

De omvang en het effect van het gebruik van mobiele apparaten in een longitudinale studie

Toon Zijlstra – Kennisinstituut voor Mobiliteitsbeleid (KiM) – toon.zijlstra@minienm.nl

Sascha Hoogendoorn-Lanser – KiM – sascha.hoogendoorn@minienm.nl

Krisje Wijgergangs – KiM – krisje.wijgergangs@minienm.nl

Bijdrage aan het Colloquium Vervoersplanologisch Speurwerk 23 en 24 november 2017, Gent

Samenvatting

Online vragenlijsten worden niet meer enkel ingevuld met traditionele apparaten, zoals de laptop en desktop. Tegenwoordig zijn online vragenlijsten per definitie mixed-mode vragenlijsten. Respondenten gebruiken tal van apparaten, waarbij de tablet en de smartphone de voornaamste nieuwkomers zijn. Er zijn echter aanmerkelijke verschillen tussen deze apparaten, zoals schermgrootte, invoermethode en connectiviteit. Bovendien zien we verschillen in het gebruik. De desktop staat vast, terwijl tablet en smartphone een mobiel karakter hebben. Dit heeft mogelijk consequenties voor de datakwantiteit, datakwaliteit en de last op de respondent. In deze bijdrage bestuderen we op gedegen wijze de impact van mode-effecten in het Mobiliteitspanel Nederland (MPN), maar de aanpak leent zich voor een brede toepassing. Impact is gedefinieerd als effect maal omvang.

Over de jaren 2013 tot 2016 zien we, qua omvang, een sterke opmars van de smartphone respondent, van 4% naar 18% van het totale aantal respondenten. Wanneer de huidige trends doorzetten verwachten we tegen 2023 50% desktop of laptop, 15% tablet en 35% smartphone respondenten.

Het profiel van de tablet en smartphone respondent wijkt duidelijk af van de traditionele online respondent. Beide kunnen zodoende niet zondermeer met elkaar vergeleken worden. Onder de smartphone respondenten zien we relatief veel vrouwen, jongeren, sociale en functionele internetgebruikers. De smartphone wordt minder snel gehanteerd door hoogopgeleiden en plattelandsbewoners. Deze inzichten worden in onze paper onder meer gebruikt om te controleren voor zelfselectie-effecten bij het modelleren van de mode-effecten. Via nearest neighbor matching worden mensen met hetzelfde profiel aan elkaar gekoppeld, waardoor de vertekening via zelfselectie geminimaliseerd wordt.

Qua mode-effecten kijken we naar vier mogelijke effecten die mogelijk relevant zijn binnen het MPN: straightlining, aantal gegeven antwoorden, ontwijkend gedrag door respondenten en de doorlooptijd. Vrijwel alle resultaten zijn niet-significant. Hetgeen impliceert dat er nauwelijks sprake is van verschillen in de data als gevolg van het gehanteerde apparaat bij het invullen van de vragenlijst. Dat is een geruststellende bevinding.

1. Een mix van traditionele en nieuwe apparaten bij online vragenlijsten

De komst van internet heeft voor een revolutie gezorgd in de wijze waarop vragenlijsten worden afgenomen. Inmiddels wordt de meerderheid van de enquêtes online afgenomen. De intrede van de tablet en de smartphone geven die revolutie een nieuw gezicht. Niet langer zijn de online vragenlijsten voorbehouden aan gebruikers van desktop en laptop. Inmiddels is één op de drie respondenten bij panels uit Noordwest Europa een tablet of smartphone respondent (Zijlstra, Wijgergangs, & Hoogendoorn-Lanser, 2017).

Onderzoekers die gebruik maken van online vragenlijsten worden geconfronteerd met een nieuwe realiteit, die van de gemengde apparaten in één enkele online steekproef (Toepoel & Lugtig, 2015). De verschillen tussen die apparaten zijn aanzienlijk. Voor desktops is een beeldscherm van een 19 tot 24 inch de standaard. Ondertussen zijn de beeldschermen van kleinere smartphones niet veel groter dan 6 inch. Tablets en smartphones kennen een aanraakscherm, terwijl desktop en laptops doorgaans bestuurd worden middel muis en toetsenbord. De traditionele PC is veelal uitgerust met een 1 Gbit LAN-kabel, terwijl de smartphone het moet hebben van 3G, 4G of wifi. Waar het gebruik van internet in de thuissituatie mogelijk afgekocht is, kan de gebruiker van de smartphone een rekening ontvangen voor iedere megabyte. De fysieke beweegruimte voor de desktop gebruiker is van nature beperkt – het is niet eenvoudig om de volledige installatie te verplaatsen, terwijl de smartphone van nature mobiel is.

Onvoldoende aandacht voor het bestaan van mobiele respondenten in online vragenlijsten heeft in het recente verleden tot zorgen geleid over de kwantiteit en kwaliteit van de ingewonnen data (Callegaro, 2010). Puur het gebruik van een smartphone of tablet leidt mogelijk al tot andere resultaten dan men verkregen zou hebben via het gebruik van traditionele apparaten. Vertekeningen in de data als gevolg van de invulwijze worden *mode-effecten* genoemd.

In deze paper bestuderen we de impact van mode effecten in vragenlijsten. Daarbij maken we de vergelijking tussen smartphone of tablet enerzijds en traditionele apparaten (desktop en laptop) anderzijds. De impact van die nieuwe apparaten in de wereld van online vragenlijsten wordt uitgedrukt als het product van het relatieve aandeel mobiele apparaten en de waargenomen veranderingen als het gevolg van het gebruik van die apparaten. Beide aspecten verdienen voldoende aandacht.

Het vervolg van dit paper is als volgt geordend: in het volgende deel (sectie 2) gaan we in op de gebruikte data en de diverse methoden en technieken. Voorts presenteren we cijfer en prognoses voor het gebruik van mobiele apparaten, kijken we naar het profiel van de mobiele respondent en geven we de resultaten m.b.t. de mode effecten. In sectie 4 sluiten we af met de voornaamste conclusies bij dit onderzoek.

2. Data en werkwijze

2.1. Data van het MobiliteitsPanel Nederland

De data voor deze studie komt uit het Mobiliteitspanel Nederland (MPN) (Hoogendoorn-Lanser, Schaap, & OldeKalter, 2015). Het MPN bestaat in de basis uit een korte huishoudvragenlijst, een uitgebreide persoonsvragenlijst en een dagboek waarin het reisgedrag over een periode van drie dagen wordt vastgelegd. De persoonsvragenlijst biedt de meeste aanknopingspunten voor de onderzoeksdoeleinden bij dit artikel. Van het MPN zijn inmiddels vier waves aan persoonsvragenlijsten beschikbaar. De huishoudvragenlijst kent ook vier waves, maar is eenvoudiger qua vraagtypes, aanzienlijk korter en bovendien hebben we hierbij minder data voor de analyses, omdat er slechts één per huishouden wordt ingevuld. KANTAR PUBLIC voert voor het MPN het veldwerk uit en heeft de diverse vragenlijsten en het dagboekje geprogrammeerd.

In de persoonsvragenlijst van het MPN zijn verschillende type vragen present (Tabel 1). Belangrijk uitgangspunt bij het opstellen van de vragenlijst was eenvoud en gemak van invullen, om de belasting op de respondent te beperken. Complexe taken worden zoveel mogelijk vermeden. Dominant in de afgelopen vier waves zijn de meerkeuzevragen met enkele of meerdere antwoordmogelijkheden, eventueel in combinatie met een open antwoordveld ('anders, namelijk ...') of een exclusief antwoord ('geen van deze opties' / 'weet ik niet'). Vragen met enkel een open veld voor het antwoord zijn schaars en moeten soms voldoen aan een specifiek format, zoals bij het invullen van het kenteken, of moeten voldoen aan een bepaalde waarde. Die vragen zijn dus eigenlijk semi-open. Bij het invullen van de vragenlijst krijgen respondenten minimaal 4 matrixvragen, afhankelijk van de routing en de wave kunnen dit er meer zijn. Bij één van de matrixvragen zijn meerdere opties mogelijk per rij, waardoor de vergelijking met andere matrices lastig is.

Tabel 1: Vraagformats binnen het MPN

Type	Subtype	Optioneel
Meerkeuze vraag	<i>één antwoordmogelijkheid</i>	<i>Escape / Anders, namelijk ...</i>
	<i>meerdere antwoordmogelijkheden</i>	<i>Escape / Anders, namelijk ...</i>
Semi-open velden	<i>open</i>	<i>Escape</i>
	<i>vast formaat</i>	<i>Escape</i>
	<i>waarden binnen bepaald interval</i>	<i>Escape</i>
Matrixvragen	<i>één antwoord per rij</i>	<i>Escape</i>
	<i>meerdere antwoorden per rij</i>	
	<i>left-right matrix</i>	

Er kan moeilijk een goed overzicht gegeven worden van de presentatie van de vragenlijst over de gehele periode, omdat deze presentatie afhankelijk is van het gebruikte apparaat, de gebruikte browser, de wave en eventuele instellingen door de respondenten zelf. De presentatiewijze is regelmatig een punt van aandacht geweest, mede vanwege de opmars van mobiele apparaten in het panel. Daarom zijn er tussen de waves ook diverse aanpassingen doorgevoerd. Dit maakt de analyses tussen waves complex. We beperken ons dan ook tot analyses binnen een wave.

In de presentatie van de vragenlijst wordt vanaf de derde wave gewerkt met een adaptieve design. De presentatiewijze voor tablet en smartphone gebruikers wijkt af van de lay-out die de desktop gebruikers te zien krijgen. In Figuur 1 wordt de presentatie van dezelfde vraag over de drie waves weergegeven, daarbij is steeds dezelfde browser en hetzelfde traditionele apparaat gebruikt.

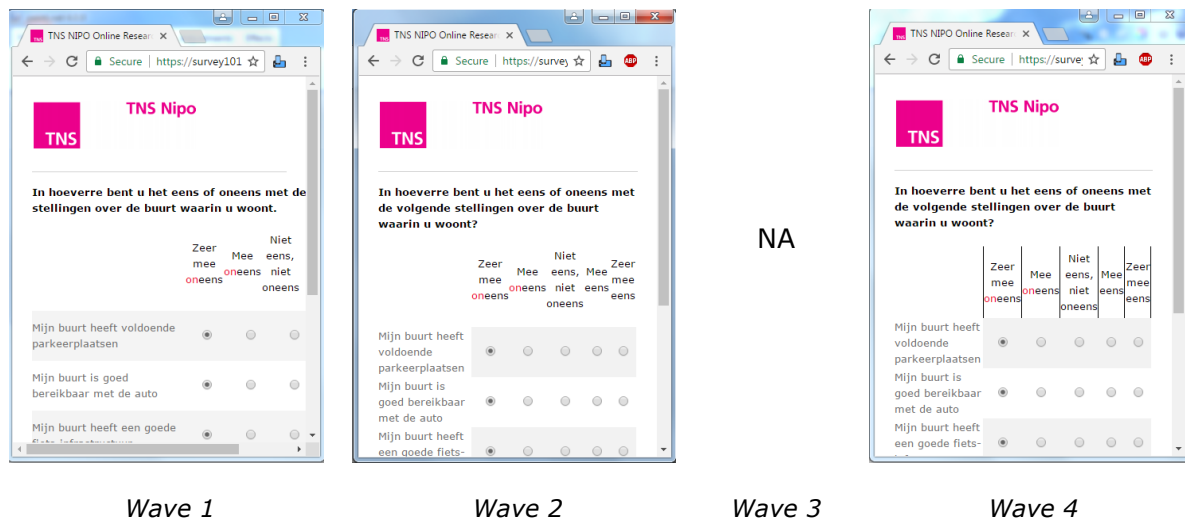


Fig. 1: Presentatiewijze van dezelfde vraag over drie waves

Het apparaat dat gebruikt wordt voor het invullen van de vragenlijst kan op verschillende manieren geïdentificeerd worden. Naast dat respondenten zelf kunnen aangeven welk apparaat men gebruikt, kan er ook informatie over het apparaat ingewonnen worden via de *user agent string (UAS)*. Hierin is onder andere informatie beschikbaar over het besturingssysteem, de webbrowsers en beeldformaat. Voor de identificatie van apparaten in het MPN wordt gekeken naar de aanwezigheid van een aanraakscherm en het beeldschermformaat via deze UAS. Bij twijfel wordt het meest waarschijnlijke apparaat geregistreerd. Zodoende kent iedere complete observatie in de dataset een apparaat. Het identificatieproces is uitgevoerd door KANTAR PUBLIC.

