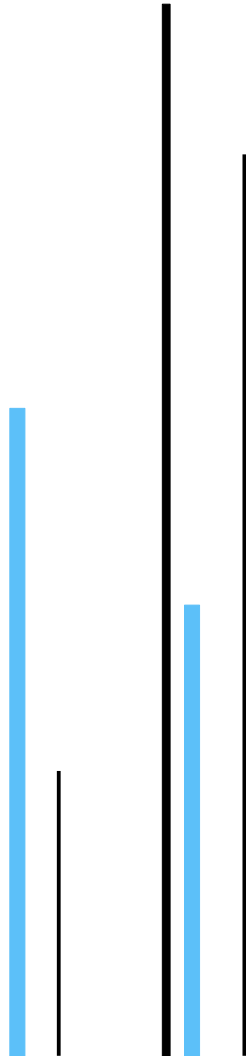


Dataprocessing WoON 2012

Onderzoeksdocumentatie



Dataprocessing WoON 2012

Onderzoeksdocumentatie

Uitgevoerd in opdracht van CBS Heerlen

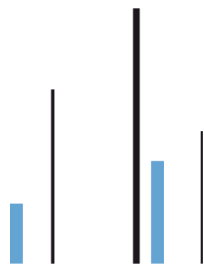
Kenneth Gopal

12 oktober 2012 | r2012-0047KG | 11180-DPC

ABF Research | Verwersdijk 8 | 2611 NH | Delft | 015 - 27 99 300

Copyright ABF Research 2012

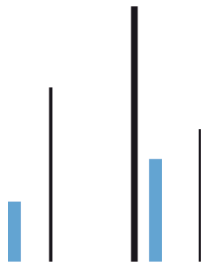
ABF Research heeft dit rapport met zorg samengesteld. ABF Research aanvaardt geen aansprakelijkheid voor type- of drukfouten of andere onvolkomenheden. Het gebruik van onderdelen van dit rapport door een andere partij dan ABF Research blijft de verantwoordelijkheid van deze andere partij en dient voorzien te zijn van duidelijke bronvermelding. In een aantal gevallen is voor gebruik van (onderdelen van) dit rapport goedkeuring door ABF Research vereist. In geval van enige twijfel dient met ABF Research hierover contact te worden opgenomen.



Inhoudsopgave

1	Inleiding.....	7
2	Werkzaamheden dataprocessing.....	9
2.1	Routingcontrole.....	9
2.2	Typeren open vragen.....	10
2.3	Consistentiecontrole	11
2.4	Imputatie.....	12
2.5	Bepalen gewenste woonmilieus.....	16
2.6	Bijkomende woonlasten	17
3	Eindproducten	19
	Bijlage 1 Consistentiecontroles.....	21
	Bijlage 2 Bijkomende woonlasten	23
	Bijlage 3 Bijschatting huurpunten.....	35

1



Inleiding

Het Woon Onderzoek Nederland (WoON) is het grootste landelijk onderzoek in Nederland op het gebied van wonen en woonomgeving. Sinds 2006 geldt het WoON als de vervanger van het Woningbehoefte Onderzoek (WBO) en de Kwalitatieve Woning Registratie (KWR).

Naast het in elkaar schuiven van deze basisonderzoeken kent het WoON een aantal belangrijke nieuwe elementen. Zo wordt er gewerkt met specifieke modules op basis van een driejaarlijkse basissteekproef van netto 43.000 respondenten. Evenals in 1998, 2002, 2006 en 2009 hebben gemeenten of regionale samenwerkingsverbanden meegedaan aan de oversampling, waarvoor nog eens 26.000 respondenten benaderd zijn. Hiermee komt het totaal aantal respondenten dat aan de basismodule van het WoON 2012 heeft deelgenomen op 69.000. Deze wordt ook wel de module Woningmarkt genoemd.

De gegevensverzameling heeft tijdens de veldwerkfase plaatsgevonden via telefonische, face-to-face en internetenquêtes. Het gehele veldwerk was in handen van het CBS.

Na de veldwerkfase volgt de fase van de dataprocessing. De term dataprocessing wordt gebruikt om de werkzaamheden samen te vatten die nodig zijn om aan de antwoorden die respondenten geven op vragen in de WoON-interviews plausibele statistische informatie te ontlenuen. Met andere woorden, er wordt in deze fase vanuit een ruw veldwerkbestand toegewerkt naar een onderzoeksbestand dat zo veel mogelijk ontdaan is van bestandsvervuiling.

De dataprocessing voor het WoON omvat de volgende onderdelen: typeren open vragen, uitvoeren consistentiecontroles, imputeren (bijschatten) item non-respons, bepalen afgeleide variabelen, draaien van inkomensprogramma, bepalen bijkomende woonlasten, wegen en ophogen, plausibiliteitcontroles en non-respons analyse.

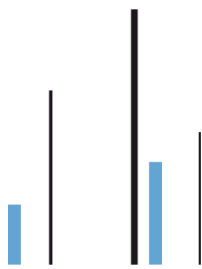
In 2002 en 2006 heeft ABF Research de volledige dataprocessing uitgevoerd van respectievelijk het WBO 2002 en het WoON 2006. Voor het WoON 2009 lag de regie over de dataprocessing echter in handen van het CBS, en is ABF Research op een aantal onderdelen ingeschakeld.

Dat is voor de nieuwste editie eveneens het geval. ABF Research heeft ten behoeve van de dataprocessing van het WoON 2012 de volgende werkzaamheden uitgevoerd:

- routingcontrole;
- typeren selectie van open vragen;
- consistentiecontrole;
- imputatie;
- bepalen gewenste woonmilieu;
- bijkomende woonlasten.

In dit rapport wordt verslag gedaan van de werkzaamheden die ABF Research uitgevoerd heeft in het kader van de dataprocessing van de module Woningmarkt van WoON 2012.

2



Werkzaamheden dataprocessing

De verschillende onderdelen van de dataprocessing van de module Woningmarkt waar ABF Research aan gewerkt heeft worden in dit hoofdstuk stapsgewijs behandeld.

2.1 Routingcontrole

Niet alle vragen hoeven door iedere respondent te worden beantwoord. Afhankelijk van bijvoorbeeld de leeftijd, de positie in het huishouden of de verhuiswens worden daarop afgestemde vragenblokken gesteld. Met ruim vierhonderd vragen verspreid over bijna veertig blokken is de routing in de vragenlijst van het WoON zeer complex. In het verleden is het verschillende malen voorgekomen dat na de afsluiting van de veldwerfphase nog fouten in de routing werden ontdekt. Bij gevolg moesten respondenten worden nagebeld of vragenblokken in het bestand extra worden geïmputeerd. Dit had direct nadelige gevolgen voor de opleveringstermijn, kwaliteit en de kosten.

Sinds het WBO 2002 is door ABF Research vóóraf een routingcontrole uitgevoerd om de kans op fouten in de routing van de vragenlijst zo klein mogelijk te houden. Ook voorafgaand aan het WoON 2012 is een dergelijke controle uitgevoerd. Hiertoe is eerst de vragenlijst door de betrokken partijen, te weten het CBS en ABF Research, onafhankelijk van elkaar geautomatiseerd. Vervolgens zijn de vragenlijsten met elkaar vergeleken waarna de gevonden verschillen zijn teruggekoppeld en zo nodig gecorrigeerd, opdat beide partijen dezelfde routing in de vragenlijst hanteren.

De routingcontrole bestaat in feite drie stappen:

1. maken geautomatiseerde vragenlijst;
2. controle via proefbestand;
3. correctie en terugkoppeling.

2.1.1 Maken computergestuurde vragenlijsten

De routing van de totale vragenlijst is door ABF Research nagebouwd in de syntax van de Surfox¹-inputfiles. De vragenlijst is ten opzichte van het WoON 2009 op veel onderdelen ongewijzigd gelaten. Bij het programmeren van de vragenlijst is daarom de geprogrammeerde vragenlijst van het WoON 2009 als uitgangspunt genomen. Voor alle vragen is goed gecontroleerd of de routing bij de betreffende vraag inderdaad ongewijzigd gebleven is.

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

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 maximumwaarden;
- De categorieën 'weet niet' en 'weigert'.

2.1.2 Controle via proefbestand

Om de vragenlijsten te kunnen testen is gewerkt met proefbestanden, ook wel dummybestanden genoemd. Deze bestaan niet uit daadwerkelijk uitgevoerde enquêtes, maar uit duizenden door de computer gegenereerde enquêtes. Ten behoeve van de toetsing van de computergestuurde vragenlijsten heeft ABF Research op diverse momenten een dummybestand van 50.000 records gegenereerd. Op dit bestand is vervolgens de routing volgens ABF Research toegepast, waarna dit gerouteerde bestand naar het CBS is verzonden.

2.1.3 Correctie en terugkoppeling

Het CBS heeft haar routing vergeleken met de routing die door ABF Research is toegepast op het dummybestand. In nauw overleg hebben beide partijen de geconstateerde verschillen besproken.

Nadat de oplossingen voor de verschillen van de routingtest gevonden waren, is op basis van de verbeterde vragenlijst door ABF Research een nieuw proefbestand verstuurd, waarna het CBS opnieuw haar controles uitgevoerd heeft. Dit proces is net zo lang herhaald totdat er geen verschillen meer optraden. Bij geen gevonden verschillen zijn immers de geprogrammeerde vragenlijsten gelijk aan elkaar.

2.2 Typeren open vragen

De module Woningmarkt kent hoofdzakelijk keuzevragen en een aantal intervalvragen, bijvoorbeeld voor jaartallen en bedragen. Voor enkele vragen is een open antwoordmogelijkheid gegeven, waarbij de enquêteur (of de respondent, in geval van een CAWI-enquête) het antwoord als tekst invoert. Deze open antwoorden moeten worden gecodeerd of getypeerd naar eenduidige codes. Meestal zijn deze open vragen voorafgegaan door een keuzevraag waarbij de categorie 'anders' een van de antwoordmogelijkheden was en ook gekozen is.

De typeringen zijn in principe een verlengstuk van de veldwerkfase en zijn vrij vroeg in het traject opgenomen, omdat zij mogelijk in het vervolg (routing, consistentie en imputatie) een rol kunnen spelen. Het typeren van de open antwoorden is eenmalig uitgevoerd, waarbij voor elke tekstregel in een apart bestand is vastgelegd wat de bijbehorende code is als deze te herleiden was naar een bestaande antwoordcode van de voorloopvraag. Indien een open antwoord niet eenduidig getypeerd kon worden is de code 'anders' in de voorloopvraag gehandhaafd, evenals de bijbehorende tekst van het open antwoord. In sommige gevallen verwijst de tekst (in)direct naar 'weet niet' of 'weigert'. Voor twee variabelen zijn deze antwoorden als 'missings' beschouwd en is de voorloopvraag vervolgens bijgeschat (waarover meer in paragraaf 2.4).

Voor het WoON 2012 zijn de volgende tekstvariabelen door ABF Research getypeerd (met tussen haakjes de codenamen van de tekstvariabele en de voorloopvragen):

- Type woning (Anders – Soortwon, SrtWoonR en AardAdr);
- Verhuurder (Omschre – Codeverh);
- Soort hypotheek (ASrtHyp – SrtCombi en SrtHypo);
- Reden om niet te verhuizen (AndRdNVH – NietVH);
- Gewenste wijk/buurt in huidige gemeente (OmsB1 – GBuurt1 en GcodeBu1);
- Gewenste wijk/buurt in andere gemeente (OmsSpecB1 – GspecBr en GcodeSp1).

In Tabel 1 is een overzicht opgenomen van het totaal aantal te typeren gevallen per variabele, gevolgd door het aantal waarvoor de typering niet respectievelijk wel gelukt is. Daarnaast is per variabele aangegeven hoeveel open antwoorden verwijzen naar de antwoordcategorieën ‘weet niet’ en ‘weigert’ en voor welke gevallen deze antwoorden zijn bijgeschat. Overigens heeft voor de variabele AndRdnNVH alleen een bijchatting plaatsgevonden van de voorloopvraag NietVH als de antwoordcategorie ‘anders’ het enige antwoord is dat de respondent opgegeven heeft.

Tabel 1 Aantal cases per open vraag met antwoord ‘anders’ dat getypeerd of bijgeschat is

Variabele	Vraag	Aantal te typeren	Niet	Wel	Weet niet/Weigert	Bijgeschat
Anders	7.6	372	0	306	66	ja
Omschre	10.2	890	364	526	-	nee
ASrtHypo	14.9	909	625	145	139	ja
AndRdNVH	25.3	2338	933	1346	59	ja
OmsB1	28.3.1	119	42	44	33	nee
OmsSpecB1	28.6.1	108	33	31	44	nee

De overige tekstvariabelen zijn door het CBS zelf getypeerd. Op verzoek is echter voor een zestal variabelen bij wijze van achtervang een ‘extra’ imputatie uitgevoerd. Deze variabelen worden in eerste instantie dus wel getypeerd, maar in het geval dit niet op een eenduidige wijze mogelijk is wordt de geïmputeerde waarde overgenomen door het CBS. Een overzicht van deze variabelen is opgenomen in Tabel 2.

Tabel 2 Variabelen die geïmputeerd zijn als achtervang voor typering

Variabele	Vraag	Imputeren/Typeren naa
NSandOP	33.3	NVOplOP, NSOplsOP
VSOplOP	33.7	VOpIOP
NSandPA	35.3	NVOplPA, NSOplsPA
VSOplPa	35.7	VOpIPA

2.3 Consistentiecontrole

Tijdens het veldwerk wordt door het CBS al rekening gehouden met de toegestane maximum- en minimumwaarden. Door een aantal waarden van variabelen te combineren kunnen andere inconsistenties aan het licht komen. Het kan hierbij onder meer gaan om de combinaties van verschillende jaartallen (voorbeeld: de respondent heeft zich bijvoorbeeld gevestigd op een adres vóór het bouwjaar van de woning). Indien een dergelijke inconsistentie gevonden wordt moet een zeer goede afweging worden

gemaakt of de gegeven antwoorden gecorrigeerd moeten worden, opdat de samenhang na de correctie wél consistent is.

Het CBS heeft in samenspraak met het Ministerie van Binnenlandse Zaken deze afweging gemaakt en vervolgens de consistentiecontroles aangeleverd. Deze zijn overwegend hetzelfde gebleven ten opzichte van de controles die in 2009 zijn uitgevoerd. ABF Research heeft de controles verwerkt in haar eigen syntax voor Surfox. Om te voorkomen dat tijdens de imputatie nieuwe inconsistenties introduceren worden de consistentiecorrecties voorafgaand aan de imputaties uitgevoerd.

In de bijlage is een overzicht te vinden van deze consistentiecontroles. Hierin is ook een overzicht opgenomen van het aantal gevonden inconsistenties per controle voor dataprocessing. Na bewerking zijn alle inconsistenties uit het bestand verwijderd.

2.4 Imputatie

Een van de belangrijkste onderdelen in het traject om van een ruw bestand tot een geschikt onderzoeksbestand te komen is de bijschatting van ontbrekende waarden, ook wel imputatie genoemd. Doorgaans geven niet alle respondenten op alle vragen een antwoord. In sommige gevallen is dit terecht als gevolg van de routing (“Niet Van Toepassing”). Echter in veel gevallen blijkt dit niet terecht waardoor men met non-respons te maken heeft (“Ten Onrechte Geen Antwoord”). Daarnaast worden voor de meeste vragen in het WoON de antwoordcategorieën ‘weigert’ en ‘weet niet’ ook als non-respons beschouwd.

Ontbrekende waarden zijn met behulp van imputatietechnieken bij te schatten. ABF Research is sinds het WBO 1998 altijd betrokken geweest bij de imputaties van ontbrekende antwoorden.

2.4.1 Waarom imputeren?

Bij bijna alle enquêtes heeft men te maken met non-respons. Indien een respondent weigert deel te nemen, spreekt men van unit non-respons. Via ophogingstechnieken worden weegvariabelen gemaakt of aangepast om dit soort non-respons te corrigeren. Hierbij gaat men er dan van uit dat de groep die meedoet hetzelfde gedrag vertoont als diegenen die wel hebben meegedaan en – in het geval van het WoON – dezelfde leeftijd, geslacht, geboorteland, positie in het huishouden en woongebied hebben. Indien eenmaal niet meer informatie bekend is van de respondenten is herweging de beste correctiemethode.

Anders ligt het bij respondenten die wel hebben meegedaan, maar die niet op alle vragen hebben geantwoord. Deze vorm van non-respons wordt item non-respons genoemd. Soms weet men het antwoord echt niet en in andere gevallen weigert men een antwoord te geven. Dit laatste gebeurt vaker bij variabelen waarbij bedragen worden gevraagd (huurprijs, woningwaarde, inkomen, hypotheek).

Indien een onderzoeksbestand item non-respons bevat liggen er twee grote gevaren op de loer. Het eerste gevaar is dat iedere gebruiker zijn eigen weg gaat om een oplossing te zoeken voor het verwijderen van de missende waarden. Voor de hand liggende methodes hiervoor zijn onder andere: herwegen per tabel, schatten met behulp van een regressiemodel, of alle ontbrekende waarden vervangen door het gemiddelde. Enerzijds kleven aan alle bovenstaande methoden bezwaren (tijdrovend, onderschatting van

de spreiding, of verkeerde veronderstelling over het gedrag van de non-respondent), en anderzijds brengen verschillende onderzoekers zodoende ook verschillende getallen naar buiten brengen voor dezelfde vraagstukken (variabelen).

Het tweede gevaar is dat een onderzoeker niets doet met de ontbrekende waarden en dus impliciet veronderstelt dat non-respons verdeeld is conform de respons – men moet zich echter wel realiseren dat niet bijschatten in feite ook een vorm van bijschatten is! Menig onderzoek heeft echter uitgewezen dat deze veronderstelling meestal onjuist is, omdat non-respons vaak selectief is. Deze hangt meestal samen met bepaalde kenmerken van de respondent. Als er voor gekozen wordt om de non-respons op een of ander manier bij te schatten, dan wordt expliciet verondersteld dat de bijschatting tot een meer betrouwbare voorspelling leidt. De methode waarmee de non-respons wordt bijgeschat bepaalt dus in hoge mate de kwaliteit van het bestand in termen van geldigheid en betrouwbaarheid.

Dé methode om de ontbrekende waarden bij te schatten is de methode van imputatie. Little en Rubin uit de Verenigde Staten hebben vanaf 1980 baanbrekend werk gedaan op dit terrein. Er bestaan diverse imputatietechnieken waarmee praktisch aan alle bezwaren van bestaande ‘oplossingen’ tegemoet wordt gekomen.

Alle technieken hebben gemeen dat de ‘verdeling’ van de geïmputeerde waarden goed blijft. Indien een persoon een vraag niet heeft beantwoord, wordt naar andere personen gezocht die de vraag wel hebben beantwoord. De personen die zoveel mogelijk dezelfde hiermee samenhangende kenmerken hebben als de persoon met het ontbrekende antwoord, vormen de donorgroep. Uit deze donorgroep wordt willekeurig een persoon getrokken. De wel ingevulde waarde door deze donor wordt bij de ontbrekende waarde ingevuld. De kenmerken op basis waarvan de donorgroep samengesteld wordt vormen een zogenaamde relatietabel.

Imputatiemethoden hebben de eigenschap dat niet getracht wordt om de beste schatting op recordniveau te vinden, maar om de verdeling van de bij te schatten doelvariabele zo veel mogelijk te handhaven gegeven de kenmerken van de donoren. Dit heeft het grote voordeel dat de variantie van de bijgeschatte variabele aanzienlijk minder onderschat wordt in vergelijking met andere technieken.

2.4.2 Voorbereidingen imputatie

Routing vragenlijst

Een belangrijk onderdeel voordat met de imputaties begonnen kan worden is een correcte routing. Door gebruik van computergestuurde vragenlijsten is het mogelijk om een zeer complexe routingstructuur in de vragenlijst aan te leggen. Voor het imputatieproces was het noodzakelijk de routing geheel en perfect na te bouwen. Voor elke niet beantwoorde vraag moet namelijk worden nagegaan of deze een NVT (Niet Van Toepassing) is of TOGA (Ten Onrechte Geen Antwoord). Een bijschatting van een routingvraag heeft consequenties voor alle daaropvolgende vragen. Deze zijn eerst NVT, maar kunnen door de bijgeschatte routingvraag TOGA worden (vervolg non-respons). Eerder is in paragraaf 2.1 beschreven op welke wijze de routing is gecontroleerd.

Registerdata

Om bij de imputaties gebruik te kunnen maken van zoveel mogelijk relevante extra informatie heeft het CBS registerdata aan het veldwerkbestand gekoppeld. Hierbij gaat het om GBA-gegevens, de WOZ-waarde, de huurtoeslag, inkomensbegrippen, energieverbruik en regionale indelingen.

2.4.3 Imputatiemethoden

De methode zoals in paragraaf 2.4.1 is uiteengezet wordt toegepast bij bijna alle imputaties. Deze methode heet Random Hotdeck Methode (RHM). In een aantal gevallen is gebruik gemaakt van simultaan imputeren. Indien een groep variabelen ongeveer hetzelfde aantal ontbrekende waarden hebben die daarbij sterk met elkaar samenhangen, dan worden deze tegelijkertijd bijgeschat. Onderwerpen die door een gewijzigde routing van 'Niet van Toepassing' naar 'Ten Onrechte Geen Antwoord' (TOGA) worden getransformeerd komen hiervoor meestal in aanmerking.

Een goed voorbeeld hiervan betreft de imputatie van de huur die een respondent voor de huidige woning betaalt. Voor sommige huurders zijn zaken als water, verwarming en elektriciteit inbegrepen en voor andere weer niet. Dit moet echter wel in de huurprijs zichtbaar blijven omdat een huurprijs inclusief andere zaken doorgaans hoger is dan exclusief. In het geval de huurprijs ten onrechte niet opgegeven is, blijven ook de mogelijk inbegrepen posten 'leeg'. Om niet tot onlogische bijschattingen voor de huurprijs te komen gelet op de inbegrepen posten, zal hier rekening mee gehouden moeten worden door de inbegrepen posten gelijktijdig eveneens bij te schatten.

De imputatie wordt in feite afgestemd op een centrale variabele (huurprijs) waarbij gebruik gemaakt wordt van RHM. Indien eenmaal een geschikte donor gevonden is, worden alle samenhangende variabelen simultaan overgenomen.

2.4.4 Uitvoering imputatie

Volgorde van bijschatting

Het WoON is zo opgezet dat de relatief eenvoudig te beantwoorden vragen in het begin staan en de privacygevoelige vragen aan het einde van de vragenlijst: eerst de demografische gegevens over het huishouden, dan achtereenvolgens de vragen over de huidige woning, eventueel de vorige woning, eventueel de gewenste woning, en tot slot werk en opleiding van zowel de ondervraagde als van de partner.

Elk blok kent zijn centrale variabelen, zoals de huur- en koopprijs bij de huidige woning. Deze vragen zijn meestal 'bedragen'. Hierin zit de meeste non-respons. De non-respons neemt toe naarmate de vragenlijst vordert. Dit heeft niet alleen te maken met afgebroken interviews maar ook met de privacygevoeligheid van de vragen. De opzet van het WoON zorgt ervoor dat sequentieel de blokken kunnen worden afgelopen.

Consistentiecontroles

Hoewel de consistentiecontroles voor de imputatie zijn uitgevoerd, kan het alsnog voorkomen dat een geïmputeerde waarde in bepaalde combinaties met andere waarden toch niet consistent is. Daarom wordt

zekerheidshalve na elke imputatie van een variabele het relevante deel van het consistentieprogramma herhaald.

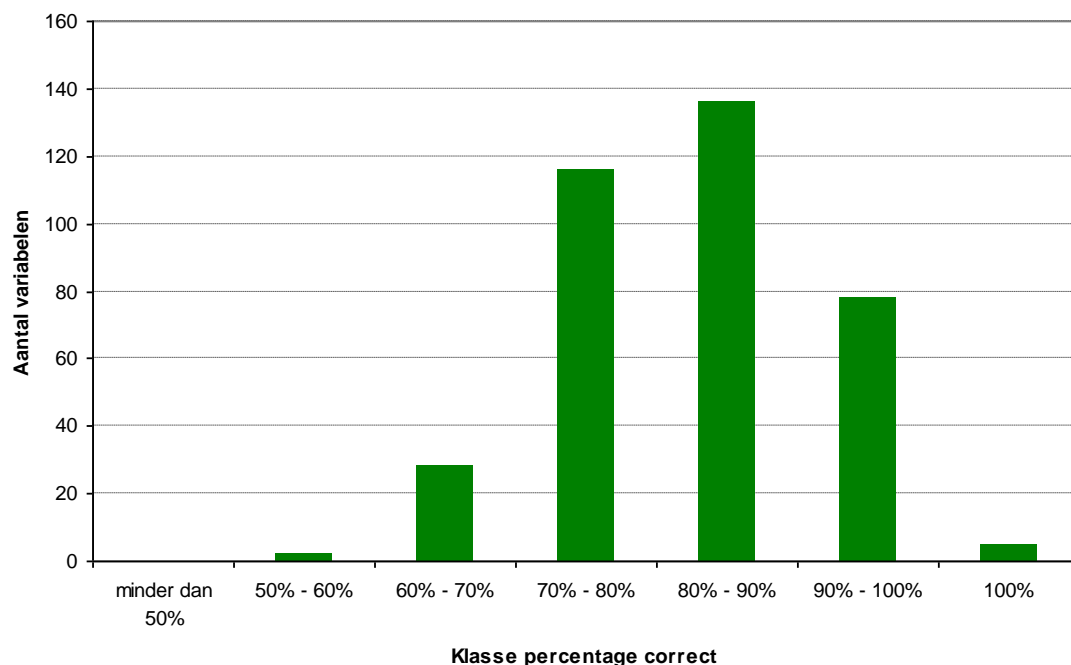
Ook extreme waarden kunnen bij een bepaalde omvang tot relevante beleidsinformatie leiden. Het probleem daarbij is dat de extreme waarden op landelijk niveau wel tot goede gemiddelden en verdelingen kunnen leiden, maar op regionaal niveau voor grote uitschieters kunnen zorgen. De uitkomsten van het WoON zullen doorgaans als gemiddelden (en niet als mediaan) gepresenteerd worden, waardoor de extreme waarden – vanuit regionaal oogpunt – extra moeten worden gecontroleerd. Om te voorkomen dat extreme waarden, waarbij bovendien geen inconsistenties optreden, als donor gaan dienen, worden deze records buiten de donorgroepen gelaten. Dit voorkomt overmatig imputatie van extreme waarden.

Gebruikte software en output

De imputaties zijn uitgevoerd met het pakket Surfox dat ABF Research zelf ontwikkeld heeft. Ook voorgaande WBO's en het WoON 2006 en 2009 zijn hiermee bijgeschat. Surfox kent een interactieve versie en een batchversie Surbox. Met Surfox zijn de imputatiemodellen gemaakt, de zogenaamde relatietabellen, waaruit de verklaringsgraad is af te lezen. De resultaten, zoals het gemiddeld aantal donorrecords en een theoretisch percentage voor correcte bijstellingen, worden bewaard in een outputfile.

In totaal zijn 365 variabelen (en hulpvariabelen) in het WoON 2012 bijgeschat. Het laagste percentage correct is 56%. Tweederde van de imputaties hebben geleid tot een percentage correct tussen 70% en 90%. Rekening houdend met het aantal records dat per variabele geïmputeerd is, komt het gewogen gemiddelde van het overall percentage correct uit op 80%.

