

De waarde van betrouwbare reistijden in personenverkeer en -vervoer in Nederland¹

Marco Kouwenhoven

Significance²

Gerard de Jong

Significance en ITS Leeds³

Paul Koster

Vrije Universiteit en Tinbergen Instituut Amsterdam⁴

Vincent van den Berg

Vrije Universiteit en Tinbergen Instituut Amsterdam⁵

Erik Verhoef

Vrije Universiteit en Tinbergen Instituut Amsterdam⁶

John Bates

John Bates Services⁷

Pim Warffemius

Kennisinstituut voor Mobiliteitsbeleid, Ministerie van Infrastructuur en Milieu⁸

¹ Dit is een bewerking en vertaling van het artikel 'New values of time and reliability in passenger transport in The Netherlands' van dezelfde auteurs, zoals dat verschenen is in Research in Transportation Economics (RETREC), Nummer 47 (2014), blz. 37-49. Er is toestemming voor publicatie van deze aangepaste vertaling in TVW verkregen van de kant van de uitgever van RETREC, Elsevier.

² kouwenhoven@significance.nl

³ dejong@significance.nl

⁴ p.r.koster@vu.nl

⁵ v.a.c.vanden.berg@vu.nl

⁶ e.t.verhoef@vu.nl

⁷ 113300.3565@compuserve.com

⁸ pim.warffemius@minienm.nl

Dit artikel presenteert nieuwe monetaire waarden voor reistijd ('value of time', VOT) en betrouwbaarheid van reistijd ('value of reliability', VOR), die bedoeld zijn voor gebruik in kosten-batenanalyses van transportprojecten in Nederland. Het gaat om één van de eerste studies wereldwijd waar deze waarden empirisch zijn bestudeerd binnen een gezamenlijk raamwerk.

De gebruikte data zijn afkomstig uit 'stated preference' (SP) enquêtes onder reizigers, waar de respondenten hypothetische alternatieven kregen voorgelegd met variatie in reistijden, reiskosten en betrouwbaarheid van reistijden. Betrouwbaarheid werd gepresenteerd in de vorm van een serie van vijf mogelijke reistijden binnen hetzelfde keuze-alternatief, met alle vijf een gelijke kans om op te treden.

In eerste instantie (2009) werden de interviews uitgevoerd via een bestaand internet panel. Later (2011) zijn ook respondenten geworven op benzinstations en parkeer-plaatsen langs snelwegen, in parkeergarages, op treinstations en bushaltes, luchthavens en jachthavens. Een belangrijke conclusie is dat het SP onderzoek via het internet panel tot aanzienlijk lagere VOT's leidt dan het SP onderzoek met werving tijdens de reis, waarschijnlijk vanwege zelfselectie van respondenten met een lage waardering van tijd in het internet panel.

Discrete keuzemodellen zijn geschat op de SP-data, waarin rekening gehouden wordt met verschillen in de VOT (en deels ook in de VOR) tussen verplaatsingen met verschillende tijd- en kostenniveaus, verschillen voor wat betreft de in de SP aangeboden tijd- en kostenveranderingen en in de waargenomen kenmerken van de respondenten (b.v. opleiding, inkomen, leeftijd). Door gebruik te maken van een 'panel latent class' model wordt ook rekening gehouden met niet-waargenomen heterogeniteit tussen respondenten en met het feit dat respondenten in het experiment herhaalde keuzes maken. De referentie VOT en VOR zijn bepaald op de steekproef uit 2011, maar het effect van de tijd- en kostenniveaus, tijd- en kostenveranderingen en socio-economische kenmerken is geschat op de data van 2009 en 2011 samen.

Trefwoorden: latent class model; reistijdbetrouwbaarheid; reistijdwaardering; stated preference; value of reliability; value of time.

1. Inleiding

De waardering van reistijd ('value of time', VOT) geeft de maatschappelijke baat van de afname van de gemiddelde reistijd of de maatschappelijke kosten van de toename daarvan. De reistijdwaardering wordt periodiek vastgesteld door een groot praktijkonderzoek onder reizigers, vervoerders en verladers. In de tijd tussen twee praktijkonderzoeken worden de waarderingsgetallen jaarlijks opgehoogd met loonontwikkeling en inflatie. Het laatste praktijkonderzoek voor het personenverkeer en -vervoer vond plaats in 1997 (Hague Consulting Group, 1998). Nu, meer dan 15 jaar later, zijn de waarderingsgetallen voor reistijd in het personenverkeer en -vervoer weer getoetst aan de praktijk en bijgesteld.

Naast reistijdwinst is de betrouwbaarheid van reistijden een belangrijke batenpost in maatschappelijke kosten-batenanalyses (MKBA's). Betrouwbaarheid is een belangrijk kwaliteitsaspect van een reis, verplaatsing of transport. Betrouwbaarheid gaat over de mate waarin de reistijd varieert, of meer specifiek: over de variatie rondom de gemiddelde reistijd. Als we praten over de betrouwbaarheid van reistijden, gaat de meeste aandacht uit naar te laat komen. Echter, ook te vroeg komen leidt tot kosten, zoals wachten op de plaats van bestemming. Bij het personenverkeer en -vervoer leiden onverwachte vertragingen tot kosten door extra wachttijd (met een hoger disnut), stress bij de reizigers, gemiste aansluitingen, gemiste afspraken en negatieve gevolgen voor de efficiëntie in de bedrijven. Om de kans op te laat komen te verkleinen, gaan reizigers vaak over tot het hanteren van veiligheidsmarges en tot het afwijken van wat zonder onzekerheid de meest geprefereerde aankomsttijden zouden zijn ('scheduling costs').

Het verbeteren van de betrouwbaarheid van reistijd betekent heel concreet het verminderen van onverwachte vertragingen. Een meer betrouwbare reistijd betekent voor de reiziger dus dat reis en activiteiten beter gepland kunnen worden. Het verbeteren van reistijdbetrouwbaarheid voor het personen- en goederenvervoer is een belangrijk onderwerp in de Structuurvisie Infrastructuur en Ruimte van het Ministerie van Infrastructuur en Milieu (2012).

De waardering van betrouwbaarheid ('value of reliability', VOR) geeft de maatschappelijke baat van het verkleinen van de spreiding (in deze studie gemeten als de standaardafwijking) van de reistijd. In 2004 zijn financiële waarderingsgetallen voor de betrouwbaarheidsverbetering vastgesteld op basis van een internationale expert meeting georganiseerd door het toenmalige ministerie van Verkeer en Waterstaat (Hamer et al., 2005; Jong, de et al., 2009). Tegelijkertijd is toen de behoefte geformuleerd aan waarderingsgetallen voor reistijdbetrouwbaarheid die zijn gebaseerd op empirisch onderzoek in Nederland. Dat onderzoek is nu uitgevoerd en de resultaten zijn beschreven in dit paper.

De standaardafwijking van de reistijdverdeling is hier gekozen als maatstaf van onbetrouwbaarheid⁹. De belangrijkste reden hiervoor is dat het opnemen van betrouwbaarheid

⁹ Voor het meten van de betrouwbaarheid van reistijden kunnen diverse maatstaven worden gebruikt. Bijvoorbeeld: de buffertijd index die laat zien hoeveel veiligheidsmarge in termen van extra reistijd een reiziger moet hanteren om 95% kans te hebben op tijd te komen, diverse statistische maten die de variatie van de reistijd rondom de verwachte reistijd in beeld brengen of schedule delay maten die gaan over het disnut als de reiziger eerder/later aankomt dan gepland.

in de bestaande transportprognosemodellen zoals LMS en NRM niet eenvoudig is, waarbij de standaardafwijking op korte en middellange termijn de meest haalbare optie is.

Het onderzoek dat in dit artikel wordt beschreven, is de eerste nationale waarderingsstudie in Nederland die zowel de VOT als de VOR in het personenverkeer en -vervoer bestudeert binnen één consistente onderzoeksopzet. Het onderzoek is uitgevoerd onder supervisie van het Kennisinstituut voor Mobiliteitsbeleid (KiM) door een breed consortium bestaande uit: Significance, Vrije Universiteit Amsterdam, John Bates Services, TNO, NEA, TNS NIPO, en PanelClix. De technische en methodologische details van het onderzoek zijn beschreven in Significance et al. (2013).

Het onderzoek levert nieuwe kengetallen voor de waardering van reistijd en betrouwbaarheid die toegepast kunnen worden in Nederlandse MKBA's van infrastructuurprojecten voor zowel het personen- als goederenvervoer. In dit artikel kijken we naar het personenverkeer en -vervoer (de uitkomsten voor de VOT en VOR in het goederenvervoer staan in Significance et al., 2013, KiM, 2013 en de Jong et al., 2014). Voor het personenverkeer en -vervoer heeft het onderzoek nieuwe VOT's en VOR's opgeleverd voor wat betreft de volgende vervoersmodaliteiten: auto, bus, tram, metro, trein, luchtvaart en recreatievaart. De verplaatsingsmotieven die zijn bestudeerd zijn: woon-werk, zakelijk en overig (incl. winkelen, onderwijs, recreatie en sociaal verkeer).

Naast hun toepassing in MKBA's kunnen de reistijd- en betrouwbaarheids-waarderingen ook gebruikt worden voor kosten-effectiviteitsanalyses waarin verschillende beleidsmaatregelen en investeringen worden vergeleken. De waarderingen kunnen ook worden gebruikt voor het berekenen van de kosten van files.

Het vervolg van dit paper bespreekt achtereenvolgens de 'stated preference' (SP) vragenlijst en experimenten (paragraaf 2), de dataverzameling en steekproeven (paragraaf 3), de modelspecificatie en schattingsresultaten (paragraaf 4) en de uitkomsten voor de VOT's en VOR's (paragraaf 5). In paragraaf 6 tenslotte worden de conclusies gepresenteerd.

2. Stated preference interviews

2.1 Vragenlijst

De vragenlijst van dit onderzoek bestond uit de volgende onderdelen (zie ook Tabel 1 voor de kenmerken die gebruikt zijn in de SP experimenten):

1. Vragen over de kenmerken van een recent ondernomen verplaatsing¹⁰, zoals reistijd en reiskosten. De waarden van deze kenmerken zijn gebruikt als de basisniveaus voor de kenmerken die in de SP experimenten zijn getoond.

¹⁰ In het onderzoek in 2009 werden de respondenten gevraagd om terug te denken aan de meest recente verplaatsing die zij voor een bepaald (voorgeselecteerd) motief hadden gemaakt. In het onderzoek in 2011 werden de respondenten gevraagd om terug te denken aan de verplaatsing die zij maakten toen ze werden geworven.