2.2. Indicatoren voor mode-effecten

Wanneer men mode-effecten wil vaststellen in observationele studies ligt het meer voor de hand om meerdere indicaties voor mode-effecten te combineren en vervolgens tot een synthese te komen. Dat is dan ook precies wat enkele voorgaande studies op dit vlak gedaan hebben (Lugtig & Toepoel, 2016; Struminskaya, Weyandt, & Bosnjak, 2015). Dat is ook wat wij doen in dit artikel.

Indicatoren voor mode-effecten zijn bijvoorbeeld de opmerkingen die respondenten maken over de vragenlijst, de waardering die men uit voor de vragenlijst, uitval tijdens invullen of de doorlooptijd. De laatst genoemde indicator geldt algemeen als een goede proxy voor de belasting op de respondent en is veelal positief gecorreleerd met uitval en non-response (Crawford, Couper, & Lamias, 2001; Galesic & Bosnjak, 2009; Rolstad, Adler, & Rydén, 2011). Wij zullen daarom de doorlooptijd onderzoeken.

Verder kijken we naar een drietal meetfouten die kunnen optreden als gevolg van mode-effecten:

- **Straightlining.** Wanneer respondenten niet differentiëren in de gegeven antwoorden, vormen deze antwoorden bij matrixvragen een rechte lijn, zodoende wordt er veelal over straightlining gesproken. Binnen de persoonsvragenlijst in het MPN zijn een beperkt aantal matrixvragen beschikbaar.
- **Ontwijkende antwoorden.** Het nalaten van het beantwoorden van een vraag is een directe meetfout, omdat er niets geregistreerd wordt. Omdat respondenten verplicht zijn om bij iedere vraag een antwoord te geven bij het MPN lijkt dit criterium niet relevant voor deze studie. Pogingen om zonder antwoorden verder te gaan wordt niet geregistreerd. Echter, respondenten hebben wel de mogelijkheid om bij meerdere vragen gebruik te maken van een 'escape' in de vorm van een optie 'dat weet ik niet / wil ik niet zeggen'. Een bovengemiddeld gebruik van die optie op een bepaald type apparaat kan worden aangemerkt als meetfout.
- **Aantal antwoorden bij meerdere opties mogelijk.** Bij vragen zoals 'welke vervoersmiddelen heeft u tot uw beschikking?' of 'welke rijbewijzen heeft u?' bestaan de antwoorden uit tal van mogelijkheden. Het is aan de respondent om alle opties te selecteren die van toepassing zijn. De volledigheid van de gegeven opties is hierbij van belang. Enkelvoudige antwoorden zijn niet noodzakelijk fout, maar kennen een grotere kans om incompleet te zijn. Kortom, het aantal gegeven antwoorden is een redelijke proxy voor de volledigheid van het antwoord.

2.3. Verwachtingen ten aanzien van de mode-effecten

We verwachten dat het gebruik van ontwijkende antwoorden onder de smartphone respondenten binnen het MPN hoger is (H1). Hier is een directe praktische reden voor: de knop met de 'weet niet' optie is in het gebruikte software pakket geplaatst aan de onderzijde van de pagina. Dat is regelmatig buiten het eerste beeld voor smartphone respondenten, terwijl diezelfde knop voor veel desktopgebruikers (en tablet gebruikers) direct zichtbaar zal zijn. Bovendien wordt de optie lang niet altijd aangeboden. Respondenten kunnen er zodoende niet op rekenen. Het gevolg is dat smartphone gebruikers zich meer gedwongen zullen voelen om te kiezen tussen de andere opties.

Twee eerder studies naar mode-effecten in online vragenlijsten van Lugtig en Toepoel (2016) en Smidt en Wenzel (2013) laten beide zien dat straightlining vaker voorkomt bij PC respondenten. Een mogelijke verklaring hiervoor is dat het trekken van een niet-rechte lijn bij onwelwillendheid voor mensen met een aanraakscherm gemakkelijker is. In lijn met deze studies verwachten wij ook om meer non-differentiatie te vinden onder de respondenten die een traditioneel apparaat gebruiken (H2).

Ten aanzien van het aantal antwoorden bij meerdere antwoorden mogelijk verwachten we dat gebruikers van apparaten met een aanraakscherm gemiddeld meer antwoorden verstrekken, mits de vragenlijst geoptimaliseerd is voor die mobiele apparaten (H3a). Het gemak waarmee opties aangetikt worden, ligt hieraan te grondslag. Wanneer de

vragenlijst niet geoptimaliseerd is, verwachten we juist minder antwoorden door mobiele respondenten (H3b). In de literatuur vonden we op dit punt weinig houvast. Eén studie naar mode-effecten onderzocht dit aspect en vond geen significant verschil tussen smartphones, tablets en traditionele apparaten (Lugtig & Toepoel, 2016).

Uit de literatuur weten we dat mobiele respondenten meer tijd gebruiken voor het invullen van de vragenlijsten. Uit een meta-analyse van Zijlstra et al. (2017) blijkt de mediale doorlooptijd voor smartphone respondenten 20% langer te zijn, wanneer die vragenlijsten reeds geschikt gemaakt zijn voor smartphones. Couper and Peterson (2017) hebben hier een drietal verklaringen voor: langere laadtijd van de pagina's, er zijn meer handelingen nodig (o.a. scrollen, zoomen of draaien), en meer afleiding vanwege het mobiele karakter van de smartphone. Wij verwachten dan ook dat smartphone respondenten een significant langere doorlooptijd hebben (H4).

2.4. Zelfselectie effecten mitigeren via matching

Het gebruik van een bepaald apparaat voor het invullen van een vragenlijst is vrijblijvend. Deelnemers kunnen zelf bepalen welk apparaat men hanteert. Binnen het MPN wordt er geen advies gegeven omtrent de invulapparaten. Er worden ook geen apparaten toegewezen of geblokkeerd. Een consequentie is dat het zeer onwaarschijnlijk is dat het gebruik van de apparaten volstrekt willekeurig verdeeld is over de bevolking. Dat wordt ook ondersteund door de literatuur (De Bruijne & Wijnant, 2014; Toepoel & Lugtig, 2014; Zijlstra et al., 2017).

Omdat het mensen vrij staat om het apparaat te hanteren van de eigen voorkeur op het moment van invullen verwachten we zelfselectie-effecten. De gebruikers van een smartphone, tablet of desktop hebben andere karakteristieken. Dit stelt ons voor het probleem dat mode-effecten niet van andere verschillen te onderscheiden zijn, zonder hiervoor de juiste technieken te gebruiken. We weten bijvoorbeeld niet of straightlining typisch iets is voor jongeren of voor smartphonegebruikers. Andere effecten moeten geneutraliseerd worden om de mode-effecten bloot te leggen. De wijze waarop we dat in deze studie doen is door het gebruik van datavoorbewerking via *nearest neighbor matching*.

Binnen een perfect geblokt en gerandomiseerd experiment zou het mogelijk zijn om mode-effecten gemakkelijk te detecteren. Deelnemers met gelijke profielen worden daarbij tegenover elkaar geplaatst en het te gebruiken apparaat wordt willekeurig opgelegd. Via matching trachten we deze ideale setting zo goed mogelijk te benaderen. In ons geval worden de gebruikers van een smartphone of tablet tegenover gebruikers van traditionele apparaten geplaatst met hetzelfde profiel. Het ultieme doel van matching is het verkrijgen van een optimaal gebalanceerde steekproef tussen de mensen met 'een behandeling' (vragenlijst invullen met smartphone/tablet) of zonder behandeling (vragenlijst invullen met traditioneel apparaat). Een voorwaarde van een effectieve match is een zo compleet mogelijk beeld van de relevante aspecten die effect hebben op de zelfselectie (Dehejia & Wahba, 1999; Ho, Imai, King, & Stuart, 2007). Verschillen in die aspecten worden bij matching zonder exacte match gebruikt als afstandsmaat.

Er zijn verschillende matchingstechnieken die allen gebruikt kunnen worden om de data voor te bewerken (Ho, Imai, King, & Stuart, 2011). Een verzamelbegrip van diverse benaderingen is propensity score matching (Caliendo & Kopeinig, 2008; Rosenbaum & Rubin, 1983). Technieken zijn o.a. full, generic en optimal matching. Die technieken zijn met name relevant wanneer de meer eenvoudige benadering van exact matching onvoldoende soelaas biedt. In ons geval is exact matching ontoereikend, omdat de pool van controle cases niet oneindig groot is, wij veel criteria gebruiken en wij intervalvariabelen hanteren als matching criteria. De kans op een exacte match wordt te klein en het verlies van observaties te groot.

Bij *nearest neighbor* matching worden mensen met een 'experimentele behandeling' (gebruik van mobiel apparaat) gekoppeld aan mensen uit de controlegroep (gebruik traditioneel apparaat) op basis van de meest sterke gelijkens voor alle gehanteerde matchingscriteria. Die meest sterke gelijkens wordt bepaald op basis van een afstandsmaat. In dit geval een logistische regressie met de matchingscriteria (§3.2). Alle mensen uit de behandelde groep worden gekoppeld. De nutteloze observaties in de controlegroep gaan verloren. Bijvoorbeeld wanneer leeftijd één van de criteria is en er geen mensen met behandeling zijn boven de 50 jaar, zullen er in de uiteindelijke dataset waarschijnlijk geen observaties zijn van personen boven de 50 jaar. Wanneer er een onbalans bestaat tussen het aantal mensen met en zonder 'experimentele behandeling' wordt deze onbalans weggewerkt middels gewichten. Het resultaat is idealiter een perfect gebalanceerde steekproef, zoals eerder gezegd. Voor het voorbereiden van de dataset gebruiken we het pakket *MatchIt* in "R" (Ho et al., 2011).

De verhouding tussen het aantal mensen met en zonder behandeling staat vrij bij nearest neighbor matching. Wij hebben gekozen voor een verhouding van één-op-één. Dat is voornamelijk ingegeven door de wens om dezelfde techniek toe te passen in iedere wave. We koppelen mensen alleen binnen dezelfde wave, omdat de verschillen tussen de waves aanzienlijk zijn. In wave 4 zijn er dermate veel gebruikers van tablet en smartphone dat andere verhoudingen, meerdere mensen uit de controlegroep per 'behandeld' persoon, niet haalbaar zijn.

3. Resultaten en discussie

3.1. Stroomtabel

In Tabel 2 geven we de transities van wave naar wave binnen het MPN. Het gebruik van de traditionele apparaten (desktop en laptop) domineert en is goed voor een duidelijke meerderheid van de participanten. Vanaf de eerste wave zien we een gestage afname van het gebruik van de traditionele apparaten: van 81% in 2013 naar 67% in 2016. Het gebruik van de tablet is gemiddeld 15% en redelijk constant over de jaren. Het gebruik van de smartphone groeit jaar op jaar ten kosten van de traditionele apparaten. In 2013 was het gebruik slechts 4% in 2016 is dit toegenomen tot 18% van de respondenten binnen het MPN.

Gebruikers van de traditionele apparaten hebben een relatief grote kans om de opeenvolgende keer opnieuw voor een desktop of laptop te kiezen. Deze kans is

bovendien opvallend stabiel (86%). Het gebruik van tablet en vooral van smartphone heeft een veel grilliger karakter. De consistentie voor smartphonegebruik wisselt tussen de 39% en 57%. Voorts zien we dat smartphone gebruikers een hogere kans op uitval vertonen (+9 procent punten). Dat houdt mogelijk verband met het profiel van deze gebruikers (§3.2).

Tabel 2: Veranderingen in apparaat gebruik tussen de waves

		Wave + 1				TOTAAL	Consistentie	
		PC	Tablet	Smartph.	NP	Abs. (%)	zelfde	ander
2013	PC	906	91	57	1619	2673 (81%)	86%	14%
	Tablet	60	136	8	275	479 (15%)	67%	33%
	Smartphone	12	3	20	109	144 (4%)	57%	43%
	NP	3065	653	249				
	TOTAAL	4043	883	334				
2014	PC	1498	134	116	2295	4043 (77%)	86%	14%
	Tablet	135	207	52	489	883 (17%)	53%	47%
	Smartphone	70	14	53	197	334 (6%)	39%	61%
	NP	978	167	195				
	TOTAAL	2681	522	416				
2015	PC	1187	93	96	1305	2681 (74%)	86%	14%
	Tablet	109	123	24	266	522 (14%)	48%	52%
	Smartphone	65	27	87	237	416 (11%)	49%	51%
	NP	3190	752	1044				
	TOTAAL	4551	995	1251				

NP = niet present; Consistentie verwijst naar gebruik van dezelfde of ander apparaat in de eerstvolgende wave

Op basis van de stroomtabel en de daaruit ontleende trends (consistentie, uitval, nieuwkomers, groei en krimp) kunnen we ook voorzichtig prognoses opstellen voor de verhoudingen tussen de invuldevices voor komende waves van het MPN. De resultaten van een dergelijk model suggereren dat we geleidelijk bewegen naar een nieuw evenwicht (Fig. 2). De nieuwe quasi-stabiele situatie treedt in vanaf wave 8 of 9. Dan hebben we 50% van de respondenten met een traditioneel apparaat, 15% kiest een tablet en de rest gebruikt een smartphone. Een meer complexe benadering van dit probleem zou ook rekenschap geven aan de onderliggende sociaal-demografische karakteristieken van de verschillende apparaatgebruikers. Die karakteristieken bespreken we in het volgende deel.

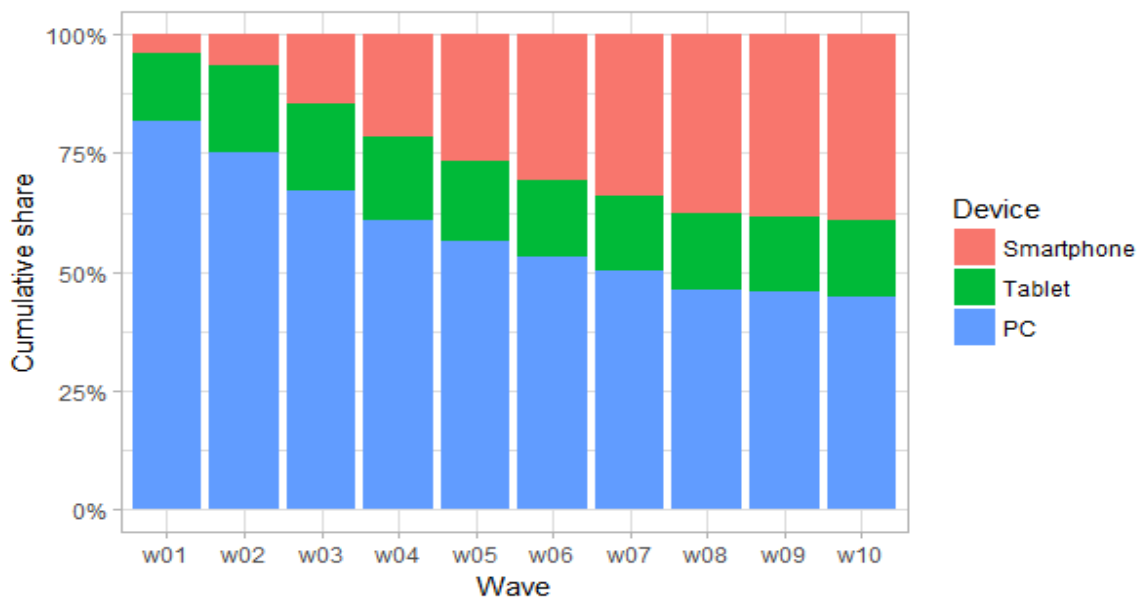


Fig. 2: Prognose voor invulapparaten binnen het MPN

3.2. Profiel van de smartphone en tablet respondent

Aan de hand van een multinomiaal logistisch (MNL) model is het mogelijk om meer grip te krijgen op de verschillen in de gekozen apparaten tussen respondenten. Dat inzicht is noodzakelijk willen we in staat zijn om te controleren voor zelfselectie.

Op basis van de internationale literatuur weten we dat leeftijd en geslacht belangrijke kenmerken zijn (Zijlstra et al., 2017). Het zijn vooral jongeren en vaker vrouwen die gebruik maken van smartphone of tablet. Educatie en stedelijkheidsgraad zijn vaak ook relevant, maar de resultaten zijn inconsistent. Toegang tot bepaalde apparaten zal een natuurlijke beperking zijn tot het gebruik, zodoende nemen we het bezit van smartphone of tablet mee, evenals het aantal traditionele apparaten (= desktops + laptops) in het huishouden. Uit eerder studies weten we dat de algemene attitude richting technologische innovaties en de meer concrete affiniteit met de nieuwe mobiele apparaten van belang is (De Bruijne, 2015). Dat inzicht is in ons model vertaald naar drie internetgebruikersprofielen. We onderscheiden sociale, plezier en functionele internetgebruikers op basis van een 11-tal vragen over gebruiksfrequenties. Dan gaat het bijvoorbeeld over het gebruik van facebook (sociaal), online video's kijken (plezier) en vergaderen op afstand (functioneel). Bijzondere bron vanuit het MPN zijn de vele gegevens over verkeer en vervoer. Hiervan hebben we diverse variabelen getoetst op relevantie. Omdat we de data uit de diverse jaren stapelen en omdat we reeds observeerden dat er aanzienlijke verschillen zijn tussen de waves, is het noodzakelijk om ook het jaar van dataverzameling op te nemen. Voor alle kenmerken geldt dat we uiteindelijk enkel de meest relevante variabelen hebben behouden in het definitieve model dat hieronder gepresenteerd wordt.

De resultaten van het multinomiaal regressie model zijn gepresenteerd in Tabel 3. De model fit (Pseudo- $p^2 = 0,11$) is goed te noemen voor een dergelijk categorisch model en een heterogene dataset en in vergelijking met soortgelijke modellen (De Bruijne &

Wijnant, 2014; Toepoel & Lugtig, 2014). Op een geaggregeerd niveau kunnen we stellen dat de toegevoegde waarde van de determinanten het grootst is bij de keuze voor de smartphone. Daar zien we de meeste significante verschillen en extremere schattingen voor de parameters. Dat impliceert dat de smartphone respondent sterker afwijkt van de desktop respondent dan de tablet respondent dat doet. Dat onderstreept nogmaals de noodzaak tot het controleren voor de zelfselectie.