Figuur 1 Aantal imputaties per klasse van percentage correct



Het maken van de relatietabellen is handwerk en moet daarom interactief gebeuren. Na de imputatie van een variabele kan de routing ervoor zorgen dat variabelen die eerst NVT waren, worden getransformeerd naar TOGA. Ook deze wijzigingen moeten handmatig worden uitgevoerd. Voor de selectie van variabelen die worden geïmputeerd wordt aangesloten bij de vragenlijst van het WoON 2009. In principe worden de vragenlijstvariabelen die in het WoON 2009 werden geïmputeerd ook nu geïmputeerd. Daarom zijn ook de relatietabellen van 2009 als uitgangspunt voor 2012 gehanteerd.

Nadat op deze wijze het hele bestand is bijgeschat, zijn alle handelingen achter elkaar gezet in een inputbestand. Dit inputbestand is geschikt voor de batchversie Surbox. Omdat na elk blok controles worden uitgevoerd met betrekking tot extreme waarden, inconsistenties en dergelijke is de inputfile in blokken opgeknipt. Bij geconstateerde fouten was het daarom niet nodig om het gehele bestand opnieuw te draaien. De output met daarin de relatietabellen inclusief verklaringsgraad en de imputatieresultaten beslaan ongeveer 50.000 regels.

Imputatie van WOZ-waarde en huurpunten

In een later stadium zijn op verzoek van het CBS en het Ministerie van Binnenlandse Zaken de ontbrekende WOZ-waarde en huurpunten (uiteraard alleen voor respondenten in een huurwoning) bijgeschat. In totaal gaat het om 1499 bijschattingen van de WOZ-waarde, waarvoor uiteindelijk een gemiddeld percentage correct van 75% is behaald. Voor de bijschatting van de huurpunten is een aparte imputatieprocedure uitgevoerd. Deze is beschreven in notitie n2012-0201BM en integraal in de bijlage van dit rapport opgenomen.

2.5 Bepalen gewenste woonmilieus

Het gewenste woonmilieu is een belangrijke variabele voor het Ministerie van Binnenlandse Zaken. Het is dan ook aan de ene kant van groot belang dat voor de berekening van deze variabele aangesloten wordt bij de werkwijze die in het verleden gehanteerd is. Aan de andere kant heeft men te maken met veranderingen in de vragenlijst om op een nog betere manier de gewenste woonmilieus af te leiden.

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 (en dus gecodeerd kan worden), dan wordt hier het (gewenste) woonmilieu aan gekoppeld. De gekoppelde woonmilieus zijn gebaseerd op de meeste recente typologie die beschikbaar is, te weten de typologie van 2009.

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 als de gewenste regio, de gewenste woonplaatsgrootte, de gewenste afstand tot het centrum van de woonplaats (plus of men binnen of buiten de bebouwde kom wil wonen), de gewenste bouwperiode van de buurt, het gewenste woningtype van de buurt en de gewenste menging in de buurt. Voor een betrouwbare imputatie is het daarom van groot belang dat zoveel mogelijk respondenten een gewenste buurt noemen.

Sinds het WoON 2009 is een aantal nieuwe vragen in de vragenlijst opgenomen met betrekking tot de kenmerken van de gewenste woonomgeving in relatie tot voorzieningen. Deze zijn betrokken bij de imputatie van het gewenste woonmilieu.

2.6 Bijkomende woonlasten

De vaste woonuitgaven betreffen de uitgaven van huurders en eigenaren aan respectievelijk huur en hypotheek. Daarnaast geven bewoners geld uit voor het feitelijk 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 nutsvoorzieningen zoals water, gas, warmte en elektriciteit.

Het WoON bevat een grote hoeveelheid aan informatie over de woonsituatie en woonuitgaven van de respondent. Daarnaast wordt een deel van de informatie uit aanvullende bronnen gehaald. Indien een bronhouder voor een gemeente, waterschap of nutsbedrijf geen tarieven heeft kunnen aanleveren, heeft ABF Research deze gegevens in eigen beheer verzameld. De aanvullende informatie betreft:

- de gemeentelijke lasten;
- de waterschapsheffingen;
- de kosten van energie en water.

2.6.1 Gemeentelijke lasten

Heffingen

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.

2.6.2 Waterschapsheffingen

Heffingen

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.

Grondslagen

De watersysteemheffing ingezetenen betreft een bedrag per woonruimte (huishouden). De watersysteemheffing gebouwd is gebaseerd op de WOZ-waarde (zie ook paragraaf 2.6.1). En de zuiveringsheffing wordt bepaald aan de hand van het aantal leden binnen het huishouden. Er geldt een tarief per vervuilingseenheid (VE). De heffing van een eenpersoonshuishouden is gebaseerd op 1 VE en die van een meerpersoonshuishouden op 3 VE.

2.6.3 Verbruik en 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 ABF Research via het CBS verkregen.

Tarieven

ABF Research heeft de tarieven van gas, warmte en elektriciteit van diverse energiebedrijven verzameld. In verband met het vrijgeven van de energiemarkt per 1 juli 2004 wordt aangaande de leveringstarieven uitgegaan van een landelijk gemiddelde.

Vewin levert de tarieven van de waterbedrijven. ABF Research houdt een overzicht bij welke waterbedrijven in welke gemeente leverancier is.

Verbruik

De verbruikgegevens zijn door het CBS gekoppeld, mits de respondent daar toestemming voor heeft gegeven én deze volgens het WoON in een woning woont. De verbruikgegevens zijn door het CBS gekoppeld, mits de respondent daar toestemming voor heeft gegeven en de respondent volgens het WoON in een woning woont. Het CBS heeft verbruikgegevens geïmputeerd als er geen verbruik beschikbaar is uit registratie of als het om een bedrijfsruimte gaat. Het gasverbruik is ook geïmputeerd als er stadsverwarming aanwezig is in de woning of als het gasverbruik nul is.

2.6.4 Overig

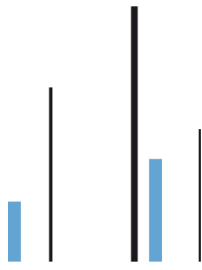
WOZ-waarden

Het CBS beschikt over de WOZ-waarde van alle adressen in Nederland. Op basis van het adres van de respondenten heeft het CBS de WOZ-waarden (peildatum 1 januari 2011) als registerdata gekoppeld aan het veldwerkbestand.

Constanten

Het Ministerie van Binnenlandse Zaken heeft een zogenaamde lijst van constanten aangeleverd die (ten dele) gebruikt zijn voor de berekening van de woonlasten. Deze zijn in de bijlage terug te vinden, evenals de algehele afleiding van de bijkomende woonlasten die rechtstreeks is overgenomen uit de syntax-files die gebruikt zijn tijdens de dataprocesing.

3

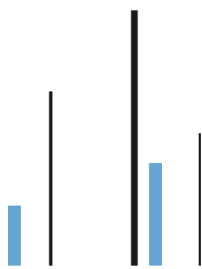


Eindproducten

Het uiteindelijke resultaat van de dataprocessing die ABF Research verricht heeft is een eindbestand dat exact hetzelfde ingericht is als het veldwerkbestand, maar dan gevuld met de bewerkte enquêtedata. Verder zijn separaat de extra aangemaakte variabelen opgeleverd, waaronder alle relevante variabelen die tijdens de berekening van de woonlasten zijn aangemaakt.

Daarnaast is ook een vlaggenbestand gemaakt. In dit bestand worden de waarden van de originele variabelen vergeleken met de waarden na de dataprocessing. Het vlaggenbestand bevat alle variabelen van de module Woningmarkt, maar elke variabele heeft een waarde die aangeeft of de waarde voor het betreffende record een ongewijzigde waarde, een consistentieaanpassing, of een bijgeschatte waarde bevat.

Tot slot is ook alle syntax beschikbaar gesteld. Dit betreft zowel de syntax voor SPSS als voor Surfox. Beide softwareprogramma's zijn gebruikt tijdens de uitvoering van de verschillende onderdelen van de dataprocessing.



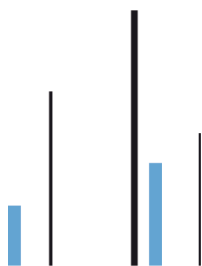
Bijlage 1 Consistentiecontroles

1. Het bouwjaar van de huidige woning ligt niet na het jaar dat de respondent op het huidige adres is komen wonen.
2. Het jaar dat de respondent op het huidige adres is komen wonen ligt niet voor het geboortjaar van de respondent.
3. Het jaar dat de huurwoning is gekocht ligt niet na het jaar dat de respondent op het huidige adres is komen wonen.
4. Het jaar dat de huurwoning is gekocht ligt niet na het geboortjaar van de respondent.
5. Het jaar dat de huurwoning is gekocht ligt niet na het bouwjaar van de woning.
6. Het verschil tussen het jaar van aankoop van de huidige woning en het jaar van komen wonen op dat adres mag niet meer dan vijf jaar zijn.
7. Het jaar dat de koopwoning is gekocht ligt niet voor het geboortjaar van de respondent.
8. Het jaar dat de koopwoning is gekocht ligt niet voor het bouwjaar van de woning.
9. Indien de woonkamer op de zevende verdieping of hoger ligt, is er altijd een lift aanwezig.
10. Indien de woonkamer op de zevende verdieping of hoger kan er geen sprake zijn van een eigen ingang aan de straat.
11. In het geval van een eenkamerwoning moet de totale woonoppervlakte groter of gelijk zijn aan de oppervlakte van de woonkamer.
12. In het geval van een eenkamerwoning moet de totale woonoppervlakte (in klassen) groter of gelijk zijn aan de oppervlakte van de woonkamer.
13. In het geval van een eenkamerwoning moet de totale woonoppervlakte groter of gelijk zijn aan de oppervlakte van de woonkamer (in klassen).
14. In het geval van een eenkamerwoning moet de totale woonoppervlakte (in klassen) groter of gelijk zijn dan de oppervlakte van de woonkamer (in klassen).
15. In het geval van een meerkamerwoning moet de totale woonoppervlakte groter of gelijk zijn aan de oppervlakte van de woonkamer.
16. In het geval van een meerkamerwoning moet de totale woonoppervlakte (in klassen) groter of gelijk zijn aan de oppervlakte van de woonkamer.
17. In het geval van een meerkamerwoning moet de totale woonoppervlakte groter of gelijk zijn aan de oppervlakte van de woonkamer (in klassen).
18. In het geval van een meerkamerwoning moet de totale woonoppervlakte (in klassen) groter of gelijk zijn dan de oppervlakte van de woonkamer (in klassen).
19. Het jaar dat de respondent op het vorige adres is gaan wonen ligt niet na het jaar dat de respondent op het huidige adres is komen wonen, en na het geboortjaar van de respondent.
20. Het bouwjaar van de vorige woning ligt niet het jaar dat de respondent op het vorige adres is gaan wonen.

21. Het bedrag aan maandelijkse hypotheekrente mag niet meer dan 10% van de totale hypotheeksom bedragen.
22. De koopprijs van een voormalige huurwoning mag niet meer dan 100 maal de (voormalige) betaalde jaarhuur is en meer dan 500.000 euro bedragen.
23. Het aantal verdiepingen met een rookmelder is niet groter dan het aantal verdiepingen van de woonruimte.
24. Het aantal verdiepingen van de woonruimte is niet groter is dan het aantal verdiepingen van het gebouw waar het deel van uitmaakt.
25. Als de huidige woning niet als bestaande woning gekocht is, mag het bouwjaar niet meer dan vijf jaar kleiner zijn dan het jaar van komen wonen op het huidige adres.
26. Als de huidige woning niet als bestaande woning gekocht is, kan het bouwjaar niet liggen in een bouwjaarklasse die meer dan vijf jaar kleiner is dan het jaar van komen wonen op het huidige adres.
27. Als onbekend is of de huidige woning via particulier opdrachtgeverschap is aangeschaft, wordt aangenomen dat dit niet geval is en dat de huidige woning als reeds bestaande woning is gekocht.

Tabel B1.3 Aantal gevonden inconsistenties per controle (voor bewerking)

Controle	Omschrijving	Aantal inconsistenties
1	Bouwjaar huidige woning	691
2	Jaar op huidig adres komen wonen	281
3	Jaar huurwoning gekocht (1)	53
4	Jaar huurwoning gekocht (2)	2
5	Jaar huurwoning gekocht (3)	75
6	Jaar koopwoning gekocht (1)	382
7	Jaar koopwoning gekocht (2)	37
8	Jaar koopwoning gekocht (3)	724
9	Lift bij hooggelegen woonkamer	13
10	Eigen ingang aan straat bij hooggelegen woonkamer	44
11	Oppervlakte woonkamer en woonoppervlak (1)	28
12	Oppervlakte woonkamer en woonoppervlak (2)	21
13	Oppervlakte woonkamer en woonoppervlak (3)	9
14	Oppervlakte woonkamer en woonoppervlak (4)	0
15	Oppervlakte woonkamer en woonoppervlak (5)	1176
16	Oppervlakte woonkamer en woonoppervlak (6)	1637
17	Oppervlakte woonkamer en woonoppervlak (7)	1503
18	Oppervlakte woonkamer en woonoppervlak (8)	1
19	Jaar op vorig adres komen wonen	48
20	Bouwjaar vorige woning	1188
21	Maandelijks bedrag hypotheekrente	278
22	Aankoopprijs voormalige huurwoning	20
23	Verdiepingen woning/rookmelders	204
24	Bouwlagen woongebouw	1004
25	Bouwjaar gekochte, niet-bestaande woonruimte (1)	1464
26	Bouwjaar gekochte, niet-bestaande woonruimte (2)	9590
27	Bestaande woning en opdrachtgeverschap	981



Bijlage 2 Bijkomende woonlasten

Tabel B2.4 Lijst van constanten van belang voor de woonlastenberekening

constante	label	waarde	
		WoON 2012	bron 2012
GRENSINK	grens modaal inkomen	32.500	CPB
MINLOON	wettelijk minimumloon (incl. 8% vakantiegeld)	18.530	Ministerie SZW
SMALL1	sociaal minimumloon alleenstaand 21-	2.730	Ministerie SZW
SMALL2	sociaal minimum alleenstaand 21+	11.062	Ministerie SZW
SMALL3	sociaal minimum alleenstaand 65+	12.131	Ministerie SZW
SMEOG1	sociaal minimum eenoudergezin 21-	5.890	Ministerie SZW
SMEOG2	sociaal minimum eenouder 21+	14.222	Ministerie SZW
SMEOG3	sociaal minimum eenouder 65+	14.658	Ministerie SZW
SMEPZ1	sociaal minimum echtpaar zonder kinderen beiden 21-	5.460	Ministerie SZW
SMEPZ2	sociaal minimum echtpaar zonder kinderen 21+/21-	10.631	Ministerie SZW
SMEPM1	sociaal minimum echtpaar beiden 21+	15.802	Ministerie SZW
SMEPM2	sociaal minimum echtpaar met kinderen beiden 21-	8.621	Ministerie SZW
SMEP3	sociaal minimum echtpaar met kinderen 21+/21-	13.792	Ministerie SZW
SMEP4	sociaal minimum echtpaar 65+	16.699	Ministerie SZW
SMEP5	sociaal minimum echtpaar 65+/65-	16.699	Ministerie SZW
	bedrag verwarming tussenwoning	89	Agentschap NL / NIBUD
	bedrag verwarming hoekwoning	102	Agentschap NL / NIBUD
	bedrag verwarming twee onder één kap	107	Agentschap NL / NIBUD
	bedrag verwarming vrijstaand	139	Agentschap NL / NIBUD
	bedrag verwarming appartement	68	Agentschap NL / NIBUD
FLBYSOCE	bedrag service soc verh eengezinswoning	5	CBS huurenquête 2009, 2010, 2011
FLBYSOCM	bedrag service soc verh meergezinswoning	30	CBS huurenquête 2009, 2010, 2011
FLBYPAVE	bedrag service part verh eengezinswoning	15	CBS huurenquête 2009, 2010, 2011
FLBYPAVM	bedrag service part verh meergezinswoning	50	CBS huurenquête 2009, 2010, 2011
FLBYPAPE	bedrag service part pers eengezinswoning	10	CBS huurenquête 2009, 2010, 2011
FLBYPAPM	bedrag service part pers meergezinswoning	50	CBS huurenquête 2009, 2010, 2011
FLBYANDE	bedrag service anders eengezins	15	CBS huurenquête 2009, 2010, 2011
FLBYANDM	bedrag service anders meergezins	50	CBS huurenquête 2009, 2010, 2011
BWATER	bedrag water	21	CBS - CPI Watervoorziening
BCAI	bedrag CAI	17	NIBUD Budgethandboek 2011
BKABEL	bedrag kabel	17	NIBUD Budgethandboek 2011
KOOPBT	bovengrens betaalbare koopwoningen	350.000	www.nhg.nl - Normen 2011
HUURKWAL	kwaliteitskortinggrens huursubsidie	361,66	MG 2011-05
HUURAFTE	onder aftoppingsgrens 1 en 2 p hh	517,64	MG 2011-05
HUURAFTM	onder aftoppingsgrens 3 en meer p hh	554,76	MG 2011-05
WIHSBEN	benedengrens huur woning IHS	208,61	MG 2010-04
WIHSBOV	bovengrens huur woning IHS	652,52	MG 2011-05
POTIHSJE	grens inkomen aandachtsgroep/IHS eenpersoons jonger dan 65	22.650	MG 2011-05
POTIHSJM	grens inkomen aandachtsgroep/IHS meerpersoons jonger dan 65	29.450	MG 2011-05
POTIHSOE	grens inkomen aandachtsgroep/IHS eenpersoons 65 jaar of ouder	21.275	MG 2011-05
POTIHSOM	grens inkomen aandachtsgroep/IHS meerpersoons 65 jaar of ouder	28.550	MG 2011-05
VERMJE	grens vermogen aandachtsgroep/IHS eenpersoons jonger dan 65	20.785	Regeling Huurtoeslag 2011
VERMJM	grens vermogen aandachtsgroep/IHS meerpersoons jonger dan 65	*	Regeling Huurtoeslag 2011
VERMOE	grens vermogen aandachtsgroep/IHS eenpersoons 65 jaar of ouder	**	Regeling Huurtoeslag 2011
VERMOM	grens vermogen aandachtsgroep/IHS meerpersoons 65 jaar of ouder	***	Regeling Huurtoeslag 2011
OPSTALP	percentage voor opstalverzekering	0,83%	NIBUD Budgethandboek 2011
TWEEGREN	grens inkomen voor tweeverdieners	100	Ministerie BZK / WBI
WOZEENH	WOZ-eenheid	2.500	Ministerie BZK / WBI

* € 20.785 + € 2.779 voor elk minderjarig kind

** (i) € 48.301 als inkomen minder dan € 14.062; (ii) € 34.543 als inkomen minimaal € 14.062 maar minder dan € 19.562; (liii) € 20.785 als inkomen meer dan € 19.562

*** (i) € 48.301 per persoon als inkomen minder dan € 14.062; (ii) € 34.543 per persoon als inkomen minimaal € 14.062 maar minder dan € 19.562; (iii) € 20.785 per persoon als inkomen meer dan € 19.562

Figuur B.2 Syntax berekening bijkomende woonlasten

```

* =====
* Aanmaken relevante variabelen bijkomende woonlasten
* =====
* Huishoudensgrootte in 4 klassen (HH4)
IF: AANTALPP=1
COMPUTE: HH4=1
ELSEIF: AANTALPP=2
COMPUTE: HH4=2
ELSEIF: AANTALPP=3
COMPUTE: HH4=3
ELSEIF: AANTALPP=4
COMPUTE: HH4=3
ELSEIF: AANTALPP>=5
COMPUTE: HH4=4
ELSE
COMPUTE: HH4=NVT
ENDIF

* BESTEEDBAAR INKOMEN VOLGENS CBS OVERNEMEN
COMPUTE: DUMMY3=Woon_bes

IF: DUMMY3<29500
COMPUTE: PINK4=1
ELSEIF: DUMMY3<1.5*29500
COMPUTE: PINK4=2
ELSEIF: DUMMY3<2*29500
COMPUTE: PINK4=3
ELSEIF: DUMMY3>=2*29500
COMPUTE: PINK4=4
ELSE
COMPUTE: PINK4=NVT
ENDIF

* Woningen in 4 typen (WON4)
IF: HVS<=3&TYPEWON=1&SOORTWON=1
COMPUTE: WON4=1
ELSEIF: HVS<=3&TYPEWON=1&SOORTWON=3
COMPUTE: WON4=1
ELSEIF: HVS<=3&TYPEWON=1&SOORTWON=4
COMPUTE: WON4=1
ELSEIF: HVS<=3&TYPEWON=1&SOORTWON>=5
COMPUTE: WON4=1
ELSEIF: HVS<=3&(TYPEWON>=2&TYPEWON<5)&SOORTWON=1
COMPUTE: WON4=2
ELSEIF: HVS<=3&(TYPEWON>=2&TYPEWON<5)&SOORTWON=3
COMPUTE: WON4=2
ELSEIF: HVS<=3&(TYPEWON>=2&TYPEWON<5)&SOORTWON=4

```



```

COMPUTE: WON4=2
ELSEIF: HVS<=3&(TYPEWON>=2&TYPEWON<5) &SOORTWON>=5
COMPUTE: WON4=2
ELSEIF: HVS<=3&(TYPEWON=6) &SOORTWON=1
COMPUTE: WON4=2
ELSEIF: HVS<=3&(TYPEWON=6) &SOORTWON=3
COMPUTE: WON4=2
ELSEIF: HVS<=3&(TYPEWON=6) &SOORTWON=4
COMPUTE: WON4=2
ELSEIF: HVS<=3&(TYPEWON=6) &SOORTWON>=5
COMPUTE: WON4=2
ELSEIF: HVS<=3&TYPEWON=NVT&SOORTWON=2
COMPUTE: WON4=3
ELSEIF: HVS<=3&TYPEWON=5
COMPUTE: WON4=3
ELSEIF: HVS<=3&TYPEWON=NVT&SOORTWON>=5
COMPUTE: WON4=3
ELSEIF: HVS>=4&HVS<=5
COMPUTE: WON4=4
ELSE
COMPUTE: WON4=NVT
ENDIF

* =====
* Importeren tarieven Gemeentelijke heffingen
* =====

IMPORT :FILE= tar2012
:DIRECT= ...\DataProc\7_Woonlasten\1Data2012
:MATCHVAR1=CBSNR
:MATCHVAR2=GEBIED
:VAR=TPEIGW
&,RIOSYS,RIO1TAR,RIOOTARE,RIOOTAR1,RIOOTAR3
&,RIO1,RIO2,RIO3,RIO4,RIO5
&,RIOVAST
&,V025,V2550,V5060,V6075,V75100,V100125
&,V125150,V150170,V170175,V175200,V200225,V225250,V250300
&,RIOPERM3,M080,M80150,M150200,M200250,M250EM
&,RIOWOZG,RIOWOZE
&,REINSYS,REINTAR1,REINTAR2,REINTAR3,REINTAR4

Compute: riomin=0

* =====
* Onroerende Zaak Belasting
* =====
* bedrag voor eigenaar woning; bedrag o.b.v. woningwaarde (WOZ)
IF: HVS<>6&HUKO=1

```

```
COMPUTE: OZB=(TPEIGW/100)*(WAARWON*1000)
ELSEIF: HVS<>6&(HUKO=2|HUKO=NVT)
COMPUTE: OZB=0
ELSE
COMPUTE: OZB=NVT
ENDIF
```

```
* =====
* Rioolrecht
* =====
```

* WATERV betreft gekoppelde verbruiken

```
IF: HVS<>6
COMPUTE: WATERV=WATERV
ELSE
COMPUTE: WATERV=NVT
ENDIF
```

IF: WATERV=NVT|WATERV=TOGA

```
COMPUTE: WATERV=0
ELSE
COMPUTE: WATERV=WATERV
ENDIF
```

* Eigenaarsdeel

```
* =====
```

```
IF: HVS<>6&HUKO=1&RIOWOZE>0
COMPUTE: ERIOOL=RIOWOZE*(WAARWON*1000)
ELSEIF: HVS<>6&HUKO=1
COMPUTE: ERIOOL=RIOOTARE
ELSEIF: HVS<>6&(HUKO=2|HUKO=NVT)
COMPUTE: ERIOOL=0
ELSE
COMPUTE: ERIOOL=NVT
ENDIF
```

* Gebruikersdeel

```
* =====
```

* RIOSYS=1 Vast bedrag gebruikersheffing

```
IF: HVS<>6&RIOSYS=1
COMPUTE: GRIOOL=RIO1TAR
ELSE
COMPUTE: GRIOOL=NVT
ENDIF
```

* RIOSYS=2 Gebruikersheffing afh van aantal personen

```
IF: HVS<>6&RIOSYS=2&AANTALPP=1
COMPUTE: GRIOOL=RIO1
ELSEIF: HVS<>6&RIOSYS=2&AANTALPP=2
COMPUTE: GRIOOL=RIO2
```

```

ELSEIF: HVS<>6&RIOSYS=2&AANTALPP=3
COMPUTE: GRIOOL=RIO3
ELSEIF: HVS<>6&RIOSYS=2&AANTALPP=4
COMPUTE: GRIOOL=RIO4
ELSEIF: HVS<>6&RIOSYS=2&AANTALPP>4
COMPUTE: GRIOOL=RIO5
ELSE
COMPUTE: GRIOOL=GRIOOL
ENDIF

* RIOSYS=3 Gebruikersheffing afh van waterverbruik
IF: HVS<>6&RIOSYS=3
COMPUTE: GRIOOL=RIOVAST
ELSE
COMPUTE: GRIOOL=GRIOOL
ENDIF

IF: HVS<>6&RIOSYS=3&WATERV<=25
COMPUTE: GRIOOL=GRIOOL+V025
ELSEIF: HVS<>6&RIOSYS=3&WATERV<=50
COMPUTE: GRIOOL=GRIOOL+V2550
ELSEIF: HVS<>6&RIOSYS=3&WATERV<=60
COMPUTE: GRIOOL=GRIOOL+V5060
ELSEIF: HVS<>6&RIOSYS=3&WATERV<=75
COMPUTE: GRIOOL=GRIOOL+V6075
ELSEIF: HVS<>6&RIOSYS=3&WATERV<=100
COMPUTE: GRIOOL=GRIOOL+V75100
ELSEIF: HVS<>6&RIOSYS=3&WATERV<=125
COMPUTE: GRIOOL=GRIOOL+V100125
ELSEIF: HVS<>6&RIOSYS=3&WATERV<=150
COMPUTE: GRIOOL=GRIOOL+V125150
ELSEIF: HVS<>6&RIOSYS=3&WATERV<=170
COMPUTE: GRIOOL=GRIOOL+V150170
ELSEIF: HVS<>6&RIOSYS=3&WATERV<=175
COMPUTE: GRIOOL=GRIOOL+V170175
ELSEIF: HVS<>6&RIOSYS=3&WATERV<=200
COMPUTE: GRIOOL=GRIOOL+V175200
ELSEIF: HVS<>6&RIOSYS=3&WATERV<=225
COMPUTE: GRIOOL=GRIOOL+V200225
ELSEIF: HVS<>6&RIOSYS=3&WATERV<=250
COMPUTE: GRIOOL=GRIOOL+V225250
ELSEIF: HVS<>6&RIOSYS=3&WATERV<=300
COMPUTE: GRIOOL=GRIOOL+V250300
ELSE
COMPUTE: GRIOOL=GRIOOL
ENDIF

IF: HVS<>6&RIOSYS=3
COMPUTE: GRIOOL=GRIOOL+WATERV*RIOPERM3

