



September 2013

Spitsmijden in de trein: Gedragseffecten

Jasper Knockaert
Stefanie Peer
Erik Verhoef
Afdeling Ruimtelijke Economie
Vrije Universiteit Amsterdam

Inhoudsopgave

1	Inleiding.....	2
2	De praktijkproef.....	4
3	Werving.....	6
4	Deelname aan Spitsmijden.....	8
4.1	Vergelijking deelnemers met niet-deelnemers.....	8
4.2	Deelname aan de praktijkproef.....	14
5	Nadere kenmerken deelnemers.....	15
6	Reisgedrag.....	18
6.1	Inleiding.....	18
6.2	Meten Reisgedrag.....	19
6.3	Geobserveerd reisgedrag.....	23
6.4	Discrete keuzemodellen: theorie.....	30
6.5	Verklarende variabelen.....	32
	Reistijd.....	32
	Afwijking van optimale vertrek- en aankomsttijd.....	32
	Reiscomfort.....	32
	Onbetrouwbaarheid.....	32
6.6	Revealed preference (RP) observaties.....	32
6.7	Stated preference (SP) observaties.....	33
6.8	Resultaten keuzemodellen.....	34
7	Evaluatie.....	38
7.1	Inleiding.....	38
7.2	Gedrag tijdens en na afloop van de proef.....	38
7.3	Interesse dalabonnement.....	39
7.4	Projectevaluatie.....	40
7.5	Afgehaakte deelnemers.....	41
8	Conclusies.....	42
9	Referenties.....	45

1 Inleiding

Piekbelasting is een fenomeen dat in vele markten voorkomt. Het ontstaat wanneer de vraag fluctueert over de tijd, terwijl capaciteitsbeperkingen zodanig zijn dat het aanbod niet in dezelfde proportie mee kan bewegen. Het treedt bijvoorbeeld op in seizoensgebonden toeristische markten, op telecomnetwerken, en ook in verkeer en vervoer. Terwijl files en congestie op de weg misschien het eerste voorbeeld vormen waaraan dan gedacht wordt, kent natuurlijk ook het spoorvervoer zijn pieken, zowel over de weken als binnen de (werk-)dag. Hoe sterker de vraag fluctueert, hoe vaker het zo zal zijn dat capaciteitsuitbreidingen maar voor een gedeelte van de tijd rendabel zijn, om de simpele reden dat een deel van de capaciteit buiten de piek ongebruikt zal blijven. Het wordt dan aantrekkelijk, zowel vanuit bedrijfseconomisch als maatschappelijk opzicht, om de vraag meer te spreiden, zodat met een lagere capaciteit kan worden volstaan en een kleiner gedeelte van de capaciteit buiten die piek ongebruikt blijft. Uiteindelijk is dit niet alleen goed voor de aanbieder, die daarmee kosten bespaart: een betere spreiding van reizigers over de tijd van de dag kan ter verbetering van het gepercipieerde reiscomfort en verbeteren van de reisbeleving en kan een goedkoper alternatief zijn voor een uitbreiding van de vervoerscapaciteit via bijvoorbeeld hogere frequenties, langere treinen, meer sporen. Het kan naast de impact op reisbeleving ook beter zijn voor de klant als we bedenken dat uiteindelijk tenminste een deel van de kosten door diezelfde klant zal moeten worden opgebracht ook bij een gesubsidieerde dienstverlening als spoorvervoer.

In dit onderzoek staat de vraag centraal of, en in welke mate, treinreizigers door financiële prikkels verleid kunnen worden om spitsreizen naar daluren te verplaatsen. Het onderzoek is onderdeel van het project *Spitsmijden in de Trein*; een project waarin ondermeer spoorvervoerders, consultants, onderzoeksinstellingen en regionale concessie verlenende overheden in opdracht van het Ministerie van I&M samenwerkten. Daar waar een parallel lopend onderzoek van TNO (Vonk Noordegraaf, de Kruijff en Vonk, 2013) zich vooral gericht heeft op de rol van werkgevers bij het vermijden van de spoorspits door hun werknemers, richt het deel van het onderzoek dat in dit rapport besproken wordt zich op de gedragsreacties van individuele deelnemers aan de proef, en op de beslissing van individuen om al dan niet deel te nemen aan een dergelijke proef. Het project *Spitsmijden in de Trein* maakt onderdeel uit van het project Actieplan Groei op het Spoor die de doel heeft meer reizigerskilometer te generen met de verschillende maatregelen, onder ander door een betere spreiding van reizigers over de tijd.

Eén van de bijzondere aspecten van de proef is dat we daadwerkelijke gedragsreacties hebben kunnen observeren van kaarthouders die een prikkel kregen om buiten de spitsen te reizen (6:30–9:00 en 16:00–18:30). Daartoe werd tussen augustus 2012 en mei 2013 een praktijkproef gehouden waaraan meer dan 1000 (abonnement-)kaarthouders deelnamen. Deze frequente treinreizigers kregen een geldelijke beloning voor het mijden van ochtend- en avondspits. Om het reisgedrag van de deelnemers vast te leggen, werd een innovatieve meetmethode ingezet, die gebruik maakte van een smartphone-app.

Het onderzoek naar *deelnamegedrag* betreft in de eerste plaats een vergelijking van de kenmerken van deelnemers en niet-deelnemers aan de praktijkproef. Specifiek wordt er bestudeerd of er een verband is tussen individu-specifieke eigenschappen en de beslissing om deel te nemen aan de proef. Daarnaast wordt nader ingegaan op de omstandigheden waaronder niet-deelnemers wel

zouden hebben willen deelnemen aan de proef, en onder welke omstandigheden deelnemers van hun deelname zouden hebben afgezien.

In de analyse van het *reisgedrag* staat de vraag centraal hoe deelnemers een vertrektijd kiezen voor hun treinreizen naar en van de werklocatie; dus, de heen- en terugreis op een typische werkdag. Eerst bestuderen we hoe de aandelen van spits- en dalreizen in het reisgedrag wijzigen tijdens de beloningsperiode, in vergelijking met de periodes voorafgaand aan en volgend op die beloningsperiode (de zogeheten voor- en nameting). Om dit reisgedrag te meten werd de genoemde smartphone-app ontwikkeld. Deze registreerde op welke momenten de telefoon bij welk treinstation gesignaleerd werd. Uit deze gegevens kunnen we afleiden in welke mate deelnemers hun vertrektijd hebben aangepast om een beloning te kunnen verdienen. In dit rapport bekijken we niet alleen of de gemiddelde deelnemer meer buiten de spits is gaan reizen, maar ook of er een (statistisch significant) verband is tussen het gemeten effect en de dagen van de week, de regio, of de hoogte van de beloning.

Vervolgens gebruiken we de observaties om gedragsmodellen te schatten die de geobserveerde tijdstipkeuzes proberen te verklaren. De daarbij geschatte coëfficiënten geven inzicht in het belang dat reizigers toekennen aan bepaalde kenmerken van de verplaatsing. Hierbij gaat het onder meer om de reisduur; het tijdstip van reizen en dan met name de mate waarin deze afwijkt van wat voor de reiziger het meest gewenste tijdstip zou zijn; de betrouwbaarheid van de verbinding (uitgedrukt als de kans op een vertraging); de bezettingsgraad van de coupés; en de prijs voor de reis, of beter gezegd: de geldelijke beloning die ontvangen wordt. Doordat we die beloning ook meenemen, kunnen we de waardering voor de andere factoren uitdrukken in een geldelijke waarde: hoeveel heeft een reiziger er voor over om dat kenmerk te verbeteren? Dergelijke waarderingen zijn een onmisbaar onderdeel voor het rationaliseren van investeringsbeslissingen in de kwaliteit van het spoorvervoer, of dat nu door een privaat bedrijf wordt gedaan dat de investering door hogere ticketprijzen en/of een toegenomen vraag zou willen terugverdienen, of door een overheid – in elk geval wanneer deze Euro's zodanig zou willen inzetten dat ze een maatschappelijke bate opleveren die hoger is dan de kosten die gemaakt worden.

De gedragsmodellen die we in dit rapport zullen presenteren zijn zogeheten discrete keuzemodellen. Deze verklaren geobserveerde tijdstipkeuzes uit objectief waarneembare factoren, maar erkennen daarnaast ook dat een deel van de overwegingen van reizigers natuurlijk niet door de onderzoeker geobserveerd kunnen worden. Voor deelnemers aan de proef kijken we daarbij zowel naar de keuzes die in het werkelijke reisgedrag zijn gemaakt (dus, gebaseerd op registraties uit de app), als naar keuzes die volgens die deelnemers in fictieve situaties zouden worden gemaakt. Deze fictieve keuzes zijn gebaseerd op enquêteresultaten. Doordat we dezelfde fictieve keuzes ook aan niet-deelnemers hebben kunnen voorleggen, kunnen we een goed beeld krijgen van de verschillen tussen die twee groepen reizigers, en daarmee voor de mate van representativiteit van de deelnemers voor de gemiddelde (kaarthoudende) spitsreiziger in de trein.

Tenslotte willen we ook inzicht bieden in de geschiktheid van een smartphone-app als instrument om reisgedrag in het OV te kunnen meten. Het gebruik van een smartphone-app in deze context was namelijk, voor zover wij weten, een primeur. Eveneens evalueren we de mogelijkheid om logboekregistraties te gebruiken om de verwachte drukte (reiscomfort) en vertraging (betrouwbaarheid) voor de verschillende aangeboden treinverbindingen te bepalen.

Het project heeft door zijn aard een enorme schat aan onderzoeksgegevens opgeleverd die zich lang niet allemaal tot in alle detail laten analyseren in het tijdsbestek tussen het beëindigen van de nametingen en het verschijnen van dit rapport. Wat hier gerapporteerd wordt zijn derhalve eerste (maar niet voorlopige) onderzoeksresultaten, die naar verwachting gevolgd zullen worden door meer gesofisticeerde analyses die vooral op wetenschappelijke tijdschriften gericht zullen zijn. De selectie van thema's in dit rapport is gemaakt in overleg met de projectpartners en de stuurgroep, en reflecteert onze inschatting van de interesses van de meeste lezers.

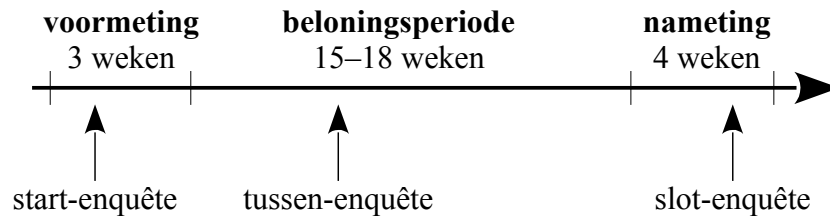
Ook al is het onderzoek in veel opzichten vernieuwend, het is niet het eerste project naar spitsmijden in het openbaar vervoer. Het onderzoek bouwt met name voort op inzichten verzameld in het eerdere *Spitsmijden in het OV* onderzoeksproject (Spitsmijden, 2009). Daaruit bleek onder meer dat een Dalurentrajectabonnement dat alleen geldig is buiten de ochtendspits (destijds gedefinieerd als 7:30–9:00), waarbij voor reizen tijdens de spits een toeslag zou moeten worden gekocht, door 80% van de respondenten minder aantrekkelijk dan het gewone Jaar- of Maandtrajectabonnement wordt gevonden. Daarnaast konden in dat project 124 deelnemers enige tijd met een 20% goedkoper dalurentrajectabonnement reizen (die korting vormde feitelijk de beloning), waarbij ze moesten bijbetalen voor een verplaatsing in de spits (dit vormde de korting op die beloning). Omdat veel van deze deelnemers ook zonder beloning al buiten de spits reisden, en daarnaast de spitsperiode relatief laat gedefinieerd was (startend om 7:30), leverde dit niet heel veel extra spitsmijdingen op. Vergeleken met de nieuwe praktijkproef die we in dit rapport presenteren, was deze eerdere praktijkproef redelijk beperkt in omvang, zowel wat betreft het deelnametraject als het aantal deelnemers, en werd er ook een eenvoudige meetmethode gebruikt (alleen logboeken; hier gebruiken we logboeken én de smartphone-app).

Dit rapport is als volgt opgebouwd. We bespreken eerst de opzet van het experiment (hoofdstuk 2), de werving van deelnemers (hoofdstuk 3), de kenmerken van de deelnemers, en hoe deze verschillen van niet-deelnemers (hoofdstukken 4 en 5). Vervolgens bestuderen we in hoofdstuk 6 het reisgedrag, hoe dit gedrag wordt aangepast om een beloning te verdienen, en hoe treinreizigers de diverse kenmerken van hun treinreis waarderen. Deze kwantitatieve waardering wordt bepaald zowel voor werkelijk reisgedrag (voor deelnemers) als voor fictieve keuzesituaties (voor deelnemers en niet-deelnemers). Hoofdstuk 7 bespreekt de belangrijkste punten die uit de evaluatie van de praktijkproef door de deelnemers naar voren kwamen. Dit betreft niet alleen een beoordeling van het ontwerp en de uitvoering van de proef, maar deelnemers delen ook hun ervaringen met het reizen tijdens de beloningsperiode en hun verwachte reisgedrag na afloop van het experiment. Er wordt daarnaast een analyse gegeven van een eenvoudig onderzoek naar de vraag onder welke (prijs-)voorwaarden deelnemers een dalabonnement zouden kopen, en we gaan in op de redenen waarom een relatief groot aantal deelnemers nog voor afloop van de proef hun deelname heeft stopgezet. Tot slot presenteren we in hoofdstuk 8 conclusies en aanbevelingen.

2 De praktijkproef

De praktijkproef is opgezet om te meten of en hoe treinreizigers hun vertrek- en aankomsttijd aanpassen als ze een geldelijke beloning ontvangen voor reizen buiten de spits. Deelname aan de praktijkproef duurt 22 tot 25 weken, bestaande uit een voormeting van 3 weken, gevolgd door een beloningsperiode van 15 tot 18 weken, en tenslotte een nameting van 4 weken. De voor- en nameting zijn nodig om gedragsveranderingen tijdens de beloningsperiode te kunnen kwantificeren

ten opzichte van een referentieperiode waarin de deelnemer zonder beloning reist.



Figuur 1: verloop deelname praktijkproef

De deelnemer installeert op zijn of haar smartphone een speciaal voor dit project ontwikkelde app om de tijdens de praktijkproef gemaakte treinreizen te kunnen laten registreren.

Bij aanmelding wordt voor elke deelnemer het deelnametraject vastgesteld. De deelnemer reist tijdens de praktijkproef in beide richtingen van dit traject tussen het beginstation, bij de woonlocatie, en het eindstation, bij de werklocatie. In principe komt dit traject overeen met het abonnement dat de deelnemer al heeft. Deelnemers met een abonnement zonder vast traject, (bijvoorbeeld OV-jaarkaart) geven het deelnametraject bij aanmelding op.

Tijdens de beloningsperiode kan de deelnemer een beloning verdienen door voor de reis op het deelnametraject een vertrektijd te kiezen buiten de ochtendspits (06:30–09:00) en avondspits (16:00–18:30).¹ Dit reisgedrag wordt in het onderzoek vergeleken met het reisgedrag tijdens de voor- en nametingen.

Om het effect van de hoogte van de beloningsprikkel te kunnen onderzoeken, wordt voor elke deelnemer halverwege de beloningsperiode die hoogte aangepast, waarbij de volgorde hoog-laag en laag-hoog willekeurig verdeeld was over deelnemers. De beloning hangt daarbij af van de lengte van het deelnametraject van de reiziger (in tariefeenheden). Alhoewel het vaak de werkgever is die het abonnement vergoedt, is het steeds de reiziger die de beloning ontvangt.

Tabel 1: Beloning per dalreis

Reisafstand (tariefeenheden)	Hoge beloning/dalreis	Lage beloning/dalreis
≤ 25 km	€ 2,50	€ 1,50
25–40 km	€ 3,50	€ 2,50
≥ 40 km	€ 4,50	€ 2,50

De deelnemer krijgt de beloning alleen voor het maken van een (door de app geregistreerde) treinreis op het deelnametraject met een vertrektijd buiten de spitsuren. Dat betekent dat de deelnemer geen beloning krijgt voor thuiswerken of het gebruik van andere vervoerswijzen. Op die manier voorkomen we dat er een prikkel ontstaat om een treinreis niet te laten registreren door de app, bijvoorbeeld door de telefoon thuis te laten of uit te zetten.

Aanvullend bij de reisregistraties uit de app was het voor de deelnemer verplicht om drie enquêtes in te vullen, op basis waarvan extra informatie verzameld wordt:

1. Start-enquête. Hierin gaat het onder meer om de socio-economische kenmerken;

¹ De beloning wordt altijd door de daadwerkelijke vertrektijd van een trein bepaald, niet de volgens de dienstregeling geplande.

randvoorwaarden bij de keuze van een vertrektijd vanuit zowel deelnemer- als werkgeverperspectief; het gebruikelijke reisgedrag in het woon-werkverkeer; en de voor de deelnemer optimale vertrek- en aankomsttijd.

2. Tussen-enquête. Deze bevat het hypothetisch keuze-experiment; de kenmerken werkgever; en de motivatie voor deelname
3. Slot-enquête. Hier wordt gevraagd naar de evaluatie van het experiment en de ervaringen met het eigen reisgedrag, en naar verwachtingen over toekomstig reisgedrag

Tenslotte vulden de deelnemers gedurende zes kalenderweken een logboek in: gedurende 1 week in de voormeting, 4 weken tijdens de beloningsperiode, en 1 week in de nameting. Hiermee wordt aanvullende informatie over hun reisgedrag geregistreerd die zich niet door de app laat meten. Voor elke weekdag (maandag t/m vrijdag) vulden deelnemers in of ze op die dag naar werk (of school) hebben gereisd en voor welke vervoerswijze zij hebben gekozen. Voor een treinreis moest de deelnemer verder aangeven of de trein vertraagd was, hoe druk het was in de trein, en hoe de beschikbaarheid van zitplaatsen was. Ook werd de deelnemer verzocht om met een rapportcijfer de algemene tevredenheid over treinreizen in de betreffende week uit te drukken. Naast de registratie van individueel reisgedrag heeft het logboek ook een rol als doorlopende monitoring van de verschillende deelnametrajecten: doordat elke kalenderweek een zekere groep deelnemers het logboek invult, hebben we doorlopende tijdreeksen met informatie over de verschillende trajecten.

De beloningen werden in 2 tranches uitbetaald: een eerste maal ongeveer halverwege de beloningsperiode, en een tweede maal na afloop van de volledige beloningsperiode. Om de opgebouwde beloningen daadwerkelijk te ontvangen moesten op de uitbetalingsdatum alle reeds opengestelde enquêtes en logboeken ingevuld zijn.

3 Werving

Om deelnemers te werven voor de praktijkproef werd gebruik gemaakt van het klantenbestand van NS.² In twee wervingsrondes ontvingen meer dan 80.000 frequente reizigers een persoonlijke uitnodiging via email. Tijdens de tweede ronde werd de werving ondersteund door posters op de drukste stations in Nederland en door een twitterbericht op het officiële NS-kanaal.

2 De NS geeft vervoerbewijzen uit die geldig zijn voor alle treinen op het Nederlandse spoorwegnet. Met de door NS uitgegeven vervoerbewijzen kunnen de reizigers ook gebruik maken van treindiensten aangeboden door andere vervoerders.

Tabel 2: Overzicht wervingsrondes

	Wervingsronde 1	Wervingsronde 2	Totaal
Doelgroep	Jaartrajectkaarthouders (B2C, B2B)	Jaartrajectkaarthouders, houders van NS-jaarkaart/Altijd Vrij of OV-jaarkaart (alleen B2B)	
Begin voormeting	30 juli 2012	5 november 2012	
Aantal uitgenodigd (persoonlijk)	33687	52378	
Spontane aanmelding mogelijk	Nee	Ja	
Aantal deelnemers met minimaal 1 app-geregistreerde treinreis (% uitgenodigd)	409 (1,2%)	761 (1,5%)	
Aantal deelnemers die deelname hebben afgemaakt, inclusief het invullen van alle enquêtes en logboeken (% van deelnemers met min 1 appregistratie)	169 (41%)	327 (43%)	
Aantal deelnemers die de start-enquête en de tussen-enquête hebben ingevuld			649
Aantal deelnemers met minimum 5 app-observaties in de voormeting en geen te grote verandering van de wekelijkse reisfrequentie			479
Aantal deelnemers uit bovenstaande groep (479 deelnemers) die ook minimum 5 app-observaties in de nameting hebben			238

De respons van 1,2% in de eerste wervingsronde, en 1,5% in de tweede wervingsronde, is laag in vergelijking met andere Spitsmijdenprojecten, onder automobilisten, waar wervingspercentages van 10% tot 40% gerealiseerd worden. Mogelijke redenen voor de lagere respons zijn de wervingsmethode (email in plaats van postaal) of zorgen op het vlak privacy (meting met smartphone-app in plaats van wegekantsysteem).

Omdat deelname niet toevallig is – reizigers die makkelijker een beloning verwachten te kunnen verdienen, hebben een grotere prikkel zich aan te melden, en voor deelnemers met een langer deelnametraject is de beloning groter – mogen resultaten voor de deelnemers zeker niet gezien worden als representatief voor de gemiddelde spitsreiziger. In de volgende sectie gaan we daarom specifiek in op de verschillen tussen deelnemers en niet-deelnemers.

Noodzakelijk voor deelname aan de praktijkproef is het bezit van een smartphone met iOS of Android, waarop de app geïnstalleerd en geactiveerd wordt: alleen automatisch geregistreerde treinreizen buiten de spits worden beloond. Een correcte en volledige reisregistratie levert de deelnemer op die manier voordeel op, hetgeen de betrouwbaarheid van de verzamelde reisregistraties ten goede komt. Als gevolg van de experimentele opzet is er geen financiële prikkel om de app of telefoon uit te houden bij reizen tijdens de spits. We zullen later zien dat er wel een prikkel is vanwege het hoge stroomgebruik van de app (hoofdstuk 7.5).

Voor de analyses die volgen, gebruiken we soms andere selecties dan de 496 die de proef hebben afgemaakt. Deze geven we in de onderste drie rijen in de tabel 2. Dit betreffen totalen van 649 die de start-enquête en de tussen-enquête hebben ingevuld, en twee totalen die alleen zijn gedefinieerd op basis van de kwaliteit van app-metingen. Dit zijn 479 deelnemers die minimaal 5 app-observaties in voormeting hadden, en geen te grote verandering van het totale aantal reizen per week (tussen 10de en 90ste percentiel); en 238 uit de laatstgenoemde groep die ook nog eens minimaal 5 app-waarnemingen in de nameting hadden. De motivatie achter deze keuzes wordt in paragraaf 6.3 nader toegelicht.

4 Deelname aan Spitsmijden

Om de beslissing over deelname aan Spitsmijden nader te onderzoeken werd onder meer een non-respons enquête uitgezet, onder niet-deelnemers. Dit maakt het mogelijk beide groepen nader te vergelijken, en zo niet alleen een indruk te krijgen van de factoren die deelname beïnvloeden, maar ook van de representativiteit van onze deelnemerpopulatie. Hierbij gaat het om representativiteit voor de groep die is uitgenodigd (abonementhouders), juist omdat we willen zien welke factoren de kans op deelname beïnvloeden gegeven dat men uitgenodigd is. Dat abonementhouders weer verschillen van de gemiddelde treinreiziger is op zichzelf geen probleem, maar men dient goed te bedenken dat de gedragsreacties die we zullen rapporteren betrekking hebben op een niet-toevallige selectie van uitgenodigden, en dat de uitgenodigden weer een niet-toevallige selectie van treinreizigers vormen.

In totaal werden op basis van de oorspronkelijke verzendlijsten 3697 niet-deelnemers via email uitgenodigd. Dit leverde 489 volledig ingevulde enquêtes op. De respondenten ontvingen een beloning (een Bol.com-bon ter waarde van 25 Euro) voor het invullen van deze enquête, wat tot een goede respons leidde. We vergelijken de resultaten uit deze een non-respons enquête met de door de deelnemers ingevulde enquête.

4.1 *Vergelijking deelnemers met niet-deelnemers*

Allereerst vergelijken we in deze sectie de deelnemers met de niet-deelnemers op basis van een groot aantal persoonlijke kenmerken.

Tabel 3 toont deze vergelijking. Voor socio-economische kenmerken zoals leeftijd, aantal personen in het huishouden en inkomen blijken deelnemers en niet-deelnemers erg vergelijkbaar te zijn, al zijn er ook enkele verschillen, zoals de provincie van het beginstation van het deelnametraject. Deelnemers hebben hun beginstation vaker in de Randstad-provincies Noord-Holland, ZuidHolland en Utrecht in vergelijking met de niet-deelnemers³, die vaker uit de provincies buiten de Randstad komen.

3 Voor niet-deelnemers is het woon-werktraject vastgesteld op basis van antwoorden uit de non-respons enquête.

Tabel 3: vergelijking deelnemers met niet-deelnemers

	Deelnemers	Niet-Deelnemers
Aantal	649	489
Gemiddelde Leeftijd	39,74 jaar	41,68 jaar
Gemiddeld aantal personen in huishouden	2,67 personen	2,66 personen
<i>Inkomen</i>		
minder dan € 1500	1,2%	0,82%
€ 1500–2000	9,1%	7,8%
€ 1500–2500	8,8%	11%
€ 2500–3500	21%	21%
€ 3500–5000	24%	23%
meer dan € 5000	6,6%	5,7%
onbekend	28%	30%
<i>Opleiding</i>		
HAVO of lager	19%	27%
HBS, VWO, Atheneum, Gymnasium	2,8%	3,7%
Hoger beroepsonderwijs	35%	42%
Wetenschappelijke onderwijs	43%	25%
overig	1,1%	1,6%
<i>Type werk</i>		
Manager	13%	10%
Onderzoeker, ingenieur, docent, (vak)specialist	54%	35%
Administratief personeel	18%	39%
Dienstverlenend personeel en verkopers	12%	15%
overig	2,9%	0,61%
<i>Sector</i>		
Financiële instellingen	28%	57%
Openbaar bestuur, overheidsdiensten, verplichte sociale verzekeringen	23%	8,0%
Onderwijs	9,7%	1,4%
overig	40%	34%
<i>Provincie (beginstation van het deelnametraject)</i>		
Noord-Holland	27%	23%
Zuid-Holland	21%	16%
Utrecht	15%	12%
Noord-Brabant	9,9%	12%
Gelderland	9,9%	9,2%
overig	17%	28%

De tabel toont ook dat deelnemers gemiddeld hoger opgeleid zijn dan niet-deelnemers, en dat ze vaker als onderzoeker, ingenieur, docent of (vak)specialist werken. Deelnemers werken daarnaast vaker bij de overheid en in het onderwijs, en minder vaak in de financiële sector. Deelnemers hebben naar alle waarschijnlijkheid mede daardoor vaker een baan met meer flexibiliteit en kunnen daarom gemakkelijker de spits mijden. Dit is precies de vorm van niet-representativiteit die we zouden verwachten, en geeft aan dat gedragsaanpassingen van onze deelnemers waarschijnlijk sterker zijn dan wat niet-deelnemers gedaan zouden hebben.

Dit beeld wordt bevestigd in tabel 4. Deze geeft aan in welke mate de werkomstandigheden flexibele werktijden mogelijk maken. Voor alle gemeten indicatoren blijken deelnemers een grotere mate van flexibiliteit te hebben. De tabel toont eveneens aan dat deelnemers ook in hun persoonlijke levenssfeer (dat wil zeggen: thuis) flexibeler zijn. De factoren die deze flexibiliteit beperken worden hier niet in tabelvorm gerapporteerd, maar bleken redelijk consistent tussen deelnemers en niet-deelnemers. In beide groepen is de zorg voor kinderen de meest belangrijke beperking in de ochtend (als er een beperking is), en activiteiten in de vrije tijd zijn het belangrijkste in de avond. Voor de ochtendspits hebben we deze cijfers kunnen vergelijken met een eerder Spitsmijdenproject op de A12, en de orde van grootte van percentages, alsmede de verschillen tussen deelnemers en niet-deelnemers, komen – hoewel niet identiek – redelijk overeen.

Tabel 4: Flexibiliteit deelnemers vs niet-deelnemers

	Deelnemers	Niet-deelnemers
<i>Flexibiliteit werkgever</i>		
Kan direct beginnen werken bij vroegere aankomst	85%	83%
Kan later beginnen werken dan gebruikelijk	82%	72%
Kan later eindigen dan gebruikelijk	92%	88%
<i>Flexibiliteit werknemer</i>		
Kan eerder dan gebruikelijk van huis vertrekken	80%	70%
Kan later dan gebruikelijk van huis vertrekken	89%	71%
Kan eerder dan gebruikelijk van het werk vertrekken	82%	61%
Kan later dan gebruikelijk van het werk vertrekken	88%	72%

In tabel 5 kijken we naar verschillen in de gebruikelijke en de optimale vertrek- en aankomsttijden (met de trein) voor deelnemers en niet-deelnemers. Zoals verwacht hebben deelnemers vaker dan niet-deelnemers een *gebruikelijke* vertrektijd voor en na de ochtend- en avondspits. Dit maakt het voor hen gemakkelijker om een vertrektijd te kiezen waarbij ze buiten de spits reizen, en daarmee wordt het dus ook makkelijker om de beloning voor dalreizen te verdienen. Verder hebben niet-deelnemers ook vaker dan deelnemers een (voor hen) *optimale* vertrektijd in de spits. Dit betekent dat niet-deelnemers een grotere voorkeur hebben om tijdens de spits te reizen. Ook dit is weer consistent met wat verwacht mag worden bij vrijwillige deelname, en vormt een andere reden waarom onze deelnemers niet representatief voor de gemiddelde reiziger zullen zijn.

Het is opvallend dat zowel deelnemers als niet-deelnemers vaker na de ochtendspits zouden willen vertrekken dan wat zij voorafgaand aan deelname gebruikelijk doen. De avondspits toont een soortgelijk beeld: gemiddeld willen deelnemers en niet-deelnemers vaker buiten (nu voor of na) de

spits reizen dan zij daadwerkelijk doen. Een mogelijke verklaring is dat deze reizigers binnen de spits reizen omdat hun werkgevers de ermee verbonden aankomsttijden op de werklocatie wenselijk vinden, terwijl dit niet aansluit bij de eigen voorkeuren van de reiziger.

Tabel 5: Gebruikelijke en optimale vertrek-/aankomsttijd voor deelnemers en niet-deelnemers

	Meest voorkomende trein vertrektijd		Meest voorkomende trein aankomsttijd		Optimale vertrektijd met de trein	
	Deelnemers	Niet-deelnemers	Deelnemers	Niet-deelnemers	Deelnemers	Niet-deelnemers
<i>Ochtend</i>						
voor 6:30	19%	15%	3,7%	1,0%	9,6%	4,5%
6:30–7:30	27%	43%	20%	26%	21%	37%
7:30–8:00	15%	21%	13%	25%	14%	25%
8:00–9:00	32%	19%	41%	41%	32%	30%
na 9:00	6,6%	1,2%	23%	7,4%	22%	3,9%
<i>Avond</i>						
voor 16:00	13%	5,3%	5,6%	2,5%	21%	9,2%
16:00–17:00	25%	37%	13%	13%	21%	38%
17:00–17:30	22%	28%	13%	23%	21%	27%
17:30–18:30	32%	28%	41%	50%	27%	23%
na 18:30	8,9%	1,8%	27%	12%	10%	2,5%

Tabel 6 toont de verschillen tussen deelnemers en niet-deelnemers voor wat betreft hun abonnement en reisgedrag. We zien dat onder deelnemers een veel hoger aandeel een abonnement voor de eerste klas heeft dan niet-deelnemers. Het wekelijkse aantal treinreizen (voorafgaand aan deelname) is iets hoger voor deelnemers dan voor niet-deelnemers, terwijl het wekelijkse aantal spitsreizen (voorafgaand aan deelname) niet erg verschilt. Dit betekent natuurlijk ook dat het aantal dalreizen iets hoger is voor deelnemers – wederom een verwacht resultaat, omdat deelnemers met een hoger aantal dalreizen voorafgaand aan de proef hun gedrag minder hoeven aan te passen dan iemand met weinig dalreizen om dezelfde beloning te verdienen.

Tabel 6: Vergelijking reisklas en afstandsklasse tussen deelnemers en niet-deelnemers

	Deelnemers	Niet-Deelnemers
Abonnement voor eerste klas	14%	5,9%
Gebruik abonnement om met de trein naar het werk te reizen	4,34 dagen/week	4,13 dagen/week
Aantal treinreizen tijdens ochtendspits (per week)	3,60 reizen/week	3,59 reizen/week
Aantal treinreizen tijdens avondspits (per week)	3,56 reizen/week	3,63 reizen/week
<i>Afstandsklasse (tariefeenheden deelnametraject)</i>		
≤ 25 km	20%	32%
25-40 km	22%	21%
≥ 40 km	58%	47%

Verder toont de tabel de reisafstanden voor beide groepen: de lengte van het deelnametraject in zogeheten tariefeenheden. Het blijkt dat onder deelnemers trajecten korter dan 25km ondervertegenwoordigd zijn, terwijl afstanden groter dan 40km oververtegenwoordigd zijn. Ook dit is conform verwachting, omdat deelnemers in een lagere afstandsklasse minder kunnen verdienen per spitsmijding, terwijl zij nog steeds hetzelfde aantal enquêtes en logboeken moeten invullen net als deelnemers in een hogere afstandsklasse.

Tenslotte presenteren we de uitkomsten van een model waarin we de waarschijnlijkheid om deel te nemen aan het proef in een (logit) model schatten. We nemen daarbij de kenmerken mee die hierboven het meest bepalend bleken. Door een model te schatten waarin die variabelen tegelijkertijd worden meegenomen (“multivariaat”), kunnen wij de zogeheten “marginale effecten” van de variabelen op de deelnamebeslissing schatten. Dat is het effect van een variabele als we de andere variabelen constant houden. Zo kan het zijn dat de effecten van “regio” verdwijnen als we tegelijkertijd rekening houden met “type werk”, en het “type werk” verschilt tussen de regio’s. Soms is het maken van een dergelijke correctie van belang, soms ook juist niet. Gemiddelde analyses, zoals hiervoor, en marginale, zoals we nu zullen geven, geven dan ook anderssoortige conclusies, en het is niet zo dat een van de twee benaderingen “beter” is dan de ander.

We drukken de effecten uit via de “odds ratio”. De “odds” is daarbij zelf ook al een ratio: die van de kans op wél-deelname versus niet-deelname. De “odds ratio” voor een groep deelt die odds door de odds voor mensen die in de referentiegroep zitten, en is dus een ratio van ratio’s. De odds ratio geeft dan aan, met welke factor de “odds” om deel te nemen verandert, als de verklarende variabele met een eenheid verandert (een coëfficiënt groter dan 1 geeft dus aan dat de toehorende variabele de kans op deelname verhoogt ten opzichte van het referentieniveau).

Net als in de univariate analyse in Tabel 3, blijkt inkomen ook in de multivariate analyse geen belangrijke rol te spelen voor de deelnamebeslissing. Terwijl Tabel 3 substantiële verschillen liet zien tussen het opleidingsniveau van deelnemers en niet-deelnemers, zien wij dat deze niet significant zijn als we controleren voor de invloed van andere variabelen zoals type werk en sector, die vaak gecorreleerd zijn met opleiding.

Voor type werk zien wij, dat *ceteris paribus* de odds dat iemand beslist om deel te nemen het kleinst zijn als deze persoon een administratieve baan heeft, vergeleken met andere beroepsgroepen. Ook de sector waarin iemand werkzaam is, is sterk bepalend voor de deelnamebeslissing. De odds voor deelname zijn het laagst voor de financiële instellingen en het hoogst voor onderwijs⁴. Ook als gecontroleerd wordt voor verschillen in type werk of opleiding, blijken de verschillen in deelnamekans tussen regio’s significant te zijn. Specifiek zien we dat treinreizigers uit de Randstad met een hogere waarschijnlijkheid deelnemen aan de proef dan reizigers van buiten de Randstad, wellicht vanwege hogere frequenties op het spoor. Tenslotte zien we dat reizigers in de laagste afstandsklasse significant lagere odds hebben om deel te nemen dan diegene met een hogere afstandsklasse, waarschijnlijk vanwege de lagere beloningen voor dalreizen.

4 Dit resultaat is misschien verrassend omdat het vaak lastig of onmogelijk zal zijn de aanvangstijd in het onderwijs aan te passen, maar het kan misschien verklaard worden door het feit dat deze mensen door het bestaande onderwijsrooster al veel dalreizen hebben, en dus zonder gedragsaanpassing beloningen kunnen verdienen.

Tabel 7: Multivariaat model van deelnamebeslissing

	Odds ratio	p-waarde
<i>Inkomen (Ref.: inkomen hoger dan 2500)</i>		
minder dan € 2500	0,80	0,22
onbekend	0,94	0,71
<i>Opleiding (Ref.: hoger beroepsonderwijs; "overig" niet in regressie meegenomen)</i>		
HAVO of lager	0,83	0,30
HBS, VWO, Atheneum, Gymnasium	0,82	0,59
Wetenschappelijke onderwijs	1,2	0,18
<i>Type werk (Ref.: Administratief personeel)</i>		
Manager	1,8	0,015
Onderzoeker, ingenieur, docent, (vak)specialist	2,0	0,00
Dienstverlenend personeel en verkopers	1,4	0,14
overig	4,0	0,034
<i>Sector (Ref.: Overig)</i>		
Financiële instellingen	0,41	0,00
Openbaar bestuur, overheidsdiensten, verplichte sociale verzekeringen	2,3	0,00
Onderwijs	4,6	0,00
<i>Provincie (Ref.: Overig)</i>		
Noord-Holland	1,4	0,045
Zuid-Holland	1,7	0,007
Utrecht	1,5	0,069
Noord-Brabant	1,3	0,33
Gelderland	1,1	0,67
<i>Afstandsklasse (Ref.: 25–40 km)</i>		
≤ 25 km	0,46	0,00
≥ 40 km	1,0	0,83

Resumerend blijken deelnemers op objectief waarneembare socio-economische kenmerken minder van niet-deelnemers te verschillen dan misschien verwacht zou mogen worden, alhoewel er significante verschillen zijn voor een aantal kenmerken. Deze verschillen (bijv. type baan of sector) hangen sterk samen met de mate van flexibiliteit die men heeft om tijdstippen aan te passen, en met gewenste aankomsttijden, en op deze minder makkelijk waarneembare kenmerken verschillen de twee groepen behoorlijk. Deze verschillen zijn zodanig dat inderdaad diegenen die makkelijker beloningen kunnen verdienen, meer geneigd zijn deel te nemen. Dit betekent weliswaar dat de propositie kennelijk duidelijk was, maar heeft als keerzijde dat de gedragsreacties die we hierna zullen rapporteren dus niet als representatief voor de gemiddelde kaarthoudende reiziger mogen worden geïnterpreteerd. Dat het gedrag verschilt, zal overigens verder worden bevestigd door de uitkomsten van het stated-choice experiment, dat door beide groepen is gedaan. Dat de

gedragsreacties niet representatief zijn voor de gemiddelde reiziger is in zekere zin overigens zelfs gewenst, aangezien de proef daarmee een betere indruk geeft van de gedragsreacties van reizigers die een grotere kans hebben om het type abonnement met tariefdifferentiatie tussen dal en spits dat hier onderzocht wordt, daadwerkelijk aan te schaffen.

4.2 Deelname aan de praktijkproef

Om een kwalitatieve indruk te krijgen van de mate waarin het toch lage deelnamepercentage bij toekomstige projecten verhoogd zou kunnen worden, hebben we niet-deelnemers gevraagd aan te geven in welke mate ze getwijfeld hebben bij hun beslissing om niet deel te nemen aan de praktijkproef. Een groot aantal niet-deelnemers heeft wel degelijk tenminste even getwijfeld over deelname, alvorens te besluiten hiervan af te zien. Dit duidt erop dat er nog een potentieel is voor Spitsmijden in de Trein. Omdat begrippen als “even” of “sterk” twijfelen erg subjectief zijn, lijkt het niet verstandig om op basis van deze antwoorden een kwantitatieve voorspelling te maken.

Tabel 8: Mate van onzekerheid over de beslissing om al dan niet deel te nemen

	Deelnemers	Niet-deelnemers
Beslissing	Deelname	Niet-deelname
Zeker	62%	54%
Even getwijfeld	36%	37%
Sterk getwijfeld	1,9%	9,8%

De genoemde hoofdredenen om niet deel te nemen (hier niet in tabelvorm getoond) zijn: te weinig mogelijkheden om buiten de spits te reizen, weinig interesse voor het onderzoek, en tenslotte de tijd die nodig is om enquêtes en logboeken in te vullen. Wanneer we aan de deelnemers vragen om welke redenen zij van deelname zouden hebben kunnen afzien, noemden deze ook nog diverse technische problemen (zoals het niet beschikken over geschikte smartphone) en onduidelijkheid over de proef en de hoogte van de beloning.

Wij hebben verder aan niet-deelnemers gevraagd om te veronderstellen dat ze tóch aan de proef zouden meedoen. Vervolgens werd hen gevraagd om voor drie verschillende hoogten van de beloning (2, 4 en 8 Euro) aan te geven hoeveel reizen ze zouden denken aan te passen om tijdens de daluren te reizen en de beloning te verdienen. Tabel 9 presenteert de gemiddelde wekelijkse zelfgerapporteerde verschuiving van spitsreizen naar dalreizen. De verschuiving hangt sterk af van de hoogte van de beloning, maar het is opvallend dat de waarden voor 4 Euro (bijna 0,4) dicht aanligt tegen de verschuiving die we later, in paragraaf 6, voor deelnemers vinden in de werkelijke proef (grofweg 10 procent van de 4,5 reizen per week verschuift van spits naar dal – dus: 0,45). Het is daarmee extra opvallend dat niet-deelnemers zo groot in getale zijn: ze verwachten zelf, bij deelname, een aanpassing van spits- naar dalreizen te kunnen maken die vergelijkbaar is met wat deelnemers daadwerkelijk hebben gedaan, ongeacht hun eerder genoemde kleinere flexibiliteit.

Tabel 9: Zelf-gerapporteerde bereidheid van niet-deelnemers tot aanpassing reisgedrag van spits naar dal (reizen/week) om een beloning te verdienen

Beloning per spitsmijding	Ochtend	Avond
2 Euro	0,22	0,23
4 Euro	0,37	0,38
8 Euro	1,16	1,16

5 Nadere kenmerken deelnemers

In deze paragraaf bespreken we enkele andere kenmerken van deelnemers in verder detail.⁵ Dit heeft in dit rapport alleen als doel om de lezer een beter beeld te geven van de kenmerken van de deelnemers; deze kenmerken zullen mogelijk – overigens net als kenmerken die in paragraaf 4 al besproken werden – pas in een latere onderzoeksfase als verklarende variabelen in gedragsmodellen worden meegenomen.

Een selectiecriteria voor deelname aan de praktijkproef was het beschikken over een treinabonnement. Voor meer dan 60% van de deelnemers is dit een Jaartrajectkaart/Altijd vrij abonnement; de rest reist met een OV- of NS-jaarkaart.⁶ De gemiddelde maandelijkse kostprijs van het abonnement bedraagt ongeveer 240 Euro. Dit bedrag wordt voor bijna 60% van de deelnemers door de werkgever betaald; 14% betaalt zelf voor het abonnement; en voor 23% van de deelnemers vergoedt de werkgever een deel van de kosten (de kleine restgroep heeft nog andere arrangementen). Bijna alle deelnemers geven aan het abonnement voornamelijk voor reizen van en naar het werk te gebruiken, sommige deelnemers geven aan het abonnement ook voor zakelijke reizen te gebruiken. Andere mogelijke reisdoelen voor het abonnement (zoals school/onderwijs, bezoeken van vrienden/familie/kennissen, vakantie/uitstapje, sport/hobby) zijn uitzonderlijk.

Gemiddeld verrichten deelnemers aan de praktijkproef 4,66 dagen per week betaald werk en daarbij hebben ze gemiddeld op één dag per week de mogelijkheid om thuis te werken. Van diegenen die thuis werken, doet bijna 20% dat op een vaste dag, de anderen hebben een flexibele weekindeling. Thuis werken gebeurt het meest op vrijdagen en het minst op dinsdagen.

De meerderheid van de deelnemers heeft een vaste werklocatie (66%) of werkt meestal op dezelfde locatie (28%). Slechts 5% hebben in een formele zin meerdere werklocaties.

Minder dan een derde van de deelnemers heeft geen alternatief vervoersmiddel (auto, motor, OV, fiets) ter beschikking voor de reis tussen woon- en werklocatie. Bijna 60% kan soms een auto gebruiken voor de woon-werk reis (heen en terug). Voor bijna 80% van de deelnemers is tenminste één auto beschikbaar in het huishouden. En meer dan 90% van de deelnemers heeft een rijbewijs voor personenauto's (categorie B).

5 De kenmerken die we hier bespreken zijn niet beschikbaar voor niet-deelnemers, bijgevolg is de vergelijking zoals die in sectie 4 gemaakt werd hier niet mogelijk.

6 De NS geeft vervoerbewijzen uit die geldig zijn voor alle treinen op het Nederlandse spoorwegnet. Met de door NS uitgegeven vervoerbewijzen kunnen de reizigers ook gebruik maken van treindiensten aangeboden door andere vervoerders.

Tabel 10: Uitgebreide kenmerken deelnemers

Deelnemers (aantal=649)	
<i>Type abonnement</i>	
OV-jaarkaart	22%
NS-jaarkaart	16%
Jaartrajectkaart/Altijd vrij	62%
Maandelijke prijs abonnement	239,55 Euro
<i>Wie betaalt abonnement?</i>	
Werknemer	14%
Werkgever	59%
Werkgever en werknemer samen	23%
anders	3,5%
Hoeveel dagen betaald werk?	4,66 dagen/week
Thuis werken	1,01 dagen/week
<i>Meerdere werklocaties</i>	
Nee	66%
Meestal op zelfde locatie	28%
Ja	5,3%
<i>Alternatief vervoer</i>	
Geen alternatief vervoer	28%
Kan een auto gebruiken	60%
Auto beschikbaar in huishoud	80%
Rijbewijs (categorie B)	92%
<i>Werken in de trein</i>	
Nooit/zelden	37%
soms	35%
regelmatig/altijd	29%

We hebben deelnemers eveneens gevraagd naar de vervoerswijze voor de verplaatsing tussen de woonlocatie en het beginstation van het deelnametraject. Daarbij kon de deelnemer meerdere alternatieven aangeven. 59% van de deelnemers gaat met de fiets naar de beginstation, 26% doet dit lopend, 17% met het OV (bus/tram/metro), 14% met de auto (als bestuurder of passagier) en 0,5% ander met een ander vervoersmiddel. Voor de verplaatsing tussen de werklocatie en het eindstation van het deelnametraject is te voet de meest genoemde verplaatsingswijze (62%). Daarna volgt het OV met 29%, de fiets met 21%, de auto met 0,06% en 1,1% andere vervoerswijzen.

Werken in de trein is nooit of zelden mogelijk voor 37% van de deelnemers. Voor 35% is het tenminste soms mogelijk, en voor 29% regelmatig of altijd. De redenen hierachter zijn niet gevraagd, maar zullen vaak samenhangen met het type werk. De belangrijkste cijfers zijn in tabel 10 samengevat.

We zijn voorts geïnteresseerd in de tevredenheid van deelnemers over hun treinreizen op het deelnametraject. Immers, die tevredenheid is medebepalend voor de keuzes die deelnemers in hun gedrag maken (trein versus alternatief vervoer, maar ook tijdstip van een treinreis als de tevredenheid sterk varieert over de dag). In elke logboekweek gaven de deelnemers met een rapportcijfer aan in welke mate ze tevreden waren over reizen met de trein in die week. Er blijkt niet veel variatie te zijn over de 6 logboekweken. De laagste gemiddelde waarde (6,78 op een schaal van 1 tot 10) wordt bereikt in de derde week, en de hoogste waarde (7,17) in de 1ste week.

Tabel 11: Gerapporteerde betrouwbaarheid en reiscomfort woon-werkreis (enquête)

	heenreis (ochtend)	terugreis (avond)
<i>Vertraging</i>		
Te vroeg, op tijd, minder dan 5 minuten	74%	72%
5–14 minuten	19%	20%
15–29 minuten	4,5%	5,4%
30–59 minuten	1,8%	2,1%
60 minuten of langer	0,39%	0,55%
<i>Drukte in de trein</i>		
Te druk	24%	26%
Druk	36%	43%
Niet druk/niet rustig	26%	23%
Rustig	14%	7,5%
Te rustig	0,63%	0,30%
<i>Beschikbaarheid zitplaatsen</i>		
Zitje van 4 voor mezelf	13%	6,7%
Zitje van 2 voor mezelf	24%	21%
Iemand zit naast	34%	39%
Niet iedereen kan zitten	15%	19%
Veel reizigers staan	14%	14%

Tabel 11 toont de in de enquête gerapporteerde vertragingen, de drukte in de trein en de beschikbaarheid van zitplaatsen. Dit zijn overigens wél kenmerken die in de gedragsmodellen in dit rapport gebruikt zullen worden.⁷ Tabel 11 laat zien dat deelnemers langere vertragingen ervaren op de terugreis (avondspits) dan op de heenreis (ochtendspits) naar de werklocatie. Ook de drukte in de trein en het ontbreken van zitplaatsen blijkt door deelnemers in sterkere mate te worden ervaren tijdens de avondspits dan in de ochtendspits.⁸

In het algemeen ervaren de deelnemers een forse meerderheid van de treinreizen als stipt of weinig

⁷ In de gedragsmodellen zullen we deze kenmerken baseren op de rapporteringen in de logboeken (zie sectie 6.6).

⁸ Het gaat hier om vertraging zoals door deelnemers gerapporteerd. Om verschillende redenen kunnen deze cijfers afwijken van de stiptheidscijfers die vervoerders meten en rapporteren. Zo gaat het hier om percepties in plaats van feitelijke metingen en daarnaast hebben stiptheidscijfers van vervoerders betrekking op treinen over heel Nederland, over de gehele dag, en over het gehele jaar; terwijl het hier over de rapportage door individuele reizigers gaat en vooral in of rond de spits op de geselecteerde trajecten en voor de duur van de proef.

vertraagd. Voor ongeveer 1 op 4 reizen rapporteren ze een vertraging van meer dan 5 minuten. Ongeveer 6 op 10 reizen wordt echter als te druk ervaren. Op bijna 3 van 10 reizen moeten tenminste enkele reizigers staan omdat er te weinig zitplekken beschikbaar zijn.

6 Reisgedrag

6.1 Inleiding

Het werkelijke reisgedrag van de deelnemers wordt, zoals gezegd, gemeten door de smartphone-app. In sectie 6.2 beschrijven we hoe we treinreizen identificeren op basis van de door de app geregistreerde “geolocaties” (dit zijn lengte- en breedtegraadmetingen), in combinatie met tijdstipindicatie. Het gemeten reisgedrag wordt vervolgens vergeleken tussen voormeting, beloningsperiode en nameting. Dit geeft een eerste indruk van de mate waarin de ingezette beloning effectief was om het aantal spitsreizen te laten afnemen. De belangrijkste aanpassing die de deelnemer kan maken in zijn of haar reisgedrag om een beloning te verdienen, is een wijziging van de vertrektijd. In onze analyse zal de verdeling van vertrektijden, over de dag, daarom centraal staan.

Naast een beschrijving van de verdeling van de vertrektijden en hoe deze wijzigt ten gevolge van de beloningsprikkel, schatten we ook gedragsmodellen die de gemaakte keuzes op individueel niveau verklaren. Deze gedragsmodellen zijn gebaseerd op discrete keuzetheorie. Centraal in een dergelijk gedragsonderzoek staat de waardering die de reiziger toekent aan de keuze van een bepaalde vertrektijd, en de daarbij horende aankomsttijd. De achterliggende gedachte is dat de spitsreiziger zonder beloningsexperiment de vertrek- of aankomsttijd kiest die hem of haar, alles afwegende en rekening houdend met de omstandigheden, het best uitkomt. Afwijken van dit optimale tijdstip wordt door de reiziger per definitie negatief gewaardeerd (anders was dat tijdstip immers niet optimaal). In de uiteindelijke keuze zal dit nadeel afgewogen worden tegen mogelijke voordelen van het reizen op een ander tijdstip, al dan niet buiten de spits. Dergelijke voordelen kunnen bestaan uit een beter zitcomfort, minder vertragingen, snellere reistijden, minder overstappen, en in de praktijkproef natuurlijk ook het verdienen van een beloning.

Uit de reeks geobserveerde keuzes voor bepaalde alternatieven, en daarmee ook observaties van alternatieven die juist niet zijn gekozen, kan worden afgeleid hoe de reiziger de verschillende kenmerken van die alternatieven tegen elkaar afweegt. Bijvoorbeeld, vindt de reiziger het aantrekkelijk om een uur eerder dan optimaal aan te komen als daarmee een zitplaats gegarandeerd is? In de te gebruiken modellen blijkt hoe deze afwegingen worden gemaakt uit de waarden van de geschatte parameters voor de verschillende kenmerken. Die parameters geven aan hoe de kans op het kiezen van een alternatief toeneemt, als het betreffende kenmerk verandert. Als dan bijvoorbeeld de parameters voor de beloning (in Euro's) en reistijd (in uren) een verhouding 10 tegen -1 hebben, betekent dit dat de reiziger een uur langere reistijd nog net accepteert als de beloning 10 Euro hoger is. We zeggen dan dat de waarde van reistijdwinst 10 Euro per uur is. Op die wijze kunnen we alle kenmerken die we observeren (inclusief tijdstip, bezetting, betrouwbaarheid) uitdrukken in waarderingen in Euro's, en dat maakt het vergelijken van kenmerken natuurlijk preciezer mogelijk dan als we bijvoorbeeld zouden weten dat betrouwbaarheid “heel belangrijk” en bezetting “enigszins belangrijk” zou zijn.

Voor het schatten van de parameters van de gedragsmodellen gebruiken we de volgende

gegevensbronnen:

- *feitelijke tijdstipkeuzes* zoals die met de app gemeten zijn (revealed preference ofwel RP); deze zijn alleen beschikbaar voor de deelnemers;
- *fictieve tijdstipkeuzes* zoals die in de tussen-enquête en de non-respons enquête zijn geregistreerd (stated preference ofwel SP); deze zijn beschikbaar voor deelnemers en niet-deelnemers.

De resultaten van deze modellen worden gepresenteerd in paragraaf 6.8.

6.2 Meten Reisgedrag

De smartphone-app werd tijdens de praktijkproef op meer dan 1000 smartphones geïnstalleerd, hetgeen meer dan 8 miljoen observaties opleverde (tijdstip en geolocatie). Voordat we deze gegevens hebben kunnen gebruiken in gedragsmodellen, hebben we de observaties opgeschoond: het verwijderen van ruis in de vorm van aantoonbaar onmogelijke verplaatsingen. Vervolgens hebben we woon-werkverplaatsingen geïdentificeerd op basis van het bij werving geregistreerde deelnametraject. Dit vormen meer dan 100.000 verplaatsingen. De trajecten van deze verplaatsingen hebben we tenslotte vergeleken met de locaties van stations om zo de treinreizen te identificeren.⁹

De app zorgt ervoor dat de positie van de smartphone wordt geregistreerd in de projectdatabase. Wanneer de smartphone een belangrijke wijziging vaststelt in de actuele locatie neemt de app via de actieve netwerkverbinding contact op met de projectdatabase, waar de nieuwe geolocatie (lengte- en breedtegraad) wordt vastgelegd samen met het tijdstip van ontvangst van deze positie. Bovendien wordt in de app een lijst met stationslocaties opgeslagen. Wanneer de smartphone een locatie binnengaat of verlaat, zal de app dit registreren in de projectdatabase. De toegepaste lijst met stationslocaties is gepersonaliseerd op basis van het deelnametraject.

Alhoewel de app het proces van reisregistraties aanstuurt, gebeurt de eigenlijke locatiemeting en de communicatie met de projectdatabase grotendeels op het niveau van het besturingssysteem (iOs of Android) en de hardware van de smartphone (bv. de GPS-chipset). Vele factoren kunnen daarbij een invloed hebben op het eindresultaat, zoals de toegepaste meetmethode (GPS, wifi-driehoeksmeting, etc), de nauwkeurigheid van de hardware, de beschikbaarheid van een netwerkverbinding, eventuele onvolkomenheden in het operating systeem, etc. Op die manier ontstaat een hoeveelheid ruis in de databaseregistraties die we eerst willen verwijderen voordat we het reisgedrag gaan bestuderen.

Een eerste stap in het opschonen van de locatieregistraties bestaat uit het verwijderen van fysiek onmogelijke opeenvolgingen van locaties in de tijd. Omdat we treinreizen bestuderen, gaan we hierbij uit van een maximale snelheid tussen twee punten van 160km/u. Waar we hogere snelheden vaststellen bekijken we een reeks opeenvolgende punten om te beoordelen welke gerapporteerde locaties de meest waarschijnlijke registratie van de beweging van de deelnemer in tijd en ruimte vastleggen. Daarbij houden we rekening met de door de smartphone gerapporteerde nauwkeurigheid van de metingen.

Nadat we de ruis verwijderd hebben, identificeren we verplaatsingen over het traject van deelname.

⁹ Merk op dat het in deze sectie gepresenteerde algoritme ontworpen is met als doel het ex-post (achteraf) analyseren van keuzes die deelnemers maken in hun reisgedrag. Dit algoritme wijkt af van het tijdens de praktijkproef toegepaste proces dat ingericht was met als doel het in real-time toekennen van beloningen.

Dit doen we door in een relatief ruime cirkel rond vertrek- en aankomststation te kijken naar reislocaties gerapporteerd door de app. Wanneer de app een verplaatsing binnen 3 uur tussen deze locaties rapporteert, gebruiken we dit als startpunt voor de identificatie van de verplaatsing. We gebruiken hierbij een op maat van het onderzoek ontwikkeld algoritme dat op zoek gaat naar begin- en eindpunt van het verplaatsingstraject. Hierbij wordt gekeken naar snelheden, nadering van begin- en eindstation van het deelnametraject en mogelijke overstaplocaties. Ook de door de smartphone gerapporteerde nauwkeurigheid van de metingen wordt hierbij betrokken. Voor elke verplaatsing bepalen we tenslotte het vertrek- en aankomsttijd, de afstand van het vertrekpunt en het eindpunt tot het beginstation resp. eindstation van het deelnametraject, en het aantal stationslocaties binnen 300m langs het geregistreerde verplaatsingstraject.

Op basis van de beschreven algoritmes identificeren we op deze wijze 90.825 verplaatsingen die gemeten zijn tijdens voormeting, beloningsperiode en nameting. Omdat we in het selectiealgoritme relatief ruime criteria gehanteerd hebben bij het vaststellen van een verplaatsing langs het deelnametraject, dienen we verder na te gaan of de reis ook werkelijk met de trein gemaakt werd. Immers, er kan ook een autorit gemaakt zijn, die in de buurt van die stations kwam. Dit doen we op basis van het aantal stationslocaties langs het traject van de geregistreerde verplaatsing (binnen 300m). Gemiddeld liggen langs elk traject 5,40 stationslocaties. Om de treinreizen te identificeren sluiten we de verplaatsingen uit met 0 stationslocaties, alsook verplaatsingen waarvoor het aantal stationslocaties langs het traject minder dan de helft bedraagt van de (individu-specifieke) mediane meting voor het deelnametraject (over alle gemeten verplaatsingen). Op die manier sluiten we 7975 verplaatsingen uit van verdere analyse omdat ze wellicht niet met de trein gemaakt zijn. Vervolgens sluiten we ook observaties tijdens het weekeinde uit, omdat de beloning alleen maar op werkdagen van toepassing was en het reisgedrag tijdens het weekeinde wellicht slechts beperkt vergelijkbaar is met dat op werkdagen (bv. andere reismotieven). Om de datakwaliteit te waarborgen sluiten we tenslotte ook de verplaatsingen uit die op meer dan 300 meter van het begin- of eindstation van het deelnametraject beginnen of eindigen. Uiteindelijk levert dit ons een tabel met 57.508 verplaatsingen met volgende karakteristieken:

- gemiddeld aantal stationslocaties langs een verplaatsingstraject: 6,08;
- gemiddelde afstand van vertrekpunt tot beginstation deelnametraject: 66,33m;
- gemiddeld afstand van eindpunt tot eindstation deelnametraject: 40,01m.

Gegeven de omvang van stations in Nederland kan een gemiddelde afstand tot begin- en eindstation van minder dan 100m als een indicatie van goede gegevenskwaliteit beschouwd worden.

Op deze wijze houden we alleen de kwalitatief goede metingen over die een nauwkeurige schatting van de gedragseffecten toestaan. Omdat er geen reden is aan te nemen dat er een systematisch verschil zit in de kwaliteit van de GPS-metingen tussen spits- en dalreizen, en als dat er al wel zou zijn er nog steeds geen reden zou zijn aan te nemen dat dit systematische effect zou verschillen tussen voor- en nameting en beloningsperiode, verwachten we niet dat deze selectie tot een vertekening in de effectschatting of de significantie daarvan leidt.

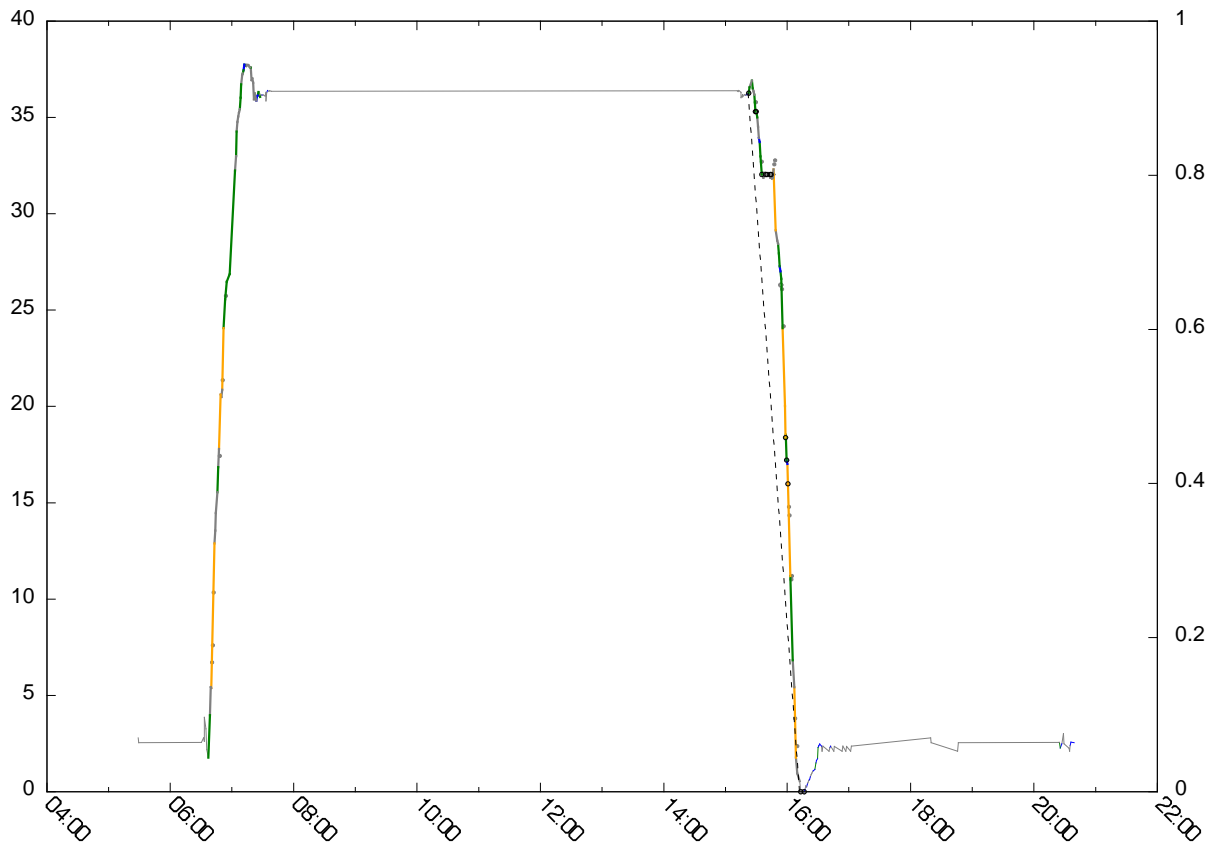
Een andere indicatie voor de gegevenskwaliteit wordt geleverd door de vergelijking tussen met de app geregistreerde treinreizen en in het logboek gerapporteerde verplaatsingen voor dezelfde persoon, op dezelfde dag en hetzelfde traject (rekening houdend met gerapporteerde vertragingen

op begin- en eindstation). Opnieuw beperken we ons tot de hierboven beschreven selectie van geobserveerde verplaatsingen.

We vinden dat beide gegevensbronnen (gemeten treinreizen en logboek) wat betreft vertrektijd voor 57% van de treinreizen minder dan 5 minuten verschillen, en voor 79% minder dan 15 minuten. Voor de aankomsttijd bedragen deze aandelen 55% en 74%. Deze aandelen stijgen nog licht als we verder rekening houden met de discipline bij het invullen van de logboeken, en de vergelijking beperken tot treinreizen die binnen de 10 dagen in het logboek werden geregistreerd.

Vervolgens verifiëren we of er een systematisch verschil bestaat tussen metingen volgens beide gegevensbronnen (bijvoorbeeld of de gemeten reistijden gemiddeld significant langer zijn dan de gerapporteerde). We vinden dat de mediaan van het verschil tussen gemeten en gerapporteerde waarden heel klein is. Voor vertrektijden is dit 84 sec later dan de gerapporteerde, en voor aankomsttijd is dit 3 sec later dan de gerapporteerde. We kunnen dus concluderen dat de vertrek en aankomsttijden voor beide databronnen verrassend dicht bij elkaar liggen. Het feit dat de mediane reistijden relatief weinig verschillen, terwijl toch ruim 40% een afwijking van meer dan 5 minuten hebben, heeft te maken met de vorm van de verdeling, en met name het feit dat deze behoorlijk symmetrisch rondom het nulpunt is.

Naast een systematische validering van de algoritmes voor het opschonen van registraties en van de identificatie van treinreizen, hebben we steekproefgewijs ook verplaatsingen visueel gevalideerd. Figuur 2 illustreert dit.



Figuur 2: Visualisatie van de afstand tot het beginstation van het deelnemetraject op basis van app-registraties (continue lijn) en logboek (streeplijn)

Tenslotte rapporteren we hier nog de reisfrequentie op enkele geselecteerde verbindingen. De meest frequent geobserveerde verbindingen zijn tussen Utrecht Centraal en Amsterdam (Zuid, Centraal, Amstel), en de verbinding tussen Utrecht Centraal en Den Haag Centraal. Een rangschikking op basis van het aantal geobserveerde treinreizen tussen begin(woon)- en eind(werk)stations van het deelnemetraject staat in tabel 12. In beide gevallen is Utrecht Centraal het meest frequente station. Voor de eindstations (bij werklocatie) is de concentratie van reizen veel groter dan voor beginstations (bij woonlocatie). Zo maakt de top-10 van de beginstations 31% van de treinreizen uit, terwijl 67% van de reizen in de top-10 van de eindstations eindigt. Dit verschil komt uiteraard doordat werklocaties in de (grote) steden geconcentreerd zijn. De Amsterdamse stations worden bijvoorbeeld veel vaker als eindstation (Centraal, Sloterdijk, Zuid, Bijlmer ArenA, Amstel) dan als beginstation gebruikt.

Tabel 12: Meest frequent begin- en eindstations voor het deelnametraject

	beginstation deelnametraject (bij woonlocatie)	eindstation deelnametraject (bij werklocatie)
Aandeel reizen top-10	31%	67%
1	Utrecht Centraal	Utrecht Centraal
2	Leiden Centraal	Amsterdam Centraal
3	Amsterdam Centraal	Den Haag Centraal
4	Den Haag Centraal	Amsterdam Sloterdijk
5	Amersfoort	Amsterdam Zuid
6	Nijmegen	Amsterdam Bijlmer Arena
7	Haarlem	Amsterdam Amstel
8	Den Bosch	Rotterdam Centraal
9	Hilversum	Eindhoven
10	Eindhoven	Leiden Centraal

6.3 Geobserveerd reisgedrag

In deze paragraaf bestuderen we hoe deelnemers door de beloning hun reisgedrag hebben aangepast. Dit doen we door het aandeel spits- en dalreizen (op basis van het met de app gemeten gedrag) te vergelijken tussen de voormeting, de beloningsperiode en de nameting. Nadat we de vergelijking voor de geaggregeerde gegevens hebben uitgevoerd, gaan we na of het algemene beeld verandert onder specifieke omstandigheden, zoals voor verschillende hoogtes van de beloning, de lengte van het deelnametraject (afstandsklasse), de regio waarin de deelnemer reist, of de dag van de week.

Het is van belang om hier kort stil te staan bij de manier waarop we de gedragsverandering uitdrukken. Dit zullen we doen door te kijken naar de aandelen van spitsreizen en dalreizen in de totale mobiliteit van de deelnemers op hun deelnametraject. Dit brengt een mogelijk dubbelzinnigheid met zich mee, die samenhangt met het verschil in veranderingen uitgedrukt in *procenten versus procentpunten*. Een voorbeeld kan dit verduidelijken. Als het aandeel van de spitsreizen zonder beloning 50% is, en mét beloning 40%, dan kunnen we ofwel zeggen dat het aandeel spitsreizen met 10 procentpunt (50% minus 40%) is afgenomen, ofwel dat het aandeel spitsreizen met 20 procent is afgenomen (de afname van 10 procentpunt vormt 20% van het oorspronkelijke aandeel van 50%). Deze tweede manier van uitdrukken heeft als voordeel dat het gerapporteerde effect op het aantal spitsreizen onafhankelijk is van hoeveel dalreizen er ook zonder beloning al gemaakt worden. Omdat vermoedelijk de meeste lezers geïnteresseerd zullen zijn in het effect van de beloning op de omvang van het spitsverkeer, zullen we als regel voor die tweede wijze van uitdrukken kiezen.

Naast een wijziging van aandelen spits- en dalreizen, wijzigt soms ook het totale aantal treinreizen dat een deelnemer maakt langs het deelnametraject (de reisfrequentie). Voor sommige deelnemers is de variatie in de gemeten reisfrequentie echter van die omvang dat deze niet verklaard kan worden door de beloningsprikkel (die in principe geen afname van de reisfrequentie stimuleert). Voor deze

deelnemers is het waarschijnlijk dat ofwel de discipline in het app-gebruik tijdens de proef is veranderd (doorgaans minder app-gebruik vanwege het effect op batterijgebruik; zie ook paragraaf 7), ofwel dat er een verandering in het persoonlijke mobiliteitsgedrag is geweest die niet met de proef samenhangt. De reismetingen van deze deelnemers sluiten we daarom uit in de verdere analyse. Daartoe berekenen we eerst voor elke deelnemer het wekelijkse aantal treinreizen voor elke week waarin tenminste één reis geregistreerd is. Vervolgens sluiten we alle deelnemers (en hun treinreizen) waarvoor de wijziging in het wekelijkse aantal reizen langs het deelnametraject tussen voormeting en beloningsperiode buiten het 10de (een grote afname) tot 90ste (een grote toename) percentiel valt, uit van verdere analyse. Dit betekent dat deelnemers met een toename van meer dan 2,00 reizen per week of een afname van meer dan 2,07 reizen per week tussen voormeting en beloningsperiode uitgesloten worden in de verdere analyse. Verder sluiten we ook nog deze deelnemers uit voor wie we minder dan 5 reisobservaties tijdens de voormeting hebben, omdat we voor deze deelnemer geen betrouwbare vergelijkingsbasis voor vertrek- en aankomsttijd kunnen vaststellen. Na deze opschoning blijven er 479 deelnemers over die we meenemen in de analyse.¹⁰

Omdat ook een deel van de dan resterende deelnemers tijdens de beloningsperiode hun deelname heeft stopgezet, of tijdens de nameting de app niet meer heeft aangezet, hebben we de vergelijking van het reisgedrag tussen de meetperiodes (voormeting, beloningsperiode en nameting) afzonderlijk uitgevoerd voor een selectie deelnemers waarvoor tenminste 5 geregistreerde reizen zijn in de nameting (238 deelnemers). De sterke daling van de groep deelnemers door deze laatste correctie lijkt erop te wijzen dat de motivatie om deel te nemen afnam tijdens de looptijd van de praktijkproef. We zullen hier nader op ingaan in de evaluatie van het project.

De volgende tabel toont de resultaten voor gedragsveranderingen tussen voormeting, beloningsperiode en nameting.

10 Zonder deze opschoning zouden we een relatief hoger aantal spitsmijdingen tijdens de beloningsperiode ten opzichte van de voormeting rapporteren. Dit kan samenhangen met het feit dat het uitzetten van de app, om batterijen te sparen, aantrekkelijker is voor een spitsreis dan voor een dalreis (waar dan een beloning wordt misgelopen). De in dit hoofdstuk gepresenteerde resultaten kunnen in dat opzicht als conservatieve schatting van het gerealiseerde aantal spitsmijdingen beschouwd worden.

Tabel 13: Reisgedrag langs het deelnametraject

	Voormeting	Beloningsperiode	Nameting	Totaal
<i>Deelnemers met minimum 5 reizen tijdens de voormeting</i>				
<i>Aantal: 479</i>				
<i>(volledige groep)</i>				
aantal reizen	5635	26216	3920	35771
reizen/week/deelnemer	4,43	4,58	4,45	4,54
aandeel ochtendspits	33%	26%	30%	28%
aandeel avondspits	27%	22%	26%	23%
aandeel dalperiode	40%	52%	44%	49%
afname van aantal spitsreizen per week t.o.v. voormeting	0	20%	6,1%	-
afname van aantal spitsreizen per week t.o.v. voormeting (lage beloningsprikkel)	-	19%	-	-
afname van aantal spitsreizen per week t.o.v. voormeting (hoge beloningsprikkel)	-	21%	-	-
<i>Deelnemers met minimum 5 reizen zowel in voormeting als in nameting</i>				
<i>Aantal: 238</i>				
<i>(beperkte groep)</i>				
aantal reizen	3098	17918	3814	24830
reizen/week/deelnemer	4,86	4,94	4,62	4,88
aandeel ochtendspits	32%	25%	31%	27%
aandeel avondspits	27%	21%	26%	23%
aandeel dalperiode	41%	54%	44%	51%
afname van aantal spitsreizen per week t.o.v. voormeting	0	22%	4,3%	-
afname van aantal spitsreizen per week t.o.v. voormeting (lage beloningsprikkel)	-	21%	-	-
afname van aantal spitsreizen per week t.o.v. voormeting (hoge beloningsprikkel)	-	23%	-	-

Uit de tabel maken we op dat voor de beperkte groep het wekelijkse aantal reizen inderdaad ongeveer constant blijft over de 3 meetperiodes, wat erop duidt dat de beloning voor hen geen of slechts een heel beperkt (stimulerend) effect heeft gehad op het totale aantal reizen. De toename in de dalperiode komt dus overeen met de afname van het aantal reizen in ochtend- en avondspits. Het gemeten aantal reizen per week is duidelijk lager dan wat deelnemers hadden aangegeven in de enquête: gemiddeld heen en terug reizen op 4,34 wekdagen, dus 8,68 reizen per week. Het lagere aantal app-geregistreerde reizen ontstaat omdat reizen niet gemeten werden (bijv. als de app uitstond of niet goed werkte), maar ook omdat we relatief strenge selectiecriteria hebben gehanteerd bij de verwerking van app-registraties (zie sectie 6.2: begin en eind van het verplaatsingstraject binnen 300m van begin- en eindstation van het deelnametraject en een minimum aantal stationslocaties binnen 300m van het verplaatsingstraject), daardoor vallen veel gemeten

verplaatsingen buiten de analyse. Aangezien de criteria in de beloningsperiode even streng zijn als daarbuiten, verwachten we niet dat dit de relatieve veranderingen tussen spits- en dalmobiliteit zal vertekenen.

De vergelijking van het reisgedrag tussen de meetperiodes geeft aan dat het aandeel reizen in de ochtend- en avondspits daalt met 12 procentpunt voor de volledige groep (479 deelnemers) en 13 procentpunt voor de beperkte groep (238 deelnemers). Deze cijfers komen overeen met een relatieve daling van respectievelijk 20% en 22%, en is statistisch significant.¹¹ Een afname van het aantal spitsreizen als gevolg van de beloning met ongeveer 20% kan als substantieel worden beschouwd. De tabel toont aan dat de afname van het aantal spitsreizen tijdens ochtend- en avondspits bijna gelijk is: voor de volledige groep bedraagt de afname tijdens zowel de ochtendspits als de avondspits 20%.

We zien slechts beperkte verschillen in reisgedrag tussen een hoge en een lage beloningsprikkel. Voor de volledige groep deelnemers daalt het aandeel spitsreizen met 19% als de lage beloningsprikkel van toepassing is (in vergelijking met de voormeting). Deze daling is voor de hoge beloningsprikkel iets groter: 21%. Vergelijkbare resultaten worden ook voor de beperkte groep gevonden. Een mogelijke reden dat dit verschil vrij laag is, kan zijn dat deelnemers hun reisgedrag voor de volledige beloningsperiode slechts éénmaal hebben gewijzigd, bijvoorbeeld omdat dit afstemming thuis en op het werk vereist. Een wijziging in de hoogte van de beloning tijdens de beloningsperiode heeft dan maar een klein gedragseffect. Voor de minder oplettende deelnemer kan het ook zo zijn dat deze zich er niet altijd bewust van was of de hoge dan wel de lage beloningsprikkel van toepassing was. Bij eerdere spitsmijdenprojecten bleek al vaker dat de hoogte van de beloning, als deze tijdens de proef gewijzigd werd, weinig effect sorteert. Juist omdat dit het gevolg kan zijn van het gemaakt hebben van afspraken en plannen voor de gehele duur van het experiment, zouden we zeker niet willen concluderen dat de hoogte van de beloning er ook op de langere termijn relatief weinig toe zou doen.

Opvallend is voorts dat het aandeel spitsreizen in de nameting lager is dan in de voormeting. Dit wijst erop dat sommige deelnemers na beëindiging van de beloningsprikkel in elk geval niet meteen terugkeren naar hun oorspronkelijke (onbeloonde) reisgedrag. Het verschil in het aandeel spitsreizen tussen voor- en nameting is 3,6 procentpunt; voor de volledige groep deelnemers is dit een relatieve afname van 6,1% (4,3% voor de beperkte groep). Ook dit is een patroon dat in eerdere spitsmijdenprojecten is gezien. Als reizigers door het experiment meer inzicht hebben gekregen in de voor- en nadelen van buiten de spits reizen, en met name wanneer dit minder onaantrekkelijk blijkt te zijn dan ze voorheen hadden gedacht, kan dit een blijvend effect zijn. Maar het lijkt ook aannemelijk dat dit effect op de wat langere duur zal verdampen. Een meer definitief antwoord kan alleen gegeven worden als een experiment een langere nameting toestaat dan tot nog toe gebruikelijk is geweest in spitsmijdenprojecten, inclusief deze.

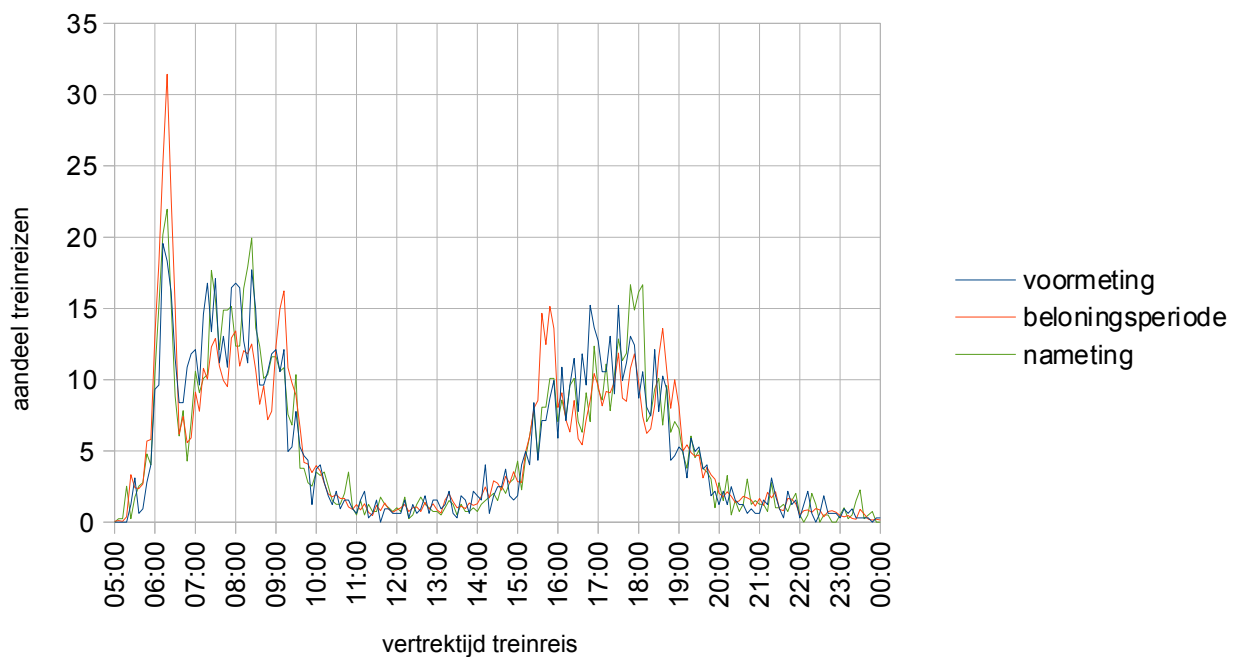
Het verschil tussen de volledige en de beperkte groep deelnemers is klein wat betreft

11 De statistische significantie is vastgesteld door het aandeel spitsreizen vast te stellen op individueel niveau voor de verschillende meetperiodes. Vervolgens hebben we gekeken wat de wijziging is in dit aandeel tussen de meetperiodes (nog steeds op individueel niveau). Door de gemiddelde wijziging in het spitsaandeel te vergelijken met de variantie ervan (de zogenaamde t-waarde) kunnen we vaststellen in welke mate deze wijziging toevallig kan zijn. De vastgestelde wijzigingen tussen meetperiodes zijn allemaal zeer significant, zowel voor de volledige als voor de beperkte groep. Voor elk van deze zes vergelijkingen werd de t-waarde vastgesteld; de kleinste t-waarde bedraagt 3,08.

gedragsveranderingen over de drie meetperiodes. Dit geeft aan dat ook deelnemers die de proef niet volledig hebben afgemaakt wel actief meededen in de weken waarin ze de app wel nog hadden aanstaan op hun smartphone. Merk op dat de beperkte groep in alle meetperiodes een hoger aantal wekelijkse reizen heeft. Een verklaring hiervoor is dat het hoger aantal reizen meer gelegenheden bood om de spits te mijden en beloningen te verdienen, bijgevolg is de prikkel kleiner om deelname aan de proef voortijdig te beëindigen.

Figuur 3 toont de verdeling van treinreizen, langs de deelnametrajecten, over de tijd van de dag voor de drie meetperiodes (alleen de reizen van de beperkte groep zijn weergegeven). De figuur toont dat tijdens de voormeting een aantal deelnemers reeds (gedeeltelijk) de spits mijdt. Het is logisch dat deze deelnemers een grote prikkel hebben om aan deze proef deel te nemen, omdat zij hun gedrag weinig of mogelijk helemaal niet hoeven aan te passen om een beloning te verdienen. Het is dus waarschijnlijk dat reizigers met dit reisgedrag oververtegenwoordigd zijn in de proef, wat aansluit bij de inzichten uit de sectie over het deelnamegedrag.

De figuur toont eveneens duidelijk dat tijdens de beloningsperiode het aantal spitsmijdingen substantieel toeneemt. Vooral het aantal spitsmijdingen net voor de ochtend- en avondspits neemt toe.



Figuur 3: Spreiding treinreizen over de dag

De algemene resultaten kunnen we vervolgens uitsplitsen. We bespreken in dit rapport uitsplitsingen naar werkdagen, regio's en afstandsklassen (zie tabel 14). We merken op dat, hoewel we verschillen vinden in effectschattingen voor subgroepen, deze bij eerste statistische testen niet of nauwelijks significant zijn, ten dele door de beperkte omvang van de uiteindelijke dataset. De resultaten geven dus alleen een eerste indicatie van de richting waarin de effecten kunnen gaan. De combinatie van statistische insignificantie van deze variabelen en behoorlijk grote variatie tussen individuen laat zien dat de verschillen binnen de sub-groepen, dus verklaard door factoren die we

niet waarnemen, groter is dan tussen de groepen.

We vinden dat het aantal treinreizen langs het deelnametraject het hoogst is op dinsdagen en het laagst op vrijdag. Het aandeel spitsreizen tijdens de voormeting was het hoogst op dinsdagen en woensdagen en het laagst op donderdagen, maar de verschillen zijn klein. Het is daarom verrassend dat de relatieve verandering in het wekelijkse aantal spitsreizen wél sterk verschilt tussen de dagen van de week: de afname in aantal spitsreizen tussen voormeting en beloningsperiode is het grootst op woensdagen (25%) en vrijdag (24%) en het kleinst op donderdagen (17%). Op woensdagen en vrijdag is ook de afname van het aantal spitsreizen in de nameting ten opzichte van de voormeting het grootst. Het lijkt erop dat de aanpassing van de vertrektijd het eenvoudigst is op deze dagen. Woensdag en vrijdag zijn ook de dagen met het kleinste aantal reizen in het geheel. Het is dus niet zo dat deelnemers hun vertrektijd het sterkst aanpassen op de drukste dagen, maar eerder dat ze de spits mijden op wekdagen waarop deelnemers sowieso al minder reizen.

Tabel 14: Uitsplitsing reisgedrag deelnemers (volledige groep: 479 deelnemers)

Deelnemers met minimum 5 reizen tijdens de voormeting aantal: 479	Totaal aantal reizen	Aandeel spitsreizen tijdens voormeting	Relatieve afname wekelijks aantal spitsreizen in beloningsperiode t.o.v. voormeting	Relatieve afname wekelijks aantal spitsreizen in nameting t.o.v. voormeting
<i>Weekdagen</i>				
Maandag	7646	60%	18%	2,0%
Dinsdag	8230	61%	19%	4,4%
Woensdag	6888	61%	25%	7,7%
Donderdag	7916	58%	17%	7,2%
Vrijdag	5091	60%	24%	10%
<i>Regio deelnametraject</i>				
Randstad	26568	58%	20%	8,1%
Nederland Noord en Oost	4712	66%	13%	-1,0%
Nederland Zuid	4491	61%	22%	4,1%
<i>Afstandsklasse deelnametraject</i>				
≤ 25 km	8392	54%	22%	2,8%
25-40 km	7565	60%	14%	-1,4%
≥ 40 km	19814	63%	22%	10%
<i>Type werk</i>				
Eigenaar van een bedrijf/vrij beroep	906	56%	16%	7,8%
Manager/Leidinggevende	4574	64%	20%	9,1%
Beleid/staf/advies/onderzoek	13732	62%	20%	2,5%
Uitvoerende/ondersteunende functie	10665	52%	25%	6,5%
Geen informatie/anders	5894	65%	10%	8,6%

Vervolgens kijken we naar de verschillen in reisgedrag over de deelnemer-specifieke kenmerken: regio's¹², afstandsklassen en type werk. In deze analyse beperken we ons tot de groep deelnemers voor wie het aandeel spitsritten afneemt: we willen namelijk nagaan of er een verband is tussen de omvang van deze daling enerzijds en kenmerken van de deelnemer anderzijds. Met eenvoudige OLS-schattingen, hebben we onderzocht welke kenmerken de sterkste verklarende waarde hebben voor de percentuele afname in de spitsreizen tussen voormeting en beloningsperiode, zowel wanneer we ze individueel (univariaat) als tegelijk (multivariaat) in de statistische analyse betrekken. De multivariate analyse beantwoordt bijvoorbeeld de vraag of er een onafhankelijk effect van reisafstand is, dan wel dat het effect voortkomt doordat deelnemers met een bepaalde reisafstand ook vaker een bepaald type werk hebben, en om die reden hun gedrag minder aanpassen. In deze analyse die gelijktijdig type werk, opleiding, inkomen, regio en afstandsklasse meeneemt voor alle deelnemers met een afname in het aandeel spitsreizen tussen voormeting en beloningsperiode, blijkt dat alleen het effect van type werk (beperkt) significant is. Verschillen die we hierboven in tabel 14 zien met betrekking tot andere deelnemer-specifieke kenmerken, ontstaan dus vooral door correlaties tussen deze kenmerken en het type werk. Daarnaast vertoont de regio Randstad significant grotere aanpassingen dan de rest van Nederland.

Tabel 15: Analyse afname aandeel spitsreizen (afhankelijke variabele: afname aandeel spitsreizen tussen voormeting en beloningsperiode; een positieve coëfficiënt betekent dat de afname groter is dan het referentieniveau)

Deelnemers met minimum 5 reizen tijdens de voormeting en met een afname in het aandeel spitsreizen tussen voormeting en beloningsperiode	Univariaat		Multivariaat	
	Coëfficiënt	p-waarde	Coëfficiënt	p-waarde
<i>Regio (referentieniveau: Nederland Noord en Oost)</i>				
Randstad	13	0,006	11	0,018
Nederland Zuid	13	0,07	9,3	0,20
<i>Afstandsklasse (referentieniveau: 25–40 km)</i>				
≤ 25 km	–0,45	0,93	–0,26	0,96
≥ 40 km	–6,3	0,19	–5,2	0,29
<i>Type werk (ref.: uitvoerende/ondersteunende functie)</i>				
Eigenaar van een bedrijf/vrij beroep	–18	0,07	–17	0,078
Manager/Leidinggevende	–11	0,04	–9,1	0,093
Beleid/staf/advies/onderzoek	–7,3	0,06	–5,8	0,15
<i>Opleiding (ref.: wetenschappelijk onderwijs of hoger)</i>				
Geen wetenschappelijk onderwijs	6,7	0,05	4,3	0,22
<i>Inkomen (referentieniveau: hoger dan 2500)</i>				
Minder dan 2500 Euro	–2,3	0,63	–2,9	0,55
Onbekend	5,7	0,19	3,79	0,38
Constante (snijpunt y-as)	Varieert over regressies		30	0

¹² Deelnametrajecten zijn daarbij toegewezen aan de provincie met het grootste aandeel in de afstand tussen begin- en eindstation (in een rechte lijn gemeten).

We zien in tabel 14 dat het aandeel spitsreizen tijdens de voormeting het kleinst is in de regio Randstad, die de provincies Noord- en Zuid-Holland, Utrecht en Flevoland omvat; misschien komt dit doordat in de Randstad relatief meer banen zijn in sectoren die meer flexibiliteit bieden dan buiten de Randstad. De wijziging van het aandeel spitsreizen is hoogst in de regio Nederland Zuid (Zeeland, Noord-Brabant en Limburg), en het laagst in Nederland Noord en Oost (Groningen, Friesland, Drenthe, Overijssel en Gelderland). Het verschil tussen de regio Randstad en Nederland Noord en Oost is in de univariate en multivariate analyse statisch significant (tabel 15). Dit betekent dat een verschil tussen de regio's blijft bestaan als men controleert voor verschillen over de regio's op het vlak van verdeling van type werk, opleiding of inkomen.

Uit tabel 14 blijkt dat de afname van het aandeel spitsreizen het kleinst is voor de middelste afstandsklasse. Uit de analyse gepresenteerd in tabel 15 blijkt echter dat voor deze afstandsklasse de afname groter zou zijn dan voor de andere afstandsklassen, alhoewel de bijdrage van de afstandsklasse aan het verklaren van deze afname niet significant blijkt. De verklaring voor de verschillende resultaten die uit beide analyses naar voren komen, is dat er in de middelste afstandsklasse relatief minder deelnemers zijn met een afname in het aandeel spitsritten. In de OLS-analyses (tabel 15) bestuderen we enkel het gedrag van deelnemers met een afname in het aandeel spitsritten tussen voormeting en beloningsperiode, terwijl tabel 14 gemiddelde waarden presenteert voor de volledige groep deelnemers.

De invloed van het type werk op de afname van het aandeel spitsreizen blijkt statistisch significant te zijn (op het 10% significantie niveau) zowel in de univariate als in de multivariate analyse (tabel 15). Deelnemers met een ondersteunende/uitvoerende functie blijken de grootste afname in het aandeel spitsreizen te realiseren. Dit is enigszins verrassend omdat deze groep de laagste kans had om deel te nemen (tabel 3), maar kan misschien juist daardoor verklaard worden: alleen diegenen met dit type werk die ook flexibel zijn (of de spits al mijden), nemen deel.

Samenvattend kunnen wij dus zeggen dat de afname in de spitsreizen vooral door het type werk bepaald wordt en of het traject in de regio Randstad ligt.

6.4 Discrete keuzemodellen: theorie

Discrete keuzetheorie¹³ bestudeert hoe een individu een keuze maakt uit verschillende alternatieven. In onze studie zijn alternatieven gedefinieerd als verschillende momenten waarop de deelnemer de reis over het deelnametraject kan maken. Centraal in de theorie staat de aanname dat het individu zijn of haar nut maximaliseert door het alternatief te kiezen dat het grootste nut heeft voor hem of haar in de beschouwde keuzesituatie (bv. de keuze tussen verschillende vertrektijden om naar het werk te reizen). Dit nut is niet direct meetbaar, maar de keuzes die gemaakt worden zijn dat wel. En door die keuzes deels te verklaren uit kenmerken van de verplaatsing – zoals tijdstip, drukte, betrouwbaarheid, al dan niet beloond – kunnen we wel kwantitatief afleiden in welke mate die kenmerken kennelijk het nut beïnvloeden, omdat ze daadwerkelijk observeerbaar de gemaakte keuzes beïnvloeden. Er staat met nadruk “deels te verklaren”: een belangrijk kenmerk van discrete-keuzemodellen is dat deze er expliciet rekening mee houden dat een deel van de overwegingen van

¹³ Enkele standaardwerken in het domein van discrete keuzetheorie zijn Anderson, Palma en Thisse (1992); Ben-Akiva en Lerman (1985); en Train (2003).

de reiziger zal verschillen tussen individuen en niet door de onderzoeker kan worden waargenomen (het “idiosyncratische (specifieke) nut”).

In een discreet keuzemodel wordt zo voor elk keuzealternatief een bepaald nutsniveau geschat, dat voor het “systematische deel” afhankelijk is van waarneembare eigenschappen van het alternatief (bv. reistijd en tarief), en voor het “idiosyncratische deel” uit een statistische verdeling komt. Hoe beter de waargenomen kenmerken de keuzes kunnen verklaren, hoe groter het relatieve gewicht van het systematische nut wordt. Geschatte coëfficiënten geven weer hoe het systematische nut van een alternatief verandert als de waarneembare eigenschap (de verklarende variabelen die in paragraaf 6.5 worden besproken) wijzigt, en zijn daarmee een maatstaf voor de voorkeuren die het individu heeft. Deze coëfficiënten worden dus geschat op basis van gemaakte keuzes, en met name op basis van de kenmerken van de wél en niet gekozen alternatieven. De coëfficiënten worden zodanig bepaald dat het model de waargenomen keuzes zo goed mogelijk voorspelt. Bijvoorbeeld: als betrouwbaarheid er niet toe zou doen, zou de voor dat kenmerk geschatte coëfficiënt niet significant afwijken van 0.

Wanneer we de coëfficiënten van het model kennen, kunnen we de afwegingen die het individu maakt kwantificeren. Wanneer bijv. reistijd (in uur) in de nutsfunctie tien keer zwaarder weegt dan tarief (in euro), betekent dit dat voor het individu een uur extra reistijd eenzelfde wijziging van het nut en de gemaakte keuzes tot gevolg heeft als een tien euro hoger tarief. Deze manier van vergelijken van coëfficiënten laat ons toe om de geldelijke waardering vast te stellen voor verschillende eigenschappen van een keuzealternatief. Dit heeft als groot voordeel dat we het belang van de verschillende kenmerken tegen dezelfde eenheden kunnen vergelijken, namelijk het aantal euro’s dat de gemiddelde individu er voor over heeft (de “betalingsbereidheid”) om dat kenmerk te verbeteren.

Een belangrijke uitbreiding van het basismodel staat toe dat niet alleen het idiosyncratische nut maar ook de coëfficiënten die het systematische nut beschrijven niet voor alle individuen gelijk zijn. De zogeheten “latent class” modellen vormen één van de mogelijke modellen die deze spreiding van voorkeuren toelaten. Voor twee of meer (gedrags)klassen schatten we dan afzonderlijke coëfficiënten. Een aanvullend model drukt de kans uit dat een individu tot een van die twee klassen behoort. Met andere woorden: ook de toekenning van individuen aan klassen is dan een statistisch proces; het is niet zo dat een gegeven individu (bij een model met twee klassen) ofwel in de ene of wel in de andere klasse zit.

Door de met de app geobserveerde treinreizen te vergelijken met niet-gekozen alternatieven op die dag, kunnen we het feitelijke keuzegedrag van de deelnemer analyseren. Dit wordt aangeduid als RP-onderzoek: Revealed Preference, ofwel gebleken voorkeuren. De niet-gekozen alternatieven en hun kenmerken kunnen we daarbij construeren uit de logboeken van alle deelnemers met hetzelfde deelnametraject. Voor niet-deelnemers hebben we dergelijke gegevens natuurlijk niet. Om toch een inschatting te kunnen maken van hun keuzegedrag, en dit te kunnen vergelijken met dat van de deelnemers, werd voor beide groepen in een enquête een reeks fictieve keuzesituaties voorgelegd. Dit vormt SP-onderzoek: Stated Preference, ofwel geuite voorkeuren. In elke SP-keuzesituatie worden twee reisalternatieven beschreven, en wordt de respondent verzocht zijn of haar voorkeur aan te geven. In secties 6.6 (RP) en 6.7 (SP) presenteren we een verdere beschrijving van beide gegevensbronnen.

6.5 Verklarende variabelen

Zoals gezegd nemen we in de modelschattingen aan dat de keuze voor een vertrektijd mede wordt bepaald door aantal waarneembare eigenschappen van de verschillende reismogelijkheden (vertrektijden). Deze vormen de verklarende variabelen in de gedragsmodellen. Deze paragraaf licht kort toe hoe deze geoperationaliseerd zijn.

Reistijd

Dit is de reistijd zoals voorzien in de dienstregeling. De reistijd op een gegeven traject kan variëren over vertrektijden door verschillen in treinsoort (Sprinter vs InterCity), het aantal overstappen, of de route die de trein volgt tussen begin- en eindstation van het deelnametraject.

Afwijking van optimale vertrek- en aankomsttijd

De aanname is dat reizigers een voorkeur hebben voor een bepaalde vertrek- of aankomsttijd. Elke minuut dat de reiziger vroeger of later vertrekt/aankomt is dan minder gewenst dan het optimale tijdstip en wordt door die reiziger negatief gewaardeerd. Meestal vinden reizigers het vervelender te laat te komen dan te vroeg. Daarom staan de modellen toe dat de negatieve waarde van een minuut “te vroeg” anders kan zijn dan van een minuut “te laat”.

De meest gewenste vertrek- of aankomsttijd voor ochtend- en avondspits kan niet rechtstreeks geobserveerd worden. Daarom is in de enquête gevraagd naar dit gewenste tijdstip. De deelnemer werd daarbij verzocht het meest optimale tijdstip aan te geven ongeacht de dienstregeling, en uitgaande van een stipte dienstuitvoering en maximaal reiscomfort, zodat de voorkeur echt alleen het tijdstip zou betreffen voor verder volstrekt vergelijkbare treinreizen.

Reiscomfort

In deze studie gebruiken we de beschikbaarheid van een zitplaats als maatstaf voor het reiscomfort. We gebruiken hiervoor de schaal met vijf niveaus zoals die in de logboeken gebruikt is:

- Er zijn zoveel vrije zitplaatsen dat ik een zitje van 4 voor mezelf kan hebben als ik dat wil
- Er zijn zoveel vrije zitplaatsen dat ik een zitje van 2 nog wel, maar van 4 niet meer, voor mezelf kan hebben als ik dat wil
- Ik kan nog wel zitten, maar er zit ook iemand in de stoel naast me
- Niet iedereen kan zitten; een paar reizigers staan
- Er moeten veel reizigers staan

Onbetrouwbaarheid

Als maatstaf voor de (verwachte) onbetrouwbaarheid gebruiken we de kans dat de reiziger met meer dan 10 minuten vertraging aankomt op het eindstation.

6.6 Revealed preference (RP) observaties

Om de RP-modellen te kunnen schatten, gebruiken we registraties uit de app, aangevuld met logboeken en enquêtes. Essentieel is dat we voor elke keuze een keuzeset construeren, dus ook

definiëren welke alternatieven de reiziger juist niet heeft gekozen. De beschikbaarheid van reismogelijkheden is voor elke verbinding verschillend, omdat de vertrekintervallen verschillen. De keuzemogelijkheid die een reiziger heeft leiden we af uit de reisregistraties in het logboek. We combineren daarbij de registraties van alle reizigers met eenzelfde deelnametraject om tot een volledig overzicht van reismogelijkheden te komen (vertrek-/aankomsttijd en reisduur).

Deelnemers registreren in het logboek niet alleen de gekozen vertrek- en aankomsttijd, maar ook het zitcomfort in de trein en eventuele vertragingen. Ook deze informatie verwerken we, zodat we voor elke reismogelijkheid een indicator krijgen voor reiscomfort en betrouwbaarheid, uitgedrukt als de kans op een vertraging.

Om voor elke reismogelijkheid te kunnen vaststellen in welke mate deze afwijkt van het gewenste tijdstip, werd in een enquête aan de deelnemers gevraagd de meest optimale vertrek- en aankomsttijd voor heen- en terugreis op te geven.

De beschikbaarheid van reisregistraties in de logboeken/enquêtes is niet voor alle deelnemers optimaal. Sommige deelnemers hebben niet alle logboeken en enquêtes ingevuld. Daarnaast is de beschikbaarheid van logboekregistraties per traject verschillend. Om voldoende betrouwbare gegevens te hebben over gemaakte en niet-gemaakte keuzes, zijn de RP onderzoeksresultaten daarom gebaseerd op een selectie van deelnemers die deelnemen op de meest frequente trajecten (127 trajecten in ochtendspits, 114 trajecten in avondspits). Omdat dit de drukstbereden trajecten zijn, veronderstellen we dat deze selectie een redelijk representatief beeld oplevert. Bovendien zal blijken dat de schattingsresultaten erg dicht aanliggen tegen de resultaten van het SP onderzoek, waaraan alle deelnemers hebben deelgenomen.

Uit de geobserveerde verplaatsingen selecteren we op die manier 7022 treinreizen die gemaakt werden door 162 verschillende deelnemers.

6.7 Stated preference (SP) observaties

In de tussen-enquête (en voor niet-deelnemers de non-respons enquête) werd aan de respondent een reeks fictieve keuzesituaties voorgelegd, met als vraag om voor elke situatie een voorkeur voor een van de omschreven reisalternatieven aan te geven.

De fictieve keuzesituaties bestonden telkens uit twee reisalternatieven voor een reis in de ochtend naar de werklocatie. Een eerste alternatief heeft steeds een aankomsttijd eerder dan of gelijk aan de (door de respondent opgegeven) gewenste aankomsttijd. Het tweede alternatief heeft daarentegen een aankomsttijd later dan (of gelijk aan) het voor de deelnemer optimale tijdstip.

Elke deelnemer aan de enquête evalueert tien keuzesituaties, die telkens verschillen in de waarden van de verklarende variabelen (zie sectie 6.5) voor beide reisalternatieven.

In totaal werden 100 verschillende keuzesituaties ontworpen, onderverdeeld in tien sets van tien keuzesituaties, waarvan elke deelnemer aan de enquête op willekeurige basis een set gepresenteerd krijgt.

Tabel 16: Voorbeeld keuzeset (voor een deelnemer die reist tussen Leiden Centraal en Amsterdam Zuid, met een gebruikelijke reistijd van 30 minuten en een optimale aankomsttijd om 8:00)

	Verbinding A	Verbinding B
Vertrek op station Leiden Centraal	6:30	8:12
Aankomst op station Amsterdam Zuid	7:00	8:30
Reistijd	30 min	18 min
Verwachte drukte	<i>Er zijn zoveel vrije zitplaatsen dat u een zitje van 2 nog wel, maar van 4 niet meer, voor uzelf kan hebben als u dat wilt.</i>	<i>U kan nog wel zitten, maar er zit ook iemand op de stoel naast u</i>
Kans op vertraging bij aankomst groter dan 10 minuten	1 op 20	1 op 30
Beloning	0 Euro	8 Euro

De context van de keuzesituatie werd voor elke deelnemer gepersonaliseerd op basis van de informatie die de respondent eerder in de enquête had opgegeven, zoals gebruikelijke reistijd, gebruikelijk vertrek- en aankomsttijd, etc. De deelnemers kregen verder de instructie om te veronderstellen dat het aantal overstappen voor beide keuze-alternatieven gelijk is aan het aantal overstappen op hun meest gebruikelijke treinreis.

De combinaties van verklarende variabelen (die de 100 verschillende keuzesituaties definiëren) werden gekozen op basis van wat wordt genoemd een “D-optimaal factorieel ontwerp”. In technische termen betekent dit dat hierbij de determinant van de informatiematrix wordt gemaximaliseerd; de consequentie is dat de antwoorden ons zo goed mogelijk (met een zo klein mogelijke onzekerheid) in staat stellen de verschillende coëfficiënten in het systematische nut te schatten.¹⁴

Er zijn meer SP dan RP observaties voor handen. In onze analyse betrekken we SP observaties voor 720 deelnemers en 520 niet-deelnemers, met telkens 10 keuzes per respondent. Het SP keuze-experiment werd uitgevoerd eind november 2012 (tussen-enquête voor eerste wervingsronde en non-respons enquête) en in januari 2013 (tussen-enquête voor tweede wervingsronde), we verwachten bijgevolg dat seizoenseffecten geen invloed hebben op geobserveerde verschillen tussen deelnemers en niet-deelnemers.

6.8 Resultaten keuzemodellen

Uit de tijdstipkeuze modellen leiden we de afweging af die treinreizigers maken bij de keuze voor een vertrek- en aankomsttijd. Doordat de geldelijke beloning een onderdeel uitmaakt van deze afweging, kunnen we de verschillende eigenschappen monetair waarderen (in geld uitdrukken; zie paragraaf 6.4). Dit doen we in een eerste stap voor de deelnemers, die het keuze-experiment als onderdeel van een enquête hebben ingevuld. Voor de praktijkproef analyseren we eveneens de gemaakte keuzes.

¹⁴ De D-fout werd hiertoe berekend voor 5000 willekeurig bepaalde keuzesituaties uitgaande van een vooraf veronderstelde verdeling van de keuzevoorkeuren (simulatie op basis van 200 draws; de vooraf veronderstelde verdeling bleek achteraf goed aan te sluiten bij de waargenomen verdeling).

Tabel 17: Waardering van deelnemers van mogelijke wijzigingen in de treinverbinding die aanleiding geven tot een lager gewaard reisalternatief

kenmerk	waardering
<i>keuze-experiment uit enquête (Stated Preference)</i>	
lager comfortniveau: (bijna) alle zitplaatsen bezet	0,6 €
lager comfortniveau: reizigers moeten staan	3,5 €
hogere kans op vertraging (>1 op 10)	3,1 €
vroeger aankomen (vóór ochtendspits)	12,9 €/h
later aankomen (ná ochtendspits)	10,5 €/h
langere reisduur	21,4 €/h
<i>praktijkproef (Revealed Preference)</i>	
lager comfortniveau: (bijna) alle zitplaatsen bezet	(-2,1 €)
lager comfortniveau: reizigers moeten staan	(-1,7 €)
hogere kans op vertraging (>1 op 40)	36,8 €
nog hogere kans op vertraging (>1 op 10)	(0,5 €)
vroeger aankomen (ochtend)	14,7 €/h
later aankomen (ochtend)	13,8 €/h
vroeger vertrekken (avond)	11,9 €/h
later vertrekken (avond)	13,8 €/h
langere reisduur	91,4 €/h

Voor de tussen haakjes weergegeven waarden is de overeenkomstige modelcoëfficiënt niet significant verschillend van 0 bij een niveau van $p=10\%$ (volgens de gegroepeerde robuuste p -waarde). Dat betekent dat er geen statistisch voldoende significante meetbaar effect van dat kenmerk op het keuzegedrag is.

De waarderingen in tabel 17 tonen voor de meeste kenmerken van een reis een realistisch beeld; alleen voor de reisduur vinden we een ongebruikelijk hoge waardering in het RP-experiment (dit resultaat verdwijnt voor het latent class model dat hierna besproken zal worden).¹⁵ Voor een aantal kenmerken is de overeenkomstige coëfficiënt niet significant verschillend van nul; dit hebben we aangegeven door de betreffende waarde tussen haakjes te plaatsen. Om deze reden zijn de negatieve waarden voor comfortniveau in het RP-model niet werkelijk problematisch; de gevonden insignificantie geeft aan dat we hier geen betekenis aan kunnen hechten. Voor de volledigheid merken we op dat we voor de beloningscoëfficiënt een behoorlijk hoge robuuste p -waarde van 10% vinden, wat op zich twijfels kan doen reizen over de significantie van deze coëfficiënt die bepalend is voor alle waarderingen. Ook hiervoor zal het latent class model overtuigender schattingen laten zien.

Omdat voorkeuren tussen deelnemers sterk verschillen, gebruiken we vervolgens een analysetechniek die deze spreiding van voorkeuren in kaart brengt (latent class: zie paragraaf 6.4). We werken daarbij met twee klassen, en hebben vastgesteld dat voor één van beide klassen (klasse

¹⁵ In eerdere Spitsmijden projecten hebben we voor autoverplaatsingen van deelnemers op een gelijkaardige manier de afweging bepaald tussen beloning, reistijd, en aanpassing tijdstip (Knockaert et al., 2012; Peer et al., 2011).

2 in tabel 18) de op basis van de praktijkproef geschatte coëfficiënten een beeld opleveren dat sterk suggereert dat de betreffende keuzes zich niet goed met het veronderstelde model laten verklaren. Zo is de beloningscoëfficiënt er negatief (wat zou impliceren dat geld een negatief nut heeft), de onbetrouwbaarheids- en discomfortcoëfficiënten zijn er positief (wat zou betekenen dat dat aantrekkelijk wordt gevonden), en enkele coëfficiënten (waaronder reistijd) zijn er niet significant. Dit soort schattingsresultaten kan onder meer voortkomen uit het feit dat de betreffende keuzesituaties niet daadwerkelijk een keuze toelieten, maar dat het alternatief om andere redenen vaststond, onafhankelijk van de kenmerken. Dit wordt bevestigd door de coëfficiënten op basis van het keuze-experiment die wél een gebruikelijk beeld opleveren voor deze klasse. De coëfficiënten voor die klasse geven dan weliswaar een wiskundig correcte schatting van de modelspecificatie op basis van de observaties uit praktijkproef, maar hebben geen verklarende waarde over het reisgedrag voor die klasse (met een aandeel van 33% van de steekproef). Het grote voordeel van het onderscheiden van die klasse is dat het wel een realistischere beschrijving toelaat van het gedrag van de andere klasse in de praktijkproef (met een aandeel van 67% van de steekproef). Vanwege deze redenen geven we hier voor klasse 2 geen waarderingen op basis van de praktijkproef.

Tabel 18: Waardering van deelnemers van mogelijke wijzigingen in de treinverbinding die aanleiding geven tot een lager gewaard reisalternatief (latent class met 2 klassen)

kenmerk	klasse 1	klasse 2
aandeel klasse in deelnemersselectie	67%	33%
<i>keuze-experiment uit enquête (Stated Preference)</i>		
lager comfortniveau: (bijna) alle zitplaatsen bezet	(0,0 €)	1,5 €
lager comfortniveau: reizigers moeten staan	3,3 €	4,1 €
hogere kans op vertraging (>1 op 10)	3,3 €	2,3 €
vroeger aankomen (vóór ochtendspits)	16,4 €/h	6,6 €/h
later aankomen (ná ochtendspits)	8,1 €/h	14,0 €/h
langere reisduur	29,6 €/h	6,8 €/h
<i>praktijkproef (Revealed Preference)</i>		
lager comfortniveau: (bijna) alle zitplaatsen bezet	0,4 €	-
lager comfortniveau: reizigers moeten staan	1,4 €	-
hogere kans op vertraging (>1 op 40)	15,7 €	-
nog hogere kans op vertraging (>1 op 10)	2,1 €	-
vroeger aankomen (ochtend)	13,5 €/h	-
later aankomen (ochtend)	6,8 €/h	-
vroeger vertrekken (avond)	6,1 €/h	-
later vertrekken (avond)	6,2 €/h	-
langere reisduur	25,9 €/h	-

Voor de tussen haakjes weergegeven waarden is de overeenkomstige modelcoëfficiënt niet significant verschillend van 0 bij een niveau van $p=10\%$ (klassieke p -waarde); voor klasse 2 zijn geen waarderingen mogelijk op basis van de praktijkproef.

De waarderingen van deelnemers uit klasse 1 geven aan dat het keuzegedrag verrassend goed

vergelijkbaar is tussen SP en RP, vooral ook kijkend naar eerdere ervaringen met gecombineerd RP en SP onderzoek. Omdat dit bijna te mooi is om waar te zijn, zal hier in vervolgonderzoek nog kritisch naar gekeken worden.

Ook de hoogten van de waarderingen lijken plausibel. Richten we ons op de RP resultaten, dan valt wellicht op dat de waardering voor een zekere zitplaats aan de lage kant is. Maar aangezien een soortgelijke waardering ook in de SP resultaten wordt gevonden, lijkt het toch robuust te zijn. Bovendien moeten we ons bedenken dat als er een aantal mensen moeten staan, er altijd nog een behoorlijke kans is dat de reiziger zelf kan zitten. De waardering voor betrouwbaarheid is vooral groot rond een grens van 2,5% kans op een vertraging (van 10 minuten of meer). Een betalingsbereidheid van 16 Euro voor een betrouwbare trein is natuurlijk hoog, maar uit de commotie die vertraagde treinen doorgaans veroorzaken, is het ook weer niet verbazingwekkend dat die waardering hoog zou zijn; vooral voor spitsreizigers zoals we hier hebben. Bij de tijdstipwaarderingen valt vooral op dat die voor aanpassingen naar eerder in de ochtend hoog zijn. Wat hier kan meespelen is dat door de opzet van de proef, tijdstipaanpassingen die kant op vaak zullen betekenen dat er erg vroeg opgestaan moet worden. De reistijdwaardering, tenslotte, lijkt vrij hoog, maar dit kan verklaard worden uit de samenstelling van de deelnemerpopulatie (relatief hoog opgeleid), en uit het feit dat een langere reistijd vaak met zich zal meebrengen dat er overgestapt moet worden; een ongemak dat zich in deze modelspecificatie in de hogere tijdwaardering uit.

Verdere conclusies zullen we in het afsluitende hoofdstuk presenteren.

Wellicht ten overvloede benadrukken we hier opnieuw dat in het latent class model elke individuele deelnemer met een statistische kans tot elke klasse kan behoren (67% voor klasse 1 en 33% voor klasse 2).¹⁶ Het is niet zo dat 33% tot één klasse behoort en 67% tot de andere en dat we kunnen vaststellen welke deelnemer tot welke klasse behoort.

Ten slotte kijken we naar de afwegingen die niet-deelnemers maken, die het keuze-experiment in de non-respons enquête ingevuld hebben. Het keuze-experiment uit de enquête wordt hierbij op dezelfde wijze geanalyseerd als voor deelnemers. De resultaten kunnen daarom goed met elkaar vergeleken worden.

Tabel 19: Waardering van niet -deelnemers van mogelijke wijzigingen in de treinverbinding die aanleiding geven tot een lager gewaard reisalternatief

kenmerk	waardering
<i>keuze-experiment uit non- respons enquête (Stated Preference)</i>	
lager comfortniveau: (bijna) alle zitplaatsen bezet	(2,1 €)
lager comfortniveau: reizigers moeten staan	11,0 €
hogere kans op vertraging (>1 op 20)	7,8 €
vroeger aankomen (vóór ochtendspits)	66,4 €/h
later aankomen (ná ochtendspits)	42,8 €/h
langere reisduur	93,2 €/h

Voor de tussen haakjes weergegeven waarden is de overeenkomstige modelcoëfficiënt niet significant verschillend van 0 bij een niveau van $p=10\%$ (gegroepeerde robuuste p -waarde)

¹⁶ In het model gepresenteerd in dit rapport is deze kansverdeling voor alle deelnemers gelijk. In vervolgonderzoek zullen we bestuderen hoe deze kansverdeling varieert op basis van deelnemer-specifieke kenmerken.

In vergelijking met deelnemers (de SP waarderingen in tabel 17) is het voor niet-deelnemers moeilijker om het tijdstip van de reis aan te passen. De verschillende waarderingen zijn voor niet-deelnemers een factor 4 hoger. Zij zullen dus pas bij veel hogere beloningen (ongeveer 4 keer zo hoog) dezelfde gedragsaanpassingen gaan maken als de deelnemers, hetgeen natuurlijk goed verklaart waarom ze besloten hebben om niet deel te nemen. De gedragseffecten van deelnemers aan de proef zijn dus niet representatief voor wat de gemiddelde spitsreiziger zou doen, maar bij vrijwillige keuze voor een dalurenabonnement verwachten we ook het type reiziger te trekken dat relatief sterk lijkt op de deelnemers aan de proef.

7 Evaluatie

7.1 Inleiding

Tijdens de nameting werd onder deelnemers een slot-enquête afgenomen, waarin een aantal onderwerpen aan bod kwam om de deelname aan het project te evalueren:

- Deelnemers werd gevraagd naar hun reiservaringen tijdens het experiment (in het bijzonder het aantal spitsmijdingen), en hun verwachtingen over het reisgedrag na afloop van de praktijkproef.
- Een klein additioneel keuze-experiment werd opgezet om na te gaan of deelnemers geïnteresseerd zijn om een dalabonnement te kopen (en onder welke prijsvoorwaarden).
- Deelnemers werd verzocht om de opzet van de praktijkproef te evalueren, aan te geven of ze opnieuw zouden deelnemen, etc.
- Deelnemers werd gevraagd naar de mogelijke redenen waarom veel (andere) deelnemers hun deelname aan de praktijkproef voortijdig beëindigd hebben.

7.2 Gedrag tijdens en na afloop van de proef

Na afloop van de praktijkproef werd een afsluitende enquête gehouden onder 604 deelnemers. In deze enquête gaven de deelnemers onder meer aan minder vaak (17%), even vaak (56%) of vaker (27%) de ochtendspits gemeden te hebben dan voorafgaand aan het experiment verwacht. Daarmee lijkt er geen systematische over- of onderschatting vooraf te zijn geweest. De hoofdreden om dit minder vaak dan verwacht te doen is een groter dan verwachte inspanning voor een aanpassing van het reisgedrag. De hoofdreden om dit vaker dan verwacht te doen is juist het onvoorziene gemak om privé-afspraken aan te passen. Met andere woorden: de afwijking van de eigen verwachting is consistent met hoe het tijdens de beloningsperiode uitpakte voor de respondent. Voor de avondspits vonden we vergelijkbare patronen.

Eveneens werd gepeild naar de verwachtingen van de deelnemers over hun reisgedrag na afloop van het experiment. Volgens de uitkomsten van de enquête wil ongeveer 70% van de deelnemers de spits vaker mijden na afloop de proef dan voorafgaand aan deelname. Dit aantal lijkt aan de hoge kant te zijn (ook vergeleken met de gemeten 6,1% daling van het aantal spitsreizen tijdens de nameting t.o.v. de voormeting: sectie 6.3). Desondanks kunnen we uit de enquête afleiden dat een positieve ervaring met het mijden van de spits, in vergelijking met wat men kennelijk verwacht had, voor deelnemers de belangrijkste reden om dat in de toekomst vaker te gaan doen. De hoofdreden

om na afloop van de proef minder vaak te rijden dan tijdens de proef is, niet verbazingwekkend, het wegvallen van de beloning.

7.3 *Interesse dalabonnement*

In de slot-enquête hebben we ook getoetst in hoeverre deelnemers bereid zijn om van hun huidige abonnement naar een dalurenabonnement over te stappen. Een dalurenabonnement is alleen geldig voor reizen tijdens de daluren. Voor reizen tijdens de ochtend- en de avondspits (6:30–9:00, 16:00–18:30) moet men dan een spitstoeslag kopen.

We hebben elke deelnemer vijf keuzes laten maken, telkens tussen hun huidige abonnement en een dalurenabonnement. Het voorgestelde dalurenabonnement varieerde daarbij in (maandelijkse) abonnementsprijs en de prijs van de spitstoeslag. De deelnemers werd hierbij gevraagd zich voor te stellen dat zij zelf voor de abonnementen zouden betalen.

Gebaseerd op de door de deelnemers in de enquête gemaakte keuzes kunnen we een eenvoudig discreet keuzemodel schatten, waar het nut van het huidige abonnement op 0 wordt gesteld¹⁷, en het alternatieve dalurenabonnement wordt gedefinieerd door twee kenmerken:

- prijsverschil ten opzichte van de huidig abonnement (maandelijks)¹⁸
- de hoogte van de spitstoeslag

De volgende tabel toont de resultaten van de schatting van het keuzemodel.

Tabel 20: Keuzemodel dalurenabonnement

	Coëfficiënt	Robuuste t-waarde
Korting op huidig abonnement	0,0175	11,55
Spitstoeslag	–0,469	–15,68
Aantal observaties	1715	
Adj. Rho square	0,183	

Uit de coëfficiënten kan worden afgeleid dat de deelnemers een toename van de prijs van een spitstoeslag met een extra Euro aanvaardden als de prijs van het abonnement met $0,469/0,0175=26,8$ Euro/maand daalt. Dit resultaat blijkt relatief onafhankelijk te zijn van de prijs van de huidige abonnement. Deze orde van grootte lijkt aannemelijk. De 8,65 trajectreizen per week die de deelnemers zeggen te maken, vermenigvuldigd met een geobserveerd spitsaandeel van 60% en 4,35 weken per maand, suggereert dat één Euro per spitsreis uitkomt op ongeveer 22,5 Euro per maand. Het iets hogere bedrag (€ 26,80 versus € 22,50) kan ofwel ruis zijn, ofwel een inschatting van de transactiekosten¹⁹ geven. Het keuzemodel geeft daarmee aan dat respondenten behoorlijk rationeel tegen een dergelijke propositie aan zullen kijken.

17 Een alternatiefspecifiek constante bleek insignificant: er is blijkbaar geen meetbaar inherent disnut door dalurenabonnement.

18 We nemen alleen diegenen mee in de analyse die de prijs van het huidige abonnement konden/wilden aangeven (343 personen).

19 De transactiekosten zijn een maat voor de moeite die men moet doen om per reis te betalen eerder dan per maand.

7.4 Projectevaluatie

Tijdens de praktijkproef had 62% van de deelnemers contact met projectbureau. Daarbij wordt de afhandeling van de vragen/klachten door 82% als zeer goed of redelijk goed beoordeeld, en de bereikbaarheid wordt door 93% als zeer goed of redelijk goed beoordeeld. 87% van de deelnemers beoordeelde de praktijkproef in het geheel als heel duidelijk of duidelijk. Gevraagd om het minst duidelijke onderdeel werden de logboeken en nieuwsbrieven genoemd. 93% ervoer de informatie over de proef (website, persoonlijke pagina, email, nieuwsbrief) in het geheel als nuttig; de nieuwsbrieven werden als het minst nuttig ervaren.

De app wordt door 60% van de deelnemers beoordeeld als positief of vrij positief; de rest van de deelnemers oordeelt (vrij) negatief. Het grootste punt van kritiek dat naar voor komt is het effect van de app op het batterijverbruik (72% van diegene die de app als (vrij) negatief beoordelen zijn het (geheel) eens met de stelling dat de levensduur van de batterij te snel omlaag gaat als de app aanstaat). 46% van de deelnemers geeft aan dat ze de app altijd hadden aanstaan. De rest van de deelnemers geeft aan dat ze de app bij gemiddeld 77% van de dalreizen en 72% van de spitsreizen aan hadden staan. Dit komt overeen met de resultaten uit sectie 6.3, waar we concluderen dat de aantal wekelijks gemeten reizen (op persoonsbasis) tijdens de voormeting redelijk vergelijkbaar is met de aantal wekelijkse reizen tijdens de beloningsperiode. Volgens 19% van de deelnemers zijn alle reizen juist geregistreerd, de rest geeft aan dat gemiddeld 71% van de reizen juist geregistreerd werd. Sommige deelnemers ondervonden technische problemen gerelateerd met het gebruik van de app en gaven aan hierdoor beloning misgelopen te zijn.

Het valt op dat bij de beoordeling van de app er verschillen zijn tussen iOS en Android gebruikers. Zo beoordeelt 67% van de iOS gebruikers de app als (vrij) positief, voor Android is dit slechts 54%. Een van de hoofdredenen daarvoor blijkt een verschil in het aantal reizen te zijn dat nauwkeurig gemeten werd: 24% van iOS gebruikers geeft aan dat alle reizen juist zijn gemeten, terwijl voor Android slechts 14% dit doet. Ook zijn 84% van de Android gebruikers het er (geheel) mee eens dat de levensduur van de batterij snel omlaag gaat als de app aanstaat, terwijl slechts 53% van de iOS het met deze uitspraak mee eens is.

Gevraagd om opnieuw deel te nemen zou 82% (vrij) zeker nog een keer meedoen aan Spitsmijden in de Trein, 13% zegt dit (vrij) zeker niet te zullen doen, en de rest weet het niet. De belangrijkste reden om opnieuw mee te doen is dat de proef een leuke mogelijkheid biedt om geld te verdienen (66%), omdat het een interessante onderzoek is (46%), en omdat de proef een aanleiding is om over het eigen mobiliteitsgedrag na te denken (36%); er waren meerdere redenen per persoon mogelijk. De opgegeven redenen om van een nieuwe deelname af te zien zijn heel divers (reisregistratie, app, moeite met invullen logboeken en enquêtes waren belangrijk).

De betrokkenheid van de werkgever bij de proef blijkt voor 81% van de deelnemers beperkt geweest te zijn, aangezien ze geen toestemming van de werkgever nodig hadden om deel te nemen kunnen aan het proef. Echter 64% geeft aan dat hun werkgever de proef als (heel) positief ziet, slechts 6% geeft een (vrij) negatief houding aan, en 30% kent de mening van de werkgever niet. Dit betekent dat, hoewel betrokkenheid van werkgever vaak als aantrekkelijk wordt gezien bij spitsmijdenprojecten, dit niet altijd een noodzakelijke voorwaarde voor slagen hoeft te zijn.

7.5 Afgehaakte deelnemers

Er zijn verschillende verklaringen waarom deelnemers hun deelname aan de proef beëindigd hebben. Om verschillende hypothesen te toetsen vergelijken we de kenmerken van deelnemers die alle logboeken hebben ingevuld met deelnemers die wel het eerste logboek hebben ingevuld maar niet meer het laatste. Deze groepen vormen een indicatie voor de deelnemers die tot het einde deelnamen aan de praktijkproef (het invullen van de logboeken was een voorwaarde tot uitbetaling van de opgebouwde beloning) enerzijds, en anderzijds de deelnemers die wel actief meededen aan de proef aan het begin (invullen van het logboek is een indicatie daarvan) maar niet meer aan het eind van de proef.

Eerst kijken we of er een verband is tussen de beslissing om deelname aan de proef te beëindigen en de afstandsklasse waartoe de deelnemer behoort, welke bepalend was voor de hoogte van de beloning. We zien dat relatief iets meer deelnemers uit de laagste afstandsklasse met de proef gestopt is – waarschijnlijk omdat hun beloningen lager zijn, terwijl ze wel dezelfde tijd en moeite moeten opbrengen om de logboeken en enquêtes in te vullen als de deelnemers uit de hogere afstandsklassen.

Wanneer we de twee groepen deelnemers vergelijken (op basis van de ingevulde logboeken), vinden we relatief duidelijke verschillen in reisgedrag tussen beide groepen. De groep deelnemers die deelname aan de proef heeft beëindigd, heeft een gemiddeld lager aantal wekelijkse treinreizen (in de weken dat ze meededen) dan de groep die de proef heeft afgemaakt (4,0 vs. 4,8). Omdat de groep die is gestopt minder blijkt te reizen, heeft ze ook minder mogelijkheden een beloning te verdienen. Hetzelfde geldt voor het aandeel dalreizen in het totale aantal treinreizen. Ook deze zijn minder voor de groep die is gestopt dan voor de groep die de proef afmaakte, hetgeen ook het verdienen van een beloning moeilijker maakt. Dit wordt ook bevestigd door de verandering in het wekelijkse aantal spitsreizen tussen deelnemers die de proef hebben afgemaakt en diegene die dit niet hebben gedaan: de daling in spitsreizen is 26% voor de eerste groep en slechts 9,9% voor de tweede groep. Dit geeft een duidelijke indicatie dat deelnemers zijn gestopt met Spitsmijden in de Trein omdat het hen niet lukte de spits te mijden, of dat de beloning niet voldoende was om de bijkomende moeite voor deelname te willen volhouden. Verder geeft het aan dat, net zoals deelname aan de proef, ook het halverwege stoppen een keuze is die op behoorlijk rationele gronden gemaakt is.

Ten slotte kijken we of het op de smartphone gebruikte platform een invloed heeft gehad op de beslissing om met de proef te stoppen. Omdat uit de eerdere analyse in paragraaf 7.4 bleek dat iOS gebruikers meer tevreden zijn met de app, zouden we kunnen verwachten dat in de groep die is gestopt het relatief aantal iOS lager is dan onder diegene die de proef afmaakten. Dit blijkt ook de geval te zijn, maar het verschil is wel erg klein.

Tabel 21: Vergelijking kenmerken deelnemers die tot einde praktijkproef meededen met deelnemers die eerder afhaakten

	Deelnemers, die alle logboeken hebben ingevuld	Deelnemers, die niet alle logboeken hebben ingevuld
Afstandsklasse deelnametraject		
≤ 25 km	18%	21%
25–40 km	23%	18%
≥ 40 km	60%	60%
Gemiddeld aantal wekelijkse treinreizen tijdens de voormeting	4,79	4,03
% spitsreizen voor de proef	56%	62%
% spitsreizen tijdens de proef	42%	56%
Afname aantal spitsreizen per week tijdens beloningsperiode t.o.v. voormeting	26%	9,9%
% iOS (eerder dan Android)	49%	48%

8 Conclusies

Dit rapport besprak het deelname- en reisgedrag in de proef Spitsmijden in de Trein. Het inzetten van (geldelijke) beloningen blijkt, net als in eerdere Spitsmijden praktijkproeven, ook hier weer een effectief instrument om reizigers te prikkelen om hun reisgedrag aan te passen. De beloning leidt tot een significante reductie van het aantal spitsreizen onder deelnemers; de orde van grootte is een afname van het aandeel spitsreizen met 22%. Dit komt overeen met een daling van het aandeel spitsreizen in het dagelijkse totaal van de deelnemers met circa 9 procentpunt. De bevindingen uit dit project kunnen daarom bijdragen aan het verminderen van capaciteitsproblemen zoals die zich voordoen in ochtend- en avondspits. Hierbij dient uiteraard bedacht te worden dat abbonementhouders niet representatief zullen zijn voor alle treinreizigers, en vrijwillige deelnemers niet voor abbonementhouders. De percentages mogen dus niet op grotere groepen worden geprojecteerd.

De effecten van de beloning verschillen over de weekdays, en zijn het sterkst op woensdagen en vrijdagen. Ze verschillen bijvoorbeeld ook tussen regio's, type werk en sectoren. Een tussentijdse verandering van de hoogte van de beloning lijkt overigens weinig effect op het gedrag te hebben. Een mogelijke reden kan zijn dat deelnemers hun reisgedrag voor de volledige beloningsperiode slechts éénmaal hebben gewijzigd, bijvoorbeeld omdat dit afstemming thuis en op het werk vereist. Een wijziging in de hoogte van de beloning tijdens de beloningsperiode heeft dan maar een klein gedragseffect. Bij eerdere spitsmijdenprojecten bleek al vaker dat de hoogte van de beloning, als deze tijdens de proef gewijzigd werd, weinig effect sorteert. Juist omdat dit het gevolg kan zijn van het gemaakt hebben van afspraken en plannen voor de gehele duur van het experiment, mogen we niet concluderen dat de hoogte van de beloning er ook op de langere termijn relatief weinig toe zou doen.

Ook opvallend is dat het aandeel spitsreizen in de niet-beloonde nameting nog steeds lager is dan in de voormeting. Dit wijst erop dat sommige deelnemers na beëindiging van de beloningsprikkel in

elk geval niet meteen terugkeren naar hun oorspronkelijke (onbeloonde) reisgedrag. Ook dit is een patroon dat in eerdere spitsmijdenprojecten is gezien. Als reizigers door het experiment meer inzicht hebben gekregen in de voor- en nadelen van buiten de spits reizen, en met name wanneer dit minder onaantrekkelijk blijkt te zijn dan ze voorheen hadden gedacht, kan dit een blijvend effect zijn. Maar het kan ook zijn dat dit effect op de wat langere duur zal verdampen. Een meer definitief antwoord vereist een experiment met een langere nameting.

Anders dan in eerdere Spitsmijden projecten hebben wij hier gebruik gemaakt van een app om het gedrag te meten. De specifiek voor het project ontwikkelde app leverde een belangrijke bijdrage aan de verzameling van gedetailleerde reisinformatie. De inzet van deze experimentele meetmethode in de praktijkproef bleek echter ook een hele uitdaging. Verder hebben wij drie enquêtes uitgevoerd onder deelnemers, en één onder niet-deelnemers, en deelnemers hebben om de 6 weken een logboek ingevuld over hun reisgedrag.

De respons op de uitnodiging om deel te nemen aan de proef was laag in vergelijking met Spitsmijdenprojecten op de weg. Het blijkt dat deelnemers en niet-deelnemers op objectieve kenmerken niet zo sterk verschillen, maar wel wat betreft type werk, sector, provincie en trajectlengte, en ook op zelf-gerapporteerde flexibiliteit. De patronen hierbij zijn goed verklaarbaar, en geven daarmee aan dat de propositie waarschijnlijk voldoende duidelijk was. Reizigers waren eerder geneigd deel te nemen als ze sowieso al vaker buiten de spits of tegen de randen van de spits reizen, en wanneer ze flexibeler zijn in tijdstipaanpassingen – wat op zijn beurt weer deels samenhangt met het type werk en opleidingsniveau. Er lijkt overigens nog wel enige rek te zitten in het deelnamepotentieel; in die zin kan het zinvol zijn om nieuwe wervingsmethodes te testen in een vervolgproject.

Tijdens de praktijkproef werd een rijke verzameling informatie over de deelnemers en hun reisgedrag opgebouwd. Naast in geld uitgedrukte waarderingen van reistijd, comfort, betrouwbaarheid en tijdstipaanpassing, blijkt hieruit dat deelnemers ook zonder beloning vaker tijdens de daluren zouden kunnen gaan reizen, mits treinverbindingen beter zijn buiten de spits. Dit kan bereikt worden op verschillende wijzen, zoals (een mix van) frequentere verbindingen, langere treinen, directe verbindingen, minder vertragingen, of meer comfort en rust. Maar het is een interessante vraag of het bieden van beloningen een meer kosteneffectieve manier is om spitscapaciteit vrij te spelen, dan deze andere manieren om dit te bereiken. En, uiteraard, kosteneffectiever dan het uitbreiden van spitscapaciteit. Het gedragseffect van beloningen dat we in deze studie gemeten hebben, stelt vervoerders beter in staat dergelijke afwegingen te maken.

Het geregistreerde reisgedrag van de deelnemers met en zonder beloning stelt ons in staat om modellen te schatten die aangeven hoe zij hun keuzes in het reisgedrag maken. Soortgelijke modellen zijn geschat op basis van enquêtes, voor zowel deelnemers als niet-deelnemers. De uitkomsten van deze modellen geven aan in welke mate verschillende kenmerken van een verplaatsing het keuzegedrag beïnvloeden. Doordat één van die kenmerken de hoogte van de beloning is, kunnen we de waardering voor de andere kenmerken in geld uitdrukken, en daarmee onderling vergelijkbaar maken. Dit levert de volgende inzichten op:

- Het werkelijke gedrag en de enquêtes leveren een onderling consistent beeld op van de afwegingen die deelnemers in hun reisgedrag maken. Die afwegingen zijn ook goed verklaarbaar (plausibel, rationeel), en geven daarmee onder meer aan dat deelnemers zich

goed bewust waren van de beloningsstructuur en daar hun gedrag door hebben laten beïnvloeden.

- Uit het werkelijke gedrag kunnen we afleiden dat deelnemers de volgende afgeronde waarderingen toekennen aan de verschillende kenmerken van een verplaatsing:
 - Een langere reisduur wordt negatief ervaren tegen een waarde van € 26 per uur
 - Eerder reizen dan op het meest gewenste moment wordt negatief ervaren tegen een waarde van € 13 per uur voor 's ochtends vroeger dan gewenst aankomen op het werk en € 6,1 per uur voor 's middags vroeger dan gewenst vertrekken van het werk
 - Later reizen dan het meest gewenste moment wordt ook negatief ervaren; in dit geval tegen waarden van € 6,8 in de ochtendspits, en € 6,2 in de avondspits
 - Grotere drukte in de trein wordt ook negatief ervaren maar slechts in beperkte mate: een bedrag van € 0,4 compenseert reeds voor een zodanige drukte dat maar net iedereen kan zitten; een bedrag van € 1,4 compenseert voor een drukte waarbij reizigers moeten staan
 - Een grotere kans op vertragingen (groter dan 10 minuten) wordt negatief gewaardeerd: € 16 voor het verschil tussen kansen onder versus boven een 1 op de 40, en nog een verdere € 2,1 voor het verschil tussen kansen onder versus boven een 1 op de 10.
- Deze waarderingen kunnen in de eerste plaats gebruikt worden voor het vormgeven van dalurenabonnementen of spitsmijdenarrangementen, en het voorspellen van de gedragseffecten. Ze geven bijvoorbeeld aan dat het veranderen van het begintijdstip van de ochtendspitsperiode een twee keer zo groot effect op gedrag zal hebben dan het veranderen van het eindtijdstip, en hoe de effecten van deze verschuivingen zich verhouden tot veranderingen in de dalbeloning. Dit alles volgt uit de vergelijking tussen € 13 met de € 6 per uur.
- De waarderingen kunnen ook gebruikt worden om investeringen in verbeteringen van het treinaanbod te rationaliseren. Inzicht in de waarde die het voor de reizigers heeft als een bepaalde verbetering wordt doorgevoerd is zowel van belang voor private investeringen, die terugverdiend zullen moeten worden, als voor publieke investeringen, die gerechtvaardigd zullen worden gevonden als de maatschappelijke bate groter is dan de kost.
- Niet-deelnemers kennen andere waarderingen dan deelnemers. Zij blijken pas bij veel hogere beloningen (ongeveer 4 keer zo hoog) dezelfde gedragsaanpassingen te gaan maken als de deelnemers, hetgeen natuurlijk goed verklaart waarom ze besloten hebben om niet deel te nemen. De gedragseffecten van deelnemers aan de proef zijn dus niet representatief voor wat de gemiddelde spitsreiziger zou doen, maar bij vrijwillige keuze voor een dalurenabonnement verwachten we ook het type reiziger te trekken dat relatief sterk lijkt op de deelnemers aan de proef.

Gedurende de praktijkproef is een aantal deelnemers afgehaakt. Als redenen hiervoor werd door hen gewezen op de inspanning om enquêtes en logboeken in te vullen, de behoefte aan meer

communicatie en ondervonden problemen met reisregistraties. In een vervolginzet van Spitsmijden in de trein zijn dit eveneens aandachtspunten bij het maximaliseren van de effectiviteit van het instrument.

9 Referenties

Arnott, R., Palma, A. de and Lindsey, R. (1993). A Structural Model of Peak-Period Congestion: A Traffic Bottleneck with Elastic Demand. *American Economic Review*, 83(1), 161–179.

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

Knockaert, J., Tseng, Y., Verhoef, E.T. en Rouwendal, J. (2012). The Spitsmijden experiment: A reward to battle congestion. *Transport Policy*, 24, 260-272.

Peer, S., Verhoef, E.T., Knockaert, J., Koster, P.R. en Tseng, Y. (2011). Long-Run vs. Short-Run Perspectives on Consumer Scheduling: Evidence from a Revealed-Preference Experiment among Peak-Hour Road Commuters. (TI Discussion Papers 11-181/3). Amsterdam: Tinbergen Institute (TI).

Spitsmijden (Augustus 2009). Effecten van belonen in Spitsmijden in het OV: Hoe verleid je OV-reizigers? (tech. rapport). Samenwerkingsverband Spitsmijden

Train, K. E. (2003). *Discrete Choice Methods with Simulation*. Cambridge, United Kingdom: Cambridge University Press.

Vonk Noordegraaf, D., de Kruijff, J. en Vonk, T. (Augustus 2013) Spitsmijden in de trein – werkgeversonderzoek (tech. rapport). Delft, Nederland: TNO.



*Spitsmijden
in de trein*