Tabel 3: Resultaten van MNL model in invulapparaat (n=13545)

Categorie	Niveau	Smartphone	Tablet
		coef. (s.f.)	coef. (s.f.)
Leeftijdsgroep	25-34	0,512 (0,098)***	-0,307 (0,106)**
	35-44	0,193 (0,098)*	-0,436 (0,098)***
	45-54	-0,573 (0,103)***	-0,277 (0,086)**
	55-64	-0,752 (0,12)***	-0,489 (0,097)***
	65-74	-1,49 (0,205)***	-0,398 (0,114)***
	75 +	-1,789 (0,299)***	-0,569 (0,141)***
Geslacht	vrouw	0,199 (0,061)**	0,147 (0,052)**
Opleidingsniveau	hoog	-0,276 (0,072)***	-0,104 (0,062).
Internetgebruiksprofiel	sociaal	0,159 (0,044)***	0,001 (0,038)
	functioneel	0,278 (0,042)***	0,072 (0,038).
Toegang tot apparaten	smartphone	0,822 (0,12)***	0,169 (0,071)*
	tablet	0,291 (0,065)***	1,481 (0,067)***
	aantal trad. in hh	-0,157 (0,025)***	-0,183 (0,022)***
Mobiliteitsaspecten	dagelijks fiets	-0,017 (0,069)	0,156 (0,057)**
	dagelijks auto	0,087 (0,072)	0,126 (0,06)*
	brommerbezit	0,251 (0,108)*	-0,095 (0,104)
	auto's in hh	0,034 (0,037)	0,141 (0,031)***
Stedelijkheidsgraad	landelijk	-0,259 (0,11)*	-0,463 (0,099)***
Wave	2014	0,573 (0,138)***	0,158 (0,079)*
	2015	1,014 (0,135)***	-0,061 (0,086)
	2016	1,545 (0,12)***	0,023 (0,075)
Constante		-3,448 (0,21)***	-2,502 (0,153)***

sig. * 0.05, ** 0.01, ***0.001

We observeren significante verschillen tussen de desktop- en smartphone-respondenten met betrekking tot geslacht, leeftijd, opleidingsniveau, internetgebruiksprofiel en toegang tot apparaten. De kans op het gebruik van de smartphone piekt onder jongvolwassenen tussen de 25 en 35 jaar. Onder oudere leeftijdsgroepen neemt het gebruik trapsgewijs snel af. Vrouwen gebruiken eerder de smartphone, maar hoger opgeleiden juist weer minder snel. Sociale internetgebruikers, maar vooral functionele internetgebruikers hebben een hogere kans op smartphonegebruik voor het invullen van de vragenlijsten. Toegang tot een smartphone is, zoals verwacht, van cruciaal belang. Echter ook toegang tot een tablet en het aantal traditionele apparaten in het huishouden hebben een significant effect: meer desktops en laptops in huis betekent minder kans op invullen met de smartphone. Van wave naar wave zien we een duidelijke jaarlijkse toename, zoals we eerder al registreerden.

Het gebruik van de tablet voor het invullen van de vragenlijsten is meer uniform over de steekproef verdeeld, maar kent het hoogste aandeel onder de referentiegroep van 16 tot 25 jaar. Opnieuw ligt het aandeel vrouwen hoger, ten opzichte van de gebruikers van traditionele apparaten. We zien weinig relevante verschillen met betrekking tot de wave en het internetgebruiksprofiel. Toegang tot de tablet is van doorslaggevend belang. En opnieuw neemt de kans op tablet gebruik af naar mate er meerdere pc's of laptops in het huishouden aanwezig zijn. Meerdere mobiliteitsaspecten blijken geassocieerd met tabletgebruik. Mensen die dagelijks de auto of de fiets nemen blijken een hogere kans op tablet gebruik te hebben, dan mensen die dit niet doen. Met het aantal auto's in het huishouden stijgt ook de kans op het gebruik van een tablet voor de vragenlijst. Een mogelijk achterliggende verklaring hiervoor is een hogere inkomen in het huishouden. Inkomen werd niet direct opgenomen, vanwege de relatief veel missende observaties. Tot slot zien we nog een scherpe afname in het tablet gebruik onder de mensen die woonachtig zijn in de meest landelijke gemeenten van Nederland.

De resultaten van het multinomiaal logistisch regressie model geven ons de set van matchingscriteria. Voor het matchen van de traditionele device respondenten met smartphone respondenten hanteren we de volgende set: leeftijdsgroep, geslacht, opleidingsniveau hoog, internetgebruikersprofielen sociaal en functioneel, smartphone bezit, tablet bezit, aantal traditionele apparaten in huishouden, brommerbezit en plattelandsbewoner. Voor de match tussen traditioneel en tablet gebruiken we de volgende set: leeftijdsgroep, geslacht, smartphone bezit, tablet bezit, aantal traditionele apparaten in huishouden, auto- en fietsgebruik en het aantal auto's in het huishouden.

3.3. Mode effecten in het MPN

In dit laatste deel met resultaten presenteren we de resultaten ten aanzien van de mode-effecten (Tabel 4). Achtereenvolgens zullen de ontwijkende antwoorden, straightlining, het aantal antwoorden en doorlooptijd behandeld worden. Per onderdeel lichten we het gebruikte model toe. In ieder model zijn alle matchingscriteria meegenomen en eventueel relevante controle variabelen, zoals het totale aantal vragen. De schattingen voor deze additionele variabelen worden niet gepresenteerd, vanwege het ruimtebeslag en de betrekkelijke relevantie. We concentreren ons op de mode-effecten hier.

Tabel 4: De schattingen van 30 regressie modellen voor mode-effects

		Wave 1 (2013)	Wave 2 (2014)	Wave 3 (2015)	Wave 4 (2016)
Apparaat		coef. (s.f.)	coef. (s.f.)	coef. (s.f.)	coef. (s.f.)
Weet ik niet	Smartphone	n.d.	-0,12 (0,76)	-0,07 (0,97)	0,17 (0,51)
	Tablet	n.d.	-0,13 (0,38)	0,28 (1,60)	0,18 (0,51)
Non-differentiatie	Smartphone	-0,23 (0,53)	-0,06 (0,31)	-0,02 (0,36)	0,17 (0,12)
	Tablet	-0,07 (0,24)	0,02 (0,17)	0,03 (0,27)	0,08 (0,14)
Aantal antwoorden	Smartphone	0,06 (0,07)	0,04 (0,47)	0,00 (0,03)	0,00 (0,02)
	Tablet	0,02 (0,04)	0,01 (0,03)	0,03 (0,03)	-0,01 (0,02)
Doorlooptijd	Smartphone	0,09 (0,08)	0,34 (0,06)***	0,09 (0,06)	0,16 (0,02)***
	Tablet	0,01 (0,04)	0,08 (0,03)*	0,01 (0,04)	0,10 (0,02) ***

n.d. = onvoldoende data; sig. * 0.05, ** 0.01, ***0.001

Omdat niet iedereen evenveel vragen gepresenteerd krijgt, door verschillen in de routing, is het aantal ontwijkende antwoorden gedeeld door het totale aantal antwoorden. De afhankelijke variabele wordt zodoende het aandeel ontwijkende antwoorden. Door middel van logistische regressiemodellen voor alle waves en beide mobiele devices is gekeken of er verschillen bestaan tussen traditionele en mobiele respondenten. Voor wave 1 bleek onvoldoende data beschikbaar, er waren nauwelijks ontwijkende antwoorden in de gehele dataset van wave 1. Voor de overige waves zien we ook een beperkt gebruik van de escape opties binnen het MPN. De schattingen zijn niet significant en bovendien inconsistent over de waves (Tabel 4). We kunnen zodoende veilig concluderen dat er geen mode-effecten zijn voor de ontwijkende antwoorden binnen het MPN (H1).

Ook voor de modellen die straightliners moeten blootleggen, werkten we met de verhouding van het aantal keer dat een rechte lijn getrokken wordt door de antwoorden in een matrixvraag ten opzichte van het totale aantal gelegenheden om dat te doen. Opnieuw is een logistisch regressiemodel gebruikt voor de schatting. Opvallende observatie is dat non-differentiatie regelmatig voorkomt. Een kleine groep trekt in de meerderheid van de matrixvragen een rechte lijn (Fig. 3) De resultaten van de regressie analyses zijn opnieuw allemaal niet-significant (Tabel 4). Hetgeen impliceert dat er geen verschillen zijn op basis van het gebruikte apparaat. Er is echter wel een beperkte en niet significante trend zichtbaar. Voor zowel de smartphone als de tabletrespondenten zien we een kanteling van negatief naar positieve coëfficiënt. Terwijl andere studies wel effecten vonden, vinden wij deze niet (H2).

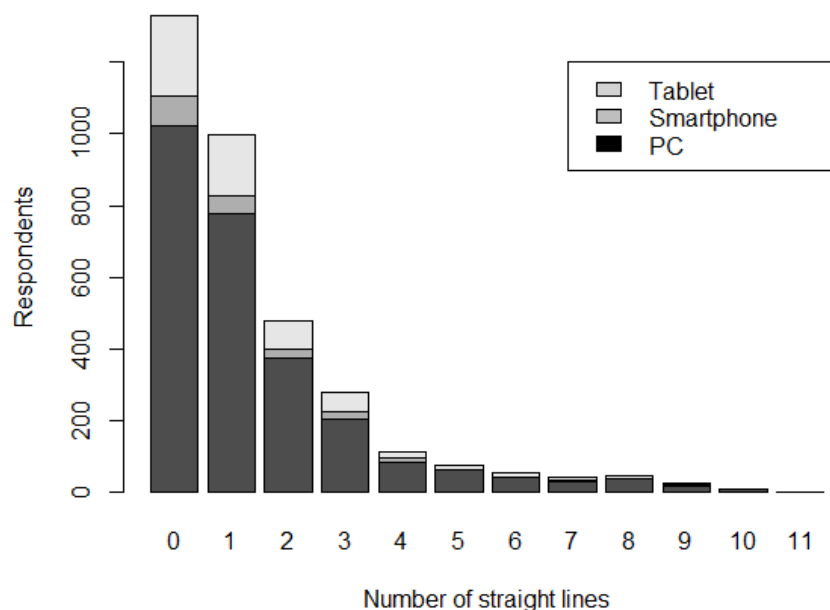


Fig. 3: Histogram met aantal keer straightlining (max.= 11)

Het aantal antwoorden bij meerdere opties mogelijk is gemodelleerd aan de hand van een tel-model (poisson regressie). Daarbij zijn alle generieke vragen met meerdere antwoorden mogelijk samengevoegd. Het resultaat is opnieuw niet significant. De schattingen liggen opvallend dicht bij de nul. Kortom, er is geen reden om aan te nemen dat smartphone of tablet respondenten meer (of minder) antwoorden verstrekken (H3a &

H3b). Deze conclusie ligt in lijn met de eerdere bevindingen van Lugtig en Toepoel (2016).

Het verschil in doorlooptijd is vastgesteld aan de hand van een regressiemodel. De afhankelijke variabele is daarbij getransformeerd tot quasi-normale verdeling door het logaritme te nemen van de doorlooptijd. In het model is gecontroleerd voor het aantal vragen. De resultaten laten zien dat mobiele respondenten langer doen over het invullen van de vragenlijsten. Alle schattingen zijn positief en een aantal sterk significant. Daarbij zien we wel dat het voornamelijk smartphonerespondenten zijn die meer tijd nodig hebben. Onze bevindingen ten aanzien van de doorlooptijd zijn in lijn met de bevindingen van andere onderzoekers (H4).

4. Conclusies

Online vragenlijsten zijn tegenwoordig per definitie mixed-mode; zowel traditionele apparaten als de laptop en desktop en moderne mobiele apparaten als de tablet en smartphone worden gebruikt voor het invullen van de vragenlijst. Dit heeft een mogelijk impact op de datakwaliteit. Die impact wordt bepaald door de omvang van de diverse groepen en de effecten op diverse criteria voor datakwaliteit. In dit paper hebben we de impact van deze recente ontwikkelingen voor het MobiliteitsPanel Nederland in kaart gebracht.

Achtereenvolgens hebben we gekeken we naar het gebruik van diverse apparaten voor het invullen van de jaarlijkse persoonsvragenlijst, de trends die zich daarbij aftekenen, het profiel van de smartphone- en tablet-respondent en de effecten die toegeschreven kunnen worden aan het gebruik van de smartphone of tablet bij het invullen van de vragenlijsten. Om de vergelijking tussen smartphone- of tablet-respondenten enerzijds en desktop-respondenten anderzijds eerlijker te maken, gebruikten we matchingtechnieken, waarbij het profiel van de smartphone- of tablet-respondent de basis vormt om respondenten aan elkaar te koppelen.

