

MIMOSIS: een microsimulatie model voor sociaal beleid in België.

**Beschrijving van het model met steekproefgegevens
1 januari 2009, inkomens 2008**



14 oktober 2011

Technische nota FOD Sociale Zekerheid

Versie: MIMOSIS_Steekproef1januari2009_NI_V1.doc

Correspondentieauteur: guy.vancamp@minsoc.fed.be

Inhoudsopgave

1	INLEIDING	1
2	BESCHRIJVING MIMOSIS STEEKPROEF	2
2.1	Constructie van een huishoudnummer en raming van de gezinsomvang	3
2.2	Constructie van een aantal karakteristieken op basis van gegevens uit het Rijksregister	4
2.3	Constructie en evaluatie van de steekproefgewichten	6
2.4	Beschrijving van een aantal variabelen in andere bronnen dan het Rijksregister	10
2.4.1	Bestanden met gegevens inzake kinderbijslag (RKW_RG, RKW_RH, RKW_BT, RSVZ_KB_RG, RSVZ_KB_RH, RSVZ_KB_BT)	14
2.4.2	Bestanden met uitkeringen ziekte, invaliditeit, arbeidsongeval en beroepsziekte (NIC, RIZIV, FAO en FBZ)	15
2.4.3	Bestand met uitkeringen ten laste van de RVA (RVA)	15
2.4.4	Bestand met uitbetaalde pensioenen (Penscad)	16
2.4.5	Bestand met gegevens inzake personen met een handicap (FOD SZ)	16
2.4.6	Bestand met sociale bijstandsgegevens (POD MI)	17
2.4.7	Bestand met gegevens zelfstandigen (RSVZ)	18
2.4.8	Bestand met loon en arbeidstijdgegevens voor RSZ-plichtigen voor het 4 ^{de} kwartaal van 2008 (RSZ_Quart)	19
2.4.9	Bestand met loon en arbeidstijdgegevens voor RSZPPO-plichtigen voor het 4 ^{de} kwartaal van 2008 (RSZPPO_Quart)	21
2.4.10	Bestanden met gegevens inzake arbeidsmarkthistoriek van loon en weddetrekkenden (LABM97, RSZPPO_20022008, RSZ_20022008 en CIMIRe)	22
2.4.11	Constructie van het laatst verdiende dagloon	24
3	WERKING VAN MIMOSIS	30
3.1	FAMREL	32
3.2	UNEM	34
3.3	SICK	36
3.4	PENSWELF	38
3.5	FAMAL	40
3.6	CONTRIB	43
3.7	PIT	46
3.8	SOCBEN	50

4	BESCHRIJVING MODULE EVAL	53
5	BESLUIT	57
6	REFERENTIES	59
	APPENDIX 1: BESCHRIJVING BEPALING STEEKPROEFGEWICHTEN	62
	APPENDIX 2: GEWOGEN RESULTATEN STEEKPROEF NAAR REGIO, LEEFTIJD EN GESLACHT	64
	APPENDIX 3: GEWOGEN RESULTATEN STEEKPROEF NAAR BURGERLIJKE STAAT, LEEFTIJD EN GESLACHT	68
	APPENDIX 4: HERWAARDERINGSFACTOREN GESCHATTE DAGLONEN	72

1 INLEIDING

Wanneer beleidsmakers bepaalde hervormingen in sociaal economische regelgeving overwegen, ontstaat vaak de behoefte om a priori een inschatting te maken van de mogelijke kostprijs en de verdelingsimpact (wie wint, wie verliest) van deze maatregelen. Deze tekst beschrijft het microsimulatiemodel MIMOSIS (MICrosimulation MOdel for Belgian Social Insurance Systems)¹, dat toelaat om dergelijke impactschattingen voor een aantal beleidsdomeinen te maken.

De beleidsdomeinen waar MIMOSIS zich in eerste instantie op richt zijn: sociale zekerheidsbijdragen, pensioenen, werkloosheidsuitkeringen, inkomensgaranties, ziekte- en invaliditeitsuitkeringen en familiale uitkeringen. Voor elke hervorming wordt via een geïntegreerde belastingmodule tevens het effect berekend op de verschuldigde personenbelasting.

Het model MIMOSIS maakt gebruik van een steekproef van gecodeerde sociale gegevens van persoonlijke aard, afkomstig uit het Datawarehouse Arbeidsmarkt en Sociale Bescherming (DWH AM&SB) van de Kruispuntbank van de Sociale Zekerheid. Het is door het gebruik van deze microgegevens dat een analyse van zowel de impact op het budget als de verdeling van inkomens mogelijk wordt.

De oorspronkelijke versie van MIMOSIS werd ontwikkeld met behulp van administratieve gegevens van 2001 (steekproef 1 januari 2002).² Deze tekst beschrijft de update van het model naar gegevens van 2008 (steekproef 1 januari 2009).³

In sectie 2 van deze nota beschrijven we de steekproef waarop het model gebouwd is. In deze sectie komt onder andere aan bod hoe gewichten geconstrueerd werden om de gegevens van de steekproef op te hogen tot het niveau van de totale populatie. In sectie 3 beschrijven we de werking van het model zelf. In deze sectie wordt voornamelijk uitgelegd hoe de inputgegevens gemanipuleerd worden om bepaalde impactschattingen te maken. In sectie 4 beschrijven we tenslotte een aantal indicatoren die door het model geproduceerd worden.

De beschrijvingen in deze tekst zijn steeds gebaseerd op de steekproef van alle beschikbare observaties. Het model kan daarnaast ook ingezet worden om de gegevens van één fictief individu of gezin te verwerken. Op deze laatste mogelijkheid gaan we in deze tekst niet verder in.

¹ MIMOSIS is één van de vele modellen die bestaan voor België. Voor een overzicht van andere modellen zie Decoster (2010) en Decoster e.a. (2008).

² Deze versie werd ontwikkeld in het kader van een aantal door Federaal Wetenschapsbeleid gefinancierde AGORA-projecten. Voor een beschrijving van de eerste versie van het model zie Decoster e.a. (2008).

³ Deze update kwam tot stand dankzij steun van Federaal Wetenschapsbeleid en de Kruispuntbank van de Sociale Zekerheid. De hulp van Chris Brijs, Hans Knapen, Anke Mutsaerts en Patrick Lusyne bij de verwerking van de gegevens wordt zeer gewaardeerd. Zij zijn uiteraard niet verantwoordelijk voor de gemaakte keuzes.

2 BESCHRIJVING MIMOSIS STEEKPROEF

Het steekproefkader⁴ wordt gevormd door alle individuen die, volgens het bestand met Rijksregistergegevens waarover de KSZ beschikt, in leven zijn op 1 januari 2009 en die verblijven in een Belgische gemeente.⁵ In deze groep zitten dus voornamelijk personen die fysiek in België verblijven. Belgen die in het buitenland verblijven, kunnen deel uitmaken van de groep maar enkel als ze een verblijfsadres in België behouden hebben. Het gaat hier dus om mensen die tijdelijk het land verlaten hebben omdat ze bijvoorbeeld op vakantie zijn of omdat ze een tijdelijke missie in het buitenland vervullen. Belgen die permanent in het buitenland verblijven worden dus niet opgenomen in ons steekproefkader.

Ieder individu in het geschetste steekproefkader kan geïdentificeerd worden op basis van een uniek identificatienummer. Uit de set van alle mogelijke nummers wordt een steekproef getrokken. Deze trekking wordt uitgevoerd in drie grote stappen.

- Stap 1: Er werd nagegaan hoeveel van de individuen, die random getrokken werden voor de vorige MIMOSIS-steekproef, nog in leven waren op 1 januari 2009.⁶ In totaal ging het om 90.641 individuen.
- Stap 2: Er werd een aanvulling van 9.359 individuen getrokken uit het geschetste steekproefkader waaruit de 90.641 nog in leven zijnde individuen van de vorige steekproef, verwijderd werden. De bijtrekking van deze 9.359 individuen gebeurt via systematische steekproeftrekking en rekening houdend met de ongelijke uitval over leeftijdsklassen.⁷
- Stap 3: Van de in totaal 100.000 individuen die na stap twee bekomen worden, worden tenslotte de identificatienummers opgevraagd van alle gezinsleden indien het getrokken individu verblijft in een privaat gezin.

Gezinnen kunnen in deze context omschreven worden als alle individuen die hetzelfde verblijfsadres hebben.⁸ Bijkomend kan men private en collectieve huishoudens onderscheiden.

Een privaat gezin, zoals gedefinieerd in het Rijksregister, vormt een benadering van wat omschreven kan worden als een sociologisch gezin. Sociologische gezinnen worden doorgaans omschreven als individuen die een verblijfsplaats delen en die gezamenlijk beslissen over het overgrote deel van hun uitgaven.⁹ Voor de rest van deze nota gebruiken we daarom de termen private en sociologische gezinnen door elkaar.

⁴ Zie Molenberghs en Gaens (1994), p. 11 voor een omschrijving van het begrip steekproefkader.

⁵ Zie DOCDWH (2011), document "fiches_beschrijving_rr_ksz.doc" voor meer uitleg over de inhoud van de Rijksregistergegevens in het DWH AM&SB.

⁶ Voor de vorige MIMOSIS-steekproef werden 100.000 individuen getrokken uit een vergelijkbaar steekproefkader op 1 januari 2002. Zie Vleminckx (2008) voor een beschrijving van de inhoud van deze steekproef.

⁷ De toegepaste principes van systematische steekproeftrekking worden uitgelegd in Vleminckx (2011), sectie 9. Het volledige steekproefmechanisme wordt geïmplementeerd in het programma SampleMIMOSIS_2009.sas (zie FOD Sociale Zekerheid(2011b)).

⁸ Zie Rijksregister (2011).

⁹ Zie Atkinson e.a. (1995), p. 16 voor meer info omtrent mogelijke omschrijvingen van het begrip huishouden.

Collectieve huishoudens omvatten samenlevingsgemeenschappen zoals bejaardentehuizen, hospitaal, gevangenis, kloosters,¹⁰

Bij het trekken van de steekproef in stap 1 en stap 2 wordt dus geen rekening gehouden met het onderscheid tussen private en collectieve huishoudens. Bij het opvragen van identificatienummers in stap 3 wel.

Van alle individuen die opgenomen worden in de uiteindelijke steekproef, worden vervolgens een aantal variabelen uit het DWH&AM&SB opgevraagd. Een volledige lijst van de opgevraagde variabelen kan men terugvinden in FOD SZ (2011a).¹¹

2.1 CONSTRUCTIE VAN EEN HUISHOUDNUMMER EN RAMING VAN DE GEZINSOMVANG

De steekproef, opgebouwd volgens de hoger beschreven principes, omvat 299.257 individuen. Van deze individuen weten we, in principe, of ze a) in een collectief huishouden verblijven en b) indien niet, wie de referentiepersoon (i.e. het gezinshoofd) van het private gezin was op 1 januari 2009 en 1 januari 2002.¹²

In totaal verblijven 1,172 individuen van deze steekproef in een collectief huishouden. Deze individuen krijgen een uniek huishoudennummer en worden in verdere analyses als "alleenstaande" beschouwd.

Van 136 individuen die niet in een collectief huishouden verblijven, was de variabele referentiepersoon op 1 januari 2009 niet ingevuld. Indien de variabele referentiepersoon op 1 januari 2002 wel ingevuld was, werd deze waarde overgenomen. In het andere geval wordt het individu beschouwd als een alleenstaande. Alle individuen uit private huishoudens, die hetzelfde identificatienummer van de referentiepersoon hebben, worden voor verdere analyses beschouwd als leden van hetzelfde huishouden.

Na controle bleek ook dat er 2 gezinnen in de steekproef opgenomen waren die respectievelijk 27 en 28 personen omvatten. Hoewel dit principieel niet onmogelijk is, neemt de kans op fouten in de gezinssamenstelling wel toe naarmate het aantal leden toeneemt. Daarom werden deze twee gezinnen verwijderd uit de steekproef die verder geanalyseerd zal worden.

Na deze ingreep houden we 299.202 individuen over. De 298.030 individuen van deze groep die in private gezinnen wonen, wonen in totaal in 97.760 verschillende gezinnen.¹³

¹⁰ Zie APS (2011).

¹¹ Het gaat hier niet noodzakelijk om variabelen die direct aanwezig zijn, maar om gemanipuleerde variabelen. Het volledige pakket programma's dat de basisgegevens van het DWH AM&SB omzet naar de gevraagde gegevens kan men vinden in FOD Sociale Zekerheid(2011c).

¹² De referentiepersoon is, in principe, de persoon die de belangen van het gezin behartigt of die voor het grootste deel in het onderhoud van het huishouden voorziet. In de praktijk kan het echter ook gaan om de persoon die zich met de administratieve zaken van het gezin bezig houdt (zie APS (2011)).

¹³ Deze manipulaties worden uitgevoerd met het programma CreatePopCharAndWeights.sas (zie FOD Sociale Zekerheid(2011c)).

2.2 CONSTRUCTIE VAN EEN AANTAL KARAKTERISTIEKEN OP BASIS VAN GEGEVENS UIT HET RIJKSREGISTER

In de set van variabelen die voor alle individuen werden opgevraagd¹⁴, nemen de variabelen uit het Rijksregisterbestand een speciale plaats in, omdat ze toelaten de gezinseenheden af te bakenen en informatie bevatten over de onderlinge relatie tussen gezinsleden.

Eén van deze Rijksregistervariabelen is de variabele NAREGNIS_RELATION die de verhouding van ieder gezinslid ten opzichte van het gezinshoofd uitdrukt met een code. De mogelijke waarden die de variabelen NAREGNIS_RELATION kan aannemen, geven we weer in Tabel 1.

Tabel 1: Mogelijke waarden van de variabele NAREGNIS_RELATION

Mogelijke waarde	Omschrijving
1	gezinshoofd
2	echtgenoot, echtgenote
3	zoon, dochter
4	schoonzoon, schoondochter
5	kleinzoon, kleindochter
6	vader, moeder
7	schoonvader, schoonmoeder
8	grootvader, grootmoeder
9	broer, zuster
10	schoonbroer, schoonzuster
11	verwant, verwante
12	zonder familieverband
13	stiefzoon, stiefdochter
14	achterkleinzoon, achterkleindochter
15	oom, tante
16	neef, nicht (verwantschap in de derde graad)
17	neef, nicht (verwantschap in de vierde graad)
20	gemeenschappen, tehuizen

In principe moet de variabele NAREGNIS_RELATION steeds een waarde 1 hebben voor de referentiepersoon van het gezin. Dit bleek in 136 gevallen niet zo te zijn. De waarde voor de variabele NAREGNIS_RELATION bleek in deze gevallen missing en werd daarom manueel op 1 geplaatst.

In principe mag er per gezin ook maar 1 individu voorkomen met de waarde 1 voor de variabele NAREGNIS_RELATION. Er bleken echter 9 gezinnen te zijn waarvoor deze variabele 2 maal een waarde 1 had binnen het gezin. In 7 van de 9 gevallen, was één van de twee geobserveerde cases reeds gezinshoofd in 2002. Voor deze gevallen werd de waarde van de

¹⁴ Zie Vlemingckx (2011) voor de volledige lijst van variabelen.

variabele NAREGNIS_RELATION niet aangepast. Het andere geval werd manueel aangepast naar een waarde 12 voor de variabele NAREGNIS_RELATION, i.e. zonder familieverband.¹⁵

Voor de twee resterende gevallen werd de waarde 1 voor NAREGNIS_RELATION behouden in de lijn waarvoor het identificatienummer van het individu gelijk was aan dat van de referentiepersoon. In het andere geval werd de waarde 1 manueel omgezet naar een waarde 12.

In principe zou de per gezin ook maar één keer een waarde 2, i.e. partner van het gezinshoofd, mogen voorkomen. Dit bleek nu niet het geval te zijn. In 7 gezinnen bleken er 2 partners te zijn en in 1 geval waren er 3 partners. Deze gevallen werden manueel gecorrigeerd na inspectie van de betreffende gezinnen. In 4 gevallen bleken er telkens 2 partners voor te komen met volledig dezelfde gegevens. Vermoedelijk gaat het hier dus om lijnen die onterecht twee maal geselecteerd geweest zijn bij het samenstellen van de steekproef en toch verschillende identificatienummers gekregen hebben. Eén van deze twee lijnen werd daarom verwijderd uit de steekproef die verder geanalyseerd zal worden.

Een variabele die we in de vorige MIMOSIS-steekproef wel direct observeerden maar nu niet, is de variabele NAREGNIS_CIVIL. We geven in Tabel 2 de mogelijke waarden die deze variabele kan aannemen.

Tabel 2: Mogelijke waarden van de variabele NAREGNIS_CIVIL

Mogelijke waarde	Omschrijving
1	Ongehuwd
2	Gehuwd
3	Weduwnaar, weduwe
4	Uit de echt gescheiden
5	Gescheiden van tafel en bed

Op basis van andere beschikbare gegevens, proberen we de waarden van NAREGNIS_CIVIL_2009 op 1 januari 2009 te reconstrueren. Indien er binnen een gezin twee individuen voorkomen met een waarde van respectievelijk 1 en 2 voor de variabele NAREGNIS_RELATION, geven we voor beide individuen de waarde 2 aan de variabele NAREGNIS_CIVIL_2009. Indien na deze operatie de variabele NAREGNIS_CIVIL_2009 nog niet is ingevuld, nemen we de waarde van NAREGNIS_CIVIL_2002 van 1 januari 2002 over indien deze waarde toen groter was dan 2. Indien na deze tweede stap de waarde van NAREGNIS_CIVIL_2009 nog niet is ingevuld, beschouwen we iemand als weduwnaar of weduwe indien een individu ouder is dan 18 en een overlevingspensioen ontvangt. Iemand wordt beschouwd als uit de echt gescheiden indien hij volgens de variabele NAREGNIS_CIVIL_2002 gehuwd was op 1 januari 2002 en er in 2008 geen overlevingspensioen werd betaald en indien er geen aanwijzing zijn van een partner op 1 januari 2009. Indien er in een gezin op 1 januari 2002 twee individuen waren met de waarde NAREGNIS_CIVIL_2002 gelijk aan 2, i.e. gehuwd, en indien voor geen van deze individuen na de voorgaande operaties een waarde werd ingevuld voor de variabele

¹⁵ De 7 observatielijnen waarvoor geen aanpassing werd uitgevoerd, waren ook lijnen waarvoor het identificatienummer van het individu gelijk was aan dat van de referentiepersoon.

NAREGNIS_CIVIL_2009, dan wordt ook op 1 januari 2009 de waarde gelijk gezet aan 2. Alle andere individuen, waarvoor in de voorgaande operaties geen waarde voor NAREGNIS_CIVIL_2009 werd ingevuld, worden beschouwd als ongehuwden.¹⁶

Naast het unieke identificatienummer van het individu, het gezin waartoe een individu behoort, de relatie van een individu met het gezinshoofd (in NAREGNIS_RELATION) en de burgerlijke staat (in NAREGNIS_CIVIL) observeren we van de individuen in de steekproef ook nog de maand en het jaar van geboorte, het geslacht, de gemeente waarin het individu woont, de LIPRO-positie, of het individu al dan niet random geselecteerd is in de eerste of de tweede stap van de steekproeftrekking en een code die het huishouden typeert.¹⁷

2.3 CONSTRUCTIE EN EVALUATIE VAN DE STEEKPROEFGEWICHTEN

In de steekproef observeren we onder andere de regio waarin de geobserveerde individuen wonen. In Tabel 3 vergelijken we de verdeling van de niet gewogen resultaten (i.e. het procentueel aandeel individuen in de totale steekproef) met de bevolkingsstatistieken op 1 januari 2009.

Tabel 3: Niet gewogen procentueel aandeel observaties naar regio¹⁸

	Werkelijk aandeel op basis van bevolkingsstatistieken 2009	Niet gewogen aandeel op basis van steekproef	Procentuele afwijking niet gewogen aandeel en werkelijk aandeel
Regio Brussel Hoofdstad	9.94	8.52	-14.26
Regio Vlaanderen	57.74	58.49	1.30
Regio Wallonië	32.32	32.98	2.03

Zoals blijkt uit de kolom met de procentuele afwijking van het niet gewogen aandeel en het werkelijke aandeel in Tabel 3, wijkt het procentueel aandeel zonder wegingsfactoren, in het bijzonder voor Brussel, aanzienlijk af van het werkelijke aandeel. Dit hoeft, gegeven de manier van steekproeftrekken, niet te verwonderen. Indien we de steekproefresultaten willen gebruiken om uitspraken te doen voor de hele populatie, lijkt het aangewezen om wegingsfactoren te construeren die corrigeren voor deze niet toevallige manier van selecteren van de observaties.

Een abstracte beschrijving van de procedure die we volgen om de steekproefgewichten te construeren, is opgenomen in Appendix 1 van deze tekst. Deze procedure wordt toegepast om zowel voor individuen als gezinnen steekproefgewichten te construeren. In beide gevallen wordt de beschreven procedure afzonderlijk toegepast voor alle gevallen per regio.¹⁹

¹⁶ Alle manipulaties in deze sectie werden uitgevoerd met het programma CreatePopCharAndWeights.sas (zie FOD Sociale Zekerheid(2011c)).

¹⁷ Meer informatie over de inhoud van deze gegevens kan men vinden in DOCDWH (2011), zie document "fiches_beschrijving_rr_ksz.doc".

¹⁸ De gegevens in kolom werkelijke verdeling zijn gebaseerd op gegevens van ADSEI, zie ADSEI (2011a): Bevolking naar geslacht, leeftijdsgroepen en leeftijdsklassen.

¹⁹ Deze procedure wordt geïmplementeerd met het programma CreatePopCharAndWeights.sas (zie FOD Sociale Zekerheid(2011c)).

In Tabel 4 geven we de verdeling van het gewogen aantal observaties naar geslacht en regio. Indien we de lijnen voor mannen en vrouwen zouden optellen, stellen we vast dat de totale aantallen per regio vrij goed benaderd worden. Dit mag echter niet verhullen dat er voor de Regio Brussel Hoofdstad een onderschatting is van het aantal mannen en een overschatting van het aantal vrouwen. Beide afwijkingen compenseren elkaar min of meer. In Appendix 2 ontbinden we deze rapportering verder naar leeftijdsgroepen van 5 jaar. Uit deze ontbinding blijkt dat de afwijking voor de Regio Brussel Hoofdstad vooral een gevolg is van afwijkingen voor mannen en vrouwen in de leeftijdsgroep van 20 tot 40 jaar.²⁰

Tabel 4: Gewogen aantal observaties op individueel niveau naar geslacht en regio²¹

		Werkelijk aantal op basis van bevolkingsstatistie ken 2009	Gewogen aantal op basis van steekproef	Procentuele afwijking gewogen aantal en werkelijk aantal
Mannen	Regio Brussel Hoofdstad	516,250	503,135	-2.6
	Regio Vlaanderen	3,064,169	3,051,317	-0.4
	Regio Wallonië	1,688,232	1,690,329	0.1
Vrouwen	Regio Brussel Hoofdstad	552,282	565,460	2.3
	Regio Vlaanderen	3,144,708	3,157,629	0.4
	Regio Wallonië	1,787,439	1,785,410	-0.1

In Tabel 5 geven we de verdeling van het gewogen aantal observaties naar geslacht en burgerlijke staat voor België. Zoals blijkt uit deze tabel wordt het aantal gehuwde en gescheiden personen aanzienlijk onderschat. Vermoedelijk worden deze personen bij gebrek aan informatie bij de constructie van de variabele NAREGNIS_CIVIL geklasseerd als ongehuwden. Ook het aantal weduwnaars wordt aanzienlijk onderschat, maar weduwen, die veruit de grootste subgroep onder de verduwden vormen, worden dan weer wel goed benaderd. We beschouwen de geobserveerde afwijkingen voor deze variabele meer als een indicator van de beperkte kwaliteit van de gereconstrueerde variabele NAREGNIS_CIVIL dan van de kwaliteit van de wegingsfactoren zelf.

²⁰ Het referentiepunt voor de bepaling van de leeftijd in alle toepassingen voor deze nota is 1 januari 2009.

²¹ De gegevens in kolom werkelijke verdeling zijn gebaseerd op gegevens van ADSEI, zie ADSEI (2011a): Bevolking naar geslacht, leeftijdsgroepen en leeftijdsklassen.

Tabel 5: Gewogen aantal observaties op individueel niveau naar geslacht en burgerlijke staat voor België²²

		Werkelijk aantal op basis van bevolkingsstatistieken 2009	Gewogen aantal op basis van steekproef	Procentuele afwijking gewogen aantal en werkelijk aantal
Mannen	Ongehuwd	2,495,157	2,698,598	7.5
	Gehuwd	2,239,410	2,112,464	-6.0
	Verweduwd	137,892	81,254	-69.7
	Gescheiden	396,175	352,368	-12.4
Vrouwen	Ongehuwd	2,195,730	2,417,431	9.2
	Gehuwd	2,238,798	2,112,415	-6.0
	Verweduwd	575,477	578,781	0.6
	Gescheiden	474,421	399,765	-18.7

Het dient ook onderstreept te worden dat voor de verdere modellering van de regelgeving in MIMOSIS het onderscheid tussen ongehuwd zijn en uit de echt gescheiden zijn niet van belang is. De aanduiding van een huwelijksband is voor latere toepassingen wel van belang.

In Appendix 3 ontbinden we de in Tabel 5 gerapporteerde gegevens, verder naar leeftijdsgroepen van 5 jaar. Uit deze ontbinding blijkt dat de afwijking voor weduwnaars tussen de steekproefverdeling en de werkelijke verdeling zich vooral voor doet in de groep weduwnaars boven de 70 jaar. Voor verdere analyses rond de combinatie van overlevingspensioenen en toegelaten arbeid op beroepsactieve leeftijd, zou deze aanduiding van burgerlijke staat dus wel bruikbaar moeten zijn.

In Tabel 6 geven we tenslotte de verdeling van de gezinseenheden, gewogen met de gereconstrueerde gezinsgewichten, naar regio en in functie van het aantal gezinsleden. Zowel de verdeling op Belgisch niveau, als verdeling op geaggregeerd niveau naar regio levert een relatief goede benadering op van de werkelijke verdeling, met uitzondering van de kleine groep van gezinnen met 7 of meer gezinsleden. Deze onderschatting van de zeer grote gezinnen doet zich voor voor alle regio's. Verder valt op dat, in lijn met de individueel gewogen resultaten, er vooral voor de Regio Brussel Hoofdstad een onderschatting is van het aantal alleenstaande mannen en een overschatting van het aantal alleenstaande vrouwen.

²² De gegevens in kolom werkelijke verdeling zijn gebaseerd op gegevens van ADSEI, zie ADSEI (2011a): Bevolking naar geslacht, leeftijdsgroepen en leeftijdsklassen.

Tabel 6: Verdeling gewogen resultaten op niveau private gezinnen naar regio en aantal gezinsleden²³

		ALLEEN WONENDE MANNEN	ALLEEN WONENDE VROUWEN	2 PERSONEN	3 PERSONEN	4 PERSONEN	5 PERSONEN	6 PERSONEN	7 PERSONEN	8 PERSONEN EN MEER	TOTAAL PRIVATE HUIS- HOUDENS
Gewogen aantal op basis van steekproef											
	België	725,182	822,762	1,463,483	701,405	583,921	214,335	64,400	19,283	11,773	4,606,544
	Regio Brussel Hoofdstad	121,054	128,787	120,790	59,909	45,394	21,032	8,855	3,398	2,239	511,458
	Regio Vlaanderen	360,678	417,671	890,689	406,879	354,812	121,376	33,421	9,595	6,145	2,601,266
	Regio Wallonië	243,450	276,304	452,004	234,617	183,715	71,927	22,124	6,290	3,389	1,493,820
Werkelijk aantal op basis van bevolkingsst atistieken 2009											
	België	710,551	837,803	1,471,120	690,936	569,434	207,817	62,850	17,196	10,271	4,577,978
	Regio Brussel Hoofdstad	108,782	135,663	123,981	59,618	44,059	22,481	9,334	3,140	2,049	509,107
	Regio Vlaanderen	356,344	429,001	894,278	401,032	342,465	116,693	31,281	8,652	4,970	2,584,716
	Regio Wallonië	245,425	273,139	452,861	230,286	182,910	68,643	22,235	5,404	3,252	1,484,155
Procentuele afwijking gewogen aantal en werkelijk aantal											
	België	-2.0	1.8	0.5	-1.5	-2.5	-3.0	-2.4	-10.8	-12.8	-0.6
	Regio Brussel Hoofdstad	-10.1	5.3	2.6	-0.5	-2.9	6.9	5.4	-7.6	-8.5	-0.5
	Regio Vlaanderen	-1.2	2.7	0.4	-1.4	-3.5	-3.9	-6.4	-9.8	-19.1	-0.6
	Regio Wallonië	0.8	-1.1	0.2	-1.8	-0.4	-4.6	0.5	-14.1	-4.0	-0.6

²³ De gegevens in kolom werkelijke verdeling zijn gebaseerd op gegevens van ADSEI. In tegenstelling tot de andere bevolkingsgegevens zijn de gegevens met betrekking tot de private gezinnen op het moment van schrijven voor 1 januari 2009 nog niet publiek beschikbaar op de ADSEI-website. Ze werden op vraag geleverd.

2.4 BESCHRIJVING VAN EEN AANTAL VARIABELEN IN ANDERE BRONNEN DAN HET RIJKSREGISTER

Van alle individuen, die geselecteerd werden voor de steekproef, werden naast de eerder geciteerde Rijksregistergegevens ook een aantal gegevens opgevraagd die door verschillende Openbare Instellingen van de Sociale Zekerheid (OISZ) geleverd worden aan het DWH AM&SB. In Tabel 7 geven we een lijst van de afkortingen die we verder gebruiken om de verschillende geconsulteerde bronnen te omschrijven en een korte omschrijving van deze bronnen. Tenzij anders vermeld zijn de gegevens uit een besproken bestand beschikbaar voor het observatiejaar 2008.

Tabel 7: Afkorting en omschrijving van de verschillende bron waaruit gegevens worden opgevraagd voor de geselecteerde individuen²⁴

Naam file	Omschrijving
CIMIRe	Een bestand met individuele gegevens over de loopbaan van werknemers. Het bestand bevat zowel gegevens van periodes waarin de werknemer beroepsactief was als periodes waarin hij niet beroepsactief was. De mogelijke observatiejaren in het bestand lopen van 1954 tot en met 2001.
NIC	Een bestand met individuele gegevens over de periode van primaire arbeidsongeschiktheid en moederschapsbescherming van verzekerden in het stelsel der werknemers en zelfstandigen.
PODMI	Een bestand met individuele gegevens over de tussenkomsten van OCMW's inzake sociale bijstand.
PensCad	Een bestand met individuele gegevens over de uitbetaalde eerste- en tweedepijlerpensioenen.
RKW_RG	Een bestand met een aanduiding dat het individu een rechtgevend kind is in de kinderbijslag in het stelsel der werknemers, het stelsel van de gewaarborgde kinderbijslag en het stelsel van de ambtenaren.
RKW_RH	Een bestand met een aanduiding dat het individu een rechthebbende is in de kinderbijslag in het stelsel der werknemers, het stelsel van de gewaarborgde kinderbijslag en het stelsel van de ambtenaren.
RKW_BT	Een bestand met een aanduiding dat het individu een bijslagtrekkende is in de kinderbijslag in het stelsel der werknemers, het stelsel van de gewaarborgde kinderbijslag en het stelsel van de ambtenaren.
RSVZ_KB_RG	Een bestand met een aanduiding dat het individu een rechtgevend kind is in de kinderbijslag in het stelsel der zelfstandigen.
RSVZ_KB_RH	Een bestand met een aanduiding dat het individu een rechthebbende is in de kinderbijslag in het stelsel der zelfstandigen.
RSVZ_KB_BT	Een bestand met een aanduiding dat het individu een bijslagtrekkende is in de kinderbijslag in het stelsel der zelfstandigen.
RVA	Een bestand met gegevens van individuen die recht hebben op een uitkering ten laste van de RVA. Het gaat hier onder andere om uitkeringsgerechtigde volledig werklozen, maar ook personen in loopbaanonderbreking of met brugpensioen.
FAO	Een bestand met gegevens over individuen waarvan de verzekering tegen arbeidsongevallen opgevolgd wordt door het Fonds voor ArbeidsOngevallen.
FBZ	Een bestand met gegevens over individuen waarvan de verzekering tegen arbeidsongevallen opgevolgd wordt door het Fonds voor Beroepsziekten.
RIZIV	Een bestand met gegevens over individuen waarvan de invaliditeit erkend is door de Geneeskundige Raad voor Invaliditeit (GRI).
FOD_SZ	Een bestand met gegevens over individuen waarvan de handicap erkend is met het oog op het bekomen van een tegemoetkoming aan personen met een handicap.
LABM97	Een bestand met een beperkt aantal arbeidsmarktgegevens van individuen voor de periode 1997 tot en met 2008.
RSVZ	Een bestand met gegevens van individuen die waren aangesloten bij een sociaal verzekeringsfonds voor zelfstandigen.
RSZPPO_QUART	Een bestand met gegevens van individuen die tewerkgesteld waren bij een lokale of provinciale overheid. Enkel gegevens voor het vierde kwartaal van 2008.
RSZPPO_20022008	Een bestand met gegevens van individuen die tewerkgesteld waren bij een lokale of provinciale overheid in de periode 2002 tot en met 2008.
RSZ_QUART	Een bestand met gegevens van individuen die tewerkgesteld waren als loontrekkende met uitzondering van deze die tewerkgesteld waren bij een lokale of provinciale overheid. Enkel gegevens voor het vierde kwartaal van 2008.
RSZ_20022008	Een bestand met gegevens van individuen die tewerkgesteld waren als loontrekkende in de periode 2002 tot en met 2008 met uitzondering van deze die in de betreffende periode tewerkgesteld waren bij een lokale of provinciale overheid.

²⁴ Meer info over de meeste van deze bestanden kan gevonden worden in DOCDWH (2011).

Een aantal gegevens over de omvang van de verschillende geciteerde bestanden zijn opgenomen in Tabel 8.

De kolom Totaal van Tabel 8 geeft het aandeel van verschillende individuen per bestand als percentage van de totale populatie.²⁵ Uit deze kolom kan onder andere afgeleid worden dat we van meer dan de helft van de populatie gegevens observeren in het loopbaanbestand van werknemers (zie de lijn CIMIRE) of in het bestand met arbeidsmarktgegevens voor de periode 1997 tot en met 2008 (zie de lijn LABM97).

In de kolom Totaal Uniek van Tabel 8 geven we het aandeel van verschillende individuen per bestand dat enkel en alleen voorkomt in het betreffende bestand als percentage van alle verschillende individuen in het bestand. Uit deze kolom blijkt onder andere dat meer dan 70% van de kinderen die recht geven op kinderbijslag en gekend bij de RKW (zie de lijn RKW_RG) en meer dan 60% van de kinderen die recht geven op kinderbijslag en gekend bij de RSVZ (zie de lijn RSVZ_KB_RG), enkel en alleen voorkomen in het bestand van de rechtgevende kinderen.

In de daaropvolgende kolommen van Tabel 8 geven we het aandeel van individuen die gelijktijdig voorkomen in het bestand dat vermeld wordt in de lijn en de kolom van de tabel, uitgedrukt als een percentage van het aantal individuen in het bestand dat vermeld wordt in de lijn. Zo blijkt uit de gegevens in Tabel 8 dat zeer veel van de kinderen die niet enkel alleen in het bestand van de rechtgevende kinderen in de kinderbijslag zitten, ook geobserveerd worden in het bestand met RSZ gegevens (zie respectievelijk de kruising van de lijn RKW_RG en kolom RSZ_20022008 en de lijn RSVZ_KB_RG en kolom RSZ_20022008).

Uit Tabel 8 kan ook afgeleid worden dat de meeste individuen in de steekproef ooit gekend geweest zijn als loontrekkende in de periode 1997 tot en met 2008 (zie kolom LABM97). Uitzonderingen komen vooral voor bij de rechtgevende kinderen (zie lijnen RKW_RG en RSVZ_KB_RG en kolom LABM97), gepensioneerden (zie de lijn PensCad en kolom LABM97) en personen met een handicap (zie de lijn FOD SZ en kolom LABM97).

Het valt tenslotte ook op dat 64% van de individuen die voorkomen in het RVA bestand ook gelijktijdig geobserveerd worden in het bestand met kwartaalgegevens van de RSZ (zie de lijn RVA en kolom RSZ_Quart van Tabel 8).

²⁵ De resultaten zijn gewogen met de individuele wegingsfactoren. Alle gepresenteerde resultaten die vanaf nu volgen zijn gewogen met de geconstrueerde wegingsfactoren.

Tabel 8: Aandeel van individuen in lijn en kolombestand als percentage van het aantal individuen in het lijnbestand

Naam file	Totaal	Totaal uniek	CIMIRE	NIC	PODMI	PensCad	RKW_RG	RKW_RH	RKW_BT	RSVZ_KB_RG	RSVZ_KB_RH	RSVZ_KB_BT	RVA	FAO	FBZ	RIZIV	FOD_SZ	LABM97	RSVZ	RSZPPO_QUART	RSZPPO_20022008	RSZ_QUART	RSZ_20022008
CIMIRE	53.4	4.5	100.0	7.5	1.0	26.8	0.7	15.4	16.4	0.1	1.1	1.1	21.3	2.2	1.1	4.3	7.0	70.4	11.0	5.1	8.2	47.1	60.1
NIC	6.3	0.0	64.0	100.0	2.1	0.9	0.9	30.3	32.8	0.1	0.5	0.6	46.6	9.0	1.1	7.2	2.6	95.5	8.0	6.8	13.1	74.1	89.5
PODMI	1.2	7.7	46.3	11.0	100.0	3.7	16.1	14.0	28.2	0.8	0.3	1.1	26.9	1.8	0.0	1.8	5.2	67.6	3.0	7.1	19.1	20.9	56.1
PensCad	19.2	16.4	74.6	0.3	0.2	100.0	0.2	1.4	1.8	0.0	0.1	0.1	2.7	0.1	2.3	0.7	14.8	31.8	4.1	0.4	2.5	7.9	21.0
RKW_RG	22.8	73.5	1.7	0.3	0.8	0.2	100.0	0.1	1.2	0.7	0.0	0.0	1.2	0.4	0.0	0.1	3.1	22.8	0.3	0.3	3.3	9.1	22.0
RKW_RH	11.3	0.0	73.1	16.9	1.5	2.4	0.1	100.0	37.8	0.0	0.3	0.4	31.6	4.3	0.5	4.0	2.4	95.8	8.0	7.6	11.0	78.9	89.2
RKW_BT	12.6	3.3	69.5	16.3	2.7	2.7	2.1	33.7	100.0	0.0	0.2	0.5	30.2	2.2	0.5	3.6	2.5	86.4	7.6	8.6	13.5	64.0	77.4
RSVZ_KB_RG	2.0	62.2	2.2	0.2	0.5	0.0	8.2	0.1	0.1	100.0	0.4	0.9	1.0	0.3	0.0	0.1	2.3	28.6	1.0	0.2	2.6	10.8	28.0
RSVZ_KB_RH	1.0	0.0	60.7	3.1	0.3	1.8	0.1	3.6	3.2	0.9	100.0	77.3	1.6	0.0	0.0	1.8	0.8	36.3	97.6	0.5	1.6	4.8	22.7
RSVZ_KB_BT	1.0	1.4	61.4	3.9	1.3	2.9	0.1	4.6	6.0	1.8	77.1	100.0	3.2	0.2	0.1	1.9	1.2	39.5	87.7	0.6	2.2	7.9	26.2
RVA	15.7	0.4	72.2	18.6	2.1	3.3	1.7	22.6	24.2	0.1	0.1	0.2	100.0	4.5	0.8	1.3	2.3	92.2	3.9	4.6	10.8	64.2	84.3
FAO	1.8	0.0	65.2	30.9	1.2	1.2	5.2	26.7	15.4	0.3	0.0	0.1	38.5	100.0	0.6	1.8	1.5	99.8	4.3	1.4	6.4	95.3	99.7
FBZ	0.7	0.0	83.8	10.1	0.1	63.6	0.6	8.7	9.0	0.0	0.1	0.1	18.6	1.5	100.0	6.4	19.3	85.0	2.5	3.0	5.4	80.8	82.4
RIZIV	2.7	0.9	85.2	16.7	0.8	5.2	1.0	16.6	17.0	0.1	0.6	0.7	7.5	1.2	1.6	100.0	32.6	67.5	8.4	2.2	4.9	34.8	54.0
FOD_SZ	5.7	4.0	65.7	2.9	1.1	50.3	12.4	4.7	5.5	0.8	0.1	0.2	6.3	0.5	2.3	15.6	100.0	25.4	2.3	0.9	1.9	12.9	19.1
LABM97	55.8	0.1	67.3	10.8	1.5	10.9	9.3	19.3	19.5	1.0	0.6	0.7	26.0	3.3	1.0	3.3	2.6	100.0	9.0	6.6	12.3	65.3	88.0
RSVZ	8.9	9.8	65.7	5.6	0.4	8.7	0.8	10.1	10.8	0.2	10.6	9.6	6.9	0.9	0.2	2.5	1.5	56.2	100.0	2.4	4.7	24.5	45.2
RSZPPO_QUART	3.7	0.0	73.9	11.7	2.3	1.9	2.2	23.3	29.7	0.1	0.1	0.2	19.9	0.7	0.6	1.6	1.3	100.0	5.8	100.0	100.0	13.5	45.0
RSZPPO_20022008	6.8	0.0	64.1	12.0	3.4	7.1	11.0	18.1	24.9	0.8	0.2	0.3	24.8	1.7	0.5	1.9	1.6	100.0	6.1	53.5	100.0	31.4	59.7
RSZ_QUART	36.5	0.0	69.0	12.8	0.7	4.1	5.7	24.3	22.2	0.6	0.1	0.2	27.7	4.8	1.5	2.6	2.0	100.0	6.0	1.4	5.9	100.0	100.0
RSZ_20022008	49.1	0.0	65.3	11.5	1.4	8.2	10.2	20.4	19.9	1.2	0.4	0.5	27.0	3.7	1.1	3.0	2.2	100.0	8.2	3.4	8.3	74.2	100.0

In de rest van deze sectie geven we een beknopt beeld van de inhoud van de beschikbare bestanden en de manipulaties die uitgevoerd zijn om de basisbestanden te herschikken tot bestanden die bruikbaar zijn als inputbestand voor het simulatiemodel.²⁶

Eén algemene regel die gevolgd werd bij de bewerking van de deelbestanden, was dat per deelbestand slechts één observatielijns per individu werd overgehouden.²⁷ In sommige bestanden kwam het immers voor dat één individu geobserveerd werd met meerdere observatielijns. Dit is bijvoorbeeld het geval wanneer individuen gedurende een jaar verschillende rechten in een bepaalde wetgeving hadden of verschillende statuten als loontrekkende of zelfstandige hadden.

Deze reductie tot één lijn gebeurde meestal door de karakteristieken te nemen van het recht of de arbeidsmarktactiviteit met de hoogste vergoeding en de uitbetaalde bedragen op te tellen over alle observatielijns. Waar het wenselijk was om meerdere rechten per individu te weerhouden, werden gegevens van bepaalde observatielijns soms geconverteerd in variabelen. Dit is onder meer gebeurd voor de betaalde pensioenen om het onderscheid tussen rust- en overlevingspensioenen te kunnen behouden. Ook bij de uitkeringen voor ziekte en invaliditeit gebeurde dit om later het onderscheid te kunnen maken tussen invaliditeits- en materniteitsuitkeringen.

Andere inhoudelijke aspecten bespreken we per bestand of per groep van bestanden.

2.4.1 BESTANDEN MET GEGEVENS INZAKE KINDERBIJSLAG (RKW_RG, RKW_RH, RKW_BT, RSVZ_KB_RG, RSVZ_KB_RH, RSVZ_KB_BT)

Deze bestanden bevatten in feite alleen identificatiegegevens. Zo kan via een individueel identificatienummer teruggevonden worden wie gedurende het jaar 2008 rechtgevend kind was, wie het recht op de kinderbijslag opende en wie de bijslag ontving. In principe moet per rechtgevend kind zowel een rechthebbende als een bijslagtrekkende gekend zijn. Vermits we met een steekproef werken komt het voor dat van bepaalde kinderen de rechthebbende of de bijslagtrekkende weliswaar gekend is in de basisgegevens van het DWH AM&SB maar dat dit individu niet is opgenomen in onze steekproef. In dat geval observeren we in de identificatievariabelen de waarde "niet in steekproef".

Naast de identificatiegegevens werd in de bestanden met kinderbijslaggegevens ook een identificator opgenomen om aan te geven of de kinderbijslag werd uitbetaald in het stelsel van de gewaarborgde kinderbijslag.

De kinderbijslagbestanden bevatten echter geen uitbetaalde bedragen, rang van het kind of statuten die het recht op een sociaal supplement of een éénoudersupplement openen. Deze concepten zullen dus, bij de latere berekening van de kinderbijslag moeten ingevuld worden op basis van andere gegevens.

²⁶ We kunnen in deze sectie uiteraard niet in detail ingaan op alle uitgevoerde datamanipulaties. Hiervoor verwijzen we naar de programma's in FOD Sociale Zekerheid(2011b) en FOD Sociale Zekerheid(2011c).

²⁷ Documentatie over de observatie-eenheden die gebruikt worden in de verschillende basisbestanden kan men vinden in DOCDWH (2011).

2.4.2 BESTANDEN MET UITKERINGEN ZIEKTE, INVALIDITEIT, ARBEIDSONGEVAL EN BEROEPSZIEKTE (NIC, RIZIV, FAO EN FBZ)

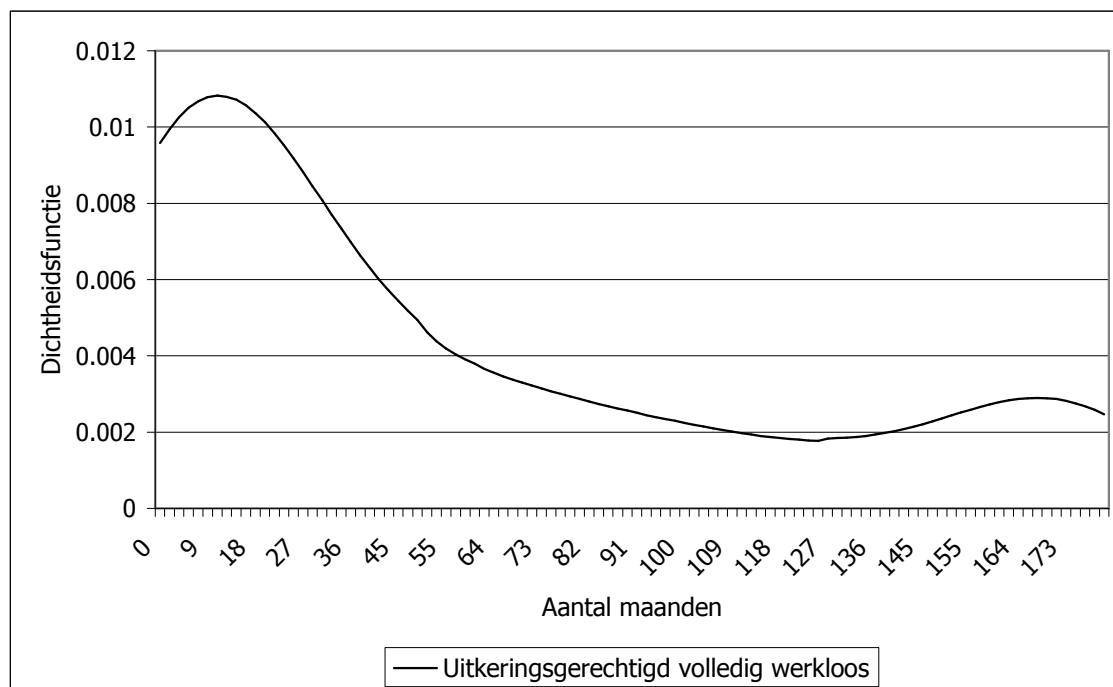
In deze bestanden observeren we voor het jaar 2008 a) een aantal niet monetaire karakteristieken van de rechten die geopend werden tengevolge de geciteerde aandoeningen, b) een aanduiding van de periode waarin de betaling gebeurde en c) een aanduiding van de hoogte van het betaalde bedrag. Alle bedragen zijn doorgaans uitgedrukt in klassen van 10 euro op jaarbasis. De tijdsaanduidingen worden net zoals voor alle andere bronnen ter beschikking gesteld onder de vorm van een maand en jaar. De precieze dag waarop een recht start of stopt of waarop een betaling gebeurde, is dus nooit gekend.

2.4.3 BESTAND MET UITKERINGEN TEN LASTE VAN DE RVA (RVA)

Het bestand met gegevens van 2008 van personen ten laste van de RVA bevat onder andere een variabele die een typering van de steuntrekker toelaat en het aantal maanden gedurende dewelke de persoon een uitkering ontvangt. Ook het uitbetaalde bedrag, ingedeeld in klassen van 10 euro, is opgenomen in het basisbestand.

Op basis van de beschikbare gegevens is het onder andere mogelijk om de groep van uitkeringsgerechtigde volledig werklozen (UVW) af te bakenen. Van deze groep tekenen we in Figuur 1 de dichtheidsfunctie van het aantal uitkeringsmaanden.²⁸ Uit deze figuur blijkt dat er tussen de 12 en 136 maanden werkloosheid een gestage uitstroom uit het uitkeringssysteem is. Na 136 maanden is er dan weer een ophoping van de zeer langdurig werklozen.

Figuur 1: Dichtheidsfunctie van het aantal maanden dat uitkeringsgerechtigde volledig werklozen een werkloosheidsuitkering ontvangen



²⁸ Alle dichtheidsfunctie in deze tekst zijn geschat met de SAS procedure, KDE, optie Univar en de default methode om de bandbreedte te schatten (i.e. de Sheather Jones Plug In methode) (zie SAS (2011)). Alle resultaten zijn ook altijd gewogen met de individuele wegingsfactoren.

2.4.4 BESTAND MET UITBETAALDE PENSIOENEN (PENSCAD)

Het pensioenkadasterbestand (Penscad) bevat voor 2008 gegevens over uitbetaalde pensioenen. Naast de bedragen, opnieuw ingedeeld in klassen van 10 euro op jaarbasis, bevat het bestand per recht ook de datum, uitgedrukt in maand en jaar, waarop het recht is ingegaan. Op basis van deze gegevens is het onder andere mogelijk om uitbetalingen aan respectievelijk rust- en overlevingspensioenen in de stelsels van werknemers, zelfstandigen en ambtenaren te onderscheiden. Ook kan geïdentificeerd worden wie gedurende het jaar recht had op een InkomensGarantie voor Ouderen.

In Tabel 9 geven we de gewogen resultaten van aantal identificeerbare pensioenstatussen op basis van de beschikbare steekproef en vergelijken we deze resultaten met deze van een externe bron.

Het valt op dat voor alle lijnen de omvang van de groep benaderd wordt, maar dat zowel het aantal begunstigden van een rust- als een overlevingspensioen bij de werknemers, onderschat wordt in vergelijking met de externe bron. Het aantal begunstigden met een rust- of overlevingspensioen als ambtenaar of zelfstandige wordt dan weer lichtjes overschat.

Tabel 9: Aantal begunstigden van een rust- of overlevingspensioen in vergelijking met een externe bron²⁹

	Aantal begunstigden op 31 december 2008 in externe bron	Gewogen aantal begunstigden in de steekproef	Procentuele afwijking gewogen aantal en werkelijk aantal
rustpensioen werknemer	1,442,487	1,403,737	-2.7
rustpensioen zelfstandige	440,907	449,792	2.0
rustpensioen overheid	393,634	410,637	4.3
overlevingspensioen werknemer	491,860	460,909	-6.3
overlevingspensioen zelfstandige	145,301	149,765	3.1
overlevingspensioen overheid	117,142	120,679	3.0

2.4.5 BESTAND MET GEGEVENS INZAKE PERSONEN MET EEN HANDICAP (FOD SZ)

Het bestand met gegevens van 2008 van personen met een handicap laat toe om personen die een erkenning van hun handicap en/of het recht op een uitkering hebben, te identificeren. Naast de effectief uitbetaalde bedragen, bevat het bestand ook een aantal karakteristieken die toelaten om de zelfredzaamheid van het individu te beoordelen en om het soort uitkering te typeren.

In Tabel 10 geven we het geschat aantal rechthebbenden met een Integratie- en/of Inkomensvervangende (IT/IVT) tegemoetkoming of een Tegemoetkoming voor Hulp aan Bejaarden (THB). Uit de vergelijking van gewogen steekproefgegevens met deze uit een externe bron, blijkt dat we het aantal gerechtigden IT/IVT met ongeveer 5% overschatten en het aantal gerechtigden THB met ongeveer 1% onderschatten.

²⁹ Voor de aanmaak van de gegevens van de steekproef maken we gebruik van de variabele soort_pensioen (zie o.a. Bijlagen van Berghman e.a. (2007)). De hier gebruikte externe gegevens zijn afkomstig uit de basistoepassingen van het DWH AM&SB (zie KSZ (2011)).

Tabel 10: Aantal rechthebbenden met recht op een Integratie- en/of Inkomensvervangende tegemoetkoming of een Tegemoetkoming voor Hulp aan Bejaarden³⁰

	Aantal begunstigden op 31 december 2008 in externe bron	Gewogen aantal begunstigden in de steekproef	Procentuele afwijking gewogen aantal en werkelijk aantal
Rechthebbend op Integratie- en/of Inkomensvervangende tegemoetkoming	143,037	149,734	4.7
Rechthebbend op een Tegemoetkoming voor Hulp aan Bejaarden	130,455	129,273	-0.9

2.4.6 BESTAND MET SOCIALE BIJSTANDSGEGEVENS (POD MI)

Het bestand met de sociale bijstandsggegevens bevat voor 2008 naast gegevens inzake leefloon of equivalent leefloon ook informatie met betrekking tot een aantal andere mogelijke interventies van OCMW's. Voor de verdere ontwikkeling van een module die sociale bijstand simuleert, focussen we voorlopig echter enkel op de gegevens die betrekking hebben op het leefloon en het equivalent leefloon.³¹ Het bestand bevat naast een aantal karakteristieken die het recht helpen identificeren ook het bedrag, ingedeeld in klassen van 10 euro op jaarbasis, dat ten laste is van de POD Maatschappelijke Integratie.

Zoals blijkt uit de lijn België van Tabel 11, benadert het gewogen totaal aantal gevallen met een leefloon of equivalent leefloon het werkelijk geregistreerd aantal gevallen vrij goed. Dit neemt niet weg dat er onder dit aggregaat regionale afwijkingen schuil gaan, die eveneens geïllustreerd worden in Tabel 11. Vooral de onderschatting van het totaal aantal gevallen in Brussel en de overschatting van het totaal aantal gevallen in Henegouwen, valt op.

³⁰ De gegevens uit de externe bron zijn afkomstig uit FOD Sociale Zekerheid (2010b).

³¹ Deze selectieregels worden nauwkeuriger gespecificeerd in Knapen (2011).

Tabel 11: Gewogen aantal observaties met recht op leefloon of equivalent leefloon³²

Regio	Provincie	Werkelijk aantal op basis van externe bron	Gewogen aantal op basis van steekproef	Procentuele afwijking gewogen aantal en werkelijk aantal
Brussel		29,393	24,080	-18.1
	Antwerpen	9,896	9,624	-2.7
	Limburg	2,483	2,945	18.6
	Oost-Vlaanderen	8,919	8,296	-7.0
	Vlaams-Brabant	4,262	3,674	-13.8
	West-Vlaanderen	4,483	6,082	35.7
	Vlaanderen	30,043	30,621	1.9
Vlaanderen	Henegouwen	15,798	20,429	29.3
	Luik	18,346	18,769	2.3
	Luxemburg	1,918	2,467	28.6
	Namen	4,742	6,358	34.1
	Waals-Brabant	2,232	1,857	-16.8
Wallonië	43,035	49,881	15.9	
België		102,471	104,582	2.1

2.4.7 BESTAND MET GEGEVENS ZELFSTANDIGEN (RSVZ)

Voor personen die in 2008 waren aangesloten bij een verzekeringsfonds voor zelfstandigen, bevat dit bestand onder andere de bijdragecode, het inkomen in klassen van 10 euro dat dient als basis voor de bepaling van de bijdragen als zelfstandigen en een aanduiding van het moment waarop de zelfstandige zich aangesloten heeft als zelfstandige.

Doorgaans worden er in het kader van deze oefening geen gegevens per deelbestand geïmputeerd. Voor personen die aangesloten zijn bij een verzekeringsfonds voor zelfstandigen wordt een uitzondering op deze regel gemaakt. Voor bepaalde bijdragegroepen wordt immers geen inkomen geregistreerd, wat met het oog op de latere constructie van een inkomensverdeling problematisch is. Indien het inkomen niet beschikbaar is, dient echter niet voor alle zelfstandigen een inkomen geschat te worden. Bepaalde bijdragecategorieën wijzen er immers op dat het om een meewerkende echtgenoot gaat waarvan het inkomen begrepen is in het inkomen van de man. Andere categorieën wijzen op het feit dat het gaat om zelfstandigen die hun rechten willen vrijwaren maar die geen activiteit meer uitvoeren.³³ Voor deze groepen wordt dan ook geen inkomen geïmputeerd. De gevallen waarvoor we wel een inkomen imputeren zijn startende zelfstandigen. Voor deze gevallen is het eerste geobserveerde inkomen normaal pas na 3 activiteitsjaren beschikbaar, hoewel er gedurende de eerste drie jaren dus vermoedelijk een inkomen verdiend wordt.

Om inkomens te kunnen imputeren, schatten we eerst het gemiddeld geobserveerd inkomen per groep van individuen die tot dezelfde bijdrageklasse en dezelfde leeftijdsklasse behoren. Indien mogelijk imputeren we dit gemiddelde bij de gevallen waarvoor we geen inkomen observeren maar die behoren tot de groep met een gegeven bijdragecode en een gegeven

³² De gegevens van de externe bron zijn afkomstig van de POD Maatschappelijke Integratie (2011).

³³ De precieze codes om deze gevallen te identificeren worden opgelijst in Knapen (2011).

leeftijdsklasse. Indien op basis van deze regel geen imputatie mogelijk is, wordt voor de overblijvende gevallen eerst het gemiddeld inkomen van dezelfde bijdragecode geïmputeerd, vervolgens (voor de nog resterende gevallen die geen geïmputeerde waarde gekregen hebben) het gemiddeld inkomen van dezelfde leeftijdsklasse en tenslotte het globaal gemiddelde. In Tabel 12 geven we een overzicht van het aantal zelfstandigen aan wie op basis van de respectievelijke regels een inkomen werd toegewezen.

Tabel 12: Regels gevolgd voor imputering jaarinkomen zelfstandigen

	Gewogen aantal zelfstandigen in de steekproef
Geen inkomen geobserveerd en geïmputeerd	136,646
Inkomen geobserveerd	606,753
Geïmputeerd inkomen is het gemiddelde van dezelfde leeftijdsgroep en de bijdragegroep als dat van de aangeslotene waarvoor het inkomen ontbreekt	154,065
Geïmputeerd inkomen is het gemiddelde van dezelfde bijdragegroep als dat van de aangeslotene waarvoor het inkomen ontbreekt	17,185
Geïmputeerd inkomen is het gemiddelde van dezelfde leeftijdsgroep als dat van de aangeslotene waarvoor het inkomen ontbreekt	6,853
Geïmputeerd inkomen is het globaal gemiddelde	250

2.4.8 BESTAND MET LOON EN ARBEIDSTIJDGEGEVENS VOOR RSZ-PLICHTIGEN VOOR HET 4^{DE} KWARTAAL VAN 2008 (RSZ_QUART)

Het bestand RSZ_Quart bevat voor het 4^{de} kwartaal van 2008 loon- en arbeidstijdgegevens van bijdrageplichtigen die aangesloten zijn bij de RSZ (i.e. werknemers in de privé en federale ambtenaren). Een individu kan in dit bestand in een gegeven kwartaal voorkomen met meerdere tewerkstellingslijnen (dit zijn alle prestaties bij een zelfde werkgever met dezelfde hoedanigheid). Tewerkstellingslijnen kunnen verder gegroepeerd worden in werknemerslijnen (dit zijn alle tewerkstellingslijnen met hetzelfde bijdrageregime) en een job (dit zijn alle tewerkstellingslijnen met hetzelfde paritair comité).³⁴

Voor later gebruik van deze data in het model, is het vooral van belang het bijdrageregime en de mogelijke bijdrageverminderingen te identificeren. Voor de verdere verwerking houden we per individu slechts één bijdrageregime over. De reductie in geval van verschillende regimes, gebeurt door de karakteristieken van de belangrijkste job over te houden. De gegevens die geaggregeerd kunnen worden zoals inkomen en arbeidsduur worden opgeteld over jobs. Dit betekent in vergelijking met het basisbestand weliswaar een verlies aan informatie, maar dit zou al bij al toch beperkt moeten zijn. Het bijdrageregime wordt immers vooral bepaald door de werknemers- en de werkgeversklasse. Ongeveer 95% van de individuen in het centrale bestand (4^{de} kwartaal van 2008) komt voor met maar 1 waarde voor de werknemersklasse en ongeveer 86% komt voor met 1 waarde van de werkgeversklasse.

³⁴ Meer uitleg over deze begrippen kan men vinden in DOCDWH (2011), documenten "fiches, beschrijving RSZ, Nederlands, versie 3-1-2011.doc" en "fiches, beschrijving RSZPPO, Nederlands, versie 20-5-2011.doc".

Voor wat de bijdrageverminderingen betreft observeren we voor alle individuen die voorkomen met een code "vermindering van de persoonlijke bijdragen" maar 1 verschillende code. Voor wat de verminderingscodes inzake bijdrageverminderingen van de werkgeversbijdragen betreft, komen bij 99,9% van de gevallen maximaal 3 verschillende verminderingscodes voor. Daarom maken we per individu, 3 bijkomende variabelen waarin deze verschillende verminderingscodes worden opgeslagen.

Naast de verminderingscodes zelf, observeren we in het RSZ bestand ook de werkelijk betaalde werkgevers- en werknemersbijdragen voor bijdrageverminderingen en verminderingen die desgevallend worden toegepast op deze respectievelijke bijdragen.³⁵

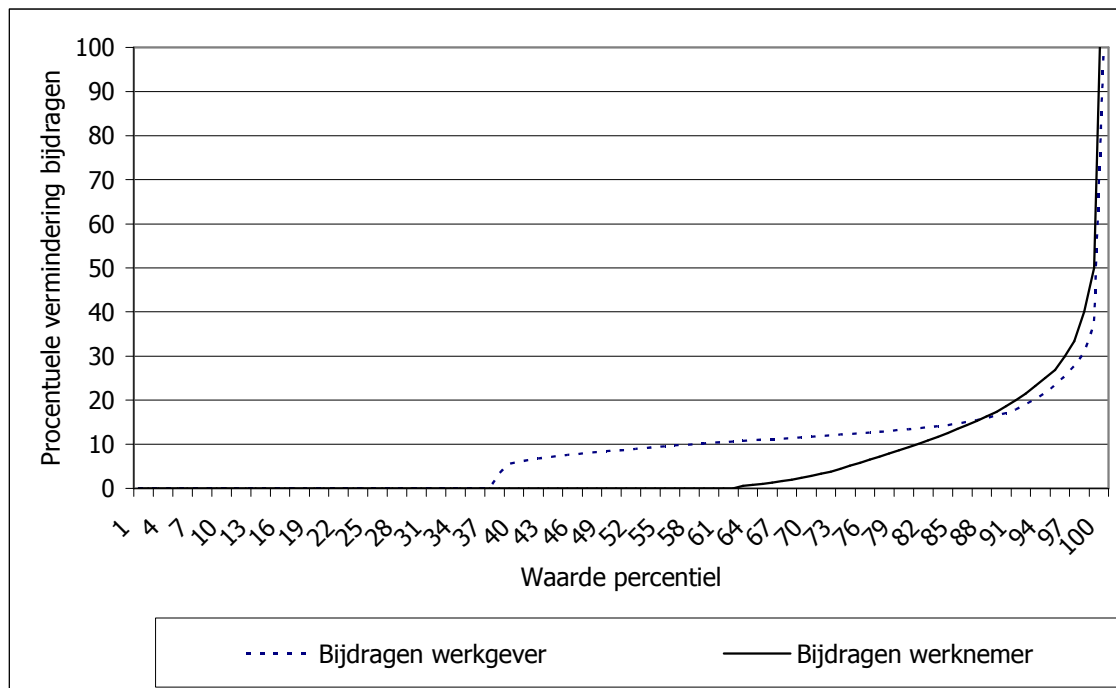
In Figuur 2 geven we bijkomende informatie over de verdeling van deze gegevens. Eerst werd de bijdragevermindering uitgedrukt als een percentage van de bijdrage voor vermindering. Vervolgens werd de verdeling van het geconstrueerde gegeven geordend van klein naar groot en ingedeeld in percentielen. In Figuur 2 wordt per percentiel van de procentuele vermindering van werkgevers- en werknemersbijdragen, het gemiddelde afgedrukt.

Uit Figuur 2 kan afgeleid worden dat ongeveer 2/3 (start rond 38^{ste} percentiel) van de geobserveerde gevallen in het RSZ bestand, geniet van een vermindering van de bijdragen van werkgevers. Voor het overgrote deel van deze gevallen, i.e. van het 38^{ste} tot het 97^{ste} percentiel, ligt deze reductie onder de 20% van het bedrag voor vermindering. Ongeveer 1/3^{de} van de geobserveerde gevallen heeft recht op een vermindering van de persoonlijke bijdragen. Ook hier ligt het gros van de verminderingswaarden onder de 20%.³⁶

³⁵ Zoals het overgrote deel van de nominale bedragen waarover we beschikken zijn deze gegevens ingedeeld in inkomensklassen van 10 euro. In dit geval gaat het om een grens toegepast op kwartaalbasis.

³⁶ De hier uitgedrukte verminderingpercentages zijn gebaseerd op de verhouding van alle verminderingen en alle bijdragen in het gegeven kwartaal. Het is dus niet uitgesloten dat deze verdeling afwijkt, van deze op het niveau van de jobs.

Figuur 2: Percentielen van de procentuele vermindering van werkgevers- en werknemersbijdragen in het bestand met RSZ gegevens



2.4.9 BESTAND MET LOON EN ARBEIDSTIJDGEGEVENS VOOR RSZPPO-PLICHTIGEN VOOR HET 4^{DE} KWARTAAL VAN 2008 (RSZPPO_QUART)

Het bestand RSZPPO_Quart bevat gegevens die vergelijkbaar zijn met deze uit het RSZ_Quart bestand maar dan voor ambtenaren en contractuelen die aangesloten zijn bij een lokale of provinciale overheid. Ook hier kunnen meerdere jobs per individu voorkomen. Een job wordt in principe op dezelfde manier bepaald als bij de RSZ maar we hebben de reductie tot één lijn per individu hier noodgedwongen anders moeten doen dan bij de RSZ. In het geval van de RSZPPO werden de gegevens van de job met het hoogste inkomen en bestaand op de laatste dag van het kwartaal weerhouden.³⁷

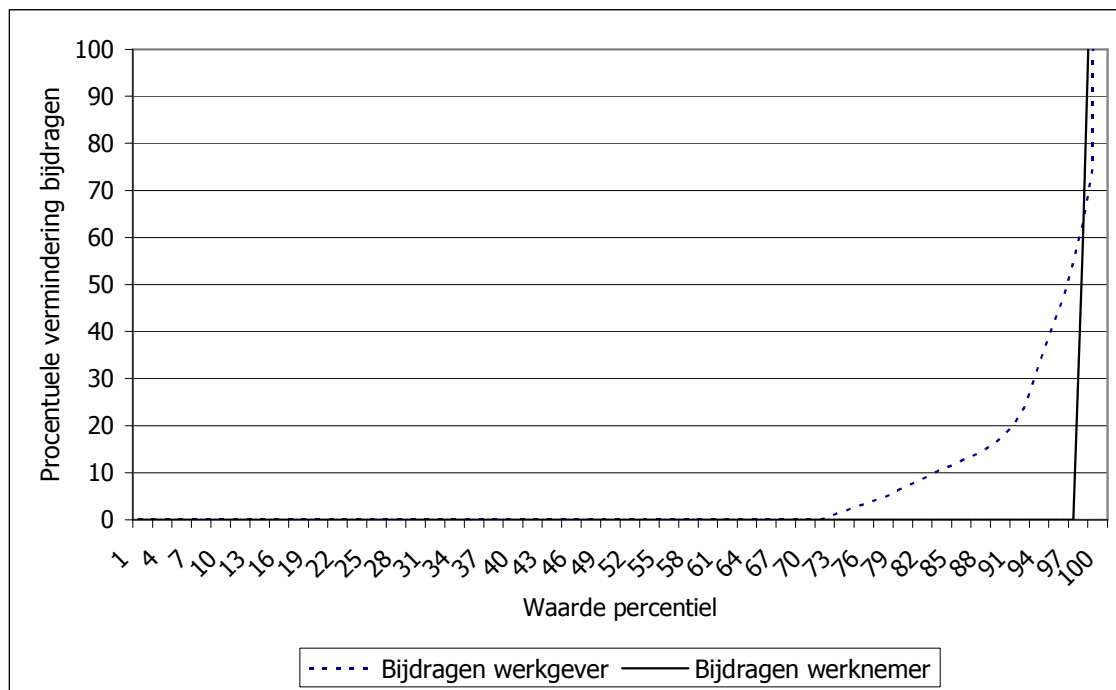
Het verlies aan informatie door deze reductie, met het oog op de latere bijdragebepaling, zou ook hier beperkt moeten zijn. Meer dan 95% van de individuen komt in het bestand van het 4^{de} kwartaal van 2008 voor met 1 werknemerscode en 1 werkgeverscode. Het lijkt dus gerechtvaardigd te veronderstellen dat individuen in een gegeven kwartaal doorgaans in hetzelfde bijdrageregime zitten.

Ook voor wat de bijdrageverminderingen betreft, zal er bij de latere modellering noodgedwongen een andere aanpak gevolgd moeten worden. Er bestaan namelijk verminderingcodes die vergelijkbaar zijn met deze van de RSZ maar deze werden niet opgenomen in de data-aanvraag.

³⁷ We volgen hier een afwijkende regel omdat het equivalent van de gegevens die bij de RSZ gebruikt werden, niet was opgenomen in de data-aanvraag. Op basis van de RSZ data kunnen we beide criteria voor de selectie van karakteristieken toepassen. Er werd vastgesteld dat in 9,5% van de gevallen de twee criteria niet resulteerden in de selectie van dezelfde lijn.

We kunnen voor observaties uit het RSZPPO bestand eventueel wel terugvallen op een raming van de procentuele vermindering van de persoonlijke- en werkgeversbijdragen. In Figuur 3 geven we, net zoals voor de RSZ gegevens, per percentiel het gemiddelde van de procentuele vermindering van werkgevers- en werknemersbijdragen. De bekomen verdeling ziet er in dit geval wel heel anders uit dan deze op basis van de RSZ gegevens. Van de geobserveerde gevallen uit het RSZPPO bestand heeft ongeveer 30% recht op een vermindering van de werkgeversbijdragen en slechts 1% op een vermindering van de werknemersbijdragen.

Figuur 3: Percentielen van de procentuele vermindering van werkgevers- en werknemersbijdragen in het bestand met RSZPPO gegevens



2.4.10 BESTANDEN MET GEGEVENS INZAKE ARBEIDSMARKTHISTORIEK VAN LOON EN WEDDETREKKENDEN (LABM97, RSZPPO_20022008, RSZ_20022008 EN CIMIRE)

Voor alle individuen uit de steekproef, die actief geweest zijn als loon en weddetrekkende in de periode 1997 tot en met 2008, kunnen we op basis van het bestand LABM97 identificeren of ze voornamelijk actief waren op de private dan wel op de publieke arbeidsmarkt. Voor de gevallen die vooral actief waren op de private arbeidsmarkt kunnen we onderscheiden of dit vooral in een arbeiders- dan wel een bediendenstatuut was. Voor gevallen op de publieke arbeidsmarkt kunnen we contractuelen van ambtenaren onderscheiden.

De bestanden RSZ_20022008 en RSZPPO_20022008 bevatten veel minder variabelen dan de bestanden RSZ_Quart en RSZPPO_Quart, maar laten toch toe om een aantal aspecten van de arbeidsmarktactiviteit van loon- en weddetrekkenden te kwantificeren voor de jaren 2002 tot en met 2008. De variabelen die voor ieder van de jaren 2002 tot en met 2008 geïdentificeerd worden, worden opgelijst in Tabel 13.

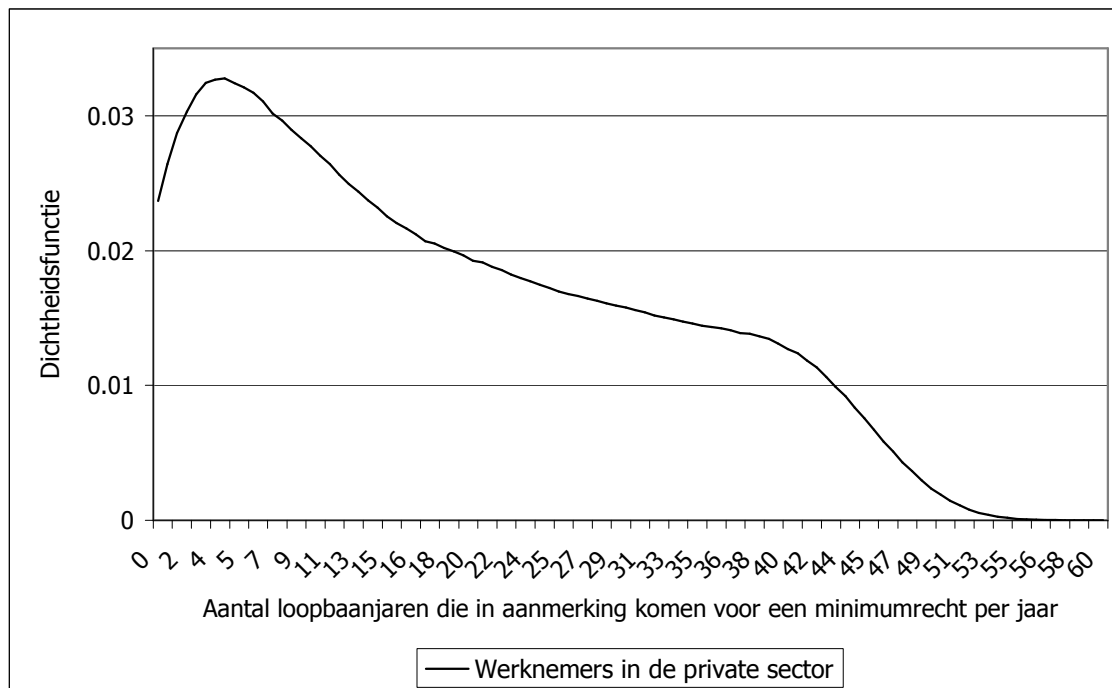
Tabel 13: Afkorting en omschrijving van variabelen die de arbeidsmarktactiviteit van loon- en weddetrekkenden kwantificeren voor de jaren 2002 tot en met 2008³⁸

Naam variabele	Omschrijving
ASSIMDAYS	Aantal gelijkgestelde dagen gedurende het jaar
REMDAYS	Aantal vergoede dagen gedurende het jaar
REMHOURS	Aantal vergoede uren gedurende het jaar
REMUN	Totaal inkomen uit arbeid op jaarbasis (in klassen van 10 Euro)
DAYS_FULL	Aantal dagen voltijdse arbeid gedurende het jaar
DAYS_PART	Aantal dagen deeltijdse arbeid gedurende het jaar
LABMSTAT_PRIV	Belangrijkste status op de private arbeidsmarkt gedurende het jaar (arbeider, bediende)
LABMSTAT_PUB	Belangrijkste status op de publieke arbeidsmarkt gedurende het jaar (contractueel, ambtenaar)
SALJRS	Dagloon

De eerste 4 variabelen uit Tabel 13 kunnen verder ook nog ingevuld worden voor de jaren dat individuen actief waren op de private arbeidsmarkt in de periode 1954 tot en met 2001.

Op basis van deze gegevens kan onder andere een beeld geschetst worden van het aantal loopbaan jaren die een minimum pensioenrecht per loopbaanjaar opleveren. In Figuur 4 schetsen we op basis van de observaties die voorkomen in het CIMIRE bestand, de dichtheidsfunctie van het aantal jaren met een minimumrecht per individu.

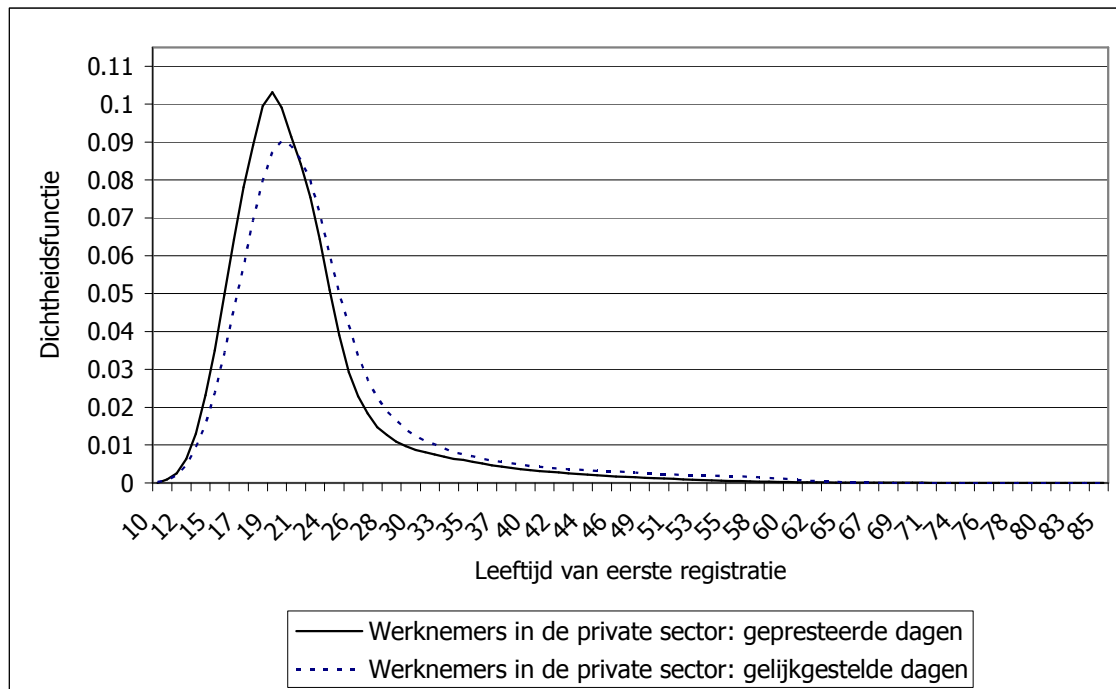
Figuur 4: Dichtheidsfunctie van het aantal loopbaanjaren dat in aanmerking komt voor een minimum pensioenrecht per loopbaanjaar voor werknemers in de private sector.



³⁸ Met de uitwerking van deze concepten bouwen we verder op het werk beschreven in Mutsaerts (2011).

Verder kan op basis van de CIMIRe gegevens ook nagegaan worden wat de leeftijd van het eerste arbeidsmarktcontact was van individuen die actief zijn of geweest zijn op de private arbeidsmarkt en van de leeftijd waarop de eerste gelijkstelling geregistreerd werd. In Figuur 5 schetsen we de dichtheid van deze leeftijdsgegevens.

Figuur 5: Dichtheidsfunctie van de leeftijd van eerste arbeidsmarktveraring en eerste gelijkstelling voor werknemers in de private sector



Zoals blijkt uit Figuur 5, levert het overgrote deel van de geobserveerden zijn eerste arbeidsmarktprestatie tussen de leeftijd van 15 en 28 jaar met een piek rond 20-21 jaar. De eerste gelijkstelling van dagen kent een gelijkaardig verloop als de gepresteerde dagen, maar piekt minder en vindt doorgaans plaats op een iets latere leeftijd dan de eerste prestatie.

2.4.11 CONSTRUCTIE VAN HET LAATST VERDIENDE DAGLOON

Een gegeven dat we voor sommige individuen direct observeren in één van de arbeidsmarktbestanden en dat we voor bepaalde andere gevallen reconstrueren, is het laatst verdiende dagloon. Dit concept wordt opgeslagen onder de naam SALJRS_LAST_CORR. Dit gegeven neemt in latere analyses een belangrijke plaats in omdat het zal dienen als input voor de herberekening van bepaalde uitkeringen en omdat het een indicatie geeft van het verdienvermogen van een individu.

Wanneer een instelling (RSZ of RSZPPO) een dagloon aanlevert, werden daarvoor een aantal vergoedingen voor gepresteerde arbeid, uitgekeerd gedurende een bepaalde periode,

gedeeld door de arbeidstijd gepresteerd gedurende dezelfde periode.³⁹ We noemen deze variabele verder dagloonconcept 1.⁴⁰

Op basis van het laatste jaar waarvoor loon- en arbeidstijdgegevens beschikbaar zijn, kunnen we ook zelf een dagloonconcept berekenen. Dit gebeurde op twee verschillende manieren:

- De geobserveerde loonmassa gedeeld door het aantal vergoede dagen (dagloonconcept 2),
- De geobserveerde loonmassa maal 7.6 gedeeld door het aantal uren werk (dagloonconcept 3).

Deze concepten worden per individu enkel onder bepaalde voorwaarden berekend. Een individu moet op jaarbasis minimaal 24 dagen of 182 uur gewerkt hebben (i.e. 4 weken in een 6 dagenweek of het equivalent in dagen van 7.6 uur per dag). De geobserveerde tijdsprestaties mogen ook niet groter zijn dan 52*6 dagen of 52*6*7.6 uren. Van observatielijnen die niet voldoen aan deze regels, worden de loon- en arbeidstijdgegevens niet gebruikt voor de reconstructie van het dagloon en wordt de waarde van het geschat dagloon op 0 gezet.

Voor ongeveer 71% (7,635,230 eenheden op een totaal van 10,753,080) van het gewogen aantal eenheden in de steekproef observeren we loon- en arbeidstijdgegevens in minstens één prestatiejaar gaande van 2008 tot 1954.

Voor ongeveer 60% (4,646,508 op een totaal van 7,635,230 eenheden) van deze groep observeren we een dagloon dat door de RSZ of RSZPPO direct wordt aangeleverd voor 2008. Voor 9% van alle gevallen is er geen dagloon voor 2008 maar wel voor een dagloonconcept dat hetzij door de RSZ of de RSZPPO direct wordt aangeleverd voor één van de jaren voorafgaand aan 2008.

Voor alle gevallen waarvoor het dagloonconcept 1 beschikbaar was, werd SALJRS_LAST_CORR ingevuld op basis van dit concept. Voor andere gevallen werd eerst de voorkeur gegeven aan dagloonconcept 2 en indien dat niet beschikbaar was werd dagloonconcept 3 gebruikt voor de invulling van het laatst verdiende dagloon. Indien SALJRS_LAST_CORR werd ingevuld op basis van gegevens van een observatiejaar voorafgaand aan 2008, werd de waarde geherwaardeerd op basis van de herwaarderingscoëfficiënten opgenomen in Appendix 4 van dit document.

Na deze ingrepen wordt gecontroleerd of de waarde van SALJRS_LAST_CORR geen outlier is ten opzichte van de verdelingsgegevens van 2008. We beschouwen een waarde als een outlier indien de gereconstrueerde waarde groter is dan het gemiddelde plus twee maal de standaardafwijking van de observaties van 2008 van het dagloonconcept 1. Concreet is de

³⁹ Meer uitleg over deze begrippen kan men vinden in DOCDWH (2011), documenten "fiches, beschrijving RSZ, Nederlands, versie 3-1-2011.doc" en "fiches, beschrijving RSZPPO, Nederlands, versie 20-5-2011.doc".

⁴⁰ In tegenstelling tot de andere nominale bedragen uit de data-aanvraag werd dit direct geobserveerde dagloonconcept niet ingedeeld in klassen.

gebruikte bovengrens gelijk aan 309.93 Euro.⁴¹ Indien voor outliers het dagloonconcept 2 of het dagloonconcept 3 geen outlier was, werd de waarde van SALJRS_LAST_CORR vervangen door één van deze waarden.

Het dagloonconcept dat we na de voorgaande stappen bekomen, blijkt in ongeveer 3% van alle gevallen waarvoor we arbeidsmarktgegevens observeren, gelijk te zijn aan 0. In een aantal gevallen moeten we ook ver teruggaan in de tijd om een waarde voor SALJRS_LAST_CORR te kunnen construeren. Het ontbreken of verouderd zijn van het gereconstrueerde gegeven is echter vooral van belang voor de groep die potentieel nog actief is op de arbeidsmarkt omdat het gereconstrueerde dagloon vooral voor simulaties met betrekking tot die groep zal ingezet worden. Daarom rapporteren we in Tabel 14 gegevens met betrekking tot het gereconstrueerde dagloon voor de groep van individuen die 18 jaar of ouder zijn en jonger dan 65.

Tabel 14 geeft meer informatie met betrekking tot de bron (i.e. uit welke bron werd het gereconstrueerde gegeven gerecupereerd), het prestatiejaar (i.e. in welk jaar werden de gegevens waarop de reconstructie gebaseerd is, geregistreerd) en de hoogte van het gereconstrueerde dagloon.

Voor de bron onderscheiden we 5 mogelijke waarden:

- Ofwel is een gegeven geen outlier volgens de hoger gedefinieerde regels en gebaseerd op gegevens van de RSZ, RSZPPO of CIMIRe. Voor de bron-waarde wordt in dit geval de afkorting van het bestand gebruikt.
- Ofwel was het niet mogelijk om een dagloon te construeren. De waarde voor bron is in dit geval "Onbekend".
- Ofwel is de observatiewaarde groter dan 309.93 Euro. De waarde voor bron is in dit geval "Aftop".

In de analyse van de hoogte van het geschatte dagloon focussen we op 4 zones:

- Ofwel is het geschatte dagloon gelijk aan 0.
- Ofwel is het geschatte dagloon kleiner dan het minimum dagloon. De waarde die we hier als grens gebruiken is 57.97 Euro.⁴²
- Ofwel is het geschatte dagloon groter of gelijk aan het minimum dagloon maar kleiner dan de grenswaarde voor outliers (i.e. 309.93 Euro).
- Ofwel is het geschatte dagloon groter of gelijk aan de grenswaarde voor outliers.

Voor latere analyses is het vooral van belang om na te gaan in welke mate er waarden onder het minimumloon of boven de maximumgrens voorkomen. Indien dit het geval is zullen deze

⁴¹ Het gemiddelde van dagloonconcept 1 dat we observeren voor 2008 is 122.05 Euro en de standaardafwijking 93.94 Euro. Dit impliceert dat we waarden vanaf 309.93 Euro klasseren als outliers.

⁴² Voor de raming van dit minimum dagloon vertrekken we van het Gewaarborgd Gemiddeld Minimum MaandInkomen (GGMMI) van een 22 jarige met 1 jaar anciënniteit. In 2008 bedroeg het GGMMI 1360.71 Euro (zie Put (2008), p. 1043). Dit bedrag wordt vermenigvuldigd met 12/13 in de veronderstelling dat dit maandbedrag het maandequivalent van een eindejaarspremie van 1 maand bevat. Vervolgens wordt dit bedrag gedeeld door $(52 \cdot 5) / 12$ of 21.66667, het verondersteld aantal gewerkte dagen per maand in een voltijdse betrekking.

waarden later namelijk ofwel verhoogd worden tot het minimum dagloon ofwel afgetopt worden op de maximumgrens. Om na te gaan of deze ingrepen verschillen over de verschillende deelpopulaties met een vervangingsinkomen of een uitkering sociale bijstand splitsen we de analyse van de variabelen bron, prestatiejaar en hoogte van het geschatte dagloon uit over de subgroepen van individuen gekend in de verschillende deelbestanden.

Tabel 14: Bron, prestatiejaar en hoogte van geschatte dagloon (GD) voor individuen gekend in bestanden met vervangingsinkomens of uitkeringen sociale bijstand van 18 jaar of ouder maar jonger dan 65

	POD_MI		NIC		PensCad		RVA		FAO		FBZ		RIZIV		FOD SZ	
	Freq	%	Freq	%	Freq	%	Freq	%	Freq	%	Freq	%	Freq	%	Freq	%
<u>Bron GD</u>																
Aftop	31	0.0	1,699	0.3	4,225	1.3	4,711	0.3	431	0.2	107	0.4	526	0.2	319	0.1
CIMIRe	20,184	20.7	21,127	3.4	110,944	33.3	204,207	12.7	143	0.1	7,512	26.3	123,241	47.2	132,488	55.8
Onbekend	16,115	16.6	6,063	1.0	19,963	6.0	28,166	1.8	74	0.0	142	0.5	11,183	4.3	26,590	11.2
RSZPPO	8,325	8.6	43,657	7.0	28,554	8.6	84,436	5.3	515	0.3	2,191	7.7	9,157	3.5	7,190	3.0
RSZ	52,737	54.2	548,587	88.3	169,403	50.9	1,285,532	80.0	184,034	99.4	18,579	65.1	117,097	44.8	70,661	29.8
Totaal	97,392	100.0	621,133	100.0	333,089	100.0	1,607,053	100.0	185,197	100.0	28,531	100.0	261,203	100.0	237,248	100.0
<u>Prestatiejaar GD</u>																
Waarde is 0	16,147	16.6	7,762	1.3	24,188	7.3	32,877	2.1	505	0.3	249	0.9	11,708	4.5	26,909	11.3
voor 1960	0	0.0	0	0.0	233	0.1	0	0.0	0	0.0	0	0.0	0	0.0	126	0.1
van 1960 tot en met 1969	1,311	1.4	0	0.0	13,452	4.0	339	0.0	0	0.0	53	0.2	517	0.2	3,379	1.4
van 1970 tot en met 1979	3,643	3.7	552	0.1	27,658	8.3	6,531	0.4	53	0.0	249	0.9	12,174	4.7	18,059	7.6
van 1980 tot en met 1989	5,115	5.3	3,570	0.6	24,704	7.4	34,287	2.1	0	0.0	1,647	5.8	27,610	10.6	26,273	11.1
van 1990 tot en met 1999	6,878	7.1	9,818	1.6	35,697	10.7	108,747	6.8	89	0.1	4,347	15.2	64,283	24.6	52,840	22.3
van 2000 tot en met 2007	25,686	26.4	69,352	11.2	136,016	40.8	354,901	22.1	1,610	0.9	7,142	25.0	107,177	41.0	67,741	28.6
2008	38,613	39.7	530,078	85.3	71,141	21.4	1,069,371	66.5	182,939	98.8	14,842	52.0	37,735	14.5	41,922	17.7
Totaal	97,392	100.0	621,133	100.0	333,089	100.0	1,607,053	100.0	185,197	100.0	28,531	100.0	261,203	100.0	237,248	100.0
<u>Hoogte geschatte waarde dagloon</u>																
GD is 0	20,317	20.9	10,206	1.6	26,699	8.0	48,597	3.0	74	0.0	794	2.8	21,591	8.3	68,404	28.8
GD < minimum	18,195	18.7	12,977	2.1	55,011	16.5	68,022	4.2	1,621	0.9	1,123	3.9	38,772	14.8	49,280	20.8
minimum < GD < toegelaten maximum	58,849	60.4	596,252	96.0	247,154	74.2	1,485,723	92.5	183,070	98.9	26,506	92.9	200,315	76.7	119,246	50.3
GD >= toegelaten maximum	31	0.0	1,699	0.3	4,225	1.3	4,711	0.3	431	0.2	107	0.4	526	0.2	319	0.1
Totaal	97,392	100.0	621,133	100.0	333,089	100.0	1,607,053	100.0	185,197	100.0	28,531	100.0	261,204	100.0	237,248	100.0

Voor wat de bron betreft kunnen we uit Tabel 14 afleiden dat het aftoppen van waarden boven de bovengrens in alle bestanden zeer weinig voorkomt. Voor alle bestanden, behalve voor individuen gekend bij het RIZIV of de FOD SZ wordt ook meer dan 60% van de geschatte daglonen gerecupereerd uit bestanden van de RSZ of de RSZPPO. Voor individuen gekend bij het RIZIV of de FOD SZ wordt het dagloon dan weer in relatief belangrijke mate geconstrueerd op basis van de CIMIRe gegevens.

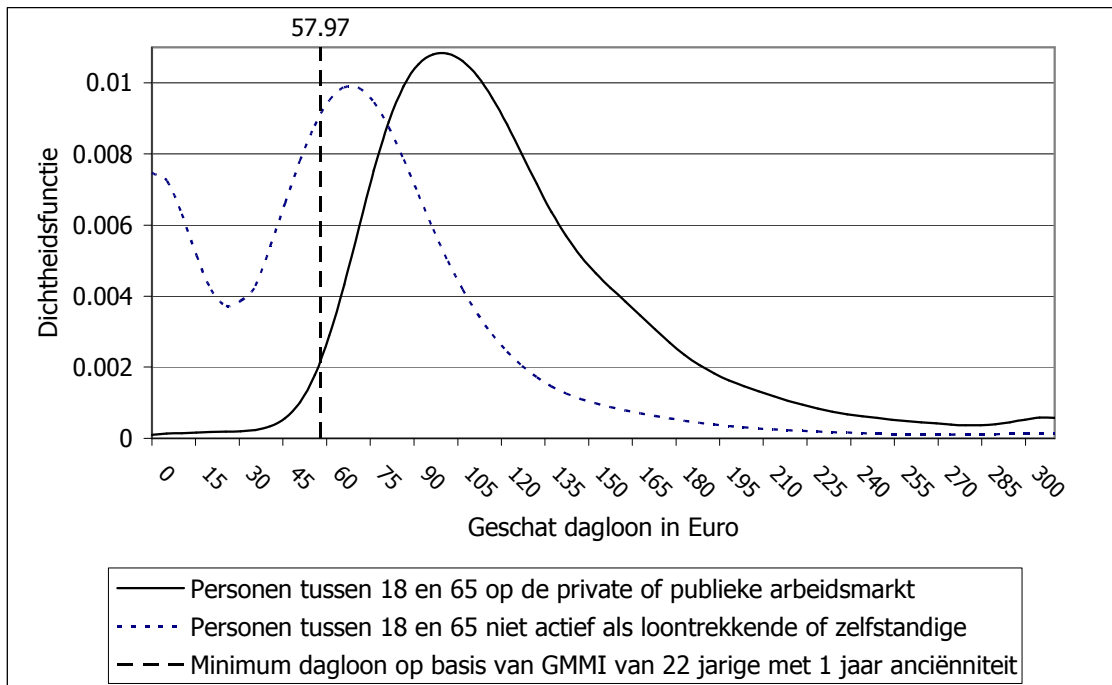
De resultaten in Tabel 14 tonen ook dat het aantal gevallen waarvan het dagloon gereconstrueerd wordt met registraties van prestatiejaren voor 1990, relatief beperkt is. Uitzondering zijn hier opnieuw de bestanden met individuen gekend bij het RIZIV en de FOD SZ en bestand met uitbetalingsgegevens van de pensioenen. In deze gevallen is het aandeel prestatiejaren van voor 1990, groter dan 15%. Voor de individuen gekend in de andere bestanden en zelfs deze gekend bij het RIZIV en in het Pensioenkadaster, is het aandeel van de laatste 9 prestatiejaren, i.e. 2000 tot en met 2008, telkens groter dan 50%. Vooral voor de populatie gekend bij de FOD SZ is het aandeel recente arbeidsmarktcontacten in vergelijking met de andere populaties beperkt.

Voor de populatie gekend bij de POD MI en bij de FOD SZ is het aandeel geschatte daglonen onder de minimumgrens relatief hoog. Dit blijkt zowel een gevolg te zijn van een relatief hoog aandeel nullen (i.e. niet reconstrueerbare waarden) als een aanzienlijk aandeel geschatte waarden onder de grens.

Zowel op basis van de bron, het prestatiejaar als de hoogte van het geschatte dagloon kan men, niet geheel onverwacht, concluderen dat de bestanden van het NIC, de RVA, het FAO en FBZ deelpopulaties bevatten die relatief, in vergelijking met de populaties van de POD MI de FOD SZ, dicht bij de arbeidsmarkt staan.

Tenslotte tekenen we in Figuur 6 de dichtheidsfunctie van het geschatte dagloon voor de groep van 18 tot 65 jarigen die respectievelijk a) actief waren als loontrekkende in het laatste kwartaal van 2008 en b) die niet actief waren als loontrekkende in het laatste kwartaal van 2008 of ingeschreven als zelfstandige in 2008.

Figuur 6: Dichtheidsfunctie van het geschat dagloon voor actieve loontrekkenden en niet actieven



Uit Figuur 6 blijkt dat personen met een laag verdienvermogen bij de groep van niet actieven relatief sterk vertegenwoordigd zijn in vergelijking met de groep van actieven. Uit deze figuur blijkt ook dat indien de niet actieven de arbeidsmarkt zouden betreden, het overgrote deel van deze groep dit vermoedelijk zal moeten doen aan een loon dat ligt tussen het minimumloon en twee maal het minimumloon (i.e. 115 euro).

3 WERKING VAN MIMOSIS

Het inputbestand van het model kan gezien worden als een matrix met 299,198 lijnen en zeer veel kolommen. De kolommen worden ingevuld op basis van de gegevens uit de verschillende extern aangeleverde bestanden die in de vorige sectie besproken werden. De gegevens uit dit inputbestand vormen de exogene variabelen van het model. Deze exogene variabelen worden samen met een reeks parameters ingelezen door het model. De parameters zijn elementen uit de verschillende wetgevingsdomeinen die door het model behandeld worden en die flexibel aanpasbaar gemaakt werden voor de gebruiker van het model. In het geval van de werkloosheidsreglementering zijn dit onder andere de percentages en de boven- en ondergrenzen voor het laatst verdiende loon die, gedifferentieerd naar gezinstype, worden toegepast voor de berekening van de werkloosheidsuitkering. De combinatie van een aantal exogene gegevens, zoals een typologie van de werkloze en zijn laatst verdiende loon, en de parameters laat dan toe om binnen het model de werkloosheidsuitkering te berekenen. Omdat de parameters eenvoudig kunnen aangepast worden, kunnen verschillende rekenscenario's relatief vlot gesimuleerd worden.

De broncode voor de verwerking van de exogene gegevens en de parameters is geschreven in FORTRAN. Dit in tegenstelling tot de broncode die gebruikt wordt voor de omzetting van de gegevens uit het DWH AM&SB naar een inputbestand voor het model. De aanmaak van dit inputbestand is volledig in SAS geprogrammeerd.

Om de programmering van het model overzichtelijk te houden, werd de broncode geprogrammeerd in verschillende modules. We kunnen in deze sectie uiteraard niet ingaan op het detail van iedere module.⁴³ In deze sectie trachten we wel een beknopt beeld te geven van wat iedere module doet en welke output iedere module genereert. In Tabel 15 geven we een lijst van de 9 modulenaamen die samen het MIMOSIS-model vormen en een beknopte omschrijving van wat ze doen.

Tabel 15: MIMOSIS-modulenaamen en korte omschrijving

Module naam	Beknopte omschrijving
FAMREL	Afbakening van families (i.e. kleinere eenheden dan het sociologisch gezin) en bepaling van relaties tussen gezinsleden
PENSWELF	Welvaartsaanpassingen aan pensioenen maar geen herberekening van pensioenen volgens de pensioenrekenregels
FAMAL	Berekening van gezinsbijslagen
CONTRIB	Berekening van bijdragen op inkomens uit arbeid en inhoudingen op vervangingsinkomens
PIT	Berekening van personenbelasting betaald per fiscaal gezin
SICK	Berekening van de uitkeringen ziekte, invaliditeit, arbeidsongeval en beroepsziekte
UNEM	Berekening van werkloosheidsuitkeringen, tijdskrediet of loopbaanonderbreking en brugpensioenen
SOCBEN	Berekening uitkeringen sociale bijstand (leefloon, IGO en uitkeringen aan personen met een handicap)
Eval	Evaluatie van de gesimuleerde output (berekening van budgettaire aggregaten en verdelingsgegevens)

Binnen één volledige run van MIMOSIS lopen de meeste modules twee keer. Eén keer in een pre run en een tweede keer in een normale run. In de pre run worden files aangemaakt die door andere modules als input gebruikt worden in de normale run. Zo heeft de module CONTRIB (bijdragen en inhoudingen) bijvoorbeeld nood aan invaliditeitsuitkeringen en pensioenen om de daarop geldende inhoudingspercentages te kunnen toepassen. Daarom genereren de modules SICK (onder andere ziekte en invaliditeitsuitkeringen) en PENSWELF (welvaartsaanpassingen pensioenen) bepaalde outputvariabelen in de prerun die dan vervolgens in de normale run door de CONTRIB module als input gebruikt worden.

In totaal worden met de 9 opgesomde modules 17 operaties sequentieel uitgevoerd. De sequentie waarin de MIMOSIS-modules lopen om één volledige run uit te voeren, is opgenomen in Tabel 16.

⁴³ Voor meer details zie FOD Sociale Zekerheid (2011d).

Tabel 16: Loopsequentie van verschillende MIMOSIS-modules

	Pre run		Normale run
1	FAMREL		
2	PENSWELF	9	PENSWELF
3	FAMAL	10	FAMAL
4	CONTRIB	11	SICK
5	PIT	12	UNEM
6	SICK	13	CONTRIB
7	UNEM	14	PIT
8	SOCBEN	16	SOBEN
		17	EVAL

Het principe bij de schikking van de verschillende modules in het model is dat eerdere runs van een module input leveren voor volgende modules. In de mate van het mogelijke wordt de input die een bepaalde module nodig heeft, gegenereerd in de normale run, maar indien dat niet kan wordt de input gegenereerd in de pre run. De input die de pre run files nodig hebben wordt ingelezen vanuit externe files. In principe is er geen reden om de werking van het model te beperken tot een pre run en een normale run. Ook voor de pre run, zou men bijvoorbeeld een pre pre run kunnen uitvoeren. Deze beperking tot twee grote stappen is echter vooral ingegeven om de looptijd van het model te beperken.⁴⁴

In de rest van deze sectie gaan we nu voor de verschillende modules, behalve de EVAL module, in op de werking van de module en de gegenereerde output.

3.1 FAMREL

De FAMREL module verwerkt voornamelijk gegevens uit het Rijksregister met het oog op de latere bepaling van relaties tussen gezinsleden.⁴⁵ In tegenstelling tot de andere modules van het model, verwerkt deze module ook niet echt regels die vervat zitten in de sociaal economische regelgeving. In feite is deze module vooral functioneel voor andere modules.

Zo wordt op basis van deze module onder andere aangeduid hoe verwantschap tot en met de derde graad tussen gezinsleden kan bepaald worden. Voor het gezinshoofd ligt dit, op basis van de variabele NAREGNIS_RELATION, min of meer voor de hand. Voor andere leden van het gezin dient voor de bepaling van deze verwantschapsgraad een conversie te gebeuren van deze geobserveerde relatievevariabele op basis van een soort conversiematrix.

Verder worden in de FAMREL module ook families afgebakend. Dit zijn kleinere eenheden dan het sociologisch gezin. Deze kleinere gezinsverbanden zijn voor bepaalde toepassingen, zoals de constructie van fiscale eenheden, een meer relevant uitgangspunt dan het ruimere concept "sociologisch gezin". Het aantal families dat op basis van deze module bepaald wordt, is in vergelijking met het aantal sociologische gezinnen echter beperkt. In totaal

⁴⁴ De huidige looptijd van een volledige run van het model op een PC met volgende eigenschappen: CPU 2.40 GHz en 2.0 GB RAM bedraagt ongeveer 20 minuten.

⁴⁵ Een beschrijving van deze module wordt gegeven in Decoster e.a. (2007b). De meest recente versie van de broncode maakt deel uit van FOD Sociale Zekerheid (2011d).

observeren we, gewogen met de steekproefgewichten voor gezinnen, 4,577,973 sociologische gezinnen en 4,606,377 families.

We hebben geen externe bronnen om de output op het niveau van de familiekeren mee te vergelijken. Om een idee te geven van de output die door de FAMREL module geproduceerd wordt, geven we in Tabel 17 toch enkele resultaten gegenereerd door deze module. De resultaten die gerapporteerd worden in deze tabel hebben betrekking op a) het familietype waartoe een individu behoort, b) de positie van de gezinsleden binnen de familiekeren en c) een typering van de koppels die hetzij gehuwd zijn of een de facto koppel vormen.

Tabel 17: Positie individuen binnen familiekeren (output van de FAMREL module)

Omschrijving	Aantal	%
Individu behoort tot		
familie van het gezinshoofd	10,676,659	99.3
familie van kinderen van het gezinshoofd	62,306	0.6
familie van ouders van het gezinshoofd of diens partner	10,888	0.1
familie van grootouders van het gezinshoofd of diens partner	31	0.0
familie van broer of zus van het gezinshoofd of diens partner	1,378	0.0
familie van kleinkinderen van het gezinshoofd of diens partner	1,752	0.0
familie van nonkel of tante van het gezinshoofd of diens partner	66	0.0
Totaal individuen	10,753,080	100.0
Positie binnen de familie		
Gezinshoofd	4,804,976	44.7
Partner van het gezinshoofd	2,584,767	24.0
Kind van het gezinshoofd of diens partner	3,206,207	29.8
Ascendent in de eerste graad van het gezinshoofd of diens partner	41,432	0.4
Ascendent in de tweede graad van het gezinshoofd of diens partner	865	0.0
Verwanten tot en met de derde graad van het gezinshoofd of diens partner	49,371	0.5
Andere familieleden	65,462	0.6
Waarde onbekende	0	0.0
Totaal individuen	10,753,080	100.0
Type relatievorm voor gehuwde of de facto koppels		
Individu behoort tot		
Gehuwd koppel: partners hebben verschillend geslacht	4,211,386	81.5
Gehuwd koppel: partners zijn beide mannen	4,513	0.1
Gehuwd koppel: partners zijn beide vrouwen	4,823	0.1
De facto koppel: partners hebben verschillend geslacht	897,299	17.4
De facto koppel: partners zijn beide mannen	28,729	0.6
De facto koppel: partners zijn beide vrouwen	22,785	0.4
Totaal individuen	5,169,534	100.0

3.2 UNEM

In de UNEM module worden de werkloosheidsuitkeringen en een aantal andere vergoedingen, ten laste van de RVA, gemodelleerd.⁴⁶ Dit gebeurt door enerzijds een aantal statusvariabelen van de RVA en het laatst verdiende loon te combineren met elementen uit de wetgeving zoals de berekeningspercentages voor werkloosheidsuitkeringen en de bijhorende onder- en bovengrenzen gedifferentieerd naar gezinstype.

Ook de typering van de gezinslast is in deze module geflexibiliseerd. Indien de inkomenstoestand van één van de gezinsleden die samenwonen met de werkloze zou veranderen, zou dit dus een weerslag kunnen hebben op het vergoedingsstatuut van de werkloze.

In een aantal uitzonderingsgevallen, zoals bij een aantal activeringsmaatregelen, worden de regels niet geïmplementeerd in deze module maar worden de werkelijk uitbetaalde bedragen gebruikt.

We kunnen de simulatieresultaten vergelijken met gegevens uit externe bronnen, maar omdat we ook de individueel uitbetaalde bedragen zelf observeren, kunnen we de simulatieresultaten ook daar mee vergelijken. We doen dit in Tabel 34 door de individueel gesimuleerde bedragen te aggregeren over bepaalde categorieën.

Tabel 18: Uitgaven werkloosheid, bruggpensioen en loopbaanonderbreking/tijdscrediet in 2008 in 1,000 Euro op basis van externe bron, simulaties UNEM en registraties in de steekproef⁴⁷

	Externe bron	Gesimuleerde bedragen	Geobserveerde bedragen in de steekproef
Werkloosheid			
Volledige werkloosheid	5,425,924	4,805,130	5,022,082
Tijdelijke werkloosheid	430,638	460,744	473,886
Activering van de werkloosheid	310,153	578,766	635,327
Bruggpensioen			
Conventioneel bruggpensioen	1,435,299	1,536,843	1,515,158
Loopbaanonderbreking en tijdscrediet	703,674	708,396	726,475
Totaal	8,305,688	8,089,880	8,372,928

De resultaten in Tabel 18 leren dat er aanzienlijke afwijkingen zijn tussen de gegevens uit de externe bron en de gesimuleerde bedragen. Vermoedelijk is dit echter meer een gevolg van classificatieproblemen (i.e. het onderbrengen van de gesimuleerde uitkering in de juiste vergelijkingsgroep) dan van de manier waarop de berekening van de uitkeringen gemodelleerd werd. De aansluiting tussen de gesimuleerde bedragen en de werkelijk geobserveerde bedragen is in alle geval beter dan deze met de gegevens uit de externe bron.

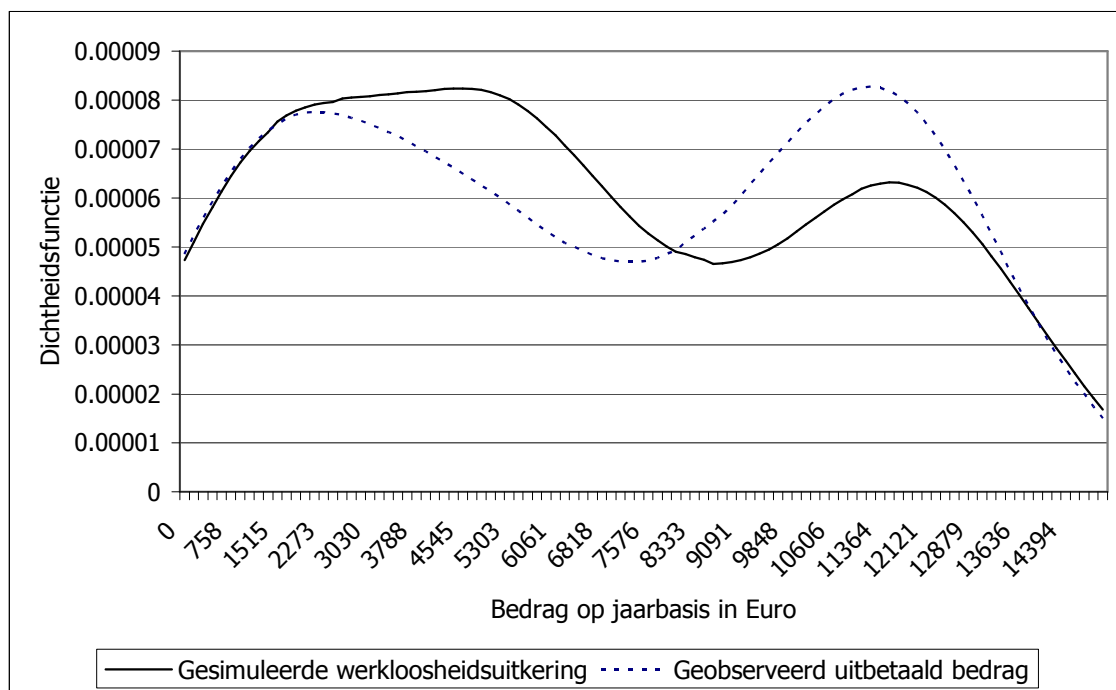
⁴⁶ De basisprincipes van deze module worden toegelicht in Rombaut e.a. (2007a). De meest recente versie van de broncode maakt deel uit van FOD Sociale Zekerheid (2011d).

⁴⁷ De gegevens van de externe bron zijn afkomstig uit FOD Sociale Zekerheid (2011a), p 188 (kolom 2008).

Toch is er ook met de externe bron een totaal verschil van ongeveer 300 miljoen Euro. Dit is vooral een gevolg van een onderschatting van de som van de gesimuleerde uitkeringen voor volledige werkloosheid. Er dient in dit verband wel opgemerkt te worden dat simulaties hier, en ook voor de andere modules, telkens worden uitgevoerd met parameters van 1 januari van een gegeven jaar, i.e. in dit geval 2008. De kolom met geobserveerde bedragen bevat dan weer uitbetalingen die gedurende heel het jaar 2008 gebeurd zijn. Het is mogelijk dat een aantal van deze betalingen geregistreerd gedurende 2008 gebaseerd zijn op basistarieven die in vergelijking met de simulatieparameters van 1 januari aangepast zijn als gevolg van een indexaanpassing of een aanpassing buiten index.

In het algemeen beschouwen we de aansluiting tussen de gesimuleerde bedragen enerzijds en de geregistreerde bedragen voor de in Tabel 18 gerapporteerde aggregaten als voldoende betrouwbaar om de module verder te gebruiken. De in Tabel 18 gerapporteerde aggregaten kunnen natuurlijk wel nog altijd individuele afwijkingen maskeren. Daarom tekenen we in Figuur 7 de dichtheidsfunctie van de gesimuleerde jaarbedragen en de werkelijk uitbetaalde bedragen voor UVW's.

Figuur 7: Dichtheidsfuncties van de gesimuleerde en geobserveerde werkloosheidsuitkeringen op jaarbasis voor uitkeringsgerechtigde volledig werklozen



Figuur 7 illustreert dat de UNEM module een groter aantal gevallen met lage werkloosheidsuitkeringen genereert en een lager aantal met een hoge uitkeringen, in vergelijking met de werkelijk geregistreerde bedragen.

Vermits de uitkeringen, waarvan de dichtheid getekend wordt in Figuur 7, uitbetalingen op jaarbasis zijn, kunnen de geobserveerde afwijkingen zowel een gevolg zijn van een verschillend aantal dagen gedurende dewelke een daguitkering wordt uitbetaald als van een verschil in de daguitkering zelf. Bij de daguitkering kunnen afwijkingen zowel een gevolg zijn van verschillen in het verloren loon als van verschillen in het gezinsstatuut van de werkloze.

De precieze oorzaak van deze afwijkingen wordt voorlopig niet verder onderzocht in deze tekst.

3.3 SICK

Met de module SICK worden uitkeringen voor ziekte- en invaliditeit, arbeidsongevallen en beroepsziekten gesimuleerd.⁴⁸ De rekenregels voor een aantal uitkeringen werden in deze module reeds geprogrammeerd maar de identificatie van het aantal dagen gedurende dewelke de verschillende uitkeringen worden uitbetaald, is onvoldoende uitgewerkt om de uitkeringen op jaarbasis betrouwbaar te ramen. Daarom werd voorlopig gefocust op de reconstructie van een aantal statuten en de bijhorende verwerking van de geobserveerde uitgekeerde bedragen.

In Tabel 19 rapporteren we het totaal aantal rechthebbenden op een aantal uitkeringen op jaarbasis en, indien beschikbaar, de gegevens die we voor een vergelijkbare groep terugvinden in externe bronnen.

Tabel 19: Rechthebbenden uitkering ziekte, invaliditeit, arbeidsongeval en beroepsziekte naar type uitkering op basis van externe bron en SICK module⁴⁹

	Externe bron	Gesimuleerd aantal	Procentuele afwijking gesimuleerd aantal en externe bron
Algemene regeling: primaire arbeidsongeschiktheid, arbeiders	NA	291,748	
Algemene regeling: primaire arbeidsongeschiktheid, bedienden	NA	113,138	
Algemene regeling: invaliditeit, arbeiders	172,565	181,701	5.3
Algemene regeling: invaliditeit, bedienden	59,588	60,158	1.0
Algemene regeling: moederschapsrust, primaire periode	NA	164,493	
Algemene regeling: moederschapsrust, invaliditeit	NA	1,266	
Regeling zelfstandigen: primaire arbeidsongeschiktheid	NA	16,489	
Regeling zelfstandigen: invaliditeit	18,552	20,304	9.4
Regeling zelfstandigen: moederschapsrust, primaire periode	NA	5,279	
Regeling zelfstandigen: moederschapsrust, invaliditeit	NA	0	
Vergoedingen tijdelijke arbeidsongeschiktheid bij arbeidsongeval	NA	98,439	
Vergoedingen permanente arbeidsongeschiktheid bij arbeidsongeval	NA	8,278	
Vergoedingen tijdelijke beroepsziekte	NA	57	
Vergoedingen permanente beroepsziekte	NA	55,909	

Uit Tabel 19 blijkt dat het aantal statuten waarvoor we vergelijkingsmateriaal hebben inzake het totaal aantal rechthebbenden, zeer beperkt is. Enkel voor werknemers en zelfstandigen

⁴⁸ De basisprincipes van deze module worden toegelicht in Rombaut e.a. (2007b). De meest recente versie van de broncode maakt deel uit van FOD Sociale Zekerheid (2011d).

⁴⁹ De gegevens van de externe bron zijn afkomstig uit FOD Sociale Zekerheid (2011a) p. 139 en 250 (kolom 2008).

in invaliditeit hebben we externe cijfers die in principe overeen zouden moeten komen met de groep die in de SICK module gereconstrueerd wordt.

De omvang van de subgroepen van individuen met een invaliditeitsuitkering wordt door de SICK module in vergelijking met de gebruikte externe bron, telkens overschat. Eén verklaring hiervoor zou kunnen zijn dat in de SICK module het gecumuleerd totaal aantal gevallen op jaarbasis geïdentificeerd wordt, terwijl de externe bron enkel de gevallen geeft op één moment in de tijd (i.e. 31 december 2008).

De bijhorende gereconstrueerde uitgaven op jaarbasis worden voor de verschillende statuten getoond in Tabel 20.

Tabel 20: Uitgaven ziekte, invaliditeit, arbeidsongeval en beroepsziekte naar type uitkering op basis van externe bron en SICK module in 1,000 Euro⁵⁰

	Externe bron	Output SICK	Procentuele afwijking bedrag SICK en externe bron
Algemene regeling: primaire arbeidsongeschiktheid, arbeiders	761,431	801,474	5.3
Algemene regeling: primaire arbeidsongeschiktheid, bedienden	346,219	310,970	-10.2
Algemene regeling: invaliditeit, arbeiders	1,914,635	1,990,606	4.0
Algemene regeling: invaliditeit, bedienden	707,159	649,563	-8.1
Algemene regeling: moederschapsrust, primaire periode	492,409	476,740	-3.2
Algemene regeling: moederschapsrust, invaliditeit	3,760	4,265	13.4
Regeling zelfstandigen: primaire arbeidsongeschiktheid	56,716	54,435	-4.0
Regeling zelfstandigen: invaliditeit	206,508	207,093	0.3
Regeling zelfstandigen: moederschapsrust, primaire periode	14,979	15,838	5.7
Regeling zelfstandigen: moederschapsrust, invaliditeit	36	0	100.0
Vergoedingen tijdelijke arbeidsongeschiktheid bij arbeidsongeval	NA	137,713	
Vergoedingen permanente arbeidsongeschiktheid bij arbeidsongeval	NA	78,991	
Vergoedingen tijdelijke beroepsziekte	NA	1,284	
Vergoedingen permanente beroepsziekte	NA	157,550	

De vergelijking van de uitgaven op jaarbasis zou minder dan de vergelijking van de statuten mogen lijden onder een vergelijkbaarheidsprobleem omdat zowel in de externe bron als in de SICK module, aggregaten op jaarbasis geproduceerd worden.

Uit Tabel 20 blijkt dat er geen uniform patroon van over- of onderschatting van de deelbudgetten is. Wel valt op dat er bij de algemene regeling vermoedelijk een indelingsprobleem is bij het toewijzen van het statuut arbeider en bediende bij de regeling primaire arbeidsongeschiktheid en de invaliditeitsregeling. Indien de bedragen primaire

⁵⁰ De gegevens van de externe bron zijn afkomstig uit FOD Sociale Zekerheid (2011a) p. 138, 140 en 252 (kolom 2008).

arbeidsongeschiktheid en invaliditeit voor arbeiders en bedienden worden opgeteld, wordt het totale budget op jaarbasis wel vrij goed benaderd.

Ook de andere gereconstrueerde deelbudgetten, waarvoor we over vergelijkingsmateriaal beschikken, sluiten nauw aan bij het extern gegeven, met uitzondering van de uitkeringen materniteit gedurende de periode van invaliditeit bij zelfstandigen. Er zitten in de steekproef geen gevallen die aan dit criterium voldoen. Gegeven de omvang van het feitelijke budget, hoeft dit ook niet echt te verwonderen.

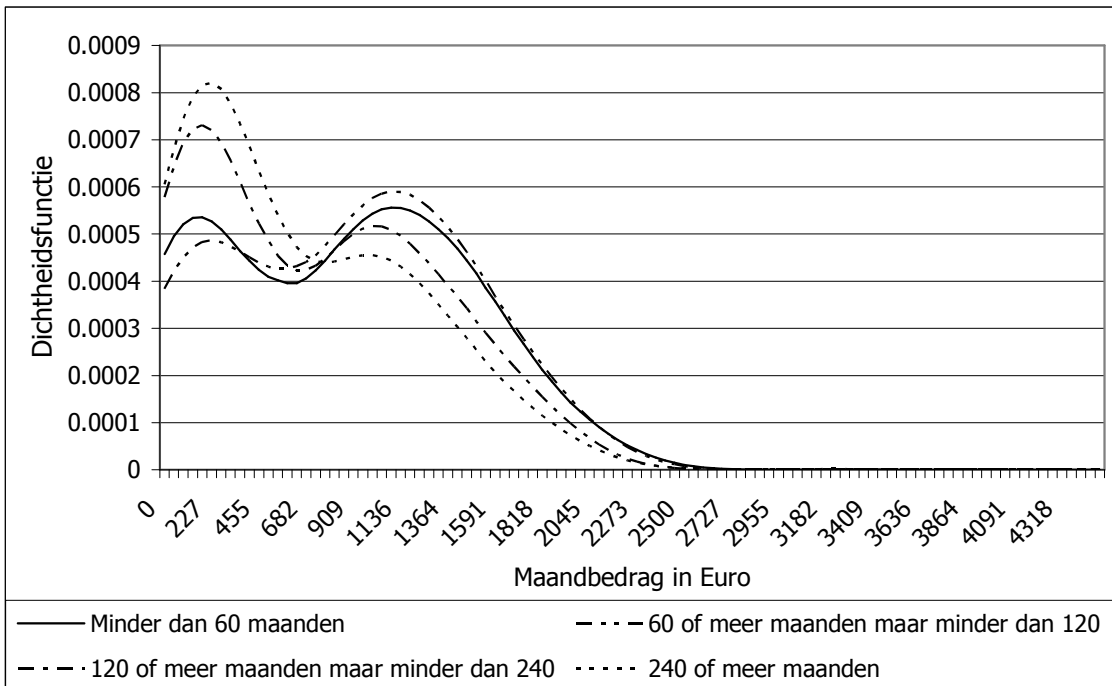
3.4 PENSWELF

Met de PENSWELF module kunnen uitbetaalde pensioenen aangepast worden.⁵¹ Dit kan gebeuren door er absolute bedragen aan toe te voegen of van af te trekken of door het basisbedrag in procentuele termen aan te passen. Deze aanpassingen kunnen bovendien gedifferentieerd worden in functie van het type pensioen en in functie van de "leeftijd" van het pensioen. In vergelijking met de andere modules die regelgevingsdomeinen behandelen, is de PENSWELF module, atypisch door het feit dat ze geobserveerde basisbedragen als gegeven beschouwt, terwijl voor andere domeinen basisbedragen in principe herberekend worden door de rekenregels toe te passen op de relevante variabelen. Het heeft voor deze module dan ook weinig zin om het basismateriaal te vergelijken met de gegevens in externe bronnen.

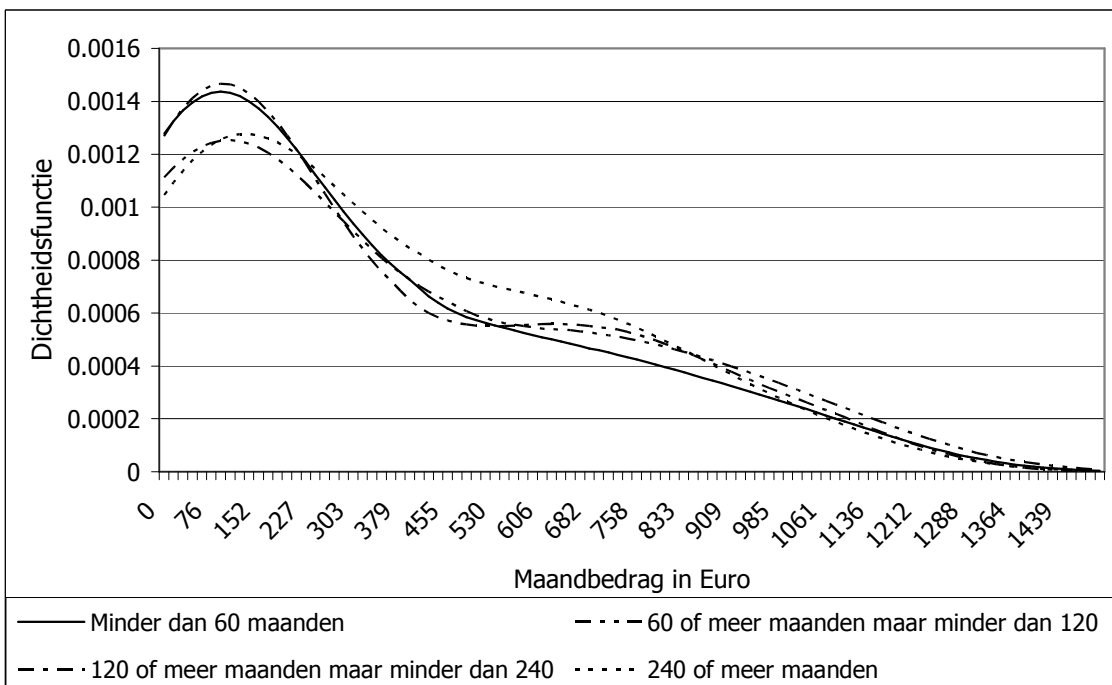
Wel kunnen we een idee geven van de basisverdelingen waarop de absolute of procentuele aanpassingen zouden kunnen toegepast worden. Dit doen we door in Figuur 8, Figuur 9 en Figuur 10 respectievelijk de dichtheidsfuncties te tekenen van de rustpensioenen van werknemers, zelfstandigen en ambtenaren, gedifferentieerd naar het aantal maanden dat de bedragen reeds betaald worden.

⁵¹ De basisprincipes van deze module worden toegelicht in Maréchal e.a. (2007). De meest recente versie van de broncode maakt deel uit van FOD Sociale Zekerheid (2011d).

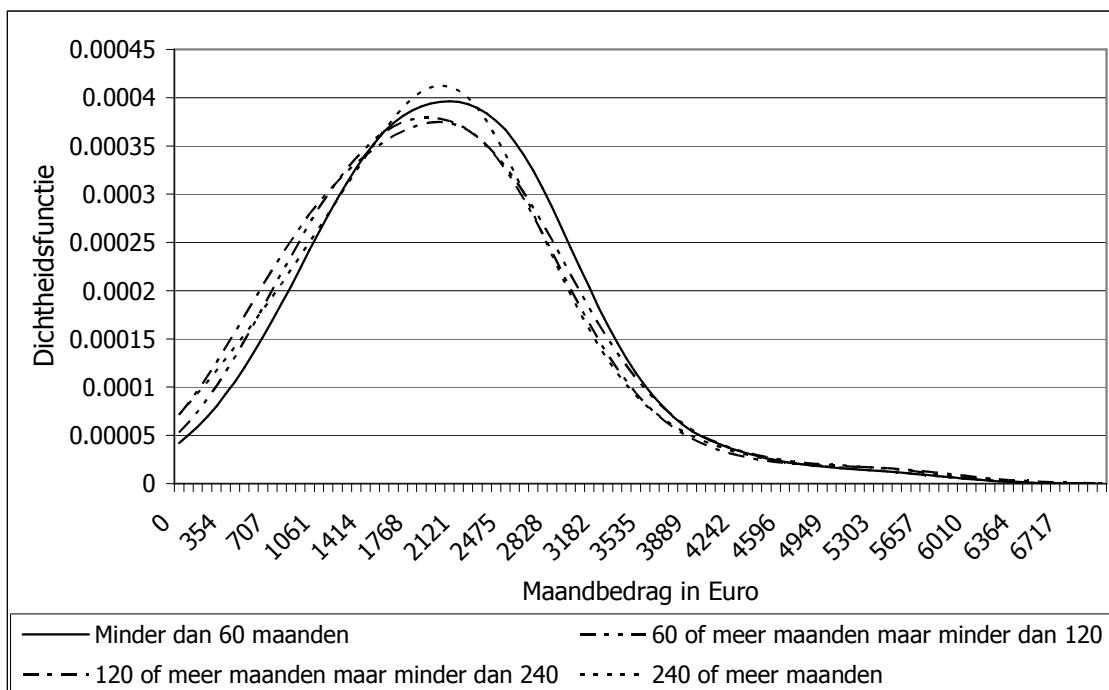
Figuur 8: Dichtheidsfuncties van de maandelijkse rustpensioenen uitbetaald in het stelsel der werknemers naar aantal maanden sinds de eerste uitbetaling van een pensioen



Figuur 9: Dichtheidsfuncties van de maandelijkse rustpensioenen uitbetaald in het stelsel der zelfstandigen naar aantal maanden sinds de eerste uitbetaling van een pensioen



Figuur 10: Dichtheidsfuncties van de maandelijkse rustpensioenen uitbetaald in het stelsel der ambtenaren naar aantal maanden sinds de eerste uitbetaling van een pensioen



De resultaten in Figuur 8, Figuur 9 en Figuur 10 suggereren alvast dat enkel in het stelsel der werknemers de verdeling van de betaalde maandbedragen aanzienlijk varieert in functie van het aantal maanden dat verstreken is sinds de eerste uitbetaling van het pensioenbedrag. Bij de werknemers neemt de concentratie lage bedragen toe, naarmate het moment van eerste uitbetaling verder terug ligt in de tijd.

3.5 FAMAL

Met de FAMAL module wordt bepaald welke kinderen in aanmerking komen voor kinderbijslag, wie voor deze kinderen respectievelijk de bijslagtrekkende en de rechthebbende is, wat de rang van ieder kind is en welke statuten eventueel een recht op een sociaal of een éénouderssupplement openen.⁵² Voor de berekening van de bijslagen wordt in de module FAMAL een onderscheid gemaakt tussen 4 stelsels: de stelsels der werknemers, zelfstandigen en ambtenaren en het stelsel van de gewaarborgde kinderbijslag.

Voor het stelsel der werknemers en zelfstandigen hebben we in externe bronnen gegevens over de verdeling naar schaal en leeftijd van de rechthebbende kinderen. We geven deze externe gegevens en deze die geproduceerd worden op basis van de FAMAL module in Tabel 21 en Tabel 22.

⁵² De basisprincipes van deze module worden toegelicht in Bay e.a. (2007b). De meest recente versie van de broncode maakt deel uit van FOD Sociale Zekerheid (2011d).

Tabel 21: Aantal rechtgevendende kinderen naar schaal en leeftijdsklasse in het stelsel der werknemers op basis van een externe bron en FAMAL module⁵³

	0-6 jaar	6-12 jaar	12-18 jaar	18-25 jaar	+ 25 jaar	Totaal
Externe bron						
Gewone schaal	478,228	442,491	435,407	234,807	16,636	1,607,569
Invalidenschaal	13,215	19,193	26,204	13,876	0	72,488
Wezenschaal	1,550	5,719	13,011	11,730	0	32,010
Schaal gepensioneerden en werklozen	57,072	58,938	59,545	29,954	0	205,509
Totaal	550,065	526,341	534,167	290,367	16,636	1,917,576
Simulatieresultaten						
Gewone schaal	448,806	461,547	450,699	238,255	13,734	1,613,042
Invalidenschaal	14,752	22,220	29,716	15,367	752	82,807
Wezenschaal	1,171	3,962	8,667	7,601	168	21,569
Schaal gepensioneerden en werklozen	56,906	64,901	61,781	33,448	312	217,348
Totaal	521,636	552,630	550,863	294,671	14,966	1,934,765

Tabel 22: Aantal rechtgevendende kinderen naar schaal en leeftijdsklasse in het stelsel der zelfstandigen op basis van een externe bron en FAMAL module⁵⁴

	0-6 jaar	6-12 jaar	12-18 jaar	18-25 jaar	+ 25 jaar	Totaal
Externe bron						
Gewone schaal	37,446	50,678	60,864	43,096	578	192,662
Invalidenschaal	182	359	675	640	66	1,922
Wezenschaal	213	895	2,222	2,675	3,435	9,440
Schaal gepensioneerden en werklozen	0	0	0	0	0	0
Totaal	37,841	51,932	63,761	46,411	4,079	204,024
Simulatieresultaten						
Gewone schaal	32,433	52,651	66,982	45,760	1,798	199,623
Invalidenschaal	308	522	1,202	939	0	2,971
Wezenschaal	131	332	899	1,465	36	2,862
Schaal gepensioneerden en werklozen	474	579	944	776	9	2,783
Totaal	33,347	54,083	70,027	48,939	1,842	208,239

Zowel in het stelsel der werknemers als in het stelsel der zelfstandigen wordt het aantal rechtgevendende kinderen door FAMAL lichtjes overschat (zie de kolom en de lijn Totaal van

⁵³ De gegevens van de externe bron zijn afkomstig uit FOD Sociale Zekerheid (2011a) p. 159 (tabel 30juni 2008).

⁵⁴ De gegevens van de externe bron zijn afkomstig uit FOD Sociale Zekerheid (2011a) p. 265 (tabel 30juni 2008). Hoewel het Rijksinstituut voor de sociale verzekeringen der zelfstandigen de gegevens met betrekking tot de pensioenschaal niet meer rapporteert, worden deze gegevens hier wel nog apart opgenomen omdat het tarief onderscheiden blijft in de wetgeving (zie Put (2008) p. 798).

Tabel 21 en Tabel 22). Deze overschatting van het aantal gevallen komt niet systematisch terug over de verschillende leeftijdsgroepen of schalen.

Het aantal gevallen in de gewone schaal wordt in beide gevallen door FAMAL wel overschat, maar ook hier lijken de afwijkingen tussen de externe bron en de output van de module, gegeven het totaal aantal gevallen dat zou moeten geproduceerd worden, relatief beperkt (zie de lijnen gewone schaal in Tabel 21 en Tabel 22).

Het aantal gevallen in de invalidenschaal en in de schaal van gepensioneerden en werklozen wordt in het stelsel der werknemers wel aanzienlijk overschat door FAMAL, wat uiteraard gevolgen heeft bij de latere berekening van de sociale supplementen.

Tenslotte valt op dat het aantal rechtgevende kinderen met wezenschaal in beide stelsels aanzienlijk onderschat wordt. Deze onderschatting komt systematisch terug bij alle leeftijdsgroepen (zie de lijnen wezenschaal in Tabel 21 en Tabel 22). Dit is mogelijk een direct gevolg van problemen bij de constructie van de variabele NAREGNIS_CIVIL (zie Tabel 5) aangezien het wezenstatuut in de module FAMAL wordt afgeleid van het weduwnaarschap van de overlevende ouder.

Voor ambtenaren beschikken we niet over externe gegevens naar schaal per leeftijdsgroep. In Tabel 23 rapporteren we enkel het totaal aantal rechtgevende kinderen in het stelsel der ambtenaren en de onderverdeling naar de schaal waarin de kinderen geklasseerd worden.

Tabel 23: Aantal rechtgevende kinderen naar schaal in het stelsel der ambtenaren op basis van een externe bron en FAMAL module⁵⁵

	Externe bron	Simulatieresultaten
Gewone schaal	427,124	414,379
Invalidenschaal	7,870	18,533
Wezenschaal	7,418	6,654
Schaal gepensioneerden en werklozen	28,923	13,933
Totaal	471,335	453,500

Het totaal aantal rechtgevende kinderen in het stelsel der ambtenaren wordt onderschat. Dit lijkt een direct gevolg van een tekort aan gevallen in de gewone schaal, want het totaal aantal gevallen in de andere schalen ligt op basis van de FAMAL gegevens dicht bij de externe cijfers (tel andere lijnen dan de gewone schaal in Tabel 23 op voor beide kolommen). Op niveau van de schalen onderling is er evenwel een aanzienlijke overschatting van het aantal gevallen in de invalidenschaal en een onderschatting van de rechthebbenden in de schaal der gepensioneerden en werklozen.

Het totaal aantal rechtgevende kinderen in het stelsel van de gewaarborgde kinderbijslag bedraagt 14,675 eenheden. Op basis van de FAMAL module worden er voor 2008, 18,802 eenheden in dit stelsel geklasseerd.

Vermoedelijk overschatten we ook het totaal aantal gevallen met recht op een éénouderssupplement. Op basis van de FAMAL module wordt dit recht aan 134,311 kinderen toegewezen. We beschikken enkel over externe gegevens omtrent dit recht voor gevallen in

⁵⁵ De gegevens van de externe bron zijn afkomstig uit FOD Sociale Zekerheid (2011a) p. 382 (tabel met toestand op 31 december 2009).

het stelsel der werknemers. In 2008 hadden 94,908 kinderen recht op een éénouderssupplement.⁵⁶ Vermoedelijk is het aantal kinderen dat dit recht opent in het stelsel der zelfstandigen, kleiner dan 40,000.

Tenslotte vergelijken we in Tabel 24 ook de totale uitgaven per stelsel die we bekomen op basis van de FAMAL module met deze gerapporteerd in een externe bron. In tegenstelling tot de gangbare praktijk in de rest van de tekst, rapporteren we in de kolom Externe bron gegevens van het budgettaire jaar 2007 omdat de meeste parameters in het kinderbijslagsysteem doorheen 2007 onveranderd gebleven zijn en er in 2008 een aantal aanpassingen van de bedragen waren. Vermits we simuleren met parameters van 1 januari 2008 veronderstellen we dat de gesimuleerde kost het best aansluit bij de uitgaven van 2007.

Tabel 24: Uitgaven kinderbijslag naar stelsel op basis van externe bron en FAMAL module in 1,000 Euro⁵⁷

	Externe bron	Gesimuleerde bedragen	Procentuele afwijking gesimuleerd bedrag en externe bron
Werknemers	3,693,798	3,674,163	-0.5
Zelfstandigen	355,031	342,131	-3.6
Ambtenaren	875,000	832,284	-4.9
Gewaarborgde gezinsbijslag	32,842	47,565	44.8

Ondanks de beperkte overschatting van het aantal rechtgevende kinderen in de stelsels der werknemers en zelfstandigen en een overschatting van het aantal rechtgevendenden op de belangrijkste sociale supplementen in het stelsel der werknemers, wordt het totale budget in beide gevallen toch lichtjes onderschat. Mogelijke oorzaken voor deze vaststellingen moeten nog verder onderzocht worden.

De onderschatting van het budget uitbetaald aan ambtenaren en de overschatting van het budget in de gewaarborgde kinderbijslag, ligt in de lijn der verwachtingen gegeven de respectievelijke onder- en overschatting van het totaal aantal rechtgevende kinderen in deze stelsels.

3.6 CONTRIB

Met de module CONTRIB worden de bijdragen en bijdrageverminderingen op inkomens uit arbeid gemodelleerd en de inhoudingen die worden ingehouden op een aantal vervangingsinkomens.⁵⁸

Voor wat de arbeidsmarktinkomens betreft, is het bijdrageregime nauw verbonden met het arbeidsmarktstatuut. In Tabel 25 geven we de aantallen werkenden naar de statuten die op basis van de CONTRIB module aangemaakt worden en deze die we observeren in een externe bron.

⁵⁶ Zie RKW (2011), p. 26.

⁵⁷ De gegevens van de externe bron zijn afkomstig uit FOD Sociale Zekerheid (2011a) p. 161 (kolom 2007), p. 267 (kolom 2007) en p. 381 (kolom 2006)

⁵⁸ De basisprincipes van deze module worden toegelicht in Bay e.a. (2007a). De meest recente versie van de broncode maakt deel uit van FOD Sociale Zekerheid (2011d).

Tabel 25: Aantal werkenden naar statuut op basis van externe bron en CONTRIB module⁵⁹

	Externe bron		Gegevens CONTRIB	
	Privé	Overheid	Privé	Overheid
Arbeiders	NA	NA	1,462,803	154,747
Bedienden	NA	NA	1,936,307	381,638
Ambtenaren	NA	NA		651,831
Zelfstandigen	923,946		921,752	

Zoals blijkt uit Tabel 25 zijn er voor de in de CONTRIB module gereconstrueerde aggregaten zeer weinig vergelijkingsbronnen. De RSZ publiceert bijvoorbeeld statusgegevens op het einde van het kwartaal maar niet de jaarcumul van alle individuen die doorheen een jaar in een gegeven statuut geklasseerd konden worden.

Vergelijking van de CONTRIB resultaten in de eerste drie lijnen met statusgegevens op het einde van het kwartaal, leert alvast dat de CONTRIB aggregaten aanzienlijk hoger liggen. De omgekeerde oefening, i.e. de aanmaak van aggregaten op basis van de individuen met een job gekend op het einde van het kwartaal, zou ook kunnen gebeuren met de beschikbare gegevens, maar dit is voorlopig nog niet gebeurd.

De enige groep waarvoor een vergelijkbaar gegeven bestaat zijn de zelfstandigen. De totale omvang van deze groep blijkt op basis van de CONTRIB raming goed benaderd te worden.

De aan bijdragen onderworpen massa's van de verschillende groepen werkenden, worden gerapporteerd in Tabel 26.

Tabel 26: Aan bijdragen onderworpen massa op basis van externe bron en CONTRIB module in 1,000 Euro⁶⁰

	Externe bron	Output CONTRIB	Procentuele afwijking bedrag CONTRIB en externe bron
Loontrekkenden private sector	76,771,400	77,301,197	0.7
Loontrekkenden publieke sector	33,256,900	34,014,882	2.3
Zelfstandigen	14,947,831	14,661,613	-1.9

De aan bijdragen onderworpen massa's van de verschillende subgroepen blijken alvast goed benaderd te worden met de beschikbare gegevens.

Het is in eerste instantie de bedoeling om bij later gebruik van de inkomenscomponenten te beschikken over bruto inkomens min persoonlijke bijdragen of inhoudingen om hierop vervolgens de relevante regels voor de bepaling van de personenbelasting te kunnen toepassen. Vermits de basis voor de bepaling van de bijdragen en personenbelasting van zelfstandigen dezelfde is⁶¹, focussen we hier voorlopig op de raming van de persoonlijke bijdragen geïnd op inkomens van loontrekkenden en de vermindering van deze bijdragen.

⁵⁹ De gegevens van de externe bron zijn afkomstig uit FOD Sociale Zekerheid (2011a) p. 242 (tabel met toestand op 31 december 2008).

⁶⁰ De gegevens van de externe bron zijn afkomstig uit FOD Sociale Zekerheid (2011a) p. 125 en p. 244 (kolom 2008).

⁶¹ Zie Knapen (2011) voor dit punt.

In Tabel 27 rapporteren we de geraamde bijdragen en bijdrageverminderingen en gegevens die ermee vergelijkbaar zouden moeten zijn op basis van een externe bron.

Tabel 27: Persoonlijke bijdragen en bijdrageverminderingen op basis van externe bron en CONTRIB module in 1,000 Euro⁶²

	Externe bron	Output CONTRIB	Procentuele afwijking bedrag CONTRIB en externe bron
Persoonlijke bijdragen voor vermindering	12,586,858	12,080,776	-4.0
Bedrag Werkbonus	698,863	754,754	8.0
Aantal voltijdsequivalenten werkbonus	1,047,113	1,062,580	1.5

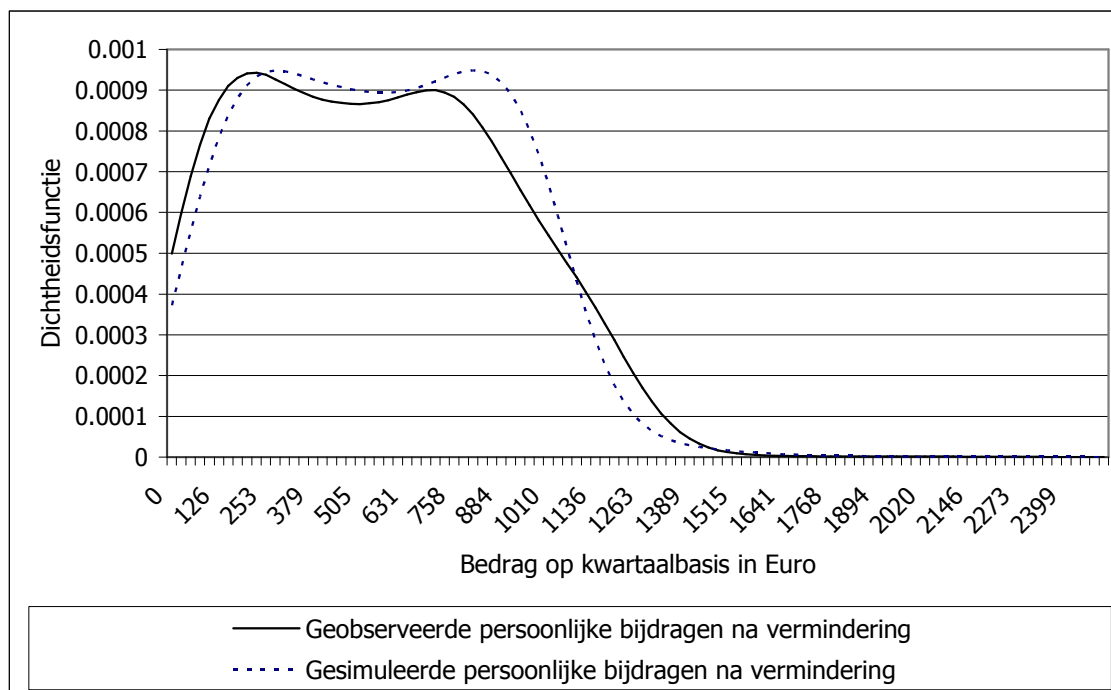
De bijdragen voor vermindering van de persoonlijke bijdragen worden onderschat.⁶³ De uitgaven aan de Werkbonus, die de mogelijke vermindering van de bijdragen omvat, worden dan weer overschat. Vermoedelijk is deze overschatting van de Werkbonus geen gevolg van een foute identificatie van de gevallen die er recht op hebben. Het gereconstrueerd en het werkelijk aantal voltijdse equivalent met een werkbonus-recht, liggen alvast dicht bij elkaar.

De rapportering over de aggregaten in Tabel 27 zegt uiteraard niets over de verdeling. Om te zien of de werkelijke en de gesimuleerde verdeling van de persoonlijke bijdragen na vermindering, van elkaar verschillen, teken we in Figuur 11 de dichtheidsfuncties van beide concepten voor werknemers uit het RSZ bestand waarvoor we een bijdragevermindering observeren.

⁶² De gegevens van de externe bron zijn afkomstig uit verschillende bronnen. Voor lijn "Persoonlijke bijdragen voor vermindering" zie FOD Sociale Zekerheid (2011a) p. 128 en p. 130. Voor lijnen "Bedrag Werkbonus" en "Aantal voltijdsequivalenten werkbonus" zie RSZ (2011), tabel 7a en 7b, Vermindering van de werknemersbijdragen met versterking van de koopkracht.

⁶³ Bij dit aggregaat uit de externe bron dient wel opgemerkt te worden dat nog verder onderzocht moet worden of het het meest geschikte vergelijkingsgegeven is.

Figuur 11: Dichtheidsfuncties van de gesimuleerde en geobserveerde persoonlijke bijdragen na bijdragevermindering (subgroep werknemers RSZ met bijdragevermindering)



Figuur 11 illustreert dat de verdeling van de werkelijke en de gesimuleerde bijdragen min bijdrageverminderingen een zeer gelijkaardig patroon kennen. Zoals te verwachten op basis van de aggregaten, zijn er meer lage gesimuleerde bedragen en minder hoge in vergelijking met de werkelijke bedragen.

3.7 PIT

Met de PIT module worden eerst fiscale eenheden gereconstrueerd, i.e. de eenheden die verondersteld worden een belastingaangifte in te dienen.⁶⁴ Vervolgens worden aan deze gereconstrueerde eenheden de geobserveerde belastbare inkomens (i.e. de bruto bedragen min betaalde persoonlijke bijdragen of inhoudingen) toegewezen. Voor iedere fiscale eenheid worden dan de verschuldigde personenbelastingen berekend, rekening houdend met de mogelijke verminderingen voor de samenstelling van het gezin, de compensaties voor personen met een handicap die ten laste zijn en de verminderingen voor vervangingsinkomens.

Het spreekt voor zich dat we op basis van de beschikbare gegevens niet heel het belastingsproces kunnen reconstrueren. Zo missen we een aantal inkomenscomponenten die het globaal of het afzonderlijk belastbaar inkomen verhogen. Componenten die we missen zijn het roerend en het onroerend inkomen en de diverse inkomsten. We missen ook alle bestedingen die mogen afgetrokken worden van het belastbaar inkomen en de bestedingen die eventueel een belastingkrediet opleveren. Voor eventuele kosten die verband houden

⁶⁴ De basisprincipes van deze module worden toegelicht in Decoster e.a. (2007a). De meest recente versie van de broncode maakt deel uit van FOD Sociale Zekerheid (2011d).

met het uitgeoefende beroep, wordt in de PIT module waar nodig een forfaitair schema toegepast.

Het gewogen aantal fiscale eenheden dat we na toepassing van deze regels bekomen is 5,973,954 eenheden. Dit is een onderschatting van het werkelijk aantal van 6,143,173 eenheden dat door ADSEI gerapporteerd wordt voor het aanslagjaar 2009 (inkomen 2008).⁶⁵ Deze onderschatting zou een gevolg kunnen zijn van het feit dat wij personen ten laste van een fiscale eenheid, niet afscheiden van deze eenheid als het inkomen niet hoog genoeg is om zelf belastingplichtig te worden. Concreet worden in het bijzonder kinderen met een eigen inkomen uit arbeid dus niet in de rekenregels opgenomen met een apart formulier indien het inkomen uit arbeid niet hoger is dan de drempel om nog kind ten laste te zijn.

De resultaten die we bekomen op basis van de PIT module kunnen niet op individueel niveau beoordeeld worden omdat er geen gegevens inzake personenbelastingen in onze steekproef zitten. In Tabel 28 geven we daarom een aantal aggregaten die door ADSEI gerapporteerd worden voor het aanslagjaar 2009 (inkomen 2008) en vervolgens de resultaten voor vergelijkbare concepten op basis van de PIT module.

⁶⁵ Zie ADSEI (2011b).

Tabel 28: Totaal belastbaar netto inkomen en belasting van het Aanslagjaar 2009 – België op basis van externe bron en PIT module (Decielenverdeling op basis van het totaal belastbaar netto-inkomen)⁶⁶

Decielen	Perctielen	Totaal bedrag belastbaar netto inkomen	In % van het totaal	Totaal bedrag belasting	In % van het totaal	Gemiddelde aanslagvoet (in %)
Externe bron						
1		881,090,524	0.5	13,646	0.0	0.0
2		5,185,634,199	3.2	72,588,052	0.2	1.4
3		7,491,804,651	4.7	161,670,093	0.4	2.2
4		9,190,184,310	5.6	708,429,848	1.8	7.7
5		11,257,736,991	6.9	1,459,065,907	3.7	13.0
6		13,558,257,618	8.3	2,535,803,998	6.5	18.7
7		16,276,426,069	9.9	3,520,213,331	9.0	21.6
8		20,367,817,827	12.4	5,018,103,088	12.8	24.6
9		27,493,259,624	16.8	7,718,778,101	19.7	28.1
	91	3,361,125,300	2.1	1,020,380,638	2.6	30.4
	92	3,520,440,662	2.1	1,090,161,732	2.8	31.0
	93	3,700,288,961	2.3	1,171,860,432	3.0	31.7
	94	3,904,114,254	2.4	1,265,583,744	3.2	32.4
	95	4,145,462,385	2.5	1,372,671,598	3.5	33.1
	96	4,447,023,528	2.7	1,510,664,604	3.9	34.0
	97	4,842,265,753	3.0	1,691,625,974	4.3	34.9
	98	5,417,005,578	3.3	1,951,061,372	5.0	36.0
	99	6,468,450,597	3.9	2,418,940,860	6.2	37.4
	100	12,650,775,385	7.7	4,423,280,218	11.3	35.0
10		52,456,952,401	32.0	17,916,231,173	45.8	34.2
Totaal		164,159,164,214	100.0	39,110,897,237	100.0	23.8
Simulatiere resultaten						
1		1,481,157,445	0.9	-78,381,271	-0.2	-5.3
2		6,015,172,194	3.7	56,799,800	0.1	0.9
3		7,710,496,535	4.7	203,224,593	0.5	2.6
4		9,393,800,038	5.7	818,103,518	2.0	8.7
5		11,728,615,659	7.2	1,864,432,608	4.5	15.9
6		14,028,393,487	8.6	2,916,179,877	7.0	20.8
7		16,683,328,420	10.2	4,118,625,210	9.9	24.7
8		20,525,635,276	12.6	5,592,131,103	13.5	27.2
9		27,445,514,717	16.8	8,155,868,551	19.6	29.7
	91	3,316,767,679	2.0	1,053,229,535	2.5	31.8
	92	3,453,291,998	2.1	1,115,829,454	2.7	32.3
	93	3,607,851,764	2.2	1,180,053,906	2.8	32.7
	94	3,789,911,459	2.3	1,261,876,021	3.0	33.3
	95	4,006,238,724	2.5	1,356,699,851	3.3	33.9
	96	4,287,523,526	2.6	1,493,248,292	3.6	34.8
	97	4,625,160,343	2.8	1,660,009,522	4.0	35.9
	98	5,139,204,547	3.1	1,908,638,226	4.6	37.1
	99	5,964,490,806	3.6	2,335,323,147	5.6	39.2
	100	10,294,527,980	6.3	4,505,272,672	10.9	43.8
10		48,484,968,826	29.7	17,870,180,626	43.0	36.9
Totaal		163,497,082,597	100.0	41,517,164,617	100.0	25.4

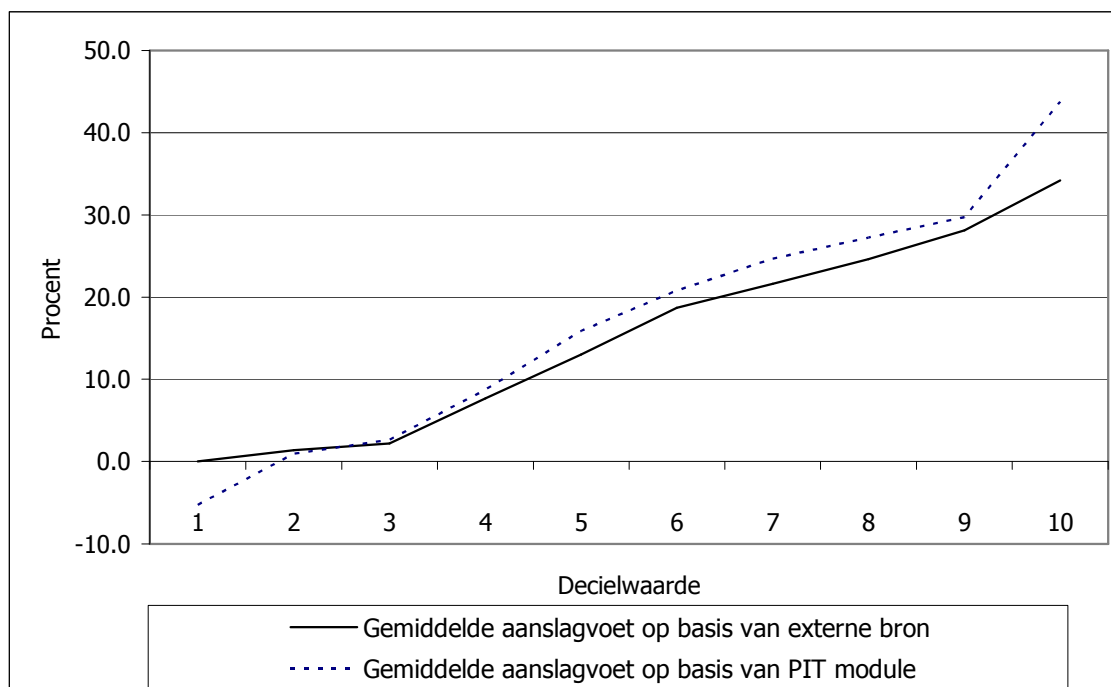
⁶⁶ De gegevens van de externe bron zijn afkomstig uit ADSEI (2011b), Tabel B.1: Decielenverdeling van het totaal netto belastbaar inkomen, de totale belasting en de gemiddelde aanslagvoet.

Tabel 28 leert dat we het totaal belastbaar inkomen onderschatten en de verschuldigde personenbelasting overschatten op basis van de PIT module (vergelijk de lijnen Totaal van de subtabellen Externe bron en Simulatieresultaten).

De onderschatting van het totaal belastbaar netto inkomen komt niet systematisch terug over alle decielen. Voor de eerste acht decielen observeren we een overschatting terwijl we voor het 9^{de} en het 10^{de} deciel een onderschatting observeren (zie kolom Totaal bedrag belastbaar netto inkomen in Tabel 28). Dit doet vermoeden dat in het bijzonder voor de hoogste decielen de ontbrekende belastbare inkomens belangrijker zijn dan de ontbrekende aftrekbare bestedingen. De overschatting van de verschuldigde personenbelasting komt dan weer systematisch terug voor alle decielen behalve voor de eerste twee en de laatste twee decielen (zie kolom Totaal bedrag belasting in Tabel 28).

De invloed van beide afwijkingen op de gemiddelde aanslagvoet is a priori niet duidelijk. Uit Figuur 12 blijkt dat deze gemiddelde aanslagvoet behalve voor de eerste twee decielen systematisch overschat wordt. Vanaf het vierde deciel ligt deze afwijking tussen de 1.0 en de 3.1 procentpunten. Ook voor de hoogste percentielen ligt de afwijking tussen beide concepten onder de 3.1 procentpunten behalve dan voor het toppercientiel (vergelijk hiervoor de lijnen percentiel van de kolom Gemiddelde aanslagvoet van de subtabellen Externe bron en Simulatieresultaten in Tabel 28).

Figuur 12: Gemiddelde aanslagvoet in procent per deciel op basis van externe bron en PIT module⁶⁷



De overschatting van de gemiddelde aanslagvoet, geïllustreerd in Figuur 12, wordt vermoedelijk mee veroorzaakt door het ontbreken van een reeks aftrekbare bestedingen.

⁶⁷ De gegevens van de externe bron zijn afkomstig uit ADSEI (2011b), Tabel B.1: Decielenverdeling van het totaal netto belastbaar inkomen, de totale belasting en de gemiddelde aanslagvoet.

Deze bestedingen zijn doorgaans ongelijker verdeeld dan het belastbaar inkomen.⁶⁸ Vermoedelijk neemt hun aandeel ook toe naarmate het inkomen stijgt.

3.8 SOCBEN

In de SOCBEN module wordt eerst het recht op een aantal sociale bijstandsuitkeringen geïdentificeerd.⁶⁹ Er wordt een onderscheid gemaakt tussen a) recht op leefloon, b) recht op een inkomensgarantie voor ouderen (IGO) of het gewaarborgd inkomen voor bejaarden (GIB) en c) de tegemoetkomingen voor personen met een handicap. De identificatie van deze rechten gebeurt voornamelijk op basis van ingelezen statusvariabelen die aangeleverd worden door de instellingen die de respectievelijke rechten uitbetalen.

Vervolgens wordt de middelentoets, die bij ieder van deze uitkeringen wordt uitgevoerd, gerepliceerd met behulp van inkomenscomponenten die door de verschillende andere modules worden aangeleverd. We beschikken evenwel niet over alle inkomenscomponenten die in rekening zouden kunnen genomen worden bij de middelentoetsen die voor de verschillende rechten worden uitgevoerd. Zo observeren we niets over het roerend en onroerend vermogen van de potentiële rechthebbende. De inkomenscomponenten waarover we niets weten worden in deze toepassing gelijk aan nul gezet.

In Tabel 29 tonen we een aantal aggregaten van de rechthebbenden die door de SOCBEN module geïdentificeerd worden na toepassing van de relevante middelentoetsen en vergelijkbare gegevens uit externe bronnen.

⁶⁸ Zie grafiek 37, p. 108 in Decoster e.a. (2011).

⁶⁹ Voor deze module werd geen nota met de basisprincipes uitgeschreven. De meest recente versie van de broncode maakt deel uit van FOD Sociale Zekerheid (2011d).

Tabel 29: Rechthebbenden naar type sociale bijstandsuitkering op basis van externe bron en SOCBEN module⁷⁰

	Externe bron	Gesimuleerd aantal	Procentuele afwijking gesimuleerd aantal en externe bron
Leefloon			
Samenwonenden met één of meerdere personen	23,137	17,592	-24.0
Alleenstaanden	37,606	37,584	-0.1
Samenwonenden met personen ten laste	22,159	22,435	1.2
Totaal	82,902	77,611	-6.4
IGO/GIB			
IGO recht en alleenstaand	54,894	51,713	-5.8
IGO recht en samenwonend	24,255	17,922	-26.1
GIB	14,468	15,729	8.7
Totaal	93,617	85,364	-8.8
Tegemoetkomingen personen met een handicap			
Inkomensvervangende en integratietegemoetkoming (IVT/IT)	139,233	111,277	-20.1
Tegemoetkoming voor hulp aan bejaarden (THB)	126,816	124,080	-2.2
Totaal	266,049	235,357	-11.5

Uit Tabel 29 kan afgeleid worden dat het aantal rechthebbenden op de verschillende bijstandsuitkeringen telkens wordt onderschat met een waarde variërend tussen 6% en 11.5% van het totaal aantal rechthebbenden.

Voor zowel het leefloon als de inkomensgarantie voor ouderen is dit vooral een gevolg van de onderschatting van het aantal samenwonende rechthebbenden. Bij de tegemoetkomingen aan personen met een handicap is dit vooral een gevolg van een onderschatting van de rechthebbenden op een inkomensvervangende- of integratietegemoetkoming.

Deze geobserveerde onderschattingen zijn vermoedelijk een gevolg van de manier waarop de middelentoetsen gerepliceerd worden in de SOCBEN module. Op basis van de aangeleverde informatie van de uitbetalende instellingen weten we namelijk vrij accuraat wie het recht effectief ontvangen heeft. De vergelijking van enerzijds het geobserveerd aantal rechthebbenden met een recht op leefloon of financiële hulp en de personen met een tegemoetkoming voor personen met een handicap en anderzijds externe gegevens (zie Tabel 10 en Tabel 11) leert ook dat deze groepen vrij goed benaderd worden op basis van de statuten die we observeren in de steekproef.

Eén verklaring voor het "falen" van de hier uitgevoerde middelentoets zou kunnen zijn dat de feitelijk uitgevoerde middelentoets moeilijk repliceerbaar is met onze gegevens. Wij gebruiken hier immers inkomens op jaarbasis en voeren met deze gegevens een middelentoets uit. Alle inkomens die na de eventuele afsluiting van een recht verdiend werden, worden daardoor in ons geval wel mee in rekening gebracht bij het uitvoeren van

⁷⁰ De gegevens van de externe bron zijn afkomstig uit FOD Sociale Zekerheid (2011a) p. 319 (Leefloon) en p. 326 (IT/IVT en THB) en FOD Sociale Zekerheid (2010) p.357 (IGO/GIB).

de middelentoets. Een individu dat gedurende de eerste twee maanden van het jaar een volledig leefloon ontving, en vervolgens dit recht gedurende de volgende tien maanden verliest wegens een inkomen uit arbeid zal op basis van onze gegevens eerst geïdentificeerd worden als een potentiële rechthebbende maar op basis van de uitgevoerde middelentoets toch geen uitkering ontvangen.

Dit argument zou echter enkel mogen spelen voor groepen waarbij het inkomen doorheen het jaar wisselt, in de eerste plaats als gevolg van wisselende arbeidsmarktactiviteiten. Voor de onderschatting van de leefloongerechtigden en in mindere mate voor de gevallen IT/IVT zou het dus een verklaring kunnen vormen. Voor de bijstandsuitkeringen aan bejaarden lijkt dit argument minder plausibel. Welke argumenten nog een verklaring kunnen vormen voor de weerkerende onderschatting van het aantal rechthebbenden, zal dan ook nog nader onderzocht moeten worden.

In Tabel 30 geven we tenslotte de totale uitgaven die we per uitkeringstype bekomen op basis van de SOCBEN module en deze die gerapporteerd worden in een externe bron.

Tabel 30: Uitgaven sociale bijstand naar stelsel op basis van externe bron en SOCBEN module in 1,000 Euro⁷¹

	Externe bron	Gesimuleerde bedragen	Procentuele afwijking gesimuleerd bedrag en externe bron
Leefloon	369,948	372,704	0.7
IGO/GIB	392,278	426,335	8.7
IT/IVT	1,013,536	802,635	-20.8
THB	416,412	347,092	-16.6

Ondanks de onderschatting van het aantal gerechtigden op een leefloon en IGO/GIB en het ontbreken van een aantal inkomenscomponenten bij de toepassing van de middelentoets wordt het uitgavenbedrag dat extern geobserveerd wordt respectievelijk vrij goed benaderd en overschat met 9%. Er is niet direct een verklaring voor deze observatie. Voor wat het leefloon betreft kan wel opgemerkt worden dat nog beter moet nagegaan worden of het extern gegeven een goede vergelijkingsbron is. Het gaat om een bedrag dat vermoedelijk ook andere kosten omvat dan enkel uitgaven aan leefloon door OCMW's. Het is op dit moment niet duidelijk hoe die andere kosten die gedragen worden door de POD Maatschappelijke Integratie zich verhouden tot de uitgaven door OCMW's die opgenomen zijn in het gesimuleerde bedrag.

De onderschatting van het budget IT/IVT ligt, op basis van het geschat aantal rechthebbenden, in de lijn der verwachtingen. De onderschatting van het THB budget is dan weer onverwacht, gegeven de beperkte afwijking tussen het extern gerapporteerd en gereconstrueerd aantal rechthebbenden. Het zoeken naar verklaringen voor deze verschillende vaststellingen dient dan ook voorwerp uit te maken van bijkomend onderzoek.

⁷¹ De gegevens van de externe bron zijn afkomstig uit FOD Sociale Zekerheid (2011a) p. 319 (Leefloon), p. 324 (IGO/GIB) en p. 327 (IT/IVT en THB).

4 BESCHRIJVING MODULE EVAL

In de evaluatiemodule (EVAL) van het model worden een aantal maatstaven voor en na een gesimuleerde hervorming samengebracht. Twee grote informatieve blokken kunnen in de door EVAL geproduceerde rapportering, onderscheiden worden.

Enerzijds worden een aantal budgettaire aggregaten samengebracht. Deze concepten worden direct overgenomen uit de verschillende deelmodules. De performantie van het model voor het basisjaar 2008 werd in de voorgaande sectie reeds besproken voor een aantal van de gerapporteerde aggregaten.

Anderzijds worden in de evaluatiemodule ook nog een aantal verdelingsmaatstaven berekend. Zo wordt de ongelijkheid van de inkomensverdeling onder andere samengevat door een ongelijkheidsmaatstaf zoals de Gini-coëfficiënt voor en na de hervorming te berekenen. Verder worden ook armoederisicopercentages en tabellen met winnaars en verliezers, ingedeeld in een aantal verschillende dimensies, berekend.

Een belangrijke plaats in deze verdelingsanalyses wordt ingenomen door het equivalent beschikbaar inkomen.⁷² Dit concept vormt onder andere de basis voor de bepaling van een courant gebruikte armoededrempel, die vastgelegd wordt als 60% van de mediaan van de nationale verdeling van het equivalent beschikbaar inkomen.

De mediaan van het equivalent beschikbaar inkomen, die we bekomen op basis van de MIMOSIS Steekproef 2009 (inkomens 2008), is 1,486 Euro per maand. Dit impliceert een geschatte waarde van de armoedenorm van 892 Euro. De "officiële" norm voor België bedraagt 966 Euro (gegeven gebaseerd op de EU_SILC 2009 (inkomens 2008)).⁷³

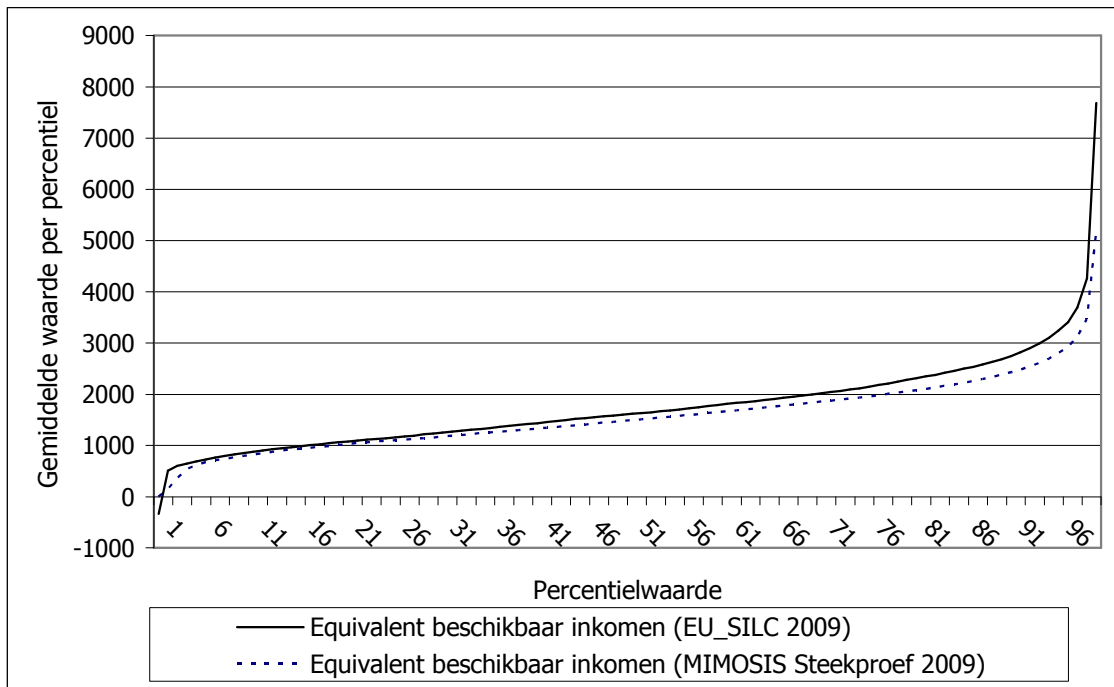
Dat we de mediaan van het equivalent beschikbaar inkomen, in vergelijking met de EU_SILC, onderschatten hoeft niet te verwonderen. Een aantal componenten zoals inkomsten uit roerende goederen en overdrachten tussen huishoudens, die gebruikt worden voor de berekening van het EU_SILC concept, zijn niet opgenomen in de MIMOSIS-steekproef. Bovendien overschatten we, door het ontbreken van een aantal gegevens die relevant zijn voor de bepaling van de personenbelasting, ook de verschuldigde personenbelastingen. Om het belang van deze verschillende componenten in de afwijking tussen beide concepten te verklaren, is meer gedetailleerd onderzoek nodig. We gaan hier niet verder in op dit punt. Wel tonen we in Figuur 13, per percentiel van het equivalent beschikbaar inkomen, de gemiddelde waarde van het equivalent beschikbaar inkomen dat geobserveerd wordt in de EU_SILC versie 2009 en op basis van de MIMOSIS steekproef 2009.⁷⁴

⁷² Het begrip beschikbaar inkomen wordt gedetailleerder toegelicht in European Commission (2003).

⁷³ Zie Eurostat (2011).

⁷⁴ De percentielwaarde in deze en volgende figuren, is telkens aangemaakt op basis van de verdeling van de gegevens in de bron zelf. De gegevens per bron worden dus telkens volledig onafhankelijk van elkaar geproduceerd.

Figuur 13: Equivalent beschikbaar inkomen in Euro per maand (gemiddelde waarde per percentiel)⁷⁵

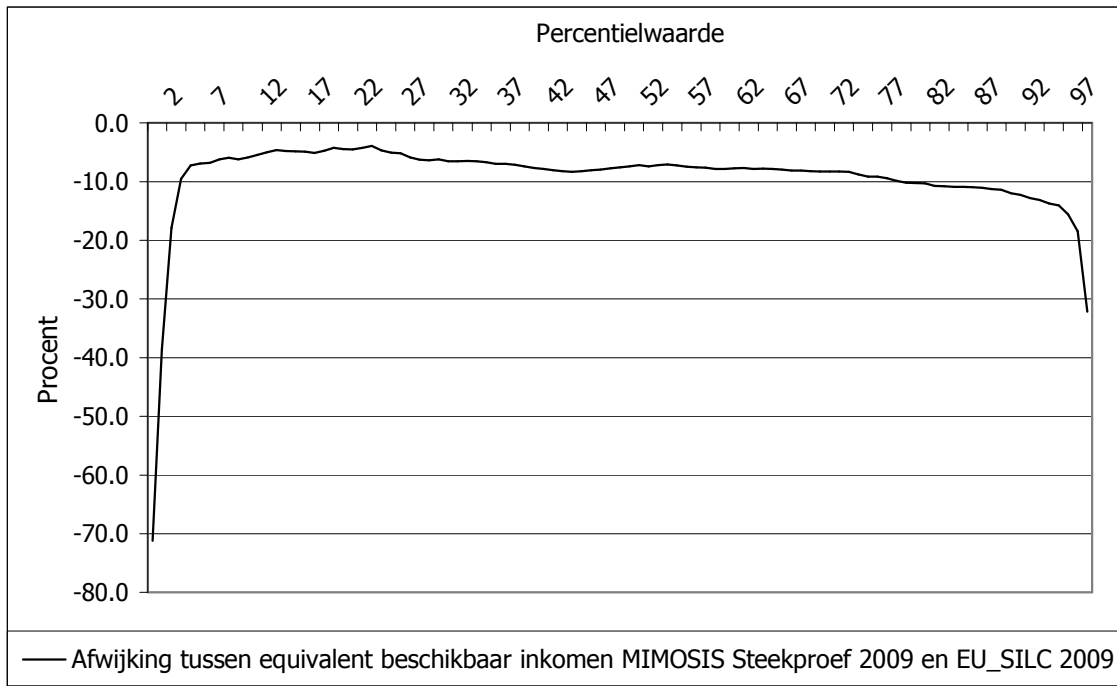


Uit Figuur 13 blijkt dat de onderschatting van het equivalent beschikbaar inkomen op basis van MIMOSIS, ten opzichte van de EU_SILC, systematisch opduikt voor alle percentielen, behalve voor het eerste percentiel. Het equivalent beschikbaar inkomen in het eerste percentiel is, op basis van de EU_SILC gegevens, negatief, terwijl op basis van de MIMOSIS gegevens een zeer klein positief bedrag geraamd wordt.

Ogenschijnlijk liggen de verdelingen van beide concepten dicht bij elkaar. Het is op basis van de gegevens in Figuur 13 echter niet te zeggen in welke mate de gemiddelde percentielwaarden van elkaar afwijken. Daarom geven we in Figuur 14, op het eerste percentiel na, de procentuele afwijking tussen beide concepten per percentiel.

⁷⁵ De percentielwaarden op basis van de EU_SILC zijn niet publiek beschikbaar. Ze werden door ADSEI geleverd op vraag.

Figuur 14: Procentuele afwijking tussen equivalent beschikbaar inkomen MIMOSIS Steekproef 2009 en EU_SILC 2009



Uit Figuur 14 blijkt dat voor de meeste percentielen het equivalent beschikbaar inkomen op basis van MIMOSIS met 10% of minder onderschat wordt. Enkel voor de twee laagste en de twee hoogste percentielen stijgt de afwijking boven de 20% uit. De concepten in beide bronnen ligt het dichtst bij elkaar tussen het 10^{de} en het 20^{ste} percentiel. Vanaf het 20^{ste} percentiel neemt de afwijking, min of meer, gestaag toe over de verschillende percentielen.

Omwille van deze ongelijke afwijkingen per percentiel en omwille van het feit dat er een aantal verklarende factoren voor deze afwijking zijn, passen we per percentiel een correctie toe op de verdeling van het equivalent beschikbaar inkomen die gesimuleerd wordt met MIMOSIS. De correctiefactor die zal worden toegepast is gelijk aan de procentuele afwijking tussen de EU_SILC waarde op basis van de EU_SILC gegevens van 2009 en de waarde die voor MIMOSIS bekomen wordt op basis van de steekproef 2009 (inkomens 2008) voor het belasting- en bijdragenjaar 2008.

De mediaan van de gesimuleerde verdeling van het equivalent beschikbaar inkomen, die we na deze correctieprocedure voor het simulatiejaar 2008 bekomen, bedraagt 1,609 Euro per maand, wat ook het bedrag is dat op basis van de EU_SILC 2009 (inkomens 2008) bekomen wordt.⁷⁶

In Tabel 31 en Tabel 32 rapporteren we voor het MIMOSIS baseline scenario van 2008, gesimuleerd met de steekproef 2009 (inkomens 2008), de armoederisicopercentages voor en na de hoger beschreven correctieprocedure van het equivalent beschikbaar inkomen opgedeeld naar respectievelijk leeftijdsklasse en regio.

⁷⁶ Zie ADSEI (2011c).

Tabel 31: Armoederisicopercentage naar leeftijdsklasse voor en na correctie equivalent beschikbaar inkomen

	Aandeel groep in populatie	Armoederisicopercentage voor correctie equivalent beschikbaar inkomen	Armoederisicopercentage na correctie equivalent beschikbaar inkomen
0 - 4 jaar	5.3	16.2	17.5
5 - 9 jaar	5.4	14.5	15.9
10 -14 jaar	5.7	14.6	16.1
15 - 19 jaar	6.1	15.9	17.5
20 - 24 jaar	5.9	12.5	13.6
25 - 29 jaar	6.3	10.5	11.2
30 - 34 jaar	6.1	10.4	11.4
35 - 39 jaar	6.8	10.8	11.7
40 - 44 jaar	7.4	11.0	12.1
45 - 49 jaar	7.6	10.2	11.2
50 - 55 jaar	7.0	10.5	11.3
56 - 59 jaar	6.5	12.1	12.8
60 - 64 jaar	6.0	14.2	15.3
65 - 69 jaar	4.4	15.7	17.5
70 - 74 jaar	4.5	18.1	20.6
75 - 79 jaar	4.0	18.3	20.9
80 - 84 jaar	2.9	17.6	20.0
85 - 89 jaar	1.6	14.4	17.0
90 - 94 jaar	0.4	11.5	14.0
>= 95 jaar	0.1	12.5	14.5
België	100.0	13.2	14.5

Tabel 32: Armoederisicopercentage naar regio voor en na correctie equivalent beschikbaar inkomen

	Aandeel groep in populatie	Armoederisicopercentage voor correctie equivalent beschikbaar inkomen	Armoederisicopercentage na correctie equivalent beschikbaar inkomen
Regio Brussel Hoofdstad	9.9	24.6	26.6
Regio Vlaanderen	57.7	9.7	10.8
Regio Wallonië	32.3	16.0	17.3
België	100.0	13.2	14.5

Het percentage van de bevolking onder de armoededrempel bedraagt, na correctie, 14.5%. Een percentage dat, gegeven de oorspronkelijke verdeling en de toegepaste correctieprocedure, niet geheel verwonderlijk, dicht aansluit bij dat van de EU_SILC. Op basis van de EU_SILC wordt dit percentage geraamd op 14.6%.⁷⁷

Een vergelijking van de verschillende subgroepresultