



Centraal Bureau
voor de Statistiek

Rapport

WoON 2018

**Onderzoeksdocumentatie en
Kwaliteitsanalyse**

Saskia Janssen-Jansen
Concept versie 0.1

Projectnummer 302697
SER sector SAL
12-03-2019

kennisgeving De in dit rapport weergegeven opvattingen zijn die van de auteurs en komen niet noodzakelijk overeen met het beleid van het Centraal Bureau voor de Statistiek.

Inhoudsopgave

1. Inleiding	5
1.1 Doel van dit document	5
2. Doel, achtergrond en geschiedenis van het onderzoek.....	6
2.1 Doel	6
2.2 Historie van WoON	6
3. Dataverzameling.....	7
3.1 Onderzoeksopzet	7
3.2 Benaderstrategie	8
3.3 Resultaten veldwerk WoON 2018	9
4. Verwerking (data processing)	10
4.1 Invoer en uniformeren veldwerkbestand	11
4.2 Contoles en consistentie checks	13
4.3 Typeren	14
4.4 Verrijken	14
4.5 Afleiden	15
4.6 Imputeren	17
5. Non-responsanalyse	18
5.1 Selectiviteit naar achtergrondkenmerken	18
5.2 Item-respons voor bepaalde vragen en thema's	20
6. Plausibiliteitsanalyse	22
7. Weging.....	23
7.1 Weging Woningmarktmodule WoON 2018	23
7.2 Extra weegfactoren	23

Supplement 1:	Onderzoeksverantwoording WoON2018
Supplement 2a:	Vragenlijst
Supplement 2b:	Vragenlijstschema
Supplement 3:	Variabelen die gevraagd zijn in WoON2015 en WoON2018
Supplement 4:	Specificaties C&C
Supplement 5:	Bestandsacceptatie WoON2018
Supplement 6:	Inkomengegevens WoON2018
Supplement 7:	Variabelen gekoppeld vanuit administratieve bronnen
Supplement 8:	Specificaties inkomens- en woonlastenafleidingen WoON2018
Supplement 9:	Specificaties huishoud en huisvesting afleidingen WoON2018
Supplement 10:	Syntax van alle afleidingen in WoON2018
Supplement 11:	Imputatie WoON2018
Supplement 12:	Plausibiliteit WoON2018
Supplement 13:	Weging WoON2018

1. Inleiding

1.1 Doel van dit document

In dit document wordt verslag gedaan van het onderzoek WoON 2018. De voorbereiding, dataverzameling en verwerking van de gegevens worden beschreven. Daarnaast dient dit document als naslagwerk voor onderzoekers die meer inzicht wensen in de achtergronden van het onderzoek en het tot stand komen van de analysebestanden. In hoofdstuk 2 wordt het doel, achtergrond en historie van het WoON beschreven. Daarna wordt in hoofdstuk 3 uitgebreid ingegaan op de dataverzameling. Deze bestaat uit een korte beschrijving van de getrokken steekproef, de verantwoording van de uitgezette steekproef, de in het veld geconstateerde kaderfouten, en de gerealiseerde responsaantallen.

Hoofdstuk 4 gaat over de in de dataprocessing en wordt beschreven hoe het veldwerkbestand wordt omgezet naar het eindbestand van WoON2018 . In hoofdstuk 5 komt de non-responsanalyse aan de orde. Daarnaast wordt nagegaan in hoeverre er sprake is van selectiviteit in de respons naar achtergrondkenmerken. Daarna wordt in hoofdstuk 6 naar de plausibiliteit gekeken door de uitkomsten met andere bronnen te vergelijken en wordt in hoofdstuk 7 ingegaan op de weging.

2. Doel, achtergrond en geschiedenis van het onderzoek

2.1 Doel

Het WoonOnderzoek Nederland (WoON) is een groot woningmarktonderzoek dat elke drie jaar resultaten levert waarmee de staat van de Nederlandse woningmarkt in beeld wordt gebracht. Enkele belangrijke thema's in het WoON zijn: de woonlasten, betaalbaarheid, woonwensen, verhuisgedrag en leefbaarheid. Het WoON is een grootschalig onderzoek dat door BZK en CBS samen wordt uitgevoerd

Het WoON kent een modulaire opzet. De hoofdmodule is de Woningmarktmodule gebaseerd op een vragenlijst met genoemde thema's die door ongeveer 60.000 respondenten wordt beantwoord. Daarnaast zijn er thematische vervolgmodes. Een belangrijke vervolgmodes is de Verhuismodule (2006 t/m 2015) waarmee op basis van registraties jaarlijks het feitelijke verhuisgedrag van de respondenten van de Woningmarktmodule wordt gemeten. Daarnaast wordt in 2018 net als in 2006 de Energiemodule uitgevoerd als een vervolgmodes op WoON.

De steekproef van de Woningmarktmodule bestaat uit een regulier deel en een oversampling. Op basis van het reguliere deel van de Woningmarktmodule kunnen betrouwbare uitspraken worden gedaan op landelijk, provinciaal en COROP niveau. Om de bruikbaarheid van het WoON voor lokale overheden te vergroten kunnen provincies, gemeenten, regio's en woningcorporaties via de oversampling meedoen met het onderzoek. De steekproef wordt opgehoogd, waardoor lokale overheden en woningcorporaties kunnen beschikken over betrouwbare beleidsinformatie voor hun interessegebied. Dit onderzoeksverslag heeft betrekking op zowel het reguliere deel als de oversampling van de Woningmarktmodule.

2.2 Historie van WoON

Het WoON heeft een traditie van meer dan tien onderzoeken, uitgevoerd sinds 1964. De huidige vorm sluit aan bij het WoON van 2006, 2009, 2012 en 2015. Voor 2006 werd het Woning Behoeftte Onderzoeken (WBO) uitgevoerd. Door de jaren heen is de inhoud van het WBO/ WoON continu verbeterd en geactualiseerd, in aansluiting op ontwikkelingen in het woonbeleid. Nieuwe thema's worden toegevoegd, andere raken op de achtergrond, terwijl voor het merendeel van de thema's de ontwikkeling in de tijd gevolgd wordt. Het onderzoek waarborgt daarmee continuïteit en blijft tegelijk inspelen op relevante thema's op een bepaald moment.

3. Dataverzameling

De dataverzameling voor het landelijke deel van de Woningmarktmodule 2018, in onderliggende document WoON-regulier genoemd, is door het CBS uitgevoerd vanaf half augustus 2017 tot en met april 2018. Het veldwerk moet aan een groot aantal randvoorwaarden voldoen. Deze randvoorwaarden zijn samen met BZK door het CBS opgesteld en bevatten onder andere afspraken over te behalen respons, het steekproefontwerp, de benaderstrategie en diverse kwaliteitseisen. Na het afstemmen van de randvoorwaarden wordt de opzet van het onderzoek vastgesteld en wordt de start van het veldwerk voorbereid. De dataverzameling voor de oversampling, in onderliggende document WoON-oversampling genoemd, is uitgevoerd vanaf begin oktober 2017 tot en met april 2018.