Wanneer we de eerste vier waves van het MPN achter elkaar plaatsen zien we de opmars van de smartphone ook terug in onze data. Die opmars gaat vooral ten koste van de traditionele apparaten, de laptops en desktop computers. Het aandeel tablet-respondenten is min of meer gelijk over de diverse waves. Op basis van de aandelen, de consistentie en de verschuivingen in het apparaat gebruik hebben we ook prognoses opgesteld. Wanneer de huidige trends zich doorzetten hebben we in de tiende wave, in 2023, 50% PC, 15% tablet en 35% smartphone.

Het gebruik van bepaalde apparaten is niet aselectief over de steekproef verdeeld. Er is sprake van zelfselectie. Smartphone gebruikers zijn onder andere jonger, vaker vrouw en gebruiken regelmatig het internet voor sociale en functionele activiteiten. Tablet-respondenten vinden we opvallend minder vaak onder plattelandsbewoners en juist meer onder bij huishoudens met meerdere auto's. Deze en andere observaties zijn gebruikt om respondenten aan elkaar te koppelen via nearest neighbor matching.

De voornaamste conclusie van deze studie is dat enige angst voor smartphone of tabletgebruik bij online vragenlijsten ongegrond is. Van de vier mode-effecten die we bestudeerden zien we slechts één indicator met een significant verschil. Smartphone-respondenten en in minder mate tablet-respondenten doen iets langer over het invullen van de vragenlijsten.

Het bescheiden aantal significante resultaten voor de mode-effecten is binnen deze context een positief signaal. Het impliceert immers dat er geen duidelijke mode-effecten zijn waarvoor gecontroleerd zou moeten worden bij inhoudelijke analyses met de data.

Daarnaast heeft deze studie meerdere andere nuttige resultaten opgeleverd. We presenteerden een multinomiaal logistische regressie model voor de keuze van het apparaat bij het invullen van de vragenlijst. Een dergelijk model hebben wij niet eerder aangetroffen in de literatuur over mixed-mode surveys. Verder stapelden we inzichten uit meerdere jaren, voegden we relevante internetgebruikersprofielen en mobiliteitsaspecten toe. Ook daarmee is dit werk uniek.

De groei van het aantal smartphone-gebruikers en de langere doorlooptijd die zijn nodig hebben is voor ons een signaal om de vragenlijsten zoveel mogelijk compact te houden. Opknippen van langere vragenlijsten is een optie die nadere afweging vraagt. De analyses van de mode-effecten hebben inzichtelijk gemaakt dat er additionele aanknopingspunten zijn om andere datasets op te schonen. Er zijn blijkbaar respondenten die in (bijna) alle matrices een rechte lijn naar beneden trekken. We kunnen redelijkerwijs aannemen dat zij ook bij andere vragen weinig inspanning leverden om tot geschikte antwoorden te komen.

Verschillen in meetfouten tussen verschillende apparaten kunnen niet zondermeer worden toegeschreven aan de verschillen tussen die apparaten zelf (mode-effecten). In dit onderzoek hebben we laten zien dat er sterke zelfselectie-effecten van kracht zijn. Ook de software is van belang. De standaard browsers op de smartphone is niet noodzakelijk hetzelfde als de gebruikte browsers op de tablet of de desktop computer, terwijl de browser ook invloed heeft op de verschijningsvorm van de vragenlijst. Verschillen tussen studies kunnen bovendien nog eens worden toegeschreven aan de uiteenlopende software voor web-based en app-based surveys. Kortom, het lijkt gepast om terughoudend te zijn wanneer het gaat om mode-effecten.

Opmerking:

Een uitgebreide Engelstalige versie van deze bijdrage werd ook ingediend voor het European Transport Conference, oktober 2017 in Barcelona, Spanje.

Referenties

- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, 22(1), 31-72.
- Couper, M. P., & Peterson, G. J. (2017). Why do web surveys take longer on smartphones? *Social Science Computer Review*, 35(3), 357-377.
- Crawford, S. D., Couper, M. P., & Lamias, M. J. (2001). Web surveys: Perceptions of burden. *Social Science Computer Review*, 19(2), 146-162.

- De Bruijne, M. (2015). *Designing web surveys for the multi-device internet*. (PhD thesis), Tilburg University, School of Economics and Management, Tilburg.
- De Bruijne, M., & Wijnant, A. (2014). Mobile response in web panels. *Social Science Computer Review*, 32(6), 728-742.
- Dehejia, R. H., & Wahba, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American statistical Association*, 94(448), 1053-1062.
- Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public opinion quarterly*, 73(2), 349-360.
- Gummer, T., & Roßmann, J. (2015). Explaining interview duration in web surveys: A multilevel approach. *Social Science Computer Review*, 33(2), 217-234.
- Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political analysis*, 15(3), 199-236.
- Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42(8), 1-28.
- Hoogendoorn-Lanser, S., Schaap, N. T., & OldeKalter, M.-J. (2015). The Netherlands Mobility Panel: An innovative design approach for web-based longitudinal travel data collection. *Transportation Research Procedia*, 11, 311-329.
- Lugtig, P., & Toepoel, V. (2016). The use of PCs, smartphones, and tablets in a probability-based panel survey: Effects on survey measurement error. *Social Science Computer Review*, 34(1), 78-94.
- Rolstad, S., Adler, J., & Rydén, A. (2011). Response burden and questionnaire length: is shorter better? A review and meta-analysis. *Value in Health*, 14(8), 1101-1108.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.
- Struminskaya, B., Weyandt, K., & Bosnjak, M. (2015). The Effects of Questionnaire Completion Using Mobile Devices on Data Quality. Evidence from a Probability-based General Population Panel. *Methods, Data, Analyses*, 9(2), 261-292. doi: 10.12758/mda.2015.014
- Toepoel, V., & Lugtig, P. (2014). What happens if you offer a mobile option to your web panel? Evidence from a probability-based panel of Internet users. *Social Science Computer Review*, 32(4), 544-560.
- Toepoel, V., & Lugtig, P. (2015). Online surveys are mixed-device surveys. Issues associated with the use of different (mobile) devices in web surveys. *Methods, Data, Analyses*, 9(2), 8.
- Zijlstra, T., Wijgergangs, K., & Hoogendoorn-Lanser, S. (2017). *Computer assisted web-interviewing with mixed-devices: a longitudinal panel study perspective*. Paper presented at the International Conference on Transport Survey Methods 2017, Estérel, Canada.