19, 1155-1162.
- Chorus, C.G., J.A. Annema, N. Mouter, en G.P. van Wee (2011). Modeling politicians' preferences for road pricing policies: A regret-based and utilitarian perspective. *Transport Policy*, 18, 856-861.
- Chorus, C.G. (2012a). Random Regret Minimization: An overview of model properties and empirical evidence. *Transport Reviews*, 32(1), 75-92.

- Chorus, C.G. (2012b). *Random regret-based discrete choice modeling: A tutorial*. Springer Briefs in Business, Springer, Heidelberg, Germany. [Online URL]: <http://www.springerlink.com/content/186k338384k06174/>
- Chorus, C.G. (2012c). What about behavior in travel demand modeling? An overview of recent progress. *Transportation Letters*, 4(2), 93-104.
- Chorus, C.G. (2012d). Logsums for utility-maximizers and regret-minimizers, and their relation with desirability and satisfaction. *Transportation Research Part A*, 46(7), 1003-1012.
- Chorus, C.G., J.M. Rose, en D.A. Hensher (in druk). Regret minimization or utility maximization: It depends on the attribute. *Environment and Planning Part B*.
- Chorus, C.G., en J.M. Rose (2012). Selecting a date: A matter of regret and compromises. In: Hess, S., en A. Daly (Eds.), *Choice modelling: the state of the art and the state of practice: Selected papers from the Second International Choice Modelling Conference*, Edward Elgar, UK.
- Chorus, C.G., T.A. Arentze, en S. Hess (in review). Incorporating need-satisfaction and regret-minimization in a discrete choice model of leisure activities.
- Chorus, C.G., en M. Bierlaire (in review). An empirical comparison of travel choice models that capture preferences for compromise alternatives.
- Chorus, C.G., M. Koetse, en A. Hoen (in review). Consumer preferences for alternative fuel vehicles: Comparing a utility maximization and a regret minimization perspective.
- Hensher, D.A., W.H. Greene, en C.G. Chorus (in druk). Random Regret Minimization or Random Utility Maximization: An exploratory analysis in the context of automobile fuel choice. *Journal of Advanced Transportation*.
- Hess, S., A. Stathopoulos, en A. Daly (2012). Mixing of behavioral processes: A modeling framework and three case-studies. *Transportation*, 39(3), 565-591.
- Kaa, E.J., van de, (2010). Prospect Theory and choice behaviour strategies: Review and synthesis of concepts from social and transport sciences. *European Journal of Transport and Infrastructure Research*, 10(4), 299-329.
- Kahneman, D., en A. Tversky (1979). Prospect Theory: an analysis of decision under risk. *Econometrica*, 47, 263-291.
- Kaplan, S., en C.G. Prato (in druk). The application of the random regret minimization model to drivers' choice of crash avoidance manoeuvres.
- Kivetz, R., O. Netzer, en V. Srinivasan (2004). Alternative models for capturing the compromise effect. *Journal of Marketing Research*, 41, 237-257.
- McFadden, D. (1974). Conditional logit analysis of qualitative choice-behaviour. In: Zarembka, P. (Eds.), *Frontiers in econometrics*, Academic Press, New York, 105-142.
- Müller, H., E.B. Kroll, en B. Vogt (2010). Fact or artifact? Empirical evidence on the robustness of compromise effects in binding and non-binding choice contexts. *Journal of Retailing and Consumer Services*, 17(5), 441-448.



Prato, C.G. (in review). Confronting the challenges of the random regret minimization model in the route choice context.

Small, K.A., en H.S. Rosen (1981). Applied welfare economics with discrete choice models. *Econometrica*, 49, 105-130.

Thiene, M., M. Boeri, en 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.

Train, K.E. (2003). *Discrete choice methods with simulations*. Cambridge University Press, Cambridge UK.

Tversky, A. (1972). Elimination by Aspects: A theory of choice. *Psychological Review*, 79, 281-299.