```

```

ELSE
COMPUTE: GRIOOL=GRIOOL
ENDIF

COMPUTE: DUM=80*M080+70*M80150
IF: WATERV<=80
COMPUTE: GRIOOL=GRIOOL+WATERV*M080
ELSEIF: WATERV<=150
COMPUTE: GRIOOL=GRIOOL+80*M080+(WATERV-80)*M80150
ELSEIF: WATERV<=200
COMPUTE: GRIOOL=GRIOOL+80*M080+70*M80150+(WATERV-150)*M150200
ELSEIF: WATERV<=250
COMPUTE: GRIOOL=GRIOOL+80*M080+70*M80150+50*M150200+(WATERV-200)*M200250
ELSEIF: WATERV>250
COMPUTE: GRIOOL=GRIOOL+DUM+50*M150200+50*M200250+(WATERV-250)*M250EM
ELSE
COMPUTE: GRIOOL=GRIOOL
ENDIF

IF: HVS<>6&RIOSYS=3&RIOMIN>0&GRIOOL<RIOMIN
COMPUTE: GRIOOL=RIOMIN
ELSE
COMPUTE: GRIOOL=GRIOOL
ENDIF

* RIOSYS=4 Gebruikersheffing afh van WOZ-waarde of Categorie Anders
IF: HVS<>6&RIOSYS=4&RIOWOZG>0
COMPUTE: GRIOOL=RIOWOZG*(WAARWON*1000)
ELSEIF: HVS<>6&RIOSYS=4&AANTALPP=1
COMPUTE: GRIOOL=RIOOTAR1
ELSEIF: HVS<>6&RIOSYS=4&AANTALPP>1
COMPUTE: GRIOOL=RIOOTAR3
ELSE
COMPUTE: GRIOOL=GRIOOL
ENDIF

* RIOSYS=5 geen gebruikersheffing
IF: HVS<>6&RIOSYS=5
COMPUTE: GRIOOL=0
ELSE
COMPUTE: GRIOOL=GRIOOL
ENDIF

* =====
* Reinigingsheffing
* =====
* Met onderscheid naar systematiek van de afvalinzameling:

```

```

* vast bedrag of afhankelijk van huishoudenomvang (REINSYS=1 of 2)
* uitgangspunt: een meerpersoonshh is alles meer dan 1 persoon
IF: HVS<>6&REINSYS<3&AANTALPP=1
COMPUTE: AFVAL=REINTAR1
ELSEIF: HVS<>6&REINSYS<3&AANTALPP=2
COMPUTE: AFVAL=REINTAR2
ELSEIF: HVS<>6&REINSYS<3&AANTALPP=3
COMPUTE: AFVAL=REINTAR3
ELSEIF: HVS<>6&REINSYS<3&AANTALPP>=4
COMPUTE: AFVAL=REINTAR4
ELSE
COMPUTE: AFVAL=NVT
ENDIF

* volume container (REINSYS=3)
* uitgangspunt: tweepersoonshh gebruiken dezelfde type containers als eenpersoonshh
IF: HVS<>6&REINSYS=3&AANTALPP<3
COMPUTE: AFVAL=REINTAR1
ELSEIF: HVS<>6&REINSYS=3&AANTALPP>=3
COMPUTE: AFVAL=REINTAR3
ELSE
COMPUTE: AFVAL=AFVAL
ENDIF

* aantal ledigingen, aantal zakken of gewicht (REINSYS=4,5,6,7,8 of 9)
* uitgangspunt: tweepersoonshh produceren twee maal zoveel afval als eenpersoonshh
(etc.)
IF: HVS<>6&(REINSYS>=4&REINSYS<=9) &AANTALPP=1
COMPUTE: AFVAL=REINTAR1
ELSEIF: HVS<>6&(REINSYS>=4&REINSYS<=9) &AANTALPP=2
COMPUTE: AFVAL=REINTAR1+(REINTAR3-REINTAR1)/2
ELSEIF: HVS<>6&(REINSYS>=4&REINSYS<=9) &AANTALPP=3
COMPUTE: AFVAL=REINTAR3
ELSEIF: HVS<>6&(REINSYS>=4&REINSYS<=9) &AANTALPP=4
COMPUTE: AFVAL=REINTAR3+(REINTAR3-REINTAR1)/2
ELSEIF: HVS<>6&(REINSYS>=4&REINSYS<=9) &AANTALPP>=5
COMPUTE: AFVAL=REINTAR3+2*(REINTAR3-REINTAR1)/2
ELSE
COMPUTE: AFVAL=AFVAL
ENDIF

* overig (REINSYS=10 of 11)
IF: HVS<>6&(REINSYS=10|REINSYS=11) &AANTALPP=1
COMPUTE: AFVAL=REINTAR1
ELSEIF: HVS<>6&(REINSYS=10|REINSYS=11) &AANTALPP>1
COMPUTE: AFVAL=REINTAR3
ELSE
COMPUTE: AFVAL=AFVAL
ENDIF

```

```

* =====
* Waterschapslasten
* =====

* Importeren tarieven waterschapslasten
IMPORT :FILE= tar2012
:DIRECT= ...\DataProc\7_Woonlasten\1Data2012
:MATCHVAR1=CBSNR
:MATCHVAR2=GEBIED
:VAR=WATHEFON,WATHEFI,WATHEFV

* Verontreinigingsheffing
* =====
* bedrag per type huishouden (eenpersoons- of meerpersoons)
IF: HVS<>6&HH4=1
COMPUTE: VERONT=WATHEFV*1
ELSEIF: HVS<>6&HH4>1
COMPUTE: VERONT=WATHEFV*3
ELSE
COMPUTE: VERONT=NVT
ENDIF

* Omslagheffing
* =====
* bedrag voor eigenaar woning; bedrag o.b.v. woningwaarde
IF: HVS<>6&HUKO=1
COMPUTE: OMSHEF=(WATHEFON/100)*(WAARWON*1000)
ELSEIF: HVS<>6&(HUKO=2|HUKO=NVT)
COMPUTE: OMSHEF=0
ELSE
COMPUTE: OMSHEF=NVT
ENDIF

* Ingezetenenheffing
* =====
* bedrag per huishouden
IF: HVS<>6
COMPUTE: INGOMS=WATHEFI
ELSE
COMPUTE: INGOMS=NVT
ENDIF

* =====
* Verbruikskosten
* =====

* Importeren tarieven verschillende type verbruikskosten
IMPORT :FILE= tar2012

```

```

:DIRECT=R:\DPC\2011\11180\Work\DataProc\7_Woonlasten\1Data2012
:MATCHVAR1=CBSNR
:MATCHVAR2=GEBIED
:VAR=VASTWAT, TARWAT, WAT0300, BTWWAT
&, GASNET, ELENET
&, GASVAST, GASLEVT
&, ELEVAST, ELELEVE, ELELEVDN, ELELEVDL

```

```

* =====
* Water
* =====

```

```

* vastrecht
* plus prijs per m3 (waterverbruik)
* plus ecotax tot 300 m3 verbruik
* plus 6% BTW
* importeren waterverbruik naar type huishouden al gedaan

```

```

* Berekenen totaal bedrag
IF: HVS<>6&WATERV>0&WATERV<=300
COMPUTE: WATERB=(VASTWAT+(TARWAT*WATERV)+(WAT0300*WATERV))*BTWWAT
ELSEIF: HVS<>6&WATERV>300
COMPUTE: WATERB=(VASTWAT+(TARWAT*WATERV)+(WAT0300*300))*BTWWAT
ELSEIF: HVS<>6
COMPUTE: WATERB=0
ELSE
COMPUTE: WATERB=NVT
ENDIF

```

```

* =====
* Elektriciteit
* =====

```

```

* netwarktief
* vastrecht stroomlevering
* prijs per kWh
* energielasting
* 19% BTW

```

```

* rekening houdend met enkeltarief en dubbeltarief
* grens: 3300 kWh
IF: ELEC>3300
COMPUTE: ELECND=ELEC/2
ELSE
COMPUTE: ELECND=0
ENDIF

```

```

IF: ELEC>3300
COMPUTE: ELECNDL=ELEC/2

```

```
ELSE
COMPUTE: ELEC DL=0
ENDIF
```

```
IF: ELEC V>3300
COMPUTE: ELECE=0
ELSE
COMPUTE: ELECE=ELEC V
ENDIF
```

```
* Berekenen bedrag stroomlevering
IF: HVS<>6&ELECE>0
COMPUTE: ELELEV=ELECE*ELELEVE
ELSEIF: HVS<>6&ELEC DN>0
COMPUTE: ELELEV=ELEC DL*ELELEV DL + ELEC DN*ELELEV DN
ELSE
COMPUTE: ELELEV=NVT
ENDIF
```

```
* Berekenen totaal bedrag aan stroomuitgaven (incl 19% BTW en energiebelasting)
* netwerktarief plus vastrecht stroomlevering plus verbruik
IF: HVS<>6&ELEC V>0
COMPUTE: ELECB= ELENET + ELEVAST + ELELEV
ELSEIF: HVS<>6
COMPUTE: ELECB= 0
ELSE
COMPUTE: ELECB= NVT
ENDIF
```

```
* =====
* Gas
* =====
* netwerktarief
* vastrecht gaslevering
* prijs per m3 gas
* energiebelasting
* 19% BTW
```

```
* Berekenen totaal bedrag aan gasuitgaven (incl 19% BTW en energiebelasting)
* netwerktarief plus vastrecht gaslevering plus verbruik
IF: HVS<>6&GAS V>0
COMPUTE: GASB= GASNET + GASVAST + GASLEVT * GASV
ELSEIF: HVS<>6
COMPUTE: GASB= 0
ELSE
COMPUTE: GASB= NVT
ENDIF
```



```

* =====
* Omzetten naar maandbedragen
* =====
* Gemeentelijke heffingen
* OZB
COMPUTE: OZBM=OZB/12

* Rioolrecht
COMPUTE: RIOOL=ERIOOL+GRIOOL
COMPUTE: RIOOLM=RIOOL/12

* Reinigingsheffing
COMPUTE: AFVALM=AFVAL/12

* Waterschapslasten
* Omslagheffing
COMPUTE: OMSHEFM=OMSHEF/12

* Ingezetenenheffing
COMPUTE: INGOMSM=INGOMS/12

* Verontreinigingsheffing
COMPUTE: VERONTM=VERONT/12

* Verbruikskosten
* Water
COMPUTE: WATERBM=WATERB/12

* Elektriciteit
COMPUTE: ELECBM=ELECB/12

* Gas
COMPUTE: GASBM=GASB/12

* =====
* Samenstellen van diverse totalen
* =====
* Totaal Gemeentelijke heffingen (bedrag per maand)
COMPUTE: TOTGEM=OZBM+RIOOLM+AFVALM

* Totaal Waterschapslasten (bedrag per maand)
COMPUTE: TOTWS=OMSHEFM+INGOMSM+VERONTM

* Totaal lasten verbruik Water&Energie (bedrag per maand)
COMPUTE: TOTENER=WATERBM+ELECBM+GASBM

* Totaal lasten Overig Publieke Lichamen (bedrag per maand)

```

COMPUTE: TOTOPL=TOTGEM+TOTWS

* Totaal Bijkomende woonlasten (bedrag per maand)

COMPUTE: TOTBIJ=TOTOPL+TOTENER

* =====

* Laatste correctieslag

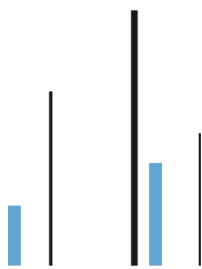
* =====

* IN VERBAND MET DUBBELTELLING OZBM IN MINDERING BRENGEN, OMDAT BIJK (VIA CBS) OOK OZBM BEVAT

COMPUTE: TOTBIJ=TOTBIJ-OZBM

* KLAAR

QUIT



Bijlage 3 Bijschatting huurpunten

1. Inleiding

In het kader van de berekeningen voor de Heroverwegingen Wonen speelt het aantal huurpunten dat toegekend wordt aan huurwoningen een belangrijke rol. Hoe hoger het aantal punten voor de huurwoning, des te hoger de huurprijs die de verhuurder kan vragen.

In het WoON 2012 is het aantal huurpunten evenals in de voorgaande edities een van de vragen die aan de huurders gesteld wordt. Deze bron van informatie wordt over het algemeen niet als betrouwbaar verondersteld, omdat huurders niet altijd even goed op de hoogte zijn van het aantal punten voor de eigen woning. Daarom is aan het WoON registerinformatie gekoppeld, die afkomstig is van NCCW en de verhuurders zelf. Hierbij is echter het probleem, dat slechts voor een beperkt aantal cases de registerdata bekend is. Daarom is besloten om de ontbrekende data bij te schatten. In deze notitie wordt de procedure voor de bijschatting van het aantal punten voor huurwoningen beschreven.

2. Het woningwaarderingstelsel

In de Handleiding Huurbeleid 2011-2012 is uiteengezet welke richtlijnen van toepassing zijn op het huren en verhuren van zelfstandige woonruimten. Daarin staat onder meer ook beschreven hoe de huurprijs tot stand komt. Deze is gebaseerd op een waarderingstelsel waarmee de kwaliteit van de woning in punten uitgedrukt wordt.

Op de volgende onderdelen worden punten toegekend aan de huurwoning:

1. Oppervlakte van vertrekken (kamers, keuken, badkamer en doucheruimte);
2. Oppervlakte overige ruimten (bijkeuken, berging, zolder, garage);
3. Verwarming;
4. Warmte-isolatie;
5. Keuken;
6. Sanitair;
7. Woonvoorzieningen voor gehandicapten
8. Privé buitenruimten;
9. Woonvorm;
10. Woonomgeving;
11. Hinderlijke situaties;
12. Bijzondere voorzieningen (alleen voor serviceflatwoningen).

Alle onderdelen zijn gemaximeerd, behalve 1 en 2. Voor punt 12 geldt – indien van toepassing – een percentuele toename (35%) op het totaal van de onderdelen 1 tot en met 11.

3. Methoden van bijschatten

Onbrekende huurpunten kunnen op verschillende manieren worden bijgeschat. Het **CPB** heeft **programmatuur** opgesteld om aan de hand van de antwoorden in het WoON het aantal punten te bepalen. Hierbij wordt voor elk van de eerder genoemde onderdelen een puntental berekend op basis van de antwoorden van de respondent en de richtlijnen zoals in de Handleiding beschreven. Deze berekening staat in principe los van de opgegeven en gekoppelde punten, en is toepasbaar op alle respondenten; met uitzondering van een selecte groep (HVS=6) voor wie essentiële informatie over de woning ontbreekt.

Een tweede manier van het bijschatten van de huurpunten is door gebruik te maken van een **imputatietechniek**. Voor de respondenten zonder punten worden deze bijgeschat door respondentgroepen met soortgelijke profielen te selecteren. Indien er sprake is van een 'match', dan wordt het antwoord van een respondent mét puntental gekopieerd naar de ontvangende respondent zonder puntental. Binnen deze techniek bestaan er verschillende manieren om een geschikte donorrecord te vinden voor een respondent waarvan het puntental ontbreekt.

Een derde manier om het ontbrekende puntental bij te schatten is door gebruik te maken van **regressie analyse**. Gegeven de personen waarvoor de huurpunten bekend zijn wordt bij deze techniek eerst door middel van lineaire regressie op basis van een aantal op te geven verklarende variabelen een formule opgesteld die het aantal huurpunten van deze personen zo goed mogelijk verklaart. Er moet dus worden gezocht naar variabelen in WoON2012 die sterk gecorreleerd zijn met het aantal huurpunten. Een voorwaarde van het gebruik van deze methode is wel dat de set van gebruikte verklarende variabelen voor alle personen waarvoor het aantal huurpunten bijgeschat dient te worden, bekend dienen te zijn.

Enkele voor- en nadelen

De drie manieren van het bijschatten van het aantal huurpunten hebben elk hun voor- en nadelen. Met de **CPB-aanpak** wordt het aantal punten benaderd op basis van de informatie die voorhanden is. Als alle benodigde informatie voorhanden is, dan levert deze methode een zeer nauwkeurige schatting van het aantal huurpunten. Feit is echter dat een groot deel van de benodigde informatie vaak ontbreekt of moeilijk te kwantificeren is op basis van de data (hoeveel punten worden bijvoorbeeld afgetrokken voor hinderlijke situaties?). Deze manier van berekenen is dan ook erg gevoelig voor 'fouten' in de antwoorden van de respondent, met name wanneer het aantal te behalen punten op een onderdeel relatief hoog is of niet gemaximeerd.

Bij toepassing van een **imputatietechniek** wordt voorbijgegaan aan de deterministische manier van bepalen zoals dat in de praktijk gebeurt. Verder blijkt op meerdere punten uit de data dat er sprake is van een scheve verhouding tussen records met punten voor de woning en records zonder punten. De groep van personen waarvoor het aantal punten moet worden bijgeschat is zeker niet aselectief, terwijl dit bij het gebruik van imputatie wel gewenst is. Zo zien we bijvoorbeeld dat de kale huur gemiddeld een stuk hoger is voor personen waarvoor het aantal huurpunten bijgeschat moet worden. Bij gebruik van de imputatietechniek zal daarom de werkelijke variantie van het puntental sterk worden onderschat. Bij imputatie worden bovendien alleen waarden bijgeschat die al voorkomen in het bestand. Dit is normaliter gewenst, omdat men er dan zeker van kan zijn dat de bijgeschatte waarden ook daadwerkelijk valide waarden zijn. In het geval van huurpunten kan dit echter als een nadeel beschouwd worden. Huurpunten

kunnen in principe immers elke waarde tussen de 40 en grofweg 300 aannemen en er is geen garantie dat al deze waarden ook voorkomen onder de groep van donoren, zeker niet binnen woningmarktgebied.

Ook aan het gebruik van **lineaire regressie** kleven enkele nadelen. Zo is het van belang dat alle personen waarvoor het aantal punten bijgeschat moet worden een geldige waarden hebben op de gebruikte verklarende variabelen. Hier staat tegenover dat vaak niet veel verklarende variabelen nodig zijn, om tot een zeer redelijke mate van verklaarde variantie te komen. Een voordeel is dat deze methode toestaat dat puntentallen geschat worden die niet in de dataset voorkomen. De spreiding van huurpunten blijft op deze manier beter gehandhaafd. Een nadeel van regressie is dat zo nu en dan extreme waarden geschat zullen worden. Als uit de regressievergelijking bijvoorbeeld blijkt dat de kale huur sterk positief gecorreleerd is met het aantal huurpunten, dan zal een ontvanger die een kale huur van €3000 heeft opgegeven op basis van de regressievergelijking een onrealistisch hoog aantal punten bijgeschat krijgen. Dit kan echter gemakkelijk worden voorkomen door het geschatte aantal punten op een bepaalde grens te maximeren.

Keuze

Gegeven bovenstaande overwegingen en de bevindingen na het toepassen van de betreffende methode is besloten om de ontbrekende huurpunten bij te schatten met behulp van **lineaire regressie**. Om te voorkomen dat extreme waarden toegekend worden nemen we alleen geschatte waarden over die zich tussen 40 en 300 punten bevinden. Voor ontvangers waarvoor het geschatte aantal punten buiten deze grenzen valt passen we de zogenaamde Predictive Mean Matching (**PMM**)-methode toe. Daarbij wordt voor een ontvanger het donorrecord bepaald waarvan het geschatte aantal huurpunten (op basis van de regressievergelijking) zo dicht mogelijk bij het geschatte aantal huurpunten van de ontvanger ligt. Vervolgens wordt dan het werkelijke puntental van deze donor gekopieerd naar de ontvanger.

Vooraf

Het WoON 2012 telt in totaal 69.339 respondenten waaronder 23.349 huurders (HUKO=2). Van deze huurders zijn er 1.519 die een huisvestings situatie anders dan “woning” hebben (HVS \neq 1). Deze groep van 1.519 zal worden uitgesloten van de analyse en er zullen voor deze groep ook geen huurpunten worden bijgeschat. Deze groep is uitgesloten vanwege de volgende oorzaken:

1. Slechts 77 van de 1.519 heeft een puntental opgegeven.
2. Voor 276 van de 1.519 is de woonoppervlakte niet bekend.
3. Voor 344 van de 1.519 is de WOZ-waarde niet bekend.
4. Voor geen van de 1.519 is de kale huur bekend.

De kale huur, de oppervlakte en de WOZ-waarde blijken juist de variabelen te zijn die het meest samenhangen met het aantal huurpunten. Voor de huurders met HVS ongelijk aan 1 kan het aantal huurpunten dus helaas niet worden bijgeschat. Er blijft dan een groep van 21.830 personen over waarvoor de huurpunten ofwel al bekend zijn (zelf opgegeven of gekoppeld vanuit NCCW) of zullen worden bijgeschat.

Ruim 12.000 ontvangers

In dit onderzoek wordt het gekoppelde aantal punten vanuit NCCW als meer betrouwbaar verondersteld dan het door de respondent zelf opgegeven aantal punten. Voor 8.904 van de 21.830 huurders kon geldige registerdata gekoppeld worden vanuit het NCCW-bestand (variabele TOTAALPU). Van de 21.830 personen hebben er 2.404 zelf een waarde opgegeven voor het aantal huurpunten (variabele AANTPUNT). Daarvan bevinden zich er echter 708 onder de 40 en 75 boven de 300 punten. Deze waarden worden als niet realistisch beoordeeld en daarom als ongeldig beschouwd.

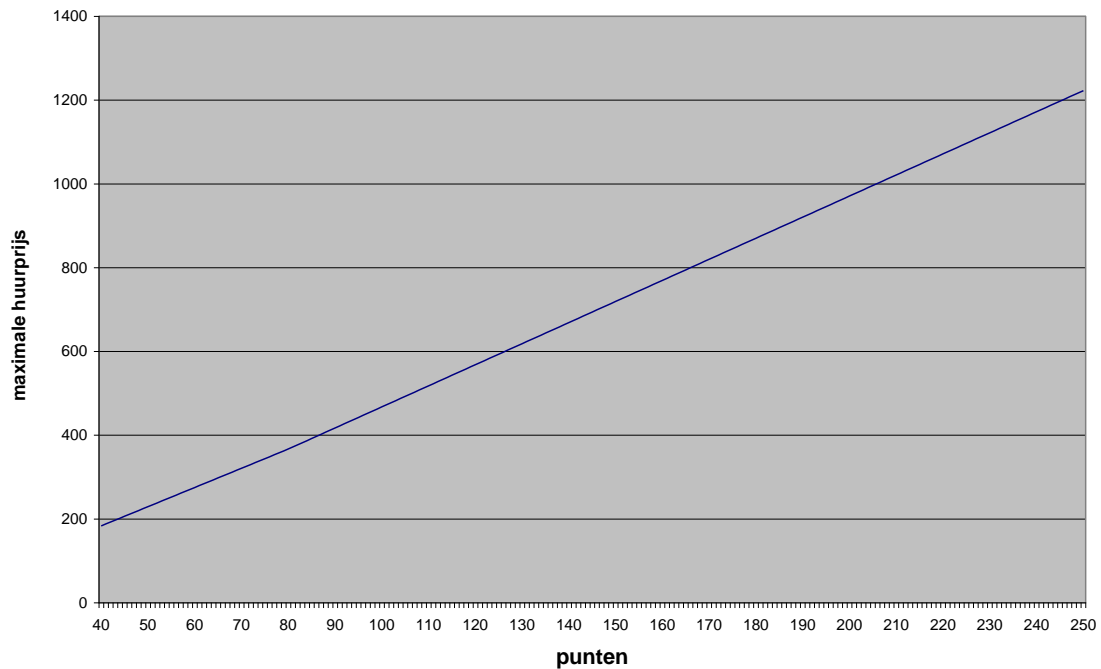
In 761 gevallen is zowel het opgegeven (tussen 40 en 300) als het gekoppelde aantal huurpunten bekend. Voor ruim 23,6% hiervan zijn beide waarden exact gelijk. 67,4% wijkt maximaal 10% af en voor 85,7% geldt dat het opgegeven aantal punten maximaal 25% afwijkt van het registraantal. We concluderen dat er een redelijke overeenkomst is tussen wat een respondent opgeeft en wat de 'officiële' bronnen melden. Gegeven de voldoende mate van overeenkomst tussen opgave en registratie is er voor gekozen om de opgave van de respondent als substituuat te gebruiken, indien geen registerinformatie bekend is. Voor personen die zelf een waarde hebben opgegeven tussen de 40 en 300 zal het aantal huurpunten dus niet hoeven worden bijgeschat. Voor 860 huurders waarvoor geen registerdata bekend is, kon op deze manier het zelf opgegeven aantal huurpunten gebruikt worden. Hiermee is voor 9.764 personen (met HUKO=2 & HVS=1) een aantal huurpunten bekend wat zich tussen 40 en 300 bevindt. Voor de overige 12.066 personen zullen de huurpunten met behulp van lineaire regressie en eventueel de PMM-methode worden bijgeschat.

Ruim 9000 bruikbare waarnemingen

Zoals hierboven beschreven zijn vóór de bijschatting voor 9.764 personen (HUKO=2 & HVS=1) het aantal huurpunten bekend. Deze groep is dan ook in principe kandidaat voor de input van de regressieanalyse. We zullen echter een deel van deze groep niet gebruiken voor de analyse. Dit heeft te maken met het optreden van onrealistische combinaties van de kale huur en het aantal huurpunten, zoals hieronder wordt toegelicht.

Op basis van het aantal huurpunten kan voor de 9.764 huurders een maximale huur bepaald worden. Dit is gebeurd op basis van de huurprijs tabel voor zelfstandige woningen per 1 juli 2012, zie Figuur B3.3.

Figuur B3.3 Maximale huurprijsgrenzen voor zelfstandige woningen per 1 juli 2012



Voor de regressieanalyse zullen alleen personen worden gebruikt waarvoor de verhouding tussen de opgegeven kale huur (KHUURII) en de hierboven bepaalde maximale huur redelijk is. We hanteren daarvoor de volgende criteria:

- Als de maximale huur meer dan 2,5 maal zo hoog is als de opgegeven kale huur, wordt de persoon niet gebruikt in de regressieanalyse.
- Als de opgegeven kale huur meer dan 10% groter is dan de maximale huur, wordt de persoon niet gebruikt in de regressieanalyse.

Met dit criterium vallen 449 van de 9.764 personen af, waarmee er 9.315 overblijven als input voor de regressieanalyse. Let wel, voor de 449 personen is het aantal huurpunten wel bekend en zal dus niet worden bijgeschat!

Als we de 9.315 personen die gebruikt zullen worden voor de regressieanalyse wegen naar huishoudens (HWEEGWON) dan vinden we een gemiddeld puntental van 135,7. Regionaal treden flinke verschillen op, zoals weergegeven in Tabel B3.5. De provincies Flevoland en Utrecht kennen onder de analysegroep het hoogste gemiddelde puntental terwijl voor de provincie Groningen het gemiddelde aantal het laagst is.

Tabel B3.5 Gemiddeld puntental analysegroep, gedifferentieerd naar gebieden

gemiddeld aantal huurpunten voor imputatie, analysegroep			
Groningen	127	Amsterdam	129
Friesland	135	's Gravenhage	120
Drenthe	142	Rotterdam	124
Overijssel	138	Utrecht	132
Flevoland	148	overige gemeenten	139
Gelderland	145		
Utrecht	148		
Noord-Holland	135		
Zuid-Holland	129		
Zeeland	137		
Noord-Brabant	144		
Limburg	138		
totaal	136	totaal	136

4. Regressie analyse

In deze paragraaf zullen de relevante aspecten van de regressieanalyse stuk voor stuk worden beschreven.

Onafhankelijke variabelen

De afwegingen bij de keuze van de te gebruiken onafhankelijke variabelen in de regressieanalyse zijn gemaakt op basis van beschikbaarheid van de variabelen, de mogelijkheden tot het eventueel bijschatten van verklaarders en zeker ook op intuïtieve juistheid van het verband met de huurpunten. Na enig experimenteren is gekozen om de kale huurprijs (variabele KHUURII), de WOZ-waarde (variabele WOZWAARD) en de woonoppervlakte van de woning (OPPTBIN) als verklaarders voor het aantal huurpunten te gebruiken. We behandelen hieronder voor elk van deze drie onafhankelijke variabelen de eventuele bewerkingen die nodig waren. Bovendien laten we voor elk van de drie variabelen de verdeling zien binnen de analysegroep en binnen de groep ontvangers. Dit levert vooraf vaak interessante inzichten op. Mocht de groep ontvangers op de gebruikte variabelen bijvoorbeeld erg afwijken van de analysegroep, dan zal dit waarschijnlijk tot gevolg hebben dat het geschatte aantal punten van de ontvangers een andere verdeling zal kennen dan het aantal punten binnen de analysegroep.

Kale huur

Zoals eerder beschreven zijn personen met een waarde 0 op KHUURII uitgesloten voor de analyse op basis van het criterium dat de maximale huur voor deze personen meer dan 2,5 maal zo hoog is als de kale huur. De verdeling van de kale huur over 12 categorieën is voor zowel de analysegroep als de groep ontvangers opgenomen in Tabel B3.6. We merken op dat voor 316 ontvangers de kale huur een waarde van maximaal 100 heeft (hiervan hebben er 245 de waarde 0). Onder de ontvangers bevinden zich bovendien 294 personen met een kale huur van groter dan 1.000, waarvan 85 met een kale huur groter dan 1.500. Onder de analysegroep komt geen enkele persoon voor met een huur lager dan 100 en komen slechts 2 personen voor met een kale huur van boven de 1.000. Gezien de verwachte positieve correlatie tussen het aantal huurpunten en de kale huur zal de verdeling van het aantal huurpunten onder de ontvangers daarom wellicht afwijken (meer extreme waarden) dan die onder de analysegroep.

Tabel B3.6 Kale huur voor ontvangers en analysegroep

kale huur	aantal		percentage		cum. percentage	
	ontvangers	analyse	ontvangers	analyse	ontvangers	analyse
[0,50]	259	0	2,15	0,00	2,15	0,00
(50,100]	47	0	0,39	0,00	2,54	0,00
(100,200]	235	36	1,95	0,39	4,48	0,39
(200,300]	721	686	5,98	7,36	10,46	7,75
(300,400]	2148	2268	17,80	24,35	28,26	32,10
(400,500]	3682	3511	30,52	37,69	58,78	69,79
(500,600]	2479	2026	20,55	21,75	79,32	91,54
(600,700]	1067	614	8,84	6,59	88,17	98,13
(700,800]	575	122	4,77	1,31	92,93	99,44
(800,900]	339	39	2,81	0,42	95,74	99,86
(900,1000]	220	11	1,82	0,12	97,56	99,98
(1000,-]	294	2	2,44	0,02	100,00	100,00

WOZ-waarde

Omdat binnen WoON12 voor een grote groep de WOZ-waarde niet bekend is (waaronder voor een aantal personen uit de analysegroep en uit de groep van ontvangers), hebben we de WOZ-waarde apart bijgeschat. Dat is gebeurd met behulp van de door ABF veel toegepaste imputatietechniek. Als een WOZ-waarde van lager dan €10.000 is opgegeven, is deze ook bijgeschat en is de originele waarde overschreven. Opgegeven waarden van meer dan €2.000.000 zijn niet als donor gebruikt, maar ook niet overschreven. In totaal is op deze manier voor 2.465 personen in WoON2012 de WOZ-waarde bijgeschat. Deze bijschatting heeft plaatsgevonden in 2 stappen. Eerst is de WOZ-waarde van de woning in 7 klassen geschat (WOZ7). De 5 leidende variabelen van de 15 variabelen in de daarbij gebruikte relatietabel zijn de provincie (PROV), het woonmilieu op buurniveau (HWMBRT), het kenmerk een/meergezinswoning (VORM), de waarde van de woning (gebaseerd op VERKP7 voor koopwoningen en BHUURII voor huurwoningen) en het type woning (HUISTYP). Het percentage correct in deze eerste bijschatting komt uit op 75,3%. Vervolgens is in een tweede imputatieronde de daadwerkelijke WOZ-waarde geschat. Daarbij is dezelfde relatietabel gebruikt, behalve dat nu de eerder bijgeschatte klasse (WOZ7) bovenin de tabel staat. In deze tweede ronde werd een percentage correct van 88,8% behaald.

Ook wat betreft de WOZ-waarde zijn er (na imputatie) verschillen tussen de analysegroep en ontvangers, zoals is weergegeven in Tabel B3.7. Zo heeft ruim 95% van de analysegroep een WOZ-waarde van maximaal 250.000 terwijl dit bij de ontvangers nog geen 89% is. Onder de ontvangers bevindt zich bovendien een groep van 40 huurders waarvan de WOZ-waarde van de woning boven de 750.000 ligt. Voor 14 daarvan bedraagt de WOZ-waarde zelfs meer dan een miljoen. Onder de analysegroep komen dergelijke waarden nauwelijks voor. Gezien de verwachte positieve correlatie tussen WOZ-waarde en huurpunten zal het geschatte puntental voor de kleine groep ontvangers met een hoge WOZ-waarde onrealistisch hoog kunnen uitkomen.

Tabel B3.7 WOZ-waarde voor ontvangers en analysegroep

WOZ	aantal		percentage		cum. percentage	
	ontvangers	analyse	ontvangers	analyse	ontvangers	analyse
[0, 50.000]	38	21	0,31	0,23	0,31	0,23
(50.000, 100.000]	949	921	7,87	9,89	8,18	10,11
(100.000, 150.000]	3701	3554	30,67	38,15	38,85	48,27
(150.000, 200.000]	4060	2956	33,65	31,73	72,50	80,00
(200.000, 250.000]	1986	1405	16,46	15,08	88,96	95,08
(250.000, 300.000]	629	357	5,21	3,83	94,17	98,92
(300.000, 400.000]	397	91	3,29	0,98	97,46	99,89
(400.000, 500.000]	156	6	1,29	0,06	98,76	99,96
(500.000, 750.000]	110	3	0,91	0,03	99,67	99,99
(750.000, 3.000.000]	40	1	0,33	0,01	100,00	100,00

Woonoppervlakte

De totale woonoppervlakte is zowel voor de analysegroep als voor de groep ontvangers gemaximeerd op 300 m². Uit Tabel B3.8 blijkt dat de orde van waarden voor oppervlakte die voorkomen onder ontvangers ook voorkomen onder de analysegroep. Wel zien we dat onder de ontvangers de binnenoppervlakte wat groter is dan onder de analysegroep. Zo heeft bijvoorbeeld 19,66% van de ontvangers een binnenoppervlakte van meer dan 100m² tegenover 12,24% van de analysegroep. Gezien de verwachte positieve correlatie tussen de oppervlakte en het aantal huurpunten zullen de ontvangers over het geheel dus wat meer huurpunten krijgen dan de analysegroep.

Tabel B3.8 Oppervlakte binnen voor ontvangers en analysegroep

oppervlakte	aantal		percentage		cum. percentage	
	ontvangers	analyse	ontvangers	analyse	ontvangers	analyse
[0,20]	56	32	0,46	0,34	0,46	0,34
(20,30]	266	215	2,20	2,31	2,67	2,65
(30,40]	624	510	5,17	5,48	7,84	8,13
(40,50]	1156	1018	9,58	10,93	17,42	19,06
(50,60]	1405	1352	11,64	14,51	29,07	33,57
(60,70]	1453	1365	12,04	14,65	41,11	48,22
(70,80]	1688	1452	13,99	15,59	55,10	63,81
(80,90]	1209	992	10,02	10,65	65,12	74,46
(90,100]	1316	904	10,91	9,70	76,02	84,17
(100,110]	521	335	4,32	3,60	80,34	87,76
(110,120]	721	414	5,98	4,44	86,32	92,21
(120,140]	637	336	5,28	3,61	91,60	95,81
(140,180]	679	291	5,63	3,12	97,22	98,94
(180,250]	231	69	1,91	0,74	99,14	99,68
(250,300]	104	30	0,86	0,32	100,00	100,00

Woningmarktgebieden

Aangezien de relatie tussen de drie verklarende variabelen en het aantal huurpunten naar verwachting regionaal flink zal differentiëren, worden per afzonderlijk woningmarktgebied de regressieparameters bepaald op basis van de analysegroep in het betreffende gebied. Voor elk van de 31 woningmarktregio's zal dus een afzonderlijke regressievergelijking met constante worden opgesteld waarmee het aantal huurpunten voor ontvangers binnen deze regio geschat kan worden met behulp van de waarden op kale

huur, oppervlakte en WOZ-waarde. De verhouding tussen de grootte van de analysegroep en de groep ontvangers per woningmarktgebied is opgenomen in Tabel B3.9. Hierin staan zowel ongewogen aantallen als gewogen aantallen. Als weegfactor is wederom het huishoudengewicht gebruikt. Deze weegfactor zal ook toegepast worden bij het opstellen van de regressievergelijkingen.

De minimale grootte van de analysegroep binnen een gebied bedraagt daarbij 23 (gebied Zuidoost- en Zuidwest-Drenthe).

Tabel B3.9 Verhouding tussen analysegroep en ontvangers per woningmarktgebied

WMG31	OMSCHRIJVING	ANALYSE		ONTVANGERS	
		ongewogen	gewogen	ongewogen	gewogen
1	Ommelanden	65	14398	150	39165
2	Stad Groningen	147	25922	96	18134
3	Noord-Friesland/Waddeneilanden	60	11460	117	30975
4	Oost-Friesland/Zuidwest-Friesland	73	17489	133	32691
5	Noord- en Midden-Drenthe	56	13687	39	10332
6	Zuidoost- en Zuidwest-Drenthe	23	4249	172	37442
7	Twente	151	26270	295	66898
8	IJsselvallei	658	36569	339	36184
9	Veluwe	515	22542	1153	68491
10	Arnhem e.o.	253	14413	1010	57938
11	Achterhoek	56	11763	170	39926
12	Nijmegen e.o. en Betuwe	223	18602	766	65230
13	Oost-Utrecht	96	21182	169	41623
14	West-Utrecht	247	37012	626	97673
15	Amsterdam	580	133879	498	121491
16	Zaanstreek/Waterland/Amstelland/Meerland	581	87766	291	61094
17	Zuid- en Midden-Kennemerland	69	19010	201	51809
18	KopNoord-Holland/West-Friesland	150	31618	214	53482
19	Bollenstreek/Leiden	146	26861	399	36937
20	s-Gravenhage/Westland	2003	151095	826	74831
21	Rijnstreek/Midden-Holland/Alblasserwaard	184	34555	264	37618
22	Drechtsteden/HoekseWaard/Goeree Overfl.	135	26613	149	34227
23	Rijnmond	1477	121554	1170	164726
24	Zeeland	172	13694	617	37606
25	Roosendaal/Breda	289	39652	705	62350
26	Midden-Brabant/Waalwijk	88	22089	172	42633
27	s-Hertogenbosch + NoordoostBrabant	32	8216	335	82618
28	Helmond/Eindhoven	214	42717	358	72101
29	Noord-Limburg	165	39085	113	25164
30	Zuid-Limburg	319	57331	318	55285
31	Flevoland	88	16829	201	38115
totaal		9315	1148122	12066	1694789

Lineaire regressiemodellen

De regressievergelijkingen per woningmarktgebied zijn opgenomen in Tabel B3.10. Voor elk van de 31 modellen geldt dat de regressiecoëfficiënten van alle variabelen significant afwijken van 0 bij een betrouwbaarheidsniveau van 95%. Bij een betrouwbaarheidsniveau van 99% wijkt alleen de regressiecoëfficiënt voor de woonoppervlakte in het model voor gebied 4 niet meer significant af van 0. Over alle 31 modellen wordt gemiddeld 59,5% van de variantie in het aantal punten verklaard door de kale huur, de WOZ-waarde en de woonoppervlakte.

Tabel B3.10 Regressievergelijkingen voor aantal punten per woningmarktgebied

	WONINGMARKTREGIO	R-kwadraat	constante	kale huur(€)	woz(€1000)	oppervlakte(m ²)
1	Ommelanden	0,360	53,707	0,063	0,460	0,094
2	Stad Groningen	0,558	12,819	0,141	0,330	0,113
3	Noord-Friesland/Waddeneilanden	0,691	18,958	0,062	0,546	0,277
4	Oost-Friesland/Zuidwest-Friesland	0,516	41,931	0,122	0,381	0,001
5	Noord- en Midden-Drenthe	0,390	57,316	0,094	0,356	0,044
6	Zuidoost- en Zuidwest-Drenthe	0,484	51,915	0,053	0,441	0,037
7	Twente	0,642	30,333	0,102	0,349	0,233
8	IJsselvallei	0,648	21,040	0,106	0,358	0,195
9	Veluwe	0,547	42,568	0,078	0,403	0,043
10	Arnhem e.o.	0,527	30,017	0,110	0,225	0,268
11	Achterhoek	0,649	10,684	0,217	0,161	0,131
12	Nijmegen e.o. en Betuwe	0,624	37,048	0,105	0,226	0,245
13	Oost-Utrecht	0,704	37,245	0,105	0,271	0,129
14	West-Utrecht	0,677	20,469	0,137	0,301	0,097
15	Amsterdam	0,581	42,974	0,105	0,110	0,287
16	Zaanstreek/Waterland/Amstelland/Meerland	0,556	43,440	0,103	0,181	0,181
17	Zuid- en Midden-Kennemerland	0,564	48,173	0,088	0,152	0,226
18	KopNoord-Holland/West-Friesland	0,685	20,593	0,137	0,278	0,231
19	Bollenstreek/Leiden	0,654	38,409	0,091	0,323	0,035
20	s-Gravenhage/Westland	0,783	28,005	0,110	0,248	0,148
21	Rijnstreek/Midden-Holland/Alblasserwaard	0,689	24,257	0,102	0,320	0,129
22	Drechtsteden/HoekseWaard/Goeree Overfl.	0,631	26,396	0,143	0,229	0,103
23	Rijnmond	0,702	34,621	0,099	0,267	0,145
24	Zeeland	0,610	39,052	0,145	0,204	0,111
25	Roosendaal/Breda	0,682	24,012	0,083	0,444	0,066
26	Midden-Brabant/Waalwijk	0,595	35,325	0,109	0,260	0,172
27	s-Hertogenbosch + NoordoostBrabant	0,525	40,631	0,133	0,252	0,112
28	Helmond/Eindhoven	0,515	42,787	0,110	0,207	0,163
29	Noord-Limburg	0,634	23,991	0,129	0,339	0,125
30	Zuid-Limburg	0,537	39,266	0,105	0,329	0,144
31	Flevoland	0,484	49,539	0,132	0,171	0,084

Met behulp van bovenstaande vergelijkingen wordt voor de 12.066 ontvangers een aantal huurpunten bepaald. Als deze waarde binnen het interval [40,300] valt, wordt de waarde rechtstreeks toegekend aan de ontvanger. Als het puntental op basis van de regressie buiten het interval [40,300] valt, dan krijgt de betreffende ontvanger het werkelijk aantal huurpunten toegewezen van een persoon uit de analysegroep van hetzelfde woningmarktgebied waarvoor het geschatte puntental het dichtst bij het geschatte puntental van de ontvanger ligt (PMM).

5. Evaluatie resultaten

Na de imputatie hebben alle 21.830 personen (met HUKO=2&HVS=1) een puntental. Het betreft dus:

- 9.315 personen met punten vooraf die gebruikt zijn voor de regressieanalyse.
- 449 personen met punten vooraf die (vanwege abnormale verhouding tussen kale huur en maximale huur) niet gebruikt zijn in de regressieanalyse.
- 12.066 personen zonder punten vooraf waarvoor het puntental met behulp lineaire regressie en eventueel de PMM-methode is bijgeschat.

Punten voor en na bijschatting

Voor toetsing van de uitkomsten worden de huurpunten vóór bijschatting vergeleken met de punten na bijschatting. Uit Tabel B3.11 valt op te maken dat door de imputatie het gemiddeld aantal huurpunten binnen elke provincie en voor elk van de vier grote steden flink stijgt. In WoON2012 kwam het landelijk gewogen gemiddelde aantal huurpunten vóór de bijschatting (dus over de 9.764 records) uit op 134,8 en na bijschatting (over 21.830 records) op 142,3. Het gemiddelde binnen de 4 grote steden lag vóór de

bijschatting op 124,6 terwijl het na de bijschatting uitkomt op 132,0. Deze uitkomst viel te verwachten, gezien de verschillen in de verdelingen van de drie onafhankelijke variabelen tussen de analysegroep en de groep ontvangers. Vooral de toename van 12 punten in de provincie Groningen en de toename van 16 punten voor de stad Utrecht springen in het oog.

Tabel B3.11 Aantal huurpunten voor en na imputatie, gewogen naar huishoudens

gemiddeld aantal huurpunten voor imputatie				gemiddeld aantal huurpunten na imputatie			
Groningen	124	Amsterdam	127	Groningen	135	Amsterdam	133
Friesland	135	's Gravenhage	119	Friesland	141	's Gravenhage	128
Drenthe	142	Rotterdam	124	Drenthe	144	Rotterdam	127
Overijssel	138	Utrecht	131	Overijssel	145	Utrecht	147
Flevoland	148	overige gemeenten	138	Flevoland	152	overige gemeenten	145
Gelderland	144			Gelderland	150		
Utrecht	147			Utrecht	153		
Noord-Holland	134			Noord-Holland	139		
Zuid-Holland	128			Zuid-Holland	134		
Zeeland	137			Zeeland	141		
Noord-Brabant	143			Noord-Brabant	150		
Limburg	138			Limburg	144		
totaal	135	totaal	135	totaal	142	totaal	142

We evalueren vervolgens hoe de huurpunten in het WoON2012 op verschillende gebiedsniveaus voor en na bijschatting zijn verdeeld zijn over de volgende vier categorieën:

- tot en met 79 punten: maximale huurprijs tot ± €362 (kwaliteitskortingsgrens)
- tussen 80 en 117 punten: maximale huurprijs tot ± €554 (aftoppingsgrens)
- tussen 118 en 137 punten: maximale huurprijs tot ± €653 (huurtoeslaggrens)
- vanaf 138 punten: maximale huurprijs vanaf ± €653

Ook uit Tabel B3.12 valt op te maken dat het gemiddeld aantal huurpunten door de bijschatting is toegenomen. Zo geldt voor alle provincies uitgezonderd Drenthe dat het aandeel woningen van meer dan 137 punten toeneemt. Landelijk gezien stijgt het percentage huurwoningen met meer dan 117 punten van 70,7% naar 77,8%.

Tabel B3.12 Procentuele puntenverdeling per provincie voor en na bijschatting WoON2012

PROV	PUNTEN VOOR BIJSCHATTING (9.764 PERS)					PUNTEN NA BIJSCHATTING (21.830 PERS)				
	t/m 79	80-117	118-137	vanaf 138	TOTAAL	t/m 79	80-117	118-137	vanaf 138	TOTAAL
1 Groningen	6,7	35,5	25,9	31,9	100,0	3,3	24,9	26,3	45,5	100,0
2 Friesland	2,5	23,7	26,7	47,2	100,0	2,1	21,3	24,2	52,5	100,0
3 Drenthe	0,0	16,9	26,3	56,8	100,0	0,0	13,7	33,8	52,5	100,0
4 Overijssel	2,9	24,2	22,9	50,0	100,0	1,4	17,1	24,2	57,3	100,0
5 Flevoland	0,0	11,3	22,5	66,1	100,0	0,0	5,9	21,4	72,6	100,0
6 Gelderland	1,6	21,7	18,3	58,5	100,0	0,9	12,9	22,0	64,2	100,0
7 Utrecht	1,9	13,6	26,7	57,8	100,0	0,6	15,0	20,9	63,5	100,0
8 Noord-Holland	2,9	24,9	29,4	42,8	100,0	1,9	23,4	27,0	47,8	100,0
9 Zuid-Holland	2,9	35,2	28,7	33,2	100,0	2,3	30,5	26,9	40,3	100,0
10 Zeeland	4,5	19,6	26,3	49,6	100,0	1,3	18,4	29,2	51,1	100,0
11 Noord-Brabant	1,1	17,7	24,4	56,8	100,0	0,8	11,8	21,6	65,8	100,0
12 Limburg	2,1	25,2	23,0	49,8	100,0	1,4	18,8	24,4	55,4	100,0
TOTAAL	2,6	26,7	26,6	44,1	100,0	1,6	20,7	25,0	52,8	100,0

In Tabel B3.13 is te zien dat het aandeel woningen met maximaal 117 punten in alleen in Rotterdam toeneemt. In de stad Utrecht stijgt het aandeel woningen met meer dan 137 punten van 38,2% naar 52,9%.

Tabel B3.13 Procentuele puntenverdeling voor 4 grote steden voor en na bijschatting WoON2012

G4_5	PUNTEN VOOR BIJSCHATTING (9.764 PERS)					PUNTEN NA BIJSCHATTING (21.830 PERS)				
	t/m 79	80-117	118-137	vanaf 138	TOTAAL	t/m 79	80-117	118-137	vanaf 138	TOTAAL
1 Amsterdam	4,4	31,9	30,3	33,4	100,0	3,2	30,5	27,9	38,4	100,0
2 's Gravenhage	6,6	47,6	22,9	22,9	100,0	4,7	39,3	22,0	34,0	100,0
3 Rotterdam	1,6	38,2	35,4	24,9	100,0	1,9	38,4	30,4	29,3	100,0
4 Utrecht	5,2	25,9	30,8	38,2	100,0	1,2	23,5	22,5	52,9	100,0
5 Overig	2,0	22,9	25,5	49,6	100,0	1,2	16,9	24,4	57,5	100,0
TOTAAL	2,6	26,7	26,6	44,1	100,0	1,6	20,7	25,0	52,8	100,0

Vergelijking met CBS-huurenquête

Voor toetsing van de uitkomsten worden de huurpunten na bijschatting in het WoON2012 vervolgens vergeleken met de huurpunten uit de CBS-huurenquête van 2011.

Het landelijk gewogen gemiddelde voor het aantal huurpunten in de huurenquête (alleen records met PCBS>0) bedraagt 138,9 punten. In het WoON2012 komt het gewogen gemiddelde puntental na bijschatting (over 21.830 records) uit op 142,3. Het gemiddeld aantal punten binnen de 4 grote steden bedraagt in de CBS huurenquête 122,2 terwijl in WoON2012 het gemiddelde na bijschatting uitkomt op 132. Kijken we naar de verdeling van huurpunten op landelijk niveau, dan zien we dat deze in de CBS-huurenquête redelijk goed spoort met het bijgeschatte WoON2012, zie rij TOTAAL in Tabel B3.14. Op het niveau van provincies valt bijvoorbeeld op dat de provincies Utrecht en Noord-Holland in het WoON2012 een hoger aandeel woningen kennen in het hoogste puntensegment dan in de CBS-huurenquête.

Tabel B3.14 BProcentuele puntenverdeling per provincie in CBS-huurenquête en in WoON2012

PROV	CBS-huurenquête					WoON2012				
	t/m 79	80-117	118-137	vanaf 138	TOTAAL	t/m 79	80-117	118-137	vanaf 138	TOTAAL
1 Groningen	1,7	25,8	30,0	42,5	100,0	3,3	24,9	26,3	45,5	100,0
2 Friesland	2,7	23,3	24,7	49,4	100,0	2,1	21,3	24,2	52,5	100,0
3 Drenthe	1,2	11,1	23,1	64,5	100,0	0,0	13,7	33,8	52,5	100,0
4 Overijssel	0,9	16,2	23,3	59,7	100,0	1,4	17,1	24,2	57,3	100,0
5 Flevoland	0,0	6,6	20,7	72,7	100,0	0,0	5,9	21,4	72,6	100,0
6 Gelderland	0,2	12,1	25,2	62,5	100,0	0,9	12,9	22,0	64,2	100,0
7 Utrecht	1,3	20,8	29,3	48,6	100,0	0,6	15,0	20,9	63,5	100,0
8 Noord-Holland	5,2	32,2	27,2	35,5	100,0	1,9	23,4	27,0	47,8	100,0
9 Zuid-Holland	3,0	28,9	28,4	39,8	100,0	2,3	30,5	26,9	40,3	100,0
10 Zeeland	0,3	19,1	24,6	55,9	100,0	1,3	18,4	29,2	51,1	100,0
11 Noord-Brabant	0,7	12,3	18,9	68,1	100,0	0,8	11,8	21,6	65,8	100,0
12 Limburg	0,6	16,2	28,9	54,3	100,0	1,4	18,8	24,4	55,4	100,0
TOTAAL	2,2	22,1	26,1	49,6	100,0	1,6	20,7	25,0	52,8	100,0

Tenslotte is in Tabel B3.15 de vergelijking gemaakt van de puntenverdeling binnen de 4 grote steden tussen de CBS-huurenquête en WoON2012. Voor Amsterdam zien we bijvoorbeeld dat het aandeel huurwoningen met maximaal 79 punten na bijschatting flink lager uitvalt dan in de huurenquête. Ondanks enkele opmerkelijke verschillen op celniveau kan worden geconcludeerd dat over het algemeen de verdelingen van de huurpunten in het geïmputeerde WoON2012 bestand redelijk overeenkomen met die in de huurenquête van het CBS.

Tabel B3.15 Procentuele puntenverdeling voor 4 grote steden in CBS-huurenquête en in WoON2012

G4_5	CBS-huurenquête					WoON2012				
	t/m 79	80-117	118-137	vanaf 138	TOTAAL	t/m 79	80-117	118-137	vanaf 138	TOTAAL
1 Amsterdam	9,7	45,8	29,2	15,2	100,0	3,2	30,5	27,9	38,4	100,0
2 's Gravenhage	3,8	46,1	22,8	27,4	100,0	4,7	39,3	22,0	34,0	100,0
3 Rotterdam	5,0	28,4	36,3	30,4	100,0	1,9	38,4	30,4	29,3	100,0
4 Utrecht	1,8	27,2	34,0	37,0	100,0	1,2	23,5	22,5	52,9	100,0
5 Overig	1,1	17,6	25,1	56,2	100,0	1,2	16,9	24,4	57,5	100,0
TOTAAL	2,2	22,1	26,1	49,6	100,0	1,6	20,7	25,0	52,8	100,0