2. Vragen over de beschikbaarheid van een andere (dan de gebruikte) vervoerwijze voor deze verplaatsing en over de waarden van kenmerken voor die vervoerwijze. (Dit leverde echter slechts een zeer beperkt aantal RP keuzen op, dat onvoldoende bleek voor de schatting van een RP model. Bovendien kunnen dergelijke schattingsresultaten vertekend worden door systematische fouten in de door respondenten gerapporteerde reistijden (zie Peer, 2013)).
3. SP experiment 1 met zes keuzesituaties waarin er steeds twee routekeuze-alternatieven tegenover elkaar worden gepresenteerd, die worden beschreven in termen van twee kenmerken: reistijd en reiskosten.
4. Een inleiding voor de respondent over wat bedoeld wordt met variabele (onbetrouwbare) reistijden.
5. SP experiment 2a met zes keuzesituaties met ieder twee routekeuze-alternatieven. De alternatieven worden nu beschreven in termen van vier kenmerken: reistijd, reiskosten, betrouwbaarheid van de reistijd en aankomsttijd. Een vijfde kenmerk, vertrektijd, wordt wel gepresenteerd aan de respondenten, maar het wordt berekend uit de andere vier kenmerken. Respondenten in de pleziervaart hebben niet aan dit experiment meegedaan.
6. SP experiment 2b met zeven keuzesituaties. Dit experiment is vergelijkbaar met experiment 2a, behalve dat er geen variatie is tussen beide alternatieven in de gebruikelijke aankomsttijd. Om een controle op consistente antwoorden te kunnen uitvoeren geldt voor een van de keuzeparen dat een van de beide alternatieven dominant is. Dit betekent dat het ene alternatief beter of even goed is op alle kenmerken vergeleken met het andere alternatief (dus een alternatief heeft dan bijvoorbeeld een lagere prijs en reistijd, en dezelfde onzekerheid van reistijd ten opzichte van het andere alternatief).
7. Vragen waarin de respondenten gevraagd wordt om de keuzen die zij in de experimenten gemaakt hebben te evalueren.
8. Vragen over de persoon (leeftijd, geslacht, enz.) en het huishouden (samenstelling, inkomen, enz.).

Tabel 1: Kenmerken die gepresenteerd worden in de SP experimenten (exclusief pleziervaart)

Kenmerk	Experiment 1	Experiment 2a	Experiment 2b
Gebruikelijke reistijd	√	√	√
Reiskosten	√	√	√
Betrouwbaarheid: vijf mogelijke reistijden		√	√
Vijf mogelijke aankomsttijden		√	√
Vertrektijd		√	√
Variatie in gebruikelijke aankomsttijd		√	

2.2 SP experimenten (met uitzondering van pleziervoort)

De keuzesituaties in alle SP experimenten betreffen keuzes tussen dezelfde vervoerwijzen ('within-mode choice'), afhankelijk van de vervoerwijze van de in werkelijkheid gemaakte verplaatsing waar het interview zich op richtte. De respondent werd steeds gevraagd voor welk van de twee alternatieven hij/zij een voorkeur had.

Voor autoverplaatsingen zijn de keuze-alternatieven gepresenteerd als twee mogelijke routes voor de verplaatsing van de respondent, terwijl voor openbaar vervoer (inclusief vliegtuig) de beide alternatieven werden voorgesteld als twee mogelijke aanbiedingen die verschillen in termen van kosten, tijden (op basis van de dienstregeling) en betrouwbaarheid.

Een voorbeeld van een feitelijk gebruikt computerscherm van een keuzesituatie staat in Figuur 1. De SP-opzet met twee alternatieven (A en B) en slechts twee kenmerken (reistijd en reiskosten) is voor experiment 1 gekozen om zo veel mogelijk consistent te zijn met de reistijdwaarderingsonderzoeken uit 1988 en 1997, zodat een vergelijking van uitkomsten in de tijd mogelijk zou zijn. Het statistische SP-ontwerp dat hier gebruikt wordt (zie Significance et al., 2007) is het zogenaamde 'Bradley design', dat nagenoeg orthogonaal is, maar zonder dat er dominant alternatieven worden getoond), en waarin alle mogelijke alternatieven voorkomen ('full design').

Welke rit heeft uw voorkeur?

Rit A	Rit B
Gebruikelijke reistijd: 60 min.	Gebruikelijke reistijd: 45 min.
Kosten: € 2.80	Kosten: € 3.60
<input type="radio"/> Voorkeur voor Rit A	<input type="radio"/> Voorkeur voor Rit B

Figuur 1: Voorbeeld van een SP keuzesituatie in experiment 1 voor automobilisten

De 'gebruikelijke reistijd' in Figuur 1 is de deur-tot-deur reistijd voor een verplaatsing in één richting. Er wordt gevarieerd rond de op het vertrektijdstip verwachte reistijd voor de verplaatsing van de respondent. 'Kosten' betreft het totaal van de kosten die de respondent moet betalen voor een verplaatsing in één richting.

Figuur 2 geeft een voorbeeld van een keuzesituatie in experiment 2a.

Welke rit heeft uw voorkeur?

3 / 6

Rit A	Rit B
Vertrektijd: 07:45	Vertrektijd: 08:25
<i>U heeft een even grote kans op elk van deze 5 reistijden en dus om op deze tijdstippen aan te komen:</i>	<i>U heeft een even grote kans op elk van deze 5 reistijden en dus om op deze tijdstippen aan te komen:</i>
Reistijd: Aankomsttijd:	Reistijd: Aankomsttijd:
50 min. -> 08:35	25 min. -> 08:50
60 min. -> 08:45	35 min. -> 09:00
60 min. -> 08:45	35 min. -> 09:00
80 min. -> 09:05	65 min. -> 09:30
100 min. -> 09:25	95 min. -> 10:00
Gebruikelijke reistijd: 60 min.	Gebruikelijke reistijd: 35 min.
Kosten: € 2.40	Kosten: € 5.80
<input type="radio"/> Voorkeur voor Rit A	<input type="radio"/> Voorkeur voor Rit B

Figuur 2: Voorbeeld van een SP keuzesituatie in experiment 2a voor automobilisten

Het moge duidelijk zijn dat vele respondenten het begrip standaardafwijking niet zullen kennen. Daarom kan dit begrip niet als zodanig getoond worden in de SP experimenten (hoewel we het wel later in de modellen zullen gebruiken). In plaats daarvan presenteren we aan de respondenten de betrouwbaarheid in de vorm van een reeks van vijf mogelijke reistijden (die alle vijf even waarschijnlijk zijn). Omdat het hier voor de respondent om een complex SP experiment gaat, is in een vooronderzoek met behulp van 'face-to-face' interviews verkend wat de beste manier is om betrouwbaarheid te presenteren aan respondenten. Hierbij bleek een verbale omschrijving (dus niet een histogram of staafdiagram) van vijf mogelijke reistijden het beste te voldoen (Significance et al., 2007; Tseng et al., 2009).

De gepresenteerde reistijdverdelingen zijn asymmetrisch, omdat dit een realistischer beeld geeft dan een symmetrische verdeling. Dit heeft tot gevolg dat het gemiddelde en de mediaan van elkaar zullen verschillen. Bovendien zijn de tweede en de derde reistijd uit de serie van vijf altijd hetzelfde, waardoor deze waarde overeenkomt met zowel de modus als de mediaan. In de latere modelschattingen hebben we modellen met het gemiddelde en modellen met de mediaan getest, maar in dit artikel presenteren we uitsluitend modellen op basis van het gemiddelde, omdat die consistent zijn met de gebruikte prognosemodellen voor verkeer en vervoer en met de KBA methodiek. Experiment 2a heeft als statistisch ontwerp een 'orthogonal factorial design' met vier kenmerken, ieder met vijf mogelijke niveaus.

Een keuzesituatie voor experiment 2b staat in Figuur 3.

Welke rit heeft uw voorkeur?

2 / 7

Rit A	Rit B
Vertrektijd: 08:10	Vertrektijd: 08:00
<i>U heeft een even grote kans op elk van deze 5 reistijden en dus om op deze tijdstippen aan te komen:</i>	
Reistijd: Aankomsttijd:	Reistijd: Aankomsttijd:
25 min. -> 08:35	35 min. -> 08:35
35 min. -> 08:45	45 min. -> 08:45
35 min. -> 08:45	45 min. -> 08:45
55 min. -> 09:05	55 min. -> 08:55
75 min. -> 09:25	65 min. -> 09:05
Gebruikelijke reistijd: 35 min.	Gebruikelijke reistijd: 45 min.
Kosten: € 1.80	Kosten: € 2.80
<input type="radio"/> Voorkeur voor Rit A	<input type="radio"/> Voorkeur voor Rit B

Figuur 3: Voorbeeld van een SP keuzesituatie in experiment 2b voor automobilisten

In dit experiment is de gebruikelijke aankomsttijd (die overeenkomt met de als tweede en derde gepresenteerde aankomsttijd in de serie van vijf) voor beide alternatieven (Rit A en Rit B) hetzelfde (08:45 uur, zie Figuur 3). In experiment 2a wijken deze juist af van elkaar (08:45 versus 09:00, zie Figuur 2). Dit is het verschil tussen de experimenten 2a en 2b: in experiment 2a is er meer ruimte voor verandering in de tijdplanning. Experiment 2b heeft een onderliggend uitgebreid 'Bradley design' met drie kenmerken met elk vijf niveaus.

2.3 SP experimenten voor pleziervaart

Bij pleziervaart (b.v. met een zeilboot of een motorjacht) is het doel van de verplaatsing meestal niet om van A naar B te komen, maar om onderweg te genieten van de reis. Daarom verwachten we dat de reistijd positief gewaardeerd wordt in plaats van negatief (hoe langer de verplaatsing, hoe attractiever). Reistijdwinst is hier dus niet relevant omdat bij pleziervaart juist aan de reis waarde wordt ontleend. De pleziervaart is een belangrijke gebruikersgroep van bruggen en sluisen. De baten die zij ontleen aan kortere wachttijden kunnen goed worden meegenomen in de MKBA's voor investeringen in deze bruggen en sluisen.

Om deze redenen hebben we voor pleziervaart een andere context gekozen voor de SP experimenten: respondenten worden gevraagd om terug te denken aan een situatie waarin zij moesten wachten bij een brug of een sluis. Omdat we er hier niet van uitgaan dat de respondenten op een bepaald tijdstip moeten vertrekken of aankomen, hebben we vertrektijd en aankomsttijd niet opgenomen. Dit heeft ook tot gevolg dat experiment 2a identiek wordt aan experiment 2b; de respondenten krijgen alleen experiment 2b voorgelegd. Voorbeelden van computerschermen van de feitelijk getoonde keuzeparen staan in Figuur 4 (experiment 1) en Figuur 5 (experiment 2b). De gebruikte kenmerken zijn:

- Gebruikelijke wachttijd voor een brug of sluis.
- De kosten die de respondent moet betalen om de brug of sluis te passeren.
- Betrouwbaarheid in de vorm van een set van vijf mogelijke wachttijden (die alle vijf even waarschijnlijk zijn).

Welke route heeft uw voorkeur?

<p>Route A</p> <p>Gebbruikelijke wachttijd voor brug/sluis 15 min.</p> <p>Kosten per passage: €0.50</p>	<p>Route B</p> <p>Gebbruikelijke wachttijd voor brug/sluis 5 min.</p> <p>Kosten per passage: €1.50</p>
<input type="radio"/> Voorkeur voor Route A	<input type="radio"/> Voorkeur voor Route B

Figuur 4: Voorbeeld van een SP keuzesituatie in experiment 1 voor pleziervaart

Welke route heeft uw voorkeur?

5 / 7

<p>Route A</p> <p>Gebbruikelijke wachttijd voor brug/sluis: 15 min.</p> <p>U heeft een even grote kans op elk van deze 5 wachttijden.</p> <p>13 min. 14 min. 15 min. 16 min. 17 min.</p> <p>Kosten: €0.50</p>	<p>Route B</p> <p>Gebbruikelijke wachttijd voor brug/sluis: 5 min.</p> <p>U heeft een even grote kans op elk van deze 5 wachttijden.</p> <p>0 min. 3 min. 5 min. 7 min. 10 min.</p> <p>Kosten: €1.50</p>
<input type="radio"/> Voorkeur voor Route A	<input type="radio"/> Voorkeur voor Route B

Figuur 5: Voorbeeld van een SP keuzesituatie in experiment 2b voor pleziervaart

3. De steekproeven van 2009 en 2011

Er zijn twee dataverzamelingen uitgevoerd. Het oorspronkelijke plan was om uitsluitend data uit internetenquêtes onder de leden van een bestaand internet panel te gebruiken. Uit initiële analyses van deze data werden echter uitkomsten verkregen die we niet als plausibel beschouwen: de reistijdwaarderingen uit multinomiale logit (MNL) modellen met tijd en kosten als verklarende variabelen op deze data waren 25-45% lager dan de waarderingen uit 1997 (zelfs zonder correctie voor inflatie). Slechts een beperkt deel van dit verschil kon verklaard worden door verschillen in de samenstellingen van de betreffende steekproeven voor wat betreft verplaatsingsafstand, inkomen en andere socio-economische variabelen en de verschillen in het statistische ontwerp van de SP.

Diverse externe experts¹¹ zijn toen geconsulteerd over de vraag hoe hiermee verder te gaan. Zowel de externe experts als het onderzoeksteam benadrukten hierbij de mogelijkheid dat de steekproef van het internet panel een systematische fout zou vertonen op het kenmerk reistijdwaardering. Hiermee wordt bedoeld dat binnen ieder segment (socio-economisch, motief, verplaatsingsafstand, vervoerwijze) de respondenten die aan zo'n internetpanel deelnemen naar verwachting een lagere reistijdwaardering hebben dan niet-deelnemers. Deelname aan een panel vergt namelijk tijd, waarvoor de deelnemers een betrekkelijke geringe geldelijke beloning ontvangen. Zelfs na herweging van de steekproef naar de populatie (op basis van populatiegegevens van het CBS) zou de gevonden reistijdwaardering nog steeds substantieel lager uitkomen dan de waarde uit 1997.

Gezien het bovenstaande is vervolgens besloten om een additionele dataverzameling te doen, waarbij de respondenten volgens dezelfde methode geworven werden als bij de vorige nationale reistijdwaarderingstudies in Nederland, dus op benzinstations/parkeerplaatsen langs snelwegen, in parkeergarages, op treinstations, bushaltes, luchthavens en jachthavens en bij sluizen. Op deze rekrutering van de respondenten volgden dan interviews met behulp van het internet (vrijwel hetzelfde 'web-based' interview als in 2009, maar met een andere werving¹²).

In de uiteindelijke modellen zijn beide datasets gebruikt. Het referentieniveau voor de reistijdwaardering is echter uitsluitend gebaseerd op de data uit 2011 (zie paragraaf 5)

Tabel 2 en 3 tonen het aantal succesvol afgeronde interviews voor de dataverzameling in 2009 en het aantal dat overbleef na het verwijderen van uitbijters (met name alle waarnemingen van personen met onverklaarbare reistijden, reiskosten, snelheid, enz.). De respondenten die hun voorkeur hadden uitgesproken voor het gedomineerde¹³ alternatief in experiment 2b zijn ook verwijderd. Respondenten die de kenmerken van de alternatieven niet tegen elkaar lijken af te ruilen omdat ze altijd voor, bijvoorbeeld, het snelste of goedkoopste alternatief kiezen ('non-traders') zijn daarentegen niet verwijderd. Hun keuzes kunnen namelijk wel degelijk een getrouwe weergave van hun preferenties weerspiegelen.

¹¹ Prof. dr. Peter Bonsall van ITS Leeds, Prof. dr. Jonas Eliasson van KTH Stockholm, en Dr. Eric Molin van de TU Delft samen met Prof. dr. Harry Timmermans van de TU Eindhoven.

¹² Er is een vraag toegevoegd aan de vragenlijst in 2011: of de respondent lid was van een internetpanel (en zo ja, van welk); zie paragraaf 5 voor de analyse van de uitkomsten hiervan.

¹³ Het gedomineerde alternatief is het tegenovergestelde van het dominante alternatief dat staat uitgelegd paragraaf 2.1 bij punt 6.

Tabel 2: Aantal succesvol afgeronde interviews in het onderzoek van 2009 (internetpanel)

Jaar		Motief			Totaal
		Woon-werk	Zakelijk	Overig	
2009	Vervoerwijzen Auto	1341	523	790	2654
	Trein	908	284	329	1521
	BTM	586	80	174	840
	Vliegtuig	0	157	374	531
	Pleziervaart	0	0	214	214
Totaal		2835	1044	1881	5760

Noot: BTM= bus, tram en metro.

Tabel 3: Aantal in de schatting gebruikte interviews uit het onderzoek van 2009 (internetpanel)

Jaar		Motief			Totaal
		Woon-werk	Zakelijk	Woon-werk	
2009	Vervoerwijzen Auto	1008	349	538	1895
	Trein	699	235	249	1183
	BTM	469	61	136	666
	Vliegtuig	0	96	297	393
	Pleziervaart	0	0	178	178
Totaal		2176	741	1398	4315

In Tabel 4 en 5 staan dezelfde gegevens voor de dataverzameling in 2011. Het aandeel dat verwijderd moest worden is in beide datasets vergelijkbaar.

Tabel 4: Aantal succesvol afgeronde interviews in het onderzoek van 2011 (rekrutering tijdens de reis)

Jaar		Motief			Totaal
		Woon-werk	Zakelijk	Other	
2011	Vervoerwijzen				
	Auto	184	305	125	614
	Trein	131	52	103	286
	BTM	125	17	91	233
	Vliegtuig	9	29	163	201
	Pleziervaart	0	0	95	95
	Totaal	449	403	577	1429

Tabel 5: Aantal in de schatting gebruikte interviews uit het onderzoek van 2011 (rekrutering tijdens de reis)

Jaar		Motief			Totaal
		Woon-werk	Zakelijk	Other	
2011	Vervoerwijzen				
	Auto	150	235	93	478
	Trein	105	41	79	225
	BTM	97	11	70	178
	Vliegtuig	7	23	152	182
	Pleziervaart	0	0	81	81
	Totaal	359	310	475	1144

4. Modelspecificatie en schattingsresultaten

4.1 Eenvoudige MNL modellen met gemiddelde en spreiding

In de literatuur over de waardering van reistijd en variabiliteit van reistijd in personenverkeer en –vervoer kunnen twee hoofdgroepen van modellen worden onderscheiden (zie de Jong et al., 2004; Batley et al., 2008; OECD, 2010; Carrion en Levinson, 2012; Significance et al., 2012):

- modellen met gemiddelde en een spreiding van reistijden;
- de tijdstipkeuze ('scheduling') aanpak.

Beide hoofdgroepen verschillen van elkaar op het gebied van de variabelen die opgenomen worden in de nutsfunctie¹⁴.

In de eerste plaats bevatten de modellen met gemiddelde en een spreidingsmaatstaf: reistijd, reiskosten en een maatstaf van de spreiding in de reistijdverdeling (meestal de standaardafwijking, soms de variantie).

In de tweede plaats bevatten de nutsfuncties van het 'scheduling' model reistijd, reiskosten en tijdstipkeuze termen voor het aantal minuten dat men eerder of later arriveert dan wat het meest geprefereerd zou zijn als alle andere kenmerken niet over de tijd zouden variëren. Deze specificatie kan worden gebaseerd op de tijdstipkeuzetheorie, zoals ontwikkeld door Vickrey (1969) en Small (1982). Een ander, gerelateerd, model, dat uitgaat van het nut dat de consument ondervindt op de herkomst en op de bestemming, is gepresenteerd door Vickrey (1973) en Tseng en Verhoef (2008).

In de derde plaats is het ook mogelijk om zowel een spreidingsterm als tijdstipkeuze variabelen in dezelfde nutsfunctie op te nemen. In dat geval heeft de spreidingsterm als functie om invloeden weer te geven die niet gerelateerd zijn aan de keuze van aankomst en vertrek op zich, zoals stress en kosten van opnieuw moeten plannen als van de verwachte reistijd wordt afgeweken, en mogelijk ook de invloed van niet-lineariteiten in het marginale nut van te vroeg of te laat aankomen.

We hebben alle drie deze specificaties op de data geschat.¹⁵ De modellen met gemiddelde en spreiding (in dit geval met de standaardafwijking) voldeden beter dat het tijdstipkeuze model voor wat betreft de model 'fit' (e.g. de log-likelihood waarde). Het gecombineerde model doet het in dit opzicht iets beter dan het model met alleen gemiddelde en spreiding. In het gecombineerde

¹⁴ Er is een theoretische equivalentie (onder de veronderstelling dat de reistijdverdeling onafhankelijk is van de dagdelen en bekend bij de reiziger en dat de reiziger zijn/haar vertrektijd optimaliseert) tussen de Vickrey/Small 'scheduling' aanpak en een model dat het gemiddelde en de standaardafwijking van de reistijd gebruikt (Bates et al, 2001; Fosgerau en Karlström, 2010). Verder is er een equivalentie tussen het Vickrey/Tseng/Verhoef 'scheduling' model en een model met het gemiddelde en de variantie van de reistijd (Fosgerau en Engelson, 2011). Hierdoor is het theoretisch mogelijk om een spreidingsmaatstaf (en dus ook een VOR) te berekenen uit een tijdstipkeuze model. Wat de beste aanpak is hangt ervan af hoe men het beste empirische data kan verzamelen en welk model het best past in het prognosesysteem voor verkeer en vervoer dat in een toepassing gebruikt wordt (Börjesson et al., 2011).

¹⁵ Alle schattingen zijn gedaan met ALOGIT en/of BIOGEME.

model gaan echter de t-waarden voor de ‘reliability ratio’ (RR) ($RR = VOR/VOT$) en voor de coëfficiënt voor te laat aankomen omlaag. In een model met enkel gemiddelde en spreiding pikt de ‘reliability ratio’ een groter deel van het effect van onbetrouwbaarheid op. Aangezien het verschil in log-likelihood klein is, er in een gecombineerd model het gevaar bestaat van interpretatieproblemen en omdat de uitkomsten moeten worden toegepast in het kader van een statisch transportmodel zonder ‘scheduling’ termen, kiezen we een model met alleen gemiddelde en spreiding, dat ook direct uitkomsten kan leveren voor de waarde van de standaardafwijking voor toepassing in de KBA (zie paragraaf 1).

Binnen de modellen met gemiddelde en spreiding starten we met een eenvoudige multinominale logit (MNL) nutsfunctie, zonder interactietermen voor socio-economische invloeden. In plaats van een aparte nutscoëfficiënt voor iedere variabele, standaardiseren we de schatting in kosteneenheden, door één schaalcoëfficiënt (β_c) voor kosten te schatten. Door deze (in de parameters) niet-lineaire specificatie kunnen de VOT en de VOR direct geschat worden, ieder als een aparte coëfficiënt, in plaats van berekend uit de verhouding van een geschatte coëfficiënt ten opzichte van de geschatte kostencoëfficiënt.

In deze schattingen combineren we beide datasets (2009 en 2011) en ook voegen we de data van alle drie de experimenten samen. We hebben data voor twee verschillende wijzen van rekruteren van de respondenten (internet en tijdens de reis) en kunnen aparte VOT’s en ‘reliability ratios’ per dataset schatten. We hebben schaalfactoren toegevoegd om mogelijk verschillen in de stochastische model component tussen beide datasets op te pikken (hierbij nemen we 2009 als referentie, dus die heeft als schaalfactor 1) en tussen de drie experimenten (met de schaalfactor voor experiment 2a op 1)¹⁶.

Het systematische of waargenomen deel van de nutsfunctie is gespecificeerd als:

$$U = (S_c^{09} \cdot \delta^{09} + S_c^{11} \cdot \delta^{11}) \cdot (S_c^{\text{exp1}} \cdot \delta^{\text{exp1}} + S_c^{\text{exp2a}} \cdot \delta^{\text{exp2a}} + S_c^{\text{exp2b}} \cdot \delta^{\text{exp2b}}) \cdot \beta_c \cdot [C + (VOT^{09} \cdot \delta^{09} + VOT^{11} \cdot \delta^{11}) \cdot (T + (RR^{09} \cdot \delta^{09} + RR^{11} \cdot \delta^{11}) \cdot \sigma)] \quad [1]$$

waarbij:

- S_c^y = de schaalfactor voor de dataset uit jaar y (2009 of 2011);
- δ^y = een dummy (0 of 1) die aangeeft of een waarneming hoort bij de dataset uit jaar y; of bij experiment 1, 2a of 2b;
- C = reiskosten; alle kostenniveaus zijn gecorrigeerd voor inflatie (prijsniveau 2010);
- T = gemiddelde reistijd;
- σ = standaardafwijking van de reistijdverdeling;
- β_c = kostenparameter (te schatten coëfficiënt; schaalfactor in geld-eenheden);
- VOT^y = reistijdwaardering uit data van jaar y (te schatten coëfficiënt);
- RR^y = Reliability ratio (=VOR/VOT) uit data van jaar y (te schatten coëfficiënt).

Om het artikel niet te lang te maken, verwijzen we hier voor schattingsresultaten voor deze relatief standaard MNL modellen naar het rapport Significance et al. (2013).

¹⁶ We hebben de hypothese getest dat experiment 1 tot hogere VOT’s zou leiden dan experimenten 2a en 2b (omdat de VOT’s uit experiment 1 iets van de VOR zouden kunnen bevatten), maar vonden hiervoor geen statistisch bewijs.

4.2 Geavanceerde MNL modellen met gemiddelde en spreiding

Nadat de eenvoudige MNL modellen met gemiddelde en spreiding met succes waren geschat, zijn we verder gegaan met meer geavanceerde modellen, die rekening houden met de gevoeligheid voor hogere basisniveaus en vervolgens ook met kleinere waarderingen voor kleinere veranderingen (zoals een lage VOT per minuut of per uur voor kleine tijdswinsten). Hierbij wordt voortgebouwd op de nutsfunctie [1] met schaalfactoren, VOT en RR. Om de uitleg hier eenvoudiger te houden laten we hieronder nutsfuncties zien zonder de schaalfactoren en de dummy-variabelen voor de jaren en experimenten, maar in de praktijk zijn deze steeds gehandhaafd. Om dezelfde reden beginnen we met een model zonder onbetrouwbaarheidsterm. Vergelijking [2] geeft dan het eenvoudige MNL model.

$$U = \beta_C \cdot [C + VOT \cdot T] \quad [2]$$

Afhankelijkheid van (waargenomen) basiskosten en basistijd

Uit eerdere studies (zoals Gunn, 2001; Mackie et al., 2003; Stathopoulos en Hess, 2011) is bekend dat de VOT in sterke mate af kan hangen van het huidige (waargenomen) niveau van reistijd en reiskosten van de respondent, zoals ook overeenkomt met de ‘prospect theory’ (Kahneman en Tversky, 1979, 1992; van de Kaa, 2008). De waarden van deze BaseCost (C_0) en BaseTime (T_0) worden in de SP experimenten gebruikt om waarden aan te bieden die rond deze tijd en kosten variëren. We hebben daarom de afhankelijkheid van de VOT van BaseTime en BaseCost opgenomen in de nutsfunctie. Hieronder wordt uitgelegd hoe dit is gedaan.

We beginnen met het formuleren van de in SP gepresenteerde variabelen in niveaus (C , T) als een basiswaarde (C_0 , T_0) plus een increment (ΔC , ΔT). De basiswaarden zijn voor alle alternatieven hetzelfde, zodat we de nutsfunctie kunnen herschrijven als:

$$U = \beta_C \cdot [\Delta C + VOT \cdot \Delta T] \quad [3]$$

Daarna hebben we, in navolging van eerdere onderzoeken (b.v. Mackie et al., 2003; Stathopoulos en Hess, 2011), een machtsfunctie gespecificeerd voor zowel BaseCost als BaseTime. Hierbij worden de incrementele kosten- en tijdtermen vermenigvuldigd met de basiswaarde verheven tot de macht λ . Bovendien delen we hierbij de basiswaarde door een arbitraire ‘referentiewaarde’. Deze is identiek voor de hele steekproef en is uitsluitend toegevoegd om de schatting te stabiliseren. De nutsfunctie wordt dan:

$$U = \beta_C \cdot \left(\left(\frac{C_0}{C_{ref}} \right)^{\lambda_C} \cdot \Delta C + VOT_{ref} \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T} \cdot \Delta T \right) \quad [4]$$

Hierbij zijn C_0 en T_0 de waargenomen basiskosten en basistijd zijn van de respondent, en C_{ref} and T_{ref} zijn de referentiewaarden voor de basiskosten en basistijd in de steekproef.

De VOT hangt nu af van C_0 en T_0 :

$$VOT(C_0, T_0) = \frac{\partial U / \partial T}{\partial U / \partial C} \Bigg|_{T=T_0, C=C_0} = VOT_{ref} \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T} / \left(\frac{C_0}{C_{ref}} \right)^{\lambda_C} \quad [5]$$

Een keuze van andere referentiepunten C_{ref} en T_{ref} zal leiden tot andere schattingen voor β_C and VOT_{ref} , maar het nut U wordt door deze keuze niet beïnvloed. Daarom hangt de $VOT(C_0, T_0)$ in [5] ook niet af van C_{ref} en T_{ref} , en kunnen we deze vrij kiezen.¹⁷

In [4] verwachten we dat de gevoeligheid voor een gegeven verandering in tijd en kosten af zal nemen met toenemende waarden voor basiskosten en –tijd, dus dat zowel λ_T als λ_C negatief zullen zijn (zie ook Stathopoulos en Hess, 2011). Hoe de VOT dan verandert met toenemende waarden voor basiskosten en tijd hangt af van deze relatieve grootte van deze twee parameters, en van de relatieve verandering T_0/C_0 , zie [5].

Afhankelijkheid van de omvang van de veranderingen in tijd en kosten zoals getoond in de SP

Naast de afhankelijkheid van de basisniveaus, hebben diverse recente VOT studies (b.v. De Borger en Fosgerau (2008) op Deense data, Ramjerdi et al. (2010) op Noorse data en Börjesson en Eliasson (2014) op Zweedse data) een afhankelijkheid van de VOT opgenomen voor de omvang van de tijdwinsten die in de SP worden aangeboden (hoewel ze aanbevelen om in de KBA dezelfde VOT per tijdseenheid te gebruiken voor grote en kleine tijdwinsten als gevolg van transportprojecten). Een uitgebreide discussie over dit onderwerp kan worden gevonden in Daly et al. (2014). We willen ook corrigeren voor de invloed van de omvang van de kosten- en tijdsveranderingen, ΔC en ΔT , zoals aangeboden in de SP. Hiervoor specificeren we nog een machtsfunctie, nu met ΔC and ΔT -termen. Net als hiervoor stabiliseren we de schatting door te delen door een arbitraire basiswaarde.

Als we beide afhankelijkheden samenvoegen leidt dat tot de volgende specificatie van de nutsfunctie:¹⁸

$$U = \beta_C \cdot \left(\left(\frac{|\Delta C|}{\Delta C_{ref}} \right)^{\gamma_C - 1} \cdot \left(\frac{C_0}{C_{ref}} \right)^{\lambda_C} \cdot \Delta C + VOT_{ref} \cdot \left(\frac{|\Delta T|}{\Delta T_{ref}} \right)^{\gamma_T - 1} \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T} \cdot \Delta T \right) \quad [6]$$

waarbij:

C_0 , en T_0	= de basiskosten en –tijd van de respondent (zoals waargenomen);
C_{ref} , en T_{ref}	= arbitraire referentiewaarden; ¹⁹
ΔC en ΔT	= kosten- en tijdveranderingen aangeboden in de SP;
ΔC_{ref} en ΔT_{ref}	= arbitraire referentiewaarden; ²⁰
γ , λ , β_C en VOT_{ref}	= te schatten coëfficiënten.

De referentiewaarden zijn uitsluitend toegevoegd om de schatting te stabiliseren.

¹⁷ De geschatte parameter VOT_{ref} is de waarde van tijd als de basiskosten en basistijd gelijk zijn aan de arbitraire referentiewaarden. Zoals later in deze paragraaf zal blijken, is in de uiteindelijke nutsfunctie [6], bij gelijkheid van de basis- aan de referentiewaarden, de VOT niet gelijk aan VOT_{ref} maar wel proportioneel ermee. Dus, in het algemeen gesproken is VOT_{ref} slechts een ondersteunende coëfficiënt zonder interpretatie.

¹⁸ Ook een ander aspect van de ‘prospect theory’, verschillende waarden voor winsten en verliezen, is getest, maar dit leidde niet tot significante verbetering van het model.

¹⁹ We kiezen voor $C_{ref} = 3$ euro, wat dicht in de buurt ligt van zowel het gemiddelde als de mediaan voor BaseCost in de data voor 2009 en 2011 voor automobilisten voor het motief woon-werk en we kiezen voor $T_{ref} = 40$ minuten, ook dicht bij het gemiddelde en de mediaan van BaseTime onder de auto/woon-werk respondenten in 2009 en 2011.

²⁰ Op basis van de gemiddelde en mediane waarden van ΔC en ΔT in onze dataset, kiezen we $\Delta C_{ref} = 1$ euro en $\Delta T_{ref} = 5$ minuten.

Onze veronderstelling is dat de reactie op een gegeven verandering in kosten en tijd toe zal nemen met toenemende omvang van de verandering in kosten en tijd. Dus verwachten we zowel γ_T als $\gamma_C > 1$ (dit komt overigens niet uit in de schattingen, zie hieronder).

We kunnen opnieuw de VOT afleiden door differentiatie van de nutsfunctie. Bij deze specificatie [6] zal de VOT afhangen van C_0 , T_0 , ΔC en ΔT :

$$VOT(C_0, T_0, \Delta C, \Delta T) = VOT_{ref} \cdot \frac{\gamma_T \cdot \left(\frac{|\Delta T|}{\Delta T_{ref}} \right)^{\gamma_T - 1} \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T}}{\gamma_C \cdot \left(\frac{|\Delta C|}{\Delta C_{ref}} \right)^{\gamma_C - 1} \cdot \left(\frac{C_0}{C_{ref}} \right)^{\lambda_C}} \quad [7]$$

Deze $VOT(C_0, T_0, \Delta T, \Delta C)$ is onafhankelijk van de gekozen referentiewaarden voor T_{ref} , C_{ref} , ΔT_{ref} en ΔC_{ref} , omdat de (geschatte) VOT_{ref} proportioneel is met $\frac{(\Delta T_{ref})^{\gamma_T - 1} \cdot (T_{ref})^{\lambda_T}}{(\Delta C_{ref})^{\gamma_C - 1} \cdot (C_{ref})^{\lambda_C}}$, en deze parameter zal veranderen om iedere verandering in de veronderstellingen over de referentiewaarden te compenseren.²¹

Betrouwbaarheid is in het model op een zelfde manier behandeld als reistijd. Het startpunt is dan het vervangen van de term $VOT \cdot T$ in [2] door $VOT \cdot (T + RR \cdot \sigma)$. Net als in [4], nemen we dan de afhankelijkheid van de RR voor de basiswaarde van de standaardafwijking op, door het hier te

specificeren als $RR_{ref} \cdot \left(\frac{\sigma_0}{\sigma_{ref}} \right)^{\lambda_R}$ waarbij σ_0 de basiswaarde is voor de standaardafwijking van de

respondent en voor σ_{ref} kiezen we de arbitraire waarde die overeenkomt met het onbetrouwbaarheidsniveau bij $T_{ref} = 40$ minuten). De “gamma” variaties uit [6] worden echter niet gebruikt voor de betrouwbaarheidsvariabele, en de variaties in VOT door verschillende niveaus van ΔT zijn geïsoleerd van de RR term. De specificatie met onbetrouwbaarheid (zonder schaafeffecten en socio-economische interacties) wordt gegeven door [8]:

$$U = \beta_C \cdot \left(\left(\frac{C_0}{C_{ref}} \right)^{\lambda_C} \cdot \left(\frac{|\Delta C|}{\Delta C_{ref}} \right)^{\gamma_C - 1} \cdot \Delta C + VOT_{ref} \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T} \cdot \left(\frac{|\Delta T|}{\Delta T_{ref}} \right)^{\gamma_T - 1} \cdot \Delta T + RR_{ref} \cdot \left(\frac{\sigma_0}{\sigma_{ref}} \right)^{\lambda_R} \cdot \Delta \sigma \right) \quad [8]$$

²¹ Nu geldt echter dat wanneer we alle argumenten van de VOT op hun referentiewaarde zetten, we niet uitkomen op VOT_{ref} maar op $(\gamma_T/\gamma_C) \cdot VOT_{ref}$. Hier geldt dus dat de geschatte parameter VOT_{ref} geïnterpreteerd kan worden als de VOT maal de ratio (γ_C/γ_T) .

De VOR kan hieruit, door differentiatie, worden afgeleid:

$$VOR(C_0, T_0, \sigma_0, \Delta C) = VOT_{ref} \cdot \frac{\left(\frac{T_0}{T_{ref}}\right)^{\lambda_T} \cdot \left(RR_{ref} \cdot \left(\frac{\sigma_0}{\sigma_{ref}}\right)^{\lambda_R}\right)}{\gamma_C \cdot \left(\frac{|\Delta C|}{\Delta C_{ref}}\right)^{\gamma_C-1} \cdot \left(\frac{C_0}{C_{ref}}\right)^{\lambda_C}} \quad [9]$$

De VOT_{ref} en de RR_{ref} in Eq. [8] worden apart geschat voor beide datasets. De coëfficiënten γ en λ worden echter gezamenlijk geschat op beide datasets, aangezien wij ervan uitgaan dat deze meer een intrinsiek effect van de bevolking weergeven, en omdat onze eerdere aparte schattingen wezen op vergelijkbare coëfficiënten in beide datasets. Op deze manier maken we optimaal gebruik van de grote dataset uit 2009 in de schatting van sommige coëfficiënten (zoals de γ and λ factors, en ook de socio-economische interactiecoëfficiënten die in sectie 4.3. aan de orde komen), zonder dat er een systematische fout wordt geïntroduceerd in de VOT en de RR door deelname aan het internetpanel.

De recente Deense, Noorse en Zweedse VOT onderzoeken op het gebied van personenverkeer en –vervoer gebruiken schatting in ‘logWTP-space’, d.w.z. zij gebruikten logaritmische nutsfuncties (zie b.v. Fosgerau, 2006; Ramjerdi et al. 2010; Börjesson et al., 2012; Börjesson en Eliasson, 2014). In onze standaard modellen met gemiddelde en spreiding vonden wij ook dat ‘logWTP’ modellen het statistisch beter deden dan modellen in ‘WTP-space’. Echter na het toevoegen van de γ and λ coëfficiënten voor de invloed van basisniveaus en de omvang van in de SP aangeboden veranderingen, voldoen modellen in ‘WTP-space’ statistisch beter dan in ‘logWTP-space’.

4.3 Geavanceerde MNL modellen met gemiddelde en spreiding en met socio-economische interactietermen

Deze modellen bouwen voort op die van sectie 4.2. Socio-economische invloeden zijn toegevoegd, via interacties van deze invloeden met VOT_{ref} in [8]. Omdat we hier een ‘reliability ratio’ (RR), schatten, zal de VOR meebewegen met de VOT en zo gelden dezelfde interacties ook voor de VOR: segmenten met een hogere VOT zullen ook een hogere VOR kennen. Er zijn significante interactie-effecten gevonden voor leeftijdsklasse, opleidingsniveau, geslacht, samenstelling van het huishouden, inkomen, verplaatsingen in de spits en voor vervoerwijze.

Het verband tussen inkomen en de VOT was vrijwel lineair (zie Significance et al., 2013). Een invloed van vervoerwijze vonden we alleen voor bus/tram/metro binnen het motief woon-werk en voor trein in het motief overig (in deze gevallen negatieve invloed, relatief ten opzichte van auto-bestuurder) Verdere segmentatie naar vervoerwijze was niet nodig, dat werd door andere variabelen opgevangen. Wel geldt dat in de berekening van de aanbevolen VOTs en VORs gebruik gemaakt wordt van ‘sample enumeration’ op de schattingsdata in combinatie met ophoogfactoren om cijfers over de mobiliteit op landelijk niveau (uit het Onderzoek Verplaatsingen in Nederland, OViN, van het CBS) zo goed mogelijk weer te geven. Een gevolg hiervan is dat verschillen in bijvoorbeeld triplengte tussen vervoerwijzen (b.v. gemiddeld langere verplaatsingen met de trein dan met andere modi) tot verschillen in de aanbevolen VOTs tussen vervoerwijzen leiden.

4.4 Latente klassen modellen

In deze sectie bespreken we modellen die voortbouwen op die uit sectie 4.3 maar die ook rekening houden met niet-waargenomen verschillen in de preferenties van de respondenten. Hiermee worden verschillen bedoeld die we niet met de beschikbare exogene variabelen kunnen verklaren, maar mogelijk wel met stochastische termen.

Als verdere uitbreiding zijn allereerst allerlei ‘mixed logit’ modellen geschat. In deze modellen werden de VOT_{ref} and RR_{ref} getrokken uit een continue kansverdeling (verschillende verdelingen zijn hierbij getest). Aanmerkelijk stabielere resultaten werden echter verkregen met ‘latente klassen’ modellen (Hess et al., 2011; Hensher et al., 2012). Deze modellen kenmerken zich door het gebruik van discrete in plaats van continue verdelingen voor bepaalde coëfficiënten in het model, zonder van te voren een bepaalde vorm voor de kansverdeling te veronderstellen. Het resultaat is een ‘histogram’ van een aantal discrete klassen met bijbehorende kansen per klasse en schattingen voor de coëfficiënten. Een model met één latente klasse is equivalent aan het standaard MNL model. We kunnen daarom statistische toetsen toepassen op de mate van niet-waargenomen heterogeniteit in de voorkeuren van respondenten en zo ook het aantal latente klassen in de data bepalen.

De SP data bevatten herhaalde metingen voor dezelfde individuen. De MNL modellen uit secties 4.1 – 4.3 veronderstellen juist dat iedere keuze op zichzelf staat (onafhankelijk is van de andere keuzen). Daarom noemen we deze modellen ‘cross-sectionele’ modellen; ze houden geen rekening met het gegeven dat een en dezelfde respondent een reeks van keuzen maakt. De panel latente klassen modellen die we hebben geschat houden wel rekening met dit gegeven door de voorkeuren van hetzelfde individu constant te houden over de reeks van zijn/haar keuzen.²²

Verdelingen met verschillende aantallen latente klassen zijn getest voor de VOT_{ref} , met 2-5 latente klassen (afhankelijk van het reismotief) als optimale uitkomst. Om zo goed mogelijk gebruik te maken van de grote steekproef uit 2009 is verondersteld dat de vorm van de verdeling hetzelfde was voor de 2009 en 2011 data, bij verschillende waarden voor de VOT voor een bepaalde klasse (VOT_{ref09_0} tot en met VOT_{ref11_4} in Tabel 6). Bovendien is verondersteld dat de diverse covariaten hetzelfde proportionele effect op de VOT en VOR hebben in 2009 en 2011.

Om het aantal te schatten parameters binnen de perken te houden, hebben we de covariaten ingebracht als interactie met de VOT_{ref} variabelen, wat tot gevolg heeft dat de vorm van de verdeling hetzelfde is voor iedere combinatie van covariaten, maar dat de gemiddelde VOT_{ref} anders is.²³ Aparte modellen zijn geschat voor de motieven woon-werk, zakelijk en overig, en voor vliegtuig-zakelijk, vliegtuig-overig en pleziervaart, waarbij de covariaten die in eerdere MNL schattingen significant bleken opnieuw zijn meegenomen. Het optimale aantal latente klassen is bepaald, aan de hand van het ‘Bayesian Information Criterion’ (BIC). Voor woon-werk is het optimum 5, voor zakelijk, overig en vliegtuig is het 4 en voor pleziervaart 3.

²² De gebruikte latente klassen modellen corrigeren niet voor autocorrelatie omdat het simuleren van een dergelijke stochastische component zeer veel rekentijd vraagt. Dit is een mogelijk interessante uitbreiding voor toekomstig onderzoek.

²³ ‘Class membership’ modellen zullen in de toekomst ook worden getest, maar deze vergen meer parameters aangezien er voor iedere covariaat N-1 variabelen aan het model moeten worden toegevoegd, waarbij N het aantal latent klassen is.

De schattingsresultaten staan in Tabel 6. De t-waarden in deze tabel zijn zogenaamde ‘robuste’ t-waarden die tot op zekere hoogte corrigeren voor specificatiefouten (Bierlaire, 2008). Voor pleziervaart konden geen stabiele schattingen voor de gamma’s worden gevonden en hebben we deze parameters op 1 gezet.

Tabel 6: Geschatte coëfficiënten en t-waarden voor geavanceerde latente klassen modellen met gemiddelde, spreiding en socio-economische interacties

Auto / Trein / BTM		Woon-werk		Zakelijk		Overig	
Fit	Observ. (resp.)	45186 (2528)		16476 (932)		20658 (1165)	
	Final log (L)	-22103.134		-7925.771		-9471.395	
	Rho ² (0)	0.33		0.362		0.156	
Coëfficiënten		Waarde	(t)	Waarde	(t)	Waarde	(t)
	BetaCost	-1.16	(-35.2)	-1.05	(-20.6)	-1.45	(-22.6)
	VOtref09_0	4.06	(19.8)	38.1	(6.4)	0.168	(0.8)
	VOtref09_1	69.3	(5.1)	9.37	(12.6)	10.5	(14.0)
	VOtref09_2	0	(0.0)	4.18	(10.9)	5.01	(15.1)
	VOtref09_3	15.3	(9.2)	0	(0.0)	163	(3.1)
	VOtref09_4	7.75	(19.8)	-		-	
	VOtref11_0	5.57	(19.3)	44.1	(5.9)	6.82	(10.3)
	VOtref11_1	57.3	(2.0)	1.31	(1.3)	39.1	(4.5)
	VOtref11_2	11.9	(10.7)	5.19	(12.3)	1.68	(1.3)
	VOtref11_3	46.3	(2.6)	12.2	(7.1)	12.4	(2.3)
	VOtref11_4	0	(0.0)	-		-	
	Group_0	0	(*)	0	(*)	0	(*)
	Group_1	-2.55	(-21.6)	1.17	(4.5)	-1.5	(-5.4)
	Group_2	-0.33	(-3.8)	2.28	(14.1)	-0.13	(-0.8)
	Group_3	-2.42	(-8.8)	1.58	(9.5)	-3.06	(-9.0)
	Group_4	-0.775	(-6.5)	-		-	
	RR09	1.17	(11.5)	1.51	(10.4)	1.2	(3.4)
	RR11	0.408	(2.2)	1.15	(6.8)	0.624	(1.3)
	facTrain					-0.106	(-1.9)
	facBTM	-0.0891	(-2.5)				
	fac3650	-0.107	(-4.1)			-0.0396	(-0.7)
	fac51pl	-0.186	(-6.5)	-0.104	(-1.9)	-0.233	(-4.8)
	facEdu1	-0.331	(-1.8)				
	facEdu2	-0.11	(-2.5)				
	facEdu12			-0.284	(-3.8)	-0.0826	(-1.0)
	facEdu34	-0.0409	(-1.5)			-0.0911	(-2.5)
	facFem			0.0062	(0.1)	0.0465	(0.8)
	facHH1	0.138	(3.4)				
	facHH12			0.165	(2.0)		
	facInc	0.0761	(8.1)	0.109	(6.3)	0.0273	(1.8)
	facPeak	0.0789	(2.8)	0.188	(3.0)	0.123	(2.4)
	gammaC	0.523	(40.9)	0.548	(28.4)	0.537	(28.6)
	gammaT	1.06	(48.8)	1.01	(31.2)	1.05	(18.8)
	lambdaC	-0.386	(-22.1)	-0.473	(-16.7)	-0.382	(-16.0)
	lambdaT	-0.526	(-14.8)	-0.515	(-10.5)	-0.559	(-7.3)
	lambdaR	-1.05	(-7.2)	-1.19	(-15.4)	-0.86	(-4.1)
	Sc11	0.723	(20.6)	0.683	(16.8)	0.636	(16.8)
	ScEx1	4.02	(24.7)	4.34	(15.1)	3.52	(14.1)
	ScEx2b	1.33	(25.1)	1.35	(15.3)	1.37	(18.6)

Vliegtuig		Woon-werk	Zakelijk / Overig	
Fit	Observ. (resp.)		9750 (575)	
	Final log (L)		-4952.262	
	Rho ² (0)		0.164	
Coëfficiënten			Waarde	(t)
	BetaCost		-3.32	(-5.1)
	VO _{Tref09_0}		3.74	(5.9)
	VO _{Tref09_1}		0	(0.0)
	VO _{Tref09_2}		1.79	(5.7)
	VO _{Tref09_3}		9.83	(4.9)
	VO _{Tref11_0}		0.593	(1.0)
	VO _{Tref11_1}		6.68	(5.7)
	VO _{Tref11_2}		2.36	(4.8)
	VO _{Tref11_3}		35	(0.6)
	Group_0		0	(*)
	Group_1		-0.0667	(-0.2)
	Group_2		0.762	(1.9)
	Group_3		-1.66	(-4.9)
	RR09		1.35	(1.1)
	RR11		0.653	(1.4)
	facBus		0.0295	(0.4)
	gammaC		0.64	(22.6)
	gammaT		1.01	(24.9)
	lambdaC		-0.722	(-13.2)
lambdaT		-0.796	(-11.6)	
lambdaR		-1.62	(-3.3)	
Sc11		0.72	(12.8)	
ScEx1		4.16	(10.7)	
ScEx2b		1.28	(12.1)	

Pleziervaart		Woon-werk	Zakelijk	Overig	
Fit	Observ. (resp.)			3102 (259)	
	Final log (L)			-1660.0	
	Rho ² (0)			0.25	
Coëfficiënten				Waarde	(t)
	BetaCost			-0.336	(-9.5)
	VO _{Tref09_0}			15.9	(8.1)
	VO _{Tref09_1}			5.15	(19.1)
	VO _{Tref09_2}			1.39	(4.2)
	VO _{Tref11_0}			0.291	(0.4)
	VO _{Tref11_1}			6.39	(13.0)
	VO _{Tref11_2}			40.8	(4.5)
	Group_0			0	(*)
	Group_1			1.54	(6.2)
	Group_2			-0.632	(-1.7)
	Sc11			1.01	(6.6)
	ScEx1			5.75	(6.9)

waarbij:

- BTM: bus, tram en metro.
- VOT_{ref}: VOT voor een van de latente klassen.
- Group: coëfficiënt voor de lidmaatschapskansen van een latente klasse.
- fac3650: reizigers in de leeftijdsklasse 36-50 hebben een lagere VOT voor woon-werk en overig dan de jongere leeftijdsklassen (die de basis of referentieklassse voor leeftijd vormen).

- fac51pl: reizigers in de leeftijdsklasse 51 jaar en ouder hebben een lagere VOT dan reizigers jonger dan 36 (en ook vergeleken met de leeftijdsklasse 36-50), voor alle drie de motieven.
- facEdu1: reizigers met basisschool als hoogste opleiding hebben een lagere VOT voor woon-werk dan reizigers met een hogere opleiding (HBO/Universiteit is de referentie).
- facEdu2: reizigers met een lagere beroepsopleiding hebben een lagere VOT voor woon-werk dan reizigers met een hogere opleiding
- facEdu34: reizigers met een middelbare of hogere middelbare schoolopleiding hebben een lagere VOT voor woon-werk en overig dan reizigers met een hogere opleiding.
- facEdu12: reizigers met basisschool of een lagere beroepsopleiding hebben een lagere VOT voor overig dan reizigers met een hogere opleiding.
- facFem: vrouwen hebben een hogere VOT voor zakelijk en overig (mogelijk omdat ze vaker te maken hebben met multi-tasking).
- facHH1: huishoudens met 1 lid hebben een hogere VOT voor woon-werk dan andere huishoudens (we verwachten dat deze huishoudens geen taken kunnen delen met andere leden, dus hebben ze hogere 'opportunity' kosten' voor reizen).
- facHH12: huishoudens met maar één volwassene, met of zonder kinderen, hebben een hogere VOT voor zakelijk dan andere huishoudens (we verwachten dat deze huishoudens meer moeite zullen hebben om taken te kunnen delen met andere leden, dus hebben ze hogere 'opportunity' kosten' voor reizen).
- facInc: lineair inkomen: hogere inkomens hebben een hogere VOT voor alle motieven.
- facPeak: verplaatsingen in de spits (middelpunt van de verplaatsing valt tussen 7-9 uur of 16-18 uur) hebben een hogere VOT voor alle motieven; dit kan verklaard worden uit het feit dat reizigers het extra onprettig vinden om te reizen als er congestie op de weg of drukte in het OV is.
- facBTM: bus/tram/metro voor woon-werk heeft een lagere VOT dan auto-bestuurder.
- facTrain: treinreizigers hebben een lagere VOT voor overig dan autobestuurders (mogelijk omdat ze hun tijd in de trein plezieriger en productiever kunnen gebruiken dan in andere vervoerwijzen door gebruik te maken van informatietechnologie).

De tekens van de geschatte coëfficiënten komen overeen met de intuïtie, hoewel sommige die in MNL nog significant waren, nu niet meer significant zijn.

Uit de schattingsresultaten kunnen we concluderen dat de parameters voor afhankelijkheid van de referentiesituatie ('reference dependence') significant zijn, behalve de $\gamma(\text{size})$ coëfficiënt voor reistijd, die niet significant van 1 verschilt in de modellen voor zakelijk en overig. Het "afstand" effect λ is tussen -0.3 en -0.6 voor zowel tijd als kosten, zodat ze beide negatief zijn, zoals verwacht. γ is tussen 0.4 en 0.6 voor kosten, en tussen 0.85 en 0.95 voor tijd. We vinden dat langere verplaatsingen een hogere VOT hebben en dat de kostenveranderingen die in de SP zijn aangeboden aanzienlijk meer effect hebben op de VOT dan de veranderingen in tijd.

In de geschatte latente klassen modellen, is de VOT_{ref} voor 2009 30% lager dan de VOT_{ref} voor 2011. Bovendien hebben we additionele modellen geschat (hier niet gepresenteerd) op alleen de 2011 data, waarbij we onderscheid maakten tussen leden van een internet panel en niet-leden. We vonden 10-30% lagere VOT's voor de leden. Dit is een duidelijke aanwijzing dat internet panels selectiviteit kunnen veroorzaken voor wat betreft de VOT.

De 'reliability ratios' voor de data sets van 2009 en 2011 verschillen significant van elkaar en zijn ook significant verschillend van 1. Voor de aanbevolen VOT en VOR gebruiken we de VOT_{ref} en

de RR_{ref} van 2011, die veel beter sporen met de voorgaande Nederlandse VOT onderzoeken en buitenlandse/internationale studies dan die voor 2009, en die niet gebukt gaan onder de ‘bias’, die naar onze mening veroorzaakt wordt door het gebruik van een internet panel.

5. Aanbevolen VoT's en VoR's

5.1 De gebruikte methode

Voor de aanbevolen VOT's en VOR's gebruiken we de schattingsresultaten van de latente klassen modellen uit paragraaf 4. Verschillende stappen waren nodig om hier vandaan te komen tot aanbevolen waarden.

Allereerst is voor het zakelijke motief de zogenaamde ‘Hensher vergelijking’ (zie Hensher 1977, Fowkes et al., 1986 en Wardman et al., 2013), gebruikt, die de VOT splitst in een werknemerscomponent en een werkgeverscomponent. De eerste komt dan uit het ‘latent class’ model op de SP data onder zakelijke reizigers, de tweede komt uit kengetallen (die zijn afgeleid uit de 2011 data) over de fractie van de reistijd die gewerkt wordt, de relatieve productiviteit van werk tijdens de reis, het deel van de tijdwinst dat gewerkt zou worden en de productieve waarde van werktijd.

Ten tweede is voor de werknemerscomponent van zakelijk en voor de andere motieven de steekproef uit 2011 gebruikt en zijn voor ieder individu uit deze steekproef de VOT en de VOR berekend, rekening houdend met zijn of haar socio-economische en verplaatsingskenmerken (respondent-specifieke basistijd en –kosten). We gaan er van uit dat de afhankelijkheid van ΔT en ΔC (gedeeltelijk) een kunstmatig gevolg is van het gebruik van de SP methode, en ook dat dit onderscheid tot praktische problemen in de KBA zou leiden (aangezien een verandering met 2 € niet langer gelijk zou zijn aan twee keer een verandering met 1 €). Daarom gebruiken we $\gamma_C = \gamma_T = 1$ om de VOT en VOR te berekenen. [6] hangt niet langer expliciet af van ΔT en ΔC , maar de waarde van VOT_{ref} hangt wel af van de voor ΔT_{ref} en ΔC_{ref} gekozen waarden, dus is een redelijke keuze van deze waarden vereist. Wij hebben dit opgelost door voor iedere respondent de gemiddelde tijd- en kostenverschillen te berekenen tussen de twee aangeboden alternatieven in de 18 niet-dominante keuzesituaties van de SP.²⁴

In de derde plaats is in een vervolgstap de gebruikte steekproef representatief gemaakt voor de mobiliteit van de Nederlandse bevolking. Hiervoor zijn de verplaatsingen in het OViN (‘Onderzoek Verplaatsingen in Nederland’), opgedeeld in klassen, gedefinieerd op basis van vijf bevolkingskenmerken (geslacht, leeftijd, inkomen, huishoudsamenstelling en opleiding) en twee verplaatsingskenmerken (dagdeel per vervoerwijze en reisduur per vervoerwijze). Het OViN is een steekproef, met ongeveer 136.000 waarnemingen. Iedere waarneming in het OViN heeft zelf een weegfactor om het OViN representatief te maken voor alle verplaatsingen van de Nederlandse bevolking in een jaar. Hierbij beperken we ons tot personen van 16 jaar en ouder, die gereisd hebben met auto, trein of bus/tram/metro. De verdeling van de verplaatsingen in onze steekproef verschilt van die voor de hele bevolking (er is b.v. een oververtegenwoordiging van woon-werk). De ‘Iterative Proportional Fitting method’ is gebruikt om nieuwe gewichten

²⁴ We hebben deze methode gevalideerd door ook geavanceerde MNL modellen te schatten met γ_C en γ_T vastgezet op 1. De gemiddelde VOT in de steekproef week hierbij minder dan 10% af van die volgens de hier beschreven methode.

voor onze steekproef te berekenen zodanig dat de gewogen steekproef naar de zeven variabelen de opgehoogde OViN verdeling goed representeert. Bij het berekenen van de finale gewichten is de VOT van iedere respondent gewogen met dit gewicht en met de reistijd.

5.2 Uitkomsten voor de VOT

De nieuwe VOT's, berekend volgens de methode beschreven in sectie 5.1, staan in Tabel 7. We kunnen deze nieuwe waarden vergelijken met de vorige officiële waarden die gebaseerd waren op het VOT onderzoek 1997-1998 (Hague Consulting Group, 1998) en vervolgens jaarlijks zijn opgehoogd met inflatie en reële inkomensgroei. Het blijkt dat de verschillen doorgaans vrij gering zijn. In het algemeen gesproken is er een daling voor auto, een stijging voor trein en een stabilisatie voor BTM. Voor een gedetailleerde vergelijking van de nieuwe VOT's met die uit de vorige studie en die uit de internationale literatuur verwijzen we naar Significance et al. (2013) en KiM (2013).

Tabel 7: Nieuwe reistijdwaarderingen (in euro's per persoon per uur, prijsniveau 2010, inclusief BTW), per vervoerwijze

	Auto	Trein	Bus, tram, metro	Landmodi (auto, trein, BTM)	Vliegtuig	Pleziervaart
Woon-werk	9,25	11,50	7,75	9,75		
<i>Zakelijk werknemer</i>	12,75	15,50	10,50	13,50	85,75	
<i>Zakelijk werkgever</i>	13,50	4,25	8,50	10,50	-	
Zakelijk	26,25	19,75	19,00	24,00	85,75	
Overig	7,50	7,00	6,00	7,00	47,00	8,25
Alle motieven	9,00	9,25	6,75	8,75	51,75	8,25

Noot: alle waarden zijn afgerond naar het dichtstbijzijnde veelvoud van € 0,25

5.3 Uitkomsten voor de RR en de VOR

De VOR voor toepassing in de KBA wordt berekend door de RR te vermenigvuldigen met de bijbehorende VOT in euro per persoon per uur. Voor zakelijke reizen is er dan een complicerende factor: de RR's zijn afkomstig van individuele reizigers, maar de zakelijke VOT bestaat uit een werkgevers- en een werknemersdeel. De vraag is of beide moeten worden gebruikt in de berekening van de VOR. Er is geen informatie beschikbaar om een aparte werkgeverscomponent te berekenen voor de zakelijke VOR. Er is voor gekozen om de zakelijke RR toe te passen op de som van de werkgevers- en werknemerscomponenten, dus op de totale zakelijke VOT. De nieuwe VOR's staan in Tabel 8. In Tabel 9 staan de gevonden RR's en worden deze vergeleken met de internationale literatuur.

Tabel 8: Nieuwe waarderingen voor betrouwbaarheid van reistijd (in euro's per persoon per uur voor de standaardafwijking, prijsniveau 2010, inclusief BTW), per vervoerwijze

	Auto	Trein	Bus, tram, metro	Landmodi (auto, trein, BTM)	Vliegtuig	Pleziervaart
Woon-werk	3,75	4,75	3,25	4,00		
<i>Zakelijk werknemer</i>	14,50	18,00	12,00	15,50	56,00	
<i>Zakelijk werkgever</i>	15,50	4,75	9,75	12,25	-	
Zakelijk	30,00	22,75	21,75	27,75	56,00	
Overig	4,75	4,50	3,75	4,50	30,75	0
Alle motieven	5,75	5,50	3,75	5,25	33,75	0

Noot: alle waarden zijn afgerond naar het dichtstbijzijnde veelvoud van € 0,25

Er zijn diverse studies in de internationale literatuur die ook uitkomsten hebben gepresenteerd over de waarde van betrouwbaarheid van reistijd. Een deel hiervan hanteert daarbij een andere operationele definitie van betrouwbaarheid dan de bij de standaardafwijking horende RR, en is daarom niet vergelijkbaar met onze studie. In Tabel 9 staan uitsluitend studies die een RR hebben berekend. In Tabel 9 zijn ook de uitkomsten opgenomen van een expert workshop uit 2004 (zie Hamer et al., 2005) waar voorlopige waarden voor de RR in de KBA in Nederland zijn bepaald, ook al gaat het hier niet om empirische uitkomsten maar om het oordeel van experts. In verschillende KBA's in Nederland zijn de betrouwbaarheidsbaten eenvoudigweg bepaald als 25% van de reistijdbaten (dit is gebaseerd op Besseling et al. (2004)).

Uit Tabel 9 concluderen we dat de nieuwe RR's goed passen binnen het bereik van waarden uit de internationale literatuur. Alle waarden die we nu verkrijgen (behalve de RR voor zakelijk auto) zijn lager dan de voorlopige waarden uit 2005, maar veel recente empirische waarden zijn lager dan die van de workshop. Voor vliegtuig hebben we slechts één andere studie gevonden die ook een RR levert (Noorwegen), en die waarde is duidelijk lager dan de onze (die meer vergelijkbaar is met de RR van de andere vervoerwijzen).

Tabel 9: *Vergelijking van de empirische uitkomsten in de literatuur over de 'reliability ratio' (voor de waarde van de standaardafwijking van reistijd ten opzichte van de waarde van de gemiddelde reistijd)*

Studie	Land	RR
Auto		
MVA (1996)	UK	0,36 - 0,78
Copley et al. (2002)	UK	Pilot onderzoek: 1,3
Hensher (2007)	Australië	0,3 - 0,4
Eliasson (2004)	Zweden	0,30 - 0,95
Mahmassani (2011)	USA	NCHRP 431: 0,80 - 1,10 SHRP 2 CO4: 0,40 - 0,90
<i>Expert workshop in 2004</i>	<i>Nederland</i>	0,8
Huidige studie	Nederland	Woon-werk: 0,4 Zakelijk: 1,1 Overig: 0,6
Trein		
ATOC (2002)	UK	0,6 - 1,5
Ramjerdi et al. (2010)	Noorwegen	Korte trips: 0,69 Lange trips: 0,54
<i>Expert workshop in 2004</i>	<i>Nederland</i>	1,4
Huidige studie	Nederland	Woon-werk: 0,4 Zakelijk: 1,1 Overig: 0,6
Bus/tram/metro		
MVA (2000)	Frankrijk	0,24
Ramjerdi et al. (2010)	Noorwegen	Korte trips: 0,69 Lange trips: 0,54
<i>Expert workshop in 2004</i>	<i>Nederland</i>	1,4
Huidige studie	Nederland	Woon-werk: 0,4 Zakelijk: 1,1 Overig: 0,6
Vliegtuig		
Ramjerdi et al. (2010)	Noorwegen	0,20
Huidige studie	Nederland	Zakelijk: 0,7 Overig: 0,7

6. Conclusies

Er zijn nieuwe waarden bepaald voor reistijd (VOT's) en betrouwbaarheid van reistijd (VOR's) in personenverkeer en -vervoer, voor gebruik in kosten-batenanalyses (KBA's) van transportprojecten in Nederland. De vorige VOT's waren gebaseerd op enquêtes die in 1997 werden uitgevoerd. Tot voor kort moesten de VOR's voor de KBA in Nederland gebaseerd worden op een inschatting van experts uit 2004, maar nu zijn er ook hiervoor waarden uit in Nederland uitgevoerd empirisch onderzoek. Het onderzoek dat in dit artikel is gerapporteerd is de eerste nationale empirische studie in Nederland die zowel de VOT als de VOR bestudeert, binnen één en hetzelfde (dus intern consistent) raamwerk.

Onbetrouwbaarheid is hier geoperationaliseerd als de standaardafwijking van reistijd. De ratio van de waarde van de reistijd ten opzichte van de VOT wordt de 'reliability ratio' (RR) genoemd. De belangrijkste reden om deze definitie te kiezen is dat andere maatstaven van onbetrouwbaarheid vaak aanzienlijk moeilijker zijn op te nemen in de nationale en regionale transportmodellen.

Een belangrijke conclusie uit deze studie is dat het 'stated preference' (SP) onderzoek uit 2009 onder leden van een internet panel beduidend lagere VOT's oplevert dan het SP onderzoek van 2011 waar reizigers tijdens hun reis werden geworven om mee te doen aan het onderzoek. De meest aannemelijke verklaring hiervoor is dat de waarden van 2011 de betere zijn en dat die van 2009 de echte VOT's onderschatten, wat dan weer veroorzaakt zou worden doordat personen met een lagere VOT (binnen ieder socio-economisch segment) een grotere kans zouden hebben om te besluiten lid te worden van een internet panel. Ook van belang (maar minder dan het genoemde zelfselectie-effect van een panel) zijn de kortere afstanden van de verplaatsingen in de steekproef van 2009 t.o.v. 2011 en de daarbij behorende kleinere aangeboden veranderingen in reistijd. We adviseren dan ook om VOT onderzoek voortaan niet uit te voeren met bestaande internet panels, omdat deze waarschijnlijk selectief zijn op het kenmerk dat men wil onderzoeken (de VOT).

De SP data zijn gebruikt voor het schatten van discrete keuzemodellen, waarin de VOT (en deels ook de VOR) verschillen tussen verplaatsingen met verschillende tijd- en kostenniveaus, verschillende in de SP aangeboden veranderingen in tijd en kosten en kernmerken van de respondenten (zoals opleiding, inkomen, leeftijd, huishoudsamenstelling). Door gebruik te maken van een 'panel latent class' model wordt ook rekening gehouden met niet-waargenomen verschillen tussen respondenten en met het feit dat respondenten een herhaalde keuze maken. In de aanbevolen waarden voor de VOT en RR gebruiken we referentiewaarden die uitsluitend gebaseerd zijn op de data uit 2011. De effecten van tijd- en kostenniveaus, tijd- en kostenveranderingen en socio-economische kenmerken zijn echter geschat op de gezamenlijke data uit 2009 en 2011.

De afhankelijkheid van de VOT en de VOR van de basisniveaus van tijd en kosten is expliciet in de geschatte modellen opgenomen. Bij de berekening van de aanbevolen waarden voor KBA hebben we dit onderscheid ingebracht door een wegingsprocedure. Een hardnekkiger probleem vormt de afhankelijkheid voor veranderingen in tijd en kosten, zoals aangeboden in de SP. Ook hier houden de geschatte modellen rekening mee. We vinden daarbij dat de invloed van kostenveranderingen op de VOT groter is dan van tijdsveranderingen. Dit heeft tot gevolg dat de keuze van de veranderingen in tijd en vooral in kosten in het SP ontwerp een invloed heeft op de

uiteindelijk gevonden waarderingen. In veel oudere studies was dit ook het geval maar bleef het impliciet. Evenals ander recent internationaal onderzoek hebben we het probleem expliciet gemaakt, maar een verdere oplossing ervan vereist nader onderzoek (zie ook Daly et al. 2014).

De voor de KBA aanbevolen waarderingen van tijd en betrouwbaarheid zijn bepaald door de geschatte modellen toe te passen op de verzamelde steekproefdata, met herweging om de verdeling van de verplaatsingen naar reistijd uit het OViN zo goed mogelijk te representeren. De nieuwe VOT's verschillen niet erg van de vorige aanbevolen waarden en de internationale literatuur.

Voor de waarde van betrouwbaarheid hebben we de 'reliability ratio' voor verschillende motieven bepaald. Hier kwamen waarden uit variërend tussen 0,4 en 1,1 wat redelijk overeenkomt met recente empirische studies uitgevoerd in andere landen.

VOT's en VOR's zijn bepaald voor auto, trein, bus/tram/metro, vliegtuig en pleziervaart. Voor fietsen en lopen ontbreken zulke waarden nog en is aanvullend onderzoek gewenst (ook internationaal is er nauwelijks onderzoek naar verricht; een uitzondering is Börjesson en Eliasson (2012)).

Voor het meenemen van reistijdverandering en betrouwbaarheidsverandering in de KBA zijn monetaire waarden nodig, zoals in dit onderzoek zijn bepaald, maar ook prognoses van de toekomstige veranderingen in reistijd en betrouwbaarheid, met name voor wat betreft de effecten van transportprojecten. Voor reistijd wordt dit al jaren geleverd door diverse transportmodellen, zoals LMS en NRM. Voor veranderingen in reistijdbetrouwbaarheid op de weg is er nu ook een prognosemodel, dat werkt als een 'post-processing' module op het LMS of NRM (zie de paper van Kouwenhoven en Mulder in dit nummer van het TVW). Maar voor alle andere vervoerwijzen (met name dus voor het OV) ontbreekt een model dat de veranderingen in de mate van onbetrouwbaarheid kan voorspellen. Zolang dit nog ontbreekt kunnen de hier gevonden monetaire waarden voor de betrouwbaarheid in het OV niet worden toegepast in de KBA.

Dankwoord

Dit artikel is gebaseerd op een studie die is uitgevoerd voor het Kennisinstituut voor Mobiliteitsbeleid van het Ministerie van Infrastructuur en Milieu. De auteurs bedanken tevens SURFsara (www.surfsara.nl) voor de ondersteuning bij het gebruik van het Lisa Computer Cluster en Jasper Knockaert van de Vrije Universiteit Amsterdam voor advies bij het gebruik van PythonBiogeme.

Referenties

Association of Train Operating Companies (ATOC) (2002) *Passenger Demand Forecasting Handbook*, London: ATOC.

Bates, J., J. Polak, P. Jones en A. Cook (2001) The valuation of reliability for personal travel, *Transportation Research E (Logistics and Transportation Review)*, 37-2/3, 191-229.

Batley, R.P., S. Grant-Muller, J. Nellthorp, G.C. de Jong, D. Watling, J.J. Bates, S. Hess en J. Polak (2008) *Multimodal Travel Time Variability*, Final Report, report for the UK Department of Transport, ITS Leeds, John Bates and Imperial College.

Besseling, P., W. de Groot en A. Verrips (2004) *Economische toets op de Nota Mobiliteit*, CPB document 65, CPB, Den Haag.

Bierlaire, M. (2008) *An introduction to BIOGEME version 1.7*, biogeme.epfl.ch.

Borger, B. de en M. Fosgerau (2008) The trade-off between money and travel time: A test of the theory of reference-dependent preferences, *Journal of Urban Economics*, 64, 101-115.

Börjesson, M. en J. Eliasson (2012), The value of time and external benefits in bicycle appraisal, *Transportation Research Part A*, 57, 73-683.

Börjesson, M., en J. Eliasson (2014) Experiences from the Swedish Value of Time study, *Transportation Research Part A* 59, 144-158.

Börjesson, M., J. Eliasson en J.P. Franklin (2012) Valuations of travel time variability in scheduling versus mean-variance models. *Transportation Research Part B: Methodological*, 46 (7), 855-873.

Carrion, C. en D. Levinson (2012) Value of travel time reliability: A review of current evidence, *Transportation Research Part A*, 46, 720-741.

Cirillo, C., A.J. Daly en K. Lindveld (2000) Eliminating bias due to the repeated measurements problem, in J. de D. Ortúzar (Ed.): *Stated Preference Modelling Techniques*, PTRC, London.

Copley, G., P. Murphy, en D. Pearce (2002) Understanding and valuing journey time variability; European Transport Conference – 2002, Cambridge.

Daly, A.J., F. Tsang en C. Rohr (2014) The value of small time savings for non-business travel, *Journal of Transport Economics and Policy*, 48-2, 205-218.

Eliasson, J. (2004) Car drivers' valuations of travel time variability, unexpected delays and queue driving, *Proceedings of the European Transport Conference 2004*, Straatsburg.

Fosgerau (2006) Investigating the distribution of the value of travel time savings, *Transportation Research Part B*, 40(8), 688-707.

Fosgerau, M. en A. Karlström (2010) The value of reliability, *Transportation Research Part B*, 44(1), 38-49.

Fosgerau, M. en L. Engelson (2011) The value of travel time variance, *Transportation Research B*, 45(1), 1-8.

Fowkes, A.S., P. Marks en C.A. Nash (1986) The value of business travel time savings, Working Paper 214, Institute for Transport studies, University of Leeds.

Hague Consulting Group (1990) The Netherlands' value of time study: final report, Report for DVK, HCG, Den Haag.

Hague Consulting Group (1998) The second Netherlands' value of time study: final report, Report 6098-1 for AVV, HCG, Den Haag.

Hamer, R., G.C. De Jong, and E.P. Kroes (2005) The value of reliability in Transport – Provisional values for the Netherlands based on expert opinion, RAND Technical Report Series, TR-240-AVV, Leiden.

HEATCO (2006) Developing Harmonised European Approaches for Transport Costing and Project Assessment, Deliverable 5, Proposal for harmonized guidelines. IER, University of Stuttgart.

Hensher, D.A. (1977) Value of business travel time, Pergamon Press.

Hensher, D.A. (2007) Valuation of Travel Time Savings, in: A. de Palma, R. Lindsey, E. Quinet and R. Vickerman (Eds.): Handbook in Transport Economics, Edward Elgar Publisher.

Hensher, D.A., J.M. Rose en Z. Li (2012) Does the choice model method and/or the data matter? *Transportation*, 39(2), 351-385.

Hess, S., M. Ben-Akiva, D. Gopinath en J. Walker (2011) Advantages of latent class over continuous mixture of logit models, Working paper, ITS, University of Leeds.

Jong, G.C. de, E.P. Kroes, R. Plasmeijer, P. Sanders en P. Warffemius (2004) The value of reliability, *Proceedings of the European Transport Conference 2004*, Straatsburg.

Jong, G.C. de, M. Kouwenhoven, E.P. Kroes, P. Rietveld en P. Warffemius (2009) Preliminary monetary values for the reliability of travel times in freight transport, in *European Journal of Transport and Infrastructure Research*, 9(2), 83-99.

Jong, G.C. de, M. Kouwenhoven, J. Bates, P. Koster, E. Verhoef, L. Tavasszy en P. Warffemius (2014) New SP-values of time and reliability for freight transport in the Netherlands, *Transportation Research Part E*, 64, 71-87.

Kaa, E.J. van de (2008) Extended prospect theory, Findings on choice behaviour from economics and the behavioural sciences and their relevance for travel behaviour, PhD Thesis, Delft University of Technology.

Kahneman, D. en A. Tversky (1979) Prospect theory: an analysis of decision under risk, *Econometrica*, 47, 263-291.

Kahneman, D. en A. Tversky (1992) Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty*, 5(4), 297-323.

Kennisinstituut voor Mobiliteitsbeleid KiM (2013) De maatschappelijke waarde van kortere en betrouwbaardere reistijden, Ministerie van Infrastructuur en Milieu, Den Haag. (tweede druk)

Mahmassani, H.S. (2011) Application to the New York metropolitan region of an integrated model of user responses to pricing and reliability with a state of the art simulation-based dynamic traffic assignment tool, Presentation to NYMTC, New York City.

Ministerie van Infrastructuur en Milieu (2012) Structuurvisie infrastructuur en ruimte: Nederland concurrerend, bereikbaar, leefbaar en veilig, Ministerie van Infrastructuur en Milieu, Den Haag.

MVA (1996) Benefits of reduced travel time variability; report to DfT; MVA.

MVA (2000) Etude de l'impact des phénomènes d'irregularité des autobus – Analyse des resultats, MVA, Paris.

Peer, S. (2013) The economics of trip scheduling, travel time variability and traffic information. PhD Thesis, VU University Amsterdam.

Ramjerdi, F., S. Flügel, H. Samstad en M. Killi (2010) Value of time, safety and environment in passenger transport – Time. TØI report 1053B/2010, Institute of Transport Economics, Oslo.

RAND Europe (2004) De Waardering van kwaliteit en betrouwbaarheid in personen- en goederen vervoer (The valuation of quality and reliability in passenger and freight transport). AVV/RAND Europe, Rotterdam.

Significance, VU University Amsterdam en John Bates (2007) The value of travel time and travel time reliability, Survey design, Final Report prepared for the Netherlands Ministry of Transport, Public Works and Water Management, Significance, Leiden.

Significance, Goudappel Coffeng en NEA (2012) Erfassung des Indikators Zuverlässigkeit des Verkehrsablaufs im Bewertungsverfahren der Bundesverkehrswegeplanung: Schlussbericht, Report for BMVBS, Significance, The Hague (zie: <http://www.bmvbs.de/SharedDocs/DE/Artikel/UI/bundesverkehrswegeplan-2015-methodische-weiterentwicklung-und-forschungsvorhaben.html>).

Significance, VU University, John Bates Services, TNO, NEA, TNS NIPO en PanelClix (2013) Values of time and reliability in passenger and freight transport in The Netherlands, Report for the Ministry of Infrastructure and the Environment, Significance, Den Haag. (zie: http://www.kimnet.nl/sites/kimnet.nl/files/filemanager/bijlagen/Bijlage_Value_of_time_and_reliability_in_passenger_and_freight_transport_in_the_Netherlands_reprint.pdf)

Small, K.A. (1982) The Scheduling of Consumer Activities: Work Trips, *American Economic Review*, 72, 467-479.

Stathopoulos, A.I. en S. Hess (2011) Revisiting reference point formation, gains-losses asymmetry and non-linear sensitivities: one size does not fit all! Paper gepresenteerd op ETC 2011, Glasgow.

Tseng, Y.Y. en E.T. Verhoef (2008) Value of time by time of day: A stated-preference study, *Transportation Research B*, 42(7-8), 607-618.

Tseng, Y.Y., E.T. Verhoef, G.C. de Jong, M. Kouwenhoven en A.I.J.M. van der Hoorn (2009) A pilot study into the perception of unreliability of travel times using in-depth interviews, *Journal of Choice Modelling*, 2(1), 8-28.

Vickrey, W.S. (1973) Pricing, metering, and efficiently using urban transport facilities, *Highway Research Record* 476, 36-48.

Vickrey, W.S. (1969) Congestion theory and transport investment, *American Economic Review (Papers and Proceedings)* 59, 251-261.

Wardman, M., R. Batley, J. Laird, P. Mackie, A.S. Fowkes, G. Lyons, J. Bates en J. Eliasson (2013) Valuation of travel time savings for business travellers, Report prepared for the Department for Transport, Institute for Transport studies, University of Leeds.