Zeelenberg, M., en R. Pieters (2007). A theory of regret regulation 1.0. *Journal of Consumer Psychology*, 17(1), 3-18.

## Appendix: het RRM-model (tekst gebaseerd op Chorus & de Jong (2010))

We veronderstellen de volgende keuzesituatie: een reiziger kiest tussen  $J$  reisopties (zoals routes, vertrektijdstoppen), die elk gedefinieerd zijn in termen van  $M$  attributen  $x_m$ . We richten ons op het voorspellen van de kans dat de reiziger alternatief  $i$  kiest uit deze keuzeset. Het nieuwe RRM-model veronderstelt dat de reiziger probeert geanticipeerde random spijt te minimaliseren, en dat - vanuit het gezichtspunt van de onderzoeker - de random spijt die geassocieerd wordt met alternatief  $i$  bestaat uit de som van een iid storingsterm  $\varepsilon_i$  en 'systematische spijt'  $R_i$ . De afkorting 'iid' staat voor 'independent and identically distributed'; dit houdt in dat de waarde van de storingsterm die bij het ene alternatief hoort, onafhankelijk is van de waarde van de storingsterm die bij elk van de andere alternatieven hoort, en dat alle storingstermen uit dezelfde verdeling getrokken zijn. We veronderstellen dat de negatieve van  $\varepsilon_i$  een zogenaamde *Extreme Value Type I* verdeling volgt.

Systematische spijt is gedefinieerd als de som van de zogenaamde bilaterale spijt die geassocieerd wordt met de vergelijking tussen  $i$  en elke andere keuzeoptie:  $R_i = \sum_{j \neq i} R_{i \leftrightarrow j}$ . Als zodanig

veronderstelt het spijtmodel dat elk alternatief dat beter presteert dan  $i$  een bijdrage levert aan de spijt die de reiziger anticipeert en wil vermijden. De bilaterale spijt die geassocieerd wordt met de vergelijking tussen  $i$  en alternatief  $j$  bestaat uit de som van de spijt die geassocieerd wordt met de vergelijking tussen de twee keuzeopties in termen van elk attribuut:  $R_{i \leftrightarrow j} = \sum_{m=1..M} R_{i \leftrightarrow j}^m$ . Tenslotte,

de spijt die geassocieerd wordt met de vergelijking tussen  $i$  en  $j$  in termen van attribuut  $m$  (in formulevorm:  $R_{i \leftrightarrow j}^m$ ) is als volgt geformuleerd:  $R_{i \leftrightarrow j}^m = \ln\left(1 + \exp\left[\beta_m \cdot (x_{jm} - x_{im})\right]\right)$ . Deze formulering impliceert dat spijt zo goed als nul is wanneer alternatief  $i$  beter presteert dan  $j$  op attribuut  $m$ , en dat spijt een bij benadering lineaire functie is van het verschil in attribuutwaarden wanneer  $j$  beter presteert dan  $i$ . In dat laatste geval is de te schatten parameter  $\beta_m$  (waarvan ook het teken wordt geschat) te interpreteren als de richtingscoëfficiënt van de bij benadering lineaire spijtfunctie.

Totale systematische spijt wordt dan:  $R_i = \sum_{j \neq i} \sum_{m=1..M} \ln\left(1 + \exp\left[\beta_m \cdot (x_{jm} - x_{im})\right]\right)$ . De kans dat

de reiziger keuzeoptie  $i$  kiest, is nu gelijk aan de kans dat de random spijt die geassocieerd wordt met  $i$  kleiner is dan de random spijt van de andere alternatieven. Uit het feit dat de negatieve van de  $\varepsilon$ -storingstermen verondersteld worden de iid-*Extreme Value Type I* verdeling te volgen, volgt dat deze kans gegeven wordt door een zogenaamde logit-formulering:

$$P_i = \exp(-R_i) / \sum_{j=1..J} \exp(-R_j).$$