In dit hoofdstuk wordt kort ingegaan op de onderzoeksopzet (3.1) en de uitvoering van het veldwerk (3.2). In supplement 1 is de dataverzamelingsproces in detail beschreven.

3.1 Onderzoeksopzet

De onderzoeksopzet wordt bepaald voorafgaand aan het veldwerk. In deze fase wordt het steekproefontwerp, de vragenlijst en de waarneemstrategie vastgesteld.

3.1.1 Doelpopulatie en Steekproefontwerp

De doelpopulatie van het WoON bestaat uit de in Nederland woonachtige personen van 18 jaar of ouder op 1 januari 2018, die deel uitmaken van particuliere huishoudens. Om de woonsituatie van de populatie goed in kaart te kunnen brengen moet het veldwerkbestand aan een groot aantal randvoorwaarden voldoen, wat leidt tot een complex steekproefontwerp. Voor WoON-regulier zijn voor 26 geografische gebieden (19 normaal verdeelde gebieden en 7 krimpgebieden) responsafspraken gemaakt. Om deze afspraken na te komen zijn er ruim 73 duizend brieven verstuurd met het verzoek aan personen om deel te nemen aan het WoON. De verwachte respons, het deel van de aangeschreven personen dat de vragenlijst invult, is 42,5 duizend. Naast WoON-regulier zijn er ook responsafspraken gemaakt voor 18 oversampling-deelnemers waarbij voor sommige deelnemers nog onderscheid gemaakt is naar een deelgebied of -populatie. Om te voldoen aan de responsafspraken van de oversampling-deelnemers zijn ongeveer 40 duizend extra personen benaderd.

In totaal zijn er ruim 110 duizend personen benaderd met het verzoek om deel te nemen aan WoON 2018. In de uitvoering van het veldwerk wordt geen onderscheid gemaakt tussen WoON-regulier en WoON-oversampling. Steekproeven voor beide delen worden daarom samengevoegd tot één steekproef.

3.1.2 Vragenlijstontwerp

De vragenlijst van WoON 2018 is gebaseerd op de vragenlijst van WoON 2015. De belangrijkste wijzigingen in de vragenlijst ten opzichte van WoON 2015 zijn de volgende:

- *Vragen gesteld aan mensen met een verhuismens*
In WoON 2015 zijn personen zonder concreet beeld over een verhuizing niet bevroegd over kenmerken van de gewenste verhuizing. Hierdoor is een trendbreuk in de cijfers ontstaan. In WoON 2018 zijn daarom vragen over een aantal basale kenmerken van de gewenste verhuizing gesteld aan alle personen met een verhuismens.

- *Vragen over de hypotheek*
De vragen in het vragenblok over hypotheek in WoON 2015 waren te moeilijk voor veel respondenten. Informatie over hypotheek is daarom beperkt uitgevraagd in WoON 2018 en zal later, indien beschikbaar, uit registers worden aangevuld.
- *Nieuwe vragenblokken: energie, woningonderhoud, zoektijden sociale huurwoningen*
In WoON 2018 is een nieuw vragenblok opgenomen over maatregelen die mensen hebben genomen (of willen nemen) om hun huis energiezuiniger te maken. Daarnaast is er een vragenblokje opgenomen over woningonderhoud en woningverbeteringen (buiten en binnen) in de afgelopen 5 jaar. Ook zijn extra vragen gesteld aan mensen die in de afgelopen 2 jaar naar een sociale huurwoning zijn verhuisd en aan mensen die op zoek zijn naar een sociale huurwoning. Deze vragen gaan over i) hoe lang ze al op zoek zijn, ii) hoe lang ze staan ingeschreven, en iii) hoe vaak ze op een woning gereageerd hebben.

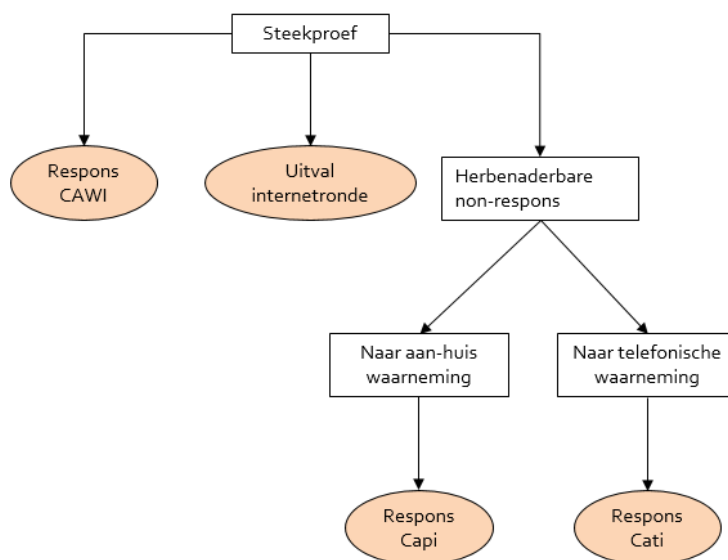
In supplementen 2a en 2b is de vragenlijst opgenomen inclusief een vragenlijstschema, waarin te volgen is welke vragen worden gesteld aan welke personen.

3.2 Benaderstrategie

De eerste stap in de dataverzameling is het versturen van een aanschrijfbrief naar de personen uit de steekproef van WoON 2018. In deze brief staat het verzoek om via internet deel te nemen aan het onderzoek (CAWI). Aan de aanschrijfbrief wordt een onvoorwaardelijke incentive toegevoegd in de vorm van een VVV-cadeaubon ter waarde van €5. Twee weken na het versturen van de aanschrijfbrief wordt een herinnering (rappebrief) verstuurd aan personen die op dat moment de vragenlijst nog niet hebben ingevuld en niet aan het CBS hebben laten weten niet te kunnen of willen deelnemen aan het onderzoek. Een week later wordt een tweede rappebrief verstuurd.

Steekproefpersonen die een maand na ontvangst van de eerste aanschrijfbrief de vragenlijst niet hebben ingevuld en niet aan het CBS hebben doorgegeven niet te kunnen of willen deelnemen worden herbenaderd. Dit gebeurt telefonisch (CATI) indien er een kwalitatief goed telefoonnummer beschikbaar is of face-to-face (CAPI) als dit niet het geval is. In figuur 1 staat de benaderstrategie schematisch weergegeven.

Figuur 1. Schematische weergave van de benaderstrategie WoON 2018



3.3 Resultaten veldwerk WoON 2018

Op 10 augustus 2017 is de dataverzameling begonnen en zijn de eerste aanschrijfbrieven voor WoON 2018 verstuurd. De dataverzameling duurde tot 30 april 2018. In tabel 1 is een overzicht gegeven van de respons op WoON 2018. Deze tabel bevat de status van alle personen uit de uitgezette steekproef op het moment dat het veldwerk werd afgesloten.

Tabel 1 responscijfers van WoON 2018 naar onderzoeksdeel.

	Totaal		Regulier		Oversampling	
	Aantal	%	Aantal	%	Aantal	%
Uitgezette stkprf	114.381	100,0	73.746	100,0	40.635	100,0
Kaderfouten	9.504	8,3	6.222	8,4	3.282	8,1
Non-respons	37.354	32,7	24.210	32,8	13.144	32,3
Volledige respons	67.523	59,0	43.314	58,7	24.209	59,6
w.v. Cawi	44.480	38,9	28.158	38,2	16.322	40,2
Cati	13.662	11,9	8.929	12,1	4.733	11,6
Capi	9.381	8,2	6.227	8,4	3.154	7,8

In totaal zijn er 44.480 vragenlijsten volledig ingevuld via internet. Dat is 65,9% van het totaal aantal gerealiseerde responsen. Via telefonische en face-to-face waarneming zijn respectievelijk 13.662 en 9.381 volledige responsen gerealiseerd, hetgeen overeenkomt met respectievelijk 20,2% en 13,9% van het de totale respons. De vooraf gemaakte responsafspraken zijn allemaal gerealiseerd. Voor een uitgebreide responsverantwoording wordt verwezen naar supplement 1.

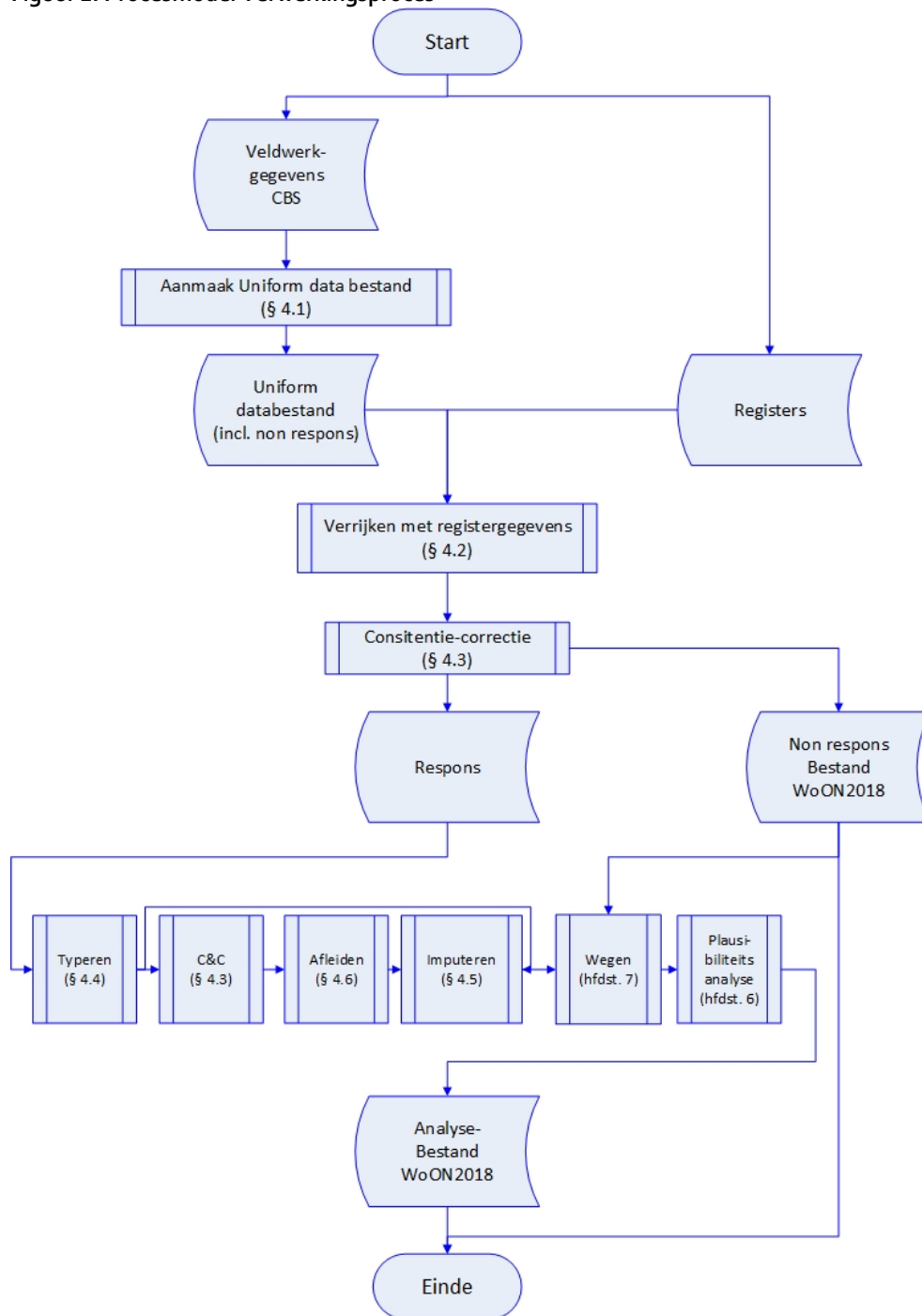
4. Verwerking (data processing)

De term dataprocessing wordt gebruikt om de werkzaamheden die nodig zijn om van het veldwerkbestand te komen tot een eindbestand, inclusief kwaliteitscontroles, afleidingen en verrijkingen. Het verwerkingsproces van WoON 2018 bestaat uit de volgende stappen:

- Invoer en uniformeren veldwerkbestand (4.1);
- Koppelen registergegevens (4.2);
- Controles (4.3);
- Typeren (4.4);
- Imputeren (4.5);
- Afleidingen (4.6).

De dataprocessing is schematisch weergegeven in figuur 2.

Figuur 2: Procesmodel Verwerkingsproces



4.1 Invoer en uniformeren veldwerkbestand

De eerste stap in de dataprocessing is het samenvoegen van drie verschillende inputbestanden. Zoals eerder beschreven (3.2) is de WoON vragenlijst op drie manieren ingevuld, via CAWI (internet), CATI (telefonisch) en CAPI (face-to-face), wat leidt tot drie verschillende inputbestanden. De inputbestanden worden samengevoegd tot één veldwerkbestand met een uniform datamodel (4.1.1). Na het aanmaken van het uniforme bestand is samen met ABF

Research een uitgebreide controle gedaan van de route die respondenten in de vragenlijst volgen (de zogenaamde routingcontrole, 4.1.2).

4.1.1 Uniform datamodel

Het uniforme datamodel staat beschreven in de vragenlijst en het vragenlijstschema. Gegeven de wijzigingen in de vragenlijst is in supplement 3 de relatie tussen WoON 2015 en WoON 2018 op variabelenniveau weergegeven. Als vragen in WoON 2018 hetzelfde zijn als in WoON 2015 is de variabelenaam niet veranderd. Als de vraag en/of de antwoordcategorieën zijn gewijzigd is in WoON 2018 gekozen voor een andere naam. In de vragenlijstverantwoording staan de wijzigingen in de vragenlijsten gedetailleerd beschreven (Cremers en Engelen, 2014). Zoals in supplement 3 is te zien zijn ten opzichte van WoON 2015 in WoON 2018 relatief veel wijzigingen doorgevoerd. Dit heeft gevolgen voor de vergelijkbaarheid tussen WoON 2018 en WoON 2015, 2012, 2009 en 2006.

4.1.2 Routingcontrole

Niet alle vragen in de vragenlijst hoeven door iedere respondent te worden beantwoord. Afhankelijk van bijvoorbeeld leeftijd, de positie in het huishouden of de verhuiswens worden daarop afgestemde vervolgvragen gesteld. Met ruim vierhonderd vragen is de routing in de vragenlijst van het WoON complex. In het verleden is het voorgekomen dat na de afsluiting van het veldwerk fouten in de routing werden gevonden. Als gevolg hiervan moest veel tijd worden geïnvesteerd in het terugbellen van respondenten of het imputeren van ontbrekende antwoorden. Dit had direct nadelige gevolgen voor de opleveringstermijn, kwaliteit en de kosten van het onderzoek. Sinds het WBO 2002 is door ABF Research vooraf een routingcontrole uitgevoerd om de kans op fouten in de route van de vragenlijst zo klein mogelijk te maken, zo ook voorafgaand aan het WoON 2018. Net als in WoON 2015 is ook de routing naar antwoordcategorieën gecontroleerd. Afhankelijk van het antwoord op een voorgaande vraag, kan het voorkomen dat omwille van consistentie niet alle antwoordmogelijkheden worden voorgelegd aan de respondent.

Deze controle bestond uit: i) onafhankelijk automatiseren van de vragenlijst door het CBS en ABF Research, ii) een vergelijking van de vragenlijsten, en iii) het bespreken en corrigeren van gevonden verschillen. In de eerste stap heeft ABF research de routing van de totale vragenlijst nagebouwd in de syntax van de Surfox¹-inputfiles. Bij het programmeren is de vragenlijst van het WoON 2015 als uitgangspunt genomen. Voor alle vragen is gecontroleerd of de routing bij de betreffende vraag ongewijzigd gebleven is. Bij de bouw van de geautomatiseerde vragenlijst is rekening gehouden met de volgende punten:

- Routing in de vragenlijst;
- Tekstvelden (deze worden overgeslagen);
- Controle op minimum- en maximum invulwaarden;
- De categorieën 'weet niet' en 'weigert'.

Om de vragenlijsten te kunnen testen is vervolgens gewerkt met proefbestanden. Deze bestaan niet uit daadwerkelijke antwoorden op de vragenlijst, maar uit duizenden computer gegenereerde enquêtes. Ten behoeve van de toetsing van de computergestuurde vragenlijsten heeft ABF Research op diverse momenten een proefbestand van 50.000 records gegenereerd. Dit bestand voldoet aan de routing zoals is geprogrammeerd door ABF Research.

¹ Surfox is een statistisch software pakket dat door ABF Research zelf is ontwikkeld.

Als laatste stap heeft het CBS een routingcontrole uitgevoerd op de proefbestanden van ABF Research en gecontroleerd of de automatisch gegenereerde records toegestaan zijn binnen het uniforme datamodel. In nauw overleg hebben ABF Research en CBS de geconstateerde verschillen besproken. Nadat de oplossingen voor de verschillen van de routingtest gevonden zijn, heeft ABF Research op basis van de verbeterde vragenlijst een nieuw proefbestand verstuurd, waarna CBS opnieuw de benodigde controles heeft uitgevoerd. Dit proces is net zo lang herhaald totdat er geen verschillen meer gevonden werden.

4.2 Contoles en consistentie checks

Bij WoON 2018 zijn diverse controle en consistentie checks uitgevoerd. Deze controles en checks kunnen worden opgesplitst in drie onderdelen:

- Controles op volledigheid;
- Controles op inconsistenties binnen de vragenlijst;
- Controles op inconsistenties ten opzichte van registergegevens.

De specificaties van de controles zijn opgenomen in supplement 4. De resultaten en de kwaliteitsbeoordeling van het veldwerkbestand staat beschreven in de bestandsacceptatie die is opgenomen in supplement 5.

4.2.1 Controleren op volledigheid

De respons van deelnemers aan het onderzoek is in het verrijkte veldwerkbestand zijn gecontroleerd op volledigheid. In totaal zijn er 3 controles op volledigheid:

- SF01b: Meer dan 15% van de vragen beantwoord met 'weet niet' of 'weigert';
- SF03_1: Steekproefpersoon is niet gelijk aan de respondent;
- SF04b: Geboortedatum respondent klopt niet;

Indien een respons niet voldoet aan één van deze controles op volledigheid, wordt de respons afgekeurd en alsnog uit het uniforme responsbestand verwijderd.

Na samenvoegen van de gegevensbestanden tot één uniform bestand bestaat het veldwerkbestand van WoON 2018 uit 70.537 responsen. In totaal voldoen 3.014 van deze respondenten niet aan één of meer van bovengenoemde controles op volledigheid (tabel 2). Veruit de meest responsen (97%) vallen af doordat meer dan 15% van de vragen is beantwoord met 'weet niet' en of 'weigert'. Dit is vooral het geval bij de internet (CAWI) responsen. In WoON 2018 zijn er 1.000 responsen meer afgevallen door dan in WoON 2015. Met name de internet (CAWI) vragenlijst is in WoON 2018 minder volledig beantwoord dan in 2015 (Janssen, SMA, 2015), respondenten geven vaker 'geen antwoord'. Na de verwijdering van de 3.014 onvolledige responsen uit het uniforme responsbestand bevat het responsbestand nog 67.523 responsen.

Tabel 2. Aantal responsen dat niet volledig is na de controle.

	Regulier	Oversampling	Totaal
Totaal volledige veldwerk responsen	45.226	25.311	70.537
Totaal onvolledige veldwerkresponsen door C&C	1.912	1.102	3.014
Totaal volledige responsen na C&C	43.314	24.209	67.523

4.2.2 Contole op consistentie respons

Op diverse plekken in het dataverwerkingsproces is het gegevensbestand gecontroleerd op inconsistente antwoorden binnen de vragenlijst. Na het samenvoegen van de inputbestanden wordt deze controle voor het eerst uitgevoerd. Vervolgens worden deze controles voorafgaand aan een nieuwe imputatieronde uitgevoerd. Dit is gedaan om inconsistenties als gevolg van imputatie te voorkomen. Nadat ABF Research een geïmputeerd veldwerkbestand heeft opgeleverd worden de controles op het CBS nog een laatste keer gedaan. Bij de interne consistentie controles wordt de onderlinge samenhang tussen antwoorden gecontroleerd. Inconsistenties zijn gesignaleerd en zoveel mogelijk hersteld. In totaal zijn er 17 controles uitgevoerd om inconsistenties van gegeven antwoorden op te sporen. Als een antwoord niet consistent is wordt deze vervangen door een consistente geïmputeerde waarde. De controles op inconsistentie zijn uitgevoerd op het uniforme responsbestand, met daarin de volledige responsen (n=67.523).

4.2.3 Controle ten opzichte van registratiegegevens.

Na samenvoegen de imputatiefase is het gegevensbestand gecontroleerd op inconsistentie ten opzichte van registergegevens. In totaal zijn er 13 controles beschreven om de consistentie van antwoorden ten opzichte van registerinformatie te controleren. Deze controles zijn uitgevoerd op registratiegegevens die doorlopen tot 1 april 2018 (=meest recent). De controles met behulp van registerinformatie zijn uitgevoerd op het uniforme responsbestand, met daarin alleen de volledige responsen (n=67.523).

4.3 Typeren

Het WoON bestaat vooral uit meerkeuzevragen en een paar intervalvragen. Bij enkele vragen is een open antwoordmogelijkheid, waarbij de enquêteur of de respondent het antwoord als tekst invoert. Meestal zijn deze open vragen voorafgegaan door een meerkeuzevraag waarbij de categorie 'anders' de gekozen antwoordmogelijkheid was. De open antwoorden moeten worden gecodeerd of getypeerd naar eenduidige codes die aansluiten op de reeds bestaande antwoordcategorieën van de betreffende vraag. Voor WoON 2018 zijn 16 variabelen getypeerd. De typering is voorafgaand aan de imputatiefase uitgevoerd omdat ze mogelijk in het vervolg (routing, consistentie en imputatie) een rol kunnen spelen. Bij het typeren van de open antwoorden is voor elke tekstregel in een apart bestand vastgelegd wat de bijbehorende code is als deze te herleiden was naar een bestaande antwoordcode van de bijbehorende vraag. Indien een open antwoord niet eenduidig getypeerd kon worden is de code 'anders' in de bijbehorende vraag gehandhaafd, evenals de bijbehorende tekst van het open antwoord. In sommige gevallen verwijst de tekst naar 'weet niet' of 'weigert'. Voor twee variabelen zijn deze antwoorden als 'missings' beschouwd en is het antwoord op de bijbehorende vraag geïmputeerd (waarover meer in paragraaf 4.6).

Tijdens het typeren is per vraag gekeken of de te typeren omschrijvingen aanleiding geven om een nieuwe antwoordcategorie aan de vraag toe te voegen aan de bijbehorende meerkeuzevraag. Dit kan het geval zijn wanneer een substantieel deel van de respondenten een vergelijkbaar antwoord niet kan uitdrukken in de bestaande antwoordmogelijkheden.

4.4 Verrijken

In WoON 2018 is ten opzichte van WoON 2015 vaker gebruik gemaakt van registerbestanden. Aan het WoON 2018 zijn variabelen toegevoegd afkomstig uit de volgende registers:

- **Basisregistratie Personen (BRP)** (voorheen de Gemeentelijke Basisadministratie (GBA)). Demografische kenmerken van de personen zijn gekoppeld en gebruikt voor de berekening van de insluitgewichten van de weging. Via deelname aan het WoON

hebben de respondenten toestemming gegeven om registergegevens te koppelen. Deze gegevens zijn vooral gebruikt in het interne dataproces. Zo zijn demografische gegevens bijvoorbeeld gebruikt in een aantal Controle en Consistentie checks. In het uiteindelijke responsbestand zelf is alleen etniciteit en de vorige woonlocatie opgenomen vanuit de BRP registratie;

- **Gegevens Belastingdienst.** De voorlopige inkomensgegevens van 2017 en de huurtoeslag gegevens van 2017 zijn afkomstig van de Belastingdienst. De inhoud van deze bestanden en de gevolgde methodiek is beschreven in supplement 6;
- **WOZ-registratie.** Vanuit de gegevens van het Kadaster zijn de WOZ waarden gekoppeld. Het gaat om WOZ gegevens uit het BAGplus bestand (peildatum 1 januari 2017) en de gegevens hebben een waardepeildatum van 1 januari 2016. Ontbrekende WOZ waarden zijn door ABF research geïmputeerd.
- **Energielabels.** Rijksdienst Voor Ondernemend Nederland (RVO) heeft in februari 2019 een bestand opgeleverd met de definitieve en de voorlopige energielabels.
- **Huurpuntenbestand (peildatum 1-1-2016).** Het CBS heeft een aantal bronnen gecombineerd om van zoveel mogelijk corporatiewoningen in de G4 de huurpunten vast te stellen.
- **Energiegegevens.** Het gas- en elektriciteitsverbruik in 2017 is bepaald op basis van de gegevens van netbedrijven. Als het verbruik niet gekoppeld is, zijn voor huishoudens in een zelfstandige woning de ontbrekende waarden geïmputeerd. In het geval van stads- en blokverwarming is het gasverbruik ook geïmputeerd. Het waterverbruik is geïmputeerd, waarbij het BudgetOnderzoek van CBS als bronbestand is gebruikt.
- **Basisregistratie Adressen (BAG-plus peilmoment 1 januari 2018).** Aan het adres van de woning waarin de respondent tijdens het interview woonde is een aantal woningkenmerken toegevoegd op basis van de BAG.

In supplement 7 staan de registervariabelen van het responsbestand nogmaals gepresenteerd naar bron en peildatum. Naast het koppelen van registergegevens is aan het responsbestand ook een aantal regionale indelingen gekoppeld. De meeste regionale indelingen worden gekoppeld op basis van het gemeentenummer (per 1-1-2018).

4.5 Afleiden

In WoON 2018 is een groot aantal variabelen afgeleid. CBS heeft de inkomens- (4.5.1), de huishoudens- en de huisvestingsafleidingen (4.5.2) uitgevoerd. Naast deze afleidingen heeft ABF het woonmilieutype en de variabelen met betrekking tot de bijkomende woonlasten vastgesteld (4.5.3 en 4.5.4).

4.5.1 Inkomensafleidingen

De inkomensgegevens zijn afkomstig van de inkomensstatistiek. Voor het responsbestand van WoON 2018 zijn ruim 67 duizend personen (c.q. huishoudens) geïnterviewd. Deze 67 duizend huishoudens bestaan uit ruim 162 duizend huishoudleden en van elk huishoudlid is het persoonlijk inkomen vastgesteld. Vervolgens is een aantal inkomensvariabelen vastgesteld op huishoudniveau (o.a. bruto, belastbaar en het VROM besteedbaar inkomen (bestedbaar inkomen gecorrigeerd voor woonlasten). Naast deze huishoudinkomens heeft ABF Research voor alle zelfstandige huishoudens de bijkomende woonlasten vastgesteld (zie 4.5.4). Als laatste zijn op basis van huishoudinkomens de bijkomende woonlasten en vragen uit de vragenlijst diverse afleidingen bepaald (o.a. aandachtsgroep huurtoeslag (aandihs), totale woonquote (nwq), netto kooplasten (nkoop) en netto huurlasten (nhuur)).

Het CBS heeft een revisie van de inkomensgegevens doorgevoerd. Omdat de inkomensgegevens die aan het WoON gekoppeld zijn zijn gebaseerd op het huishouden zoals door de respondent is opgegeven, zijn de gevolgen van de inkomensrevisie klein. Toch is besloten om voor WoON 2012 en 2015 ook over te gaan op de gereviseerde inkomencijfers. Naast een revisie van de inkomensgegevens zijn in WoON 2018 ook de woonlasten gerevisieerd. Na revisie van de woonlasten zijn de gegevens over woonlasten uit registratie leidend geworden. Als de gegevens in registratie beschikbaar zijn dan is dit overgenomen in plaats van het antwoord van de respondent. Door deze andere benadering zijn de woonlasten plaatjes waarbij ook inkomen een belangrijke rol speelt consistent. In supplement 6 staat beschreven hoe de inkomensgegevens aan het WoON zijn gekoppeld. De specificaties van de inkomensafleidingen zijn gepresenteerd in supplement 8. Supplement 10 bevat alle relevante inkomensafleidingen.

4.5.2 Huishoudens- en huisvestingsafleidingen

Bij het maken van de huishoudens- en de huisvestingsafleidingen is WoON 2015 het uitgangspunt geweest. In supplement 9 staan de afleidingen gespecificeerd die zijn aangemaakt voor de huishoudens- en huisvestingsvariabelen. Afleidingen worden uitgevoerd op een volledig geïmputeerd veldwerkbestand dat verrijkt is met registratie gegevens. Ook zijn de typering al uitgevoerd. Tijdens het proces van imputeren worden de afleidingen ook uitgevoerd, de resultaten van deze afleidingen worden niet meegenomen in het productieproces.

In de huishoudensvariabelen worden belangrijke kenmerken van het huishouden vastgesteld. Het gaat hier onder andere om de aanmaak van leeftijdsklassen, opleiding, samenstelling van het huishouden en de positie van de respondent in het huishouden. Bij de huisvestingsvariabelen wordt de woonsituatie vastgesteld in een aantal afleidingen. Het gaat dan vooral om afleidingen voor het type woning, het eigendomstype van de woning en of respondent in een zelfstandige woning woont. Naast de woonsituatie in de huidige woning wordt ook de woonsituatie in de vorige woning (indien de respondent in de afgelopen 2 jaar verhuisd is) en van de gewenste woning (respondent heeft een verhuiscens) bepaald. De programmatuur van de huishoudens- en huisvestingsvariabelen zijn opgenomen in supplement 10.

4.5.3 Woonmilieutype

In het databestand van WoON 2018 is door ABF Research het woonmilieu zowel voor de huidige woonomgeving (HWMBRT) als voor de gewenste woonomgeving (GWMBRT) vastgesteld. Zowel HWMBRT en GWMBRT hebben de categorieën:

1. centrum-stedelijk
2. buiten-centrum
3. groen-stedelijk
4. centrum-dorps
5. landelijk wonen

Het woonmilieu is afgeleid op basis van de wijk- of buurtcode waarin de respondent woont. De gekoppelde woonmilieus zijn gebaseerd op de meest recente typologie die beschikbaar is, te weten de typologie van 2018. Voor de huidige woning is altijd een woonmilieu vastgesteld. Het gewenste woonmilieu is deels afgeleid en deels geïmputeerd. Aan iedereen die wil of moet verhuizen is gevraagd naar welke woonplaats en naar welke wijk of buurt binnen die woonplaats men wil verhuizen. Als een gewenste wijk of buurt bekend is, dan wordt hier het

(gewenste) woonmilieu aan gekoppeld. Is de buurt niet bekend en ook niet te typeren naar een bestaande wijk- of buurtcode, dan wordt het woonmilieu geïmputeerd op basis van de antwoorden op vragen over de gewenste regio, de gewenste woonplaatsgrootte, de gewenste afstand tot het centrum van de woonplaats, de gewenste bouwperiode van de buurt, het gewenste woningtype van de buurt en de gewenste bevolkingssamenstelling in de buurt.

4.5.4 Bijkomende woonlasten

Naast de woonmilieus heeft ABF Research ook de bijkomende woonlasten bepaald. Woonuitgaven betreffen de uitgaven van huurders en eigenaren aan huur en hypotheekrente. Daarnaast geven bewoners geld uit voor het feitelijke gebruik van de woning. Deze uitgaven worden aangeduid als de bijkomende woonlasten. Het gaat hierbij om heffingen van gemeenten en waterschappen (heffingen van Openbaar Publiekrechtelijke Lichamen ofwel OPL) en uitgaven aan water, gas, warmte en elektriciteit (nutsvoorzieningen). Het WoON 2018 bevat een grote hoeveelheid aan informatie over de woonuitgaven van de respondent. Daarnaast wordt een deel van de informatie met aanvullende bronnen bepaald. Indien een bronhouder voor een gemeente, waterschap of nutsbedrijf geen tarieven heeft kunnen aanleveren, heeft ABF Research deze gegevens in eigen beheer verzameld. De informatie betreft:

- de gemeentelijke lasten. ABF Research heeft per gemeente de tarieven verzameld van de relevante gemeentelijke heffingen, te weten onroerend zaakbelasting, rioolrechten en afvalstoffenheffing. Deze gegevens zijn bewerkt zodat deze gebruikt konden worden in de rekenprogrammatuur;
- de waterschapsheffingen. ABF Research heeft per waterschap de tarieven verzameld van de relevante waterschapsheffingen, te weten watersysteemheffing ingezetenen, watersysteemheffing gebouwd en zuiveringsheffing. Op basis van het werkgebied en het takenpakket van het waterschap berekent ABF Research per gemeente de omvang van de tarieven. Deze gegevens zijn bewerkt zodat deze gebruikt konden worden in de rekenprogrammatuur;
- de kosten van energie en water. De tarieven en heffingen heeft ABF Research in eigen beheer verzameld. De verbruikgegevens met betrekking tot energie (elektriciteit, gas en warmte) heeft CBS aangeleverd.

4.6 Imputeren

Eén van de onderdelen in de dataprocessing is de imputatie-fase. De respondent kan een vraag weigeren te beantwoorden, vergeten te beantwoorden, of het antwoord op de vraag niet weten. Daarnaast kan er ook een ontbrekende waarde zijn door een inconsistentie in de vragenlijst. Deze vorm van ontbrekende respons wordt item non-respons genoemd. Een weigering van een antwoord gebeurt vaker bij variabelen waarbij bedragen worden gevraagd zoals huurprijs, woningwaarde en hypotheek. De ontbrekende antwoorden worden geïmputeerd met een waarde die het beste past bij de desbetreffende persoon. Dit gebeurt met behulp van de Random Hotdeck Methode. Voor het WoON 2018 heeft ABF Research de imputatie uitgevoerd. In supplement 11 wordt het imputatieproces in detail beschreven.

5. Non-responsanalyse

5.1 Selectiviteit naar achtergrondkenmerken

Bereidheid tot deelname aan het WoON 2018 verschilt per bevolkingsgroep. Tabel 3 toont het responspercentage naar leeftijd, geslacht, burgerlijke staat, stedelijkheid woongemeente, landsdeel woongemeente, herkomstgeneratie, plaats in het huishouden, vermogenspositie, inkomensbron, besteedbaar huishoudinkomen en woningkenmerken.

Tabel 3. Responspercentages naar achtergrondkenmerken.

	Totaal	Regulier	Oversampling
%			
Totaal	59	59	60
Leeftijd			
18 tot en met 24 jaar	53	53	53
25 tot en met 34 jaar	51	51	51
35 tot en met 44 jaar	56	55	57
45 tot en met 54 jaar	59	59	59
55 tot en met 64 jaar	66	66	67
65 tot en met 74 jaar	70	69	71
75 plus	59	59	59
Geslacht			
man	57	57	58
vrouw	61	60	61
Burgerlijkestaat			
Ongehuwd	53	53	53
Gehuwd/partnerschap	66	66	67
Verweduwd	61	61	61
Gescheiden	57	57	57
Stedelijkheid van de buurt			
Zeer sterk (>=2.500 omgevingsadressen/km ²)	52	52	52
Sterk (1.500 tot 2.500 omgevingsadressen/km ²)	59	58	59
Matig (1.000 tot 1.500 omgevingsadressen/km ²)	62	62	63
Weinig (500 tot 1.000 omgevingsadressen/km ²)	64	64	63
Niet (<500 omgevingsadressen/km ²)	62	62	63
Landsdeel			
Noord-Nederland (Gr, Fr, Dr)	60	60	-
Oost-Nederland (Ov, Gld, Fl)	61	62	59
West-Nederland (NH, ZH, Ut, Zl)	57	56	58
Zuid-Nederland (NB, Li)	62	61	64
G4			
Niet behorend tot G4	61	61	61
Wel behorend tot G4	48	49	48
Herkomst			
Autochtonen	63	63	64
Niet-westerse allochtonen	38	38	38
Westerse allochtonen	50	50	49

	Totaal	Regulier	Oversampling
Generatie			
autochtoon	63	63	64
eerste generatie allochtoon	38	38	37
tweede generatie allochtoon	53	53	53
Positie in het huishouden			
kind	54	54	54
alleenstaande	54	54	55
partners zonder kinderen	68	68	68
partners met kinderen	63	62	64
een-ouder gezin	54	53	54
overig huishouden	40	40	39
institutioneel huishouden	24	27	17
Quintiel besteedbaar inkomen			
Laagste	50	50	50
Tweede	56	56	55
Derde	61	61	62
Vierde	64	63	64
Hoogste	65	65	66
Quintiel welvaart (combinatie van inkomen en vermogen)			
Laagste	48	48	48
Tweede	56	55	56
Derde	61	61	61
Vierde	65	65	65
Hoogste	68	67	69
Belangrijkste inkomensbron van het huishouden			
Loon	61	61	61
Loon directeur-groootaandeelhouder	57	57	58
Winst zelfstandig ondernemer	52	52	53
Inkomen overige zelfstandige	46	45	47
Werkloosheidsuitkering	60	60	60
Bijstandsuitkering	43	44	41
Uitkering sociale voorziening overig	46	47	45
Uitkering ziekte/arbeidsongeschiktheid	52	52	53
Pensioenuitkering	65	65	66
Studiefinanciering	55	55	53
Inkomen uit vermogen	57	54	62
Eigendomstype			
EigenaarGebruiker	65	64	66
Woningcorporatie	53	53	53
Verhuurder anders dan Woningcorporatie	50	51	49
WOZ waarde			
minder dan 150.000 Euro	52	53	51
150.000-199.999 Euro	60	59	61
200.000-249.999 Euro	63	63	64
250.000-299.999 Euro	66	66	66
300.000-399.999 Euro	66	65	67
400.000-499.999 Euro	66	65	68
500.000 Euro of meer	59	59	60

Uit tabel 3 kan worden geconcludeerd dat mensen van 18 tot 35 jaar minder responderen dan andere leeftijdsgroepen. Vrouwen responderen vaker dan mannen. Gehuwden en personen in partnerschap doen vaker mee dan degenen die ongehuwd, gescheiden of verweduwd zijn. In zeer sterk stedelijke gebieden blijft de respons achter op de overige woongebieden en personen die in het westen van Nederland wonen hebben minder vaak mee gedaan dan bewoners van de drie andere landsdelen. Responspercentages in de 4 grote steden (Amsterdam, Rotterdam, Utrecht, Den Haag) zijn een stuk lager dan in de rest van Nederland. De responsverschillen naar herkomst en generatie zijn eveneens aanzienlijk. Vooral niet-westers allochtonen doen met 38 procent beduidend minder vaak mee aan WoON 2018 dan westers allochtonen (50 procent) en autochtonen (63 procent). Het blijken voornamelijk allochtonen van de eerste generatie te zijn die minder responderen. Vergelijken met WoON 2015 is het responspercentage onder niet-westers allochtonen 3 procentpunt hoger en onder westers allochtonen 4 procentpunt hoger.

Als gekeken wordt naar de positie in het huishouden van de respondent, dan responderen kinderen, alleenstaanden en ouders uit één oudergezinnen minder goed. Overige huishoudleden en personen die tot een institutioneel huishouden behoorden responderen het slechtst. Steekproefpersonen in het laagste inkomens- en welvaartsquintiel responderen minder vaak dan personen uit overige inkomensgroepen. Op woningniveau is te zien dat steekproefpersonen die in een eigen woning wonen het beste responderen. Tenslotte responderen steekproefpersonen die in een woning wonen met een WOZ die lager is dan 150 duizend euro minder goed dan personen woonachtig in een woning met een hogere WOZ – waarde.

5.2 Item-respons voor bepaalde vragen en thema's

In de projectovereenkomst van WoON 2018 zijn afspraken gemaakt over minimale item-respons voor de volgende vragen: netto maandhuur (huurders), gewenste gemeente (verhuigeneigden), gewenste wijk/buurt (verhuigeneigden), tevredenheid woning, tevredenheid woonomgeving, betrokken bij leefbaarheid van de buurt, interesse in particulier opdrachtgeverschap, interesse in kopen eigen huurwoning (huurders), opleiding van respondent en zijn of haar partner. In tabel 4 wordt de behaalde respons op deze vragen samengevat.

Tabel 4. Item respons van specifieke vraagthema's naar onderzoeksdeel en waarneemmethode

(Reg=Reguliere woningmarktmodule, Over=oversampling).

Vraagthema	Eis		Totaal	Deel		Methode		
	BZK			Reg	Over	CAPI	CATI	CAWI
	%	N*	%	%	%	%	%	%
Netto maandhuur	95	19.620	93	93	92	99	94	90
Tevredenheid woning	95	59.073	100	100	100	100	100	100
Tevredenheid woonomgeving	95	67.498	100	100	100	100	100	100
Betrokken bij leefbaarheid van de buurt	95	66.922	99	99	99	100	100	99
Interesse in particulier opdrachtgeverschap	95	57.308	95	95	95	100	100	93
Interesse in kopen eigen huurwoning	95	20.757	98	98	98	100	100	97
Opleiding OP	95	65.204	97	97	96	100	99	95
Opleiding PA	95	33.067	94	95	94	99	98	93
Gewenste gemeente	95	25.487	98	98	98	100	99	98
Gewenste wijk / buurt / kern	90	24.798	96	96	96	99	96	95

*Het aantal respondenten dat volgens de routing het vraagthema moet beantwoorden.

De item-responspercentages van de vragen naar 'Netto maandhuur' ligt 2 procentpunten lager dan de eis van BZK. Dit komt met name door de waarneming via internet (CAWI) en via telefoon (CATI) waar bij het percentage item-respons op dit onderdeel respectievelijk 90% en 94% is. De item-responspercentage voor de opleiding van de partner ligt ook 1% onder de eis van BZK. Dit wordt veroorzaakt door de respondenten die via internet (CAWI) hebben deelgenomen. Voor de overige vraagthema's is de item-respons hoger dan het door BZK geëiste percentage.

De item-responspercentages verschilt nauwelijks tussen het reguliere deel en de oversampling. Bij de item-responspercentages naar waarneemmethode is het percentage het hoogst bij de face-to-face (CAPI) waarneming, gevolgd door de telefonische (CATI) waarneming. Bij de interent waarneming (CAWI) is het percentage item-respons duidelijk het laagst.

6. Plausibiliteitsanalyse

In supplement 12 is een gedetailleerde plausibiliteitsrapportage opgenomen. In dit supplement wordt voor een aantal kenmerken gekeken naar de plausibiliteit van de uitkomsten van het WoON 2018. Door een overzicht te maken van de uitkomsten van op een aantal belangrijke variabelen kan snel getoetst worden of het bestand voldoet aan de eisen en de verwachting.

7. Weging

Non-respons is een belangrijk probleem bij steekproefonderzoek. Tijdens de dataverzameling wordt gestreefd naar een zo hoog mogelijke respons is en een zo goed mogelijke representatie van de doelpopulatie. Door selectieve non-respons zijn bepaalde groepen onder- of oververtegenwoordigd waardoor de respondenten geen juiste afspiegeling van de populatie vormen. Dit kan leiden tot vertekening in de uitkomsten van het onderzoek. Weging is een techniek om vertekening door non-respons te reduceren. Met het toekennen van gewichten aan de respondenten wordt de representativiteit zo goed mogelijk hersteld.

Een tweede reden om gewichten toe te kennen aan respondenten is het verkrijgen van consistentie met registerinformatie of publicatietabellen van andere onderzoeken. Wanneer op basis van de verzamelde data een tabel wordt gemaakt voor een bepaalde doelvariabele is het vaak gewenst dat totalen overeenkomen met registergegevens. Neem bijvoorbeeld de doelvariabele 'verhuisgeneigdheid'. In een tabel waarin zowel voor vrouwen als voor mannen het aantal verhuisgeneigden en het aantal niet- verhuisgeneigden personen is opgenomen is het gewenst dat het totale aantal vrouwen en het totale aantal mannen overeenkomt met de aantallen mannen en vrouwen in de Basisregistratie Personen.

7.1 Weging Woningmarktmodule WoON 2018

De weging van het WoON 2018 is zoveel als mogelijk opgezet conform de weging van het WoON 2015, zodat de continuïteit in de methodiek gewaarborgd blijft. De weging levert een weefactor voor personen (=WEEGPERS) en een weefactor voor huishoudens (HWEEGWON) op. In supplement 13 staat de weging van WoON 2018 in detail beschreven.

7.2 Extra weefactoren

Naast een weefactor voor personen en huishoudens wordt voor de gewenste woonsituatie eveneens een weefactor (GWEEGWON) samengesteld. Op basis van deze drie weefactoren zijn er 6 afleidingen aangemaakt waarmee vraag en aanbod van de woningmarkt zijn vertegenwoordigd:

- **WEEGV_T.** Weefactor vraag alle woningzoekers. De V staat voor Vraag; de T voor Totaal. Door te wegen met deze variabele ontstaat een beeld van de vraag die verhuisgeneigden uitoefenen. Deze vraag kan worden geconfronteerd met het aanbod dat deze groep achterlaat als ze verhuizen; de woning waar ze nu wonen;
- **WEEGA_T.** Weefactor aanbod alle woningzoekers. De A staat voor Aanbod; de T voor Totaal. Door te wegen met deze variabele ontstaat een beeld van het aanbod dat de groep vragers achterlaat als ze verhuizen; de woning waar ze nu wonen;
- **WEEGV_M.** Weefactor vraag woningzoekers 'in de markt'. De V staat voor Vraag; de M voor Markt. Door te wegen met deze variabele ontstaat een beeld van de vraag die actieve verhuisgeneigden uitoefenen. Dit zijn verhuisgeneigde huishoudens die al een woning hebben gevonden of al actief aan het zoeken zijn. Deze vraag kan worden geconfronteerd met het aanbod dat deze groep achterlaat als ze verhuizen; de woning waar ze nu wonen.
- **WEEGA_M.** Weefactor aanbod woningzoekers 'in de markt'. De A staat voor Aanbod; de M voor Markt. Door te wegen met deze variabele ontstaat een beeld van het aanbod dat de groep actieve vragers achterlaat als ze verhuizen; de woning waar ze nu wonen.
- **WEEGV_U.** Weefactor vraag urgente woningzoekers. De V staat voor Vraag; de U voor Urgent. Door te wegen met deze variabele ontstaat een beeld van de vraag die urgente verhuisgeneigden uitoefenen. Deze vraag kan worden geconfronteerd met het aanbod dat

deze groep achterlaat als ze verhuizen; de woning dus waar ze nu wonen. Urgente verhuisgeneigden zijn woningzoekers in de markt, die aangeven direct actie te ondernemen als men de gewenste woning in het aanbod aantreft;

- **WEEGA_U**. Weegfactor aanbod urgente woningzoekers. De A staat voor Aanbod; de U voor Urgent. Door te wegen met deze variabele ontstaat een beeld van het aanbod dat de groep urgente vragers achterlaat als ze verhuizen; de woning waar ze nu wonen.

De SPSS syntax van de afleidingen worden gepresenteerd in supplement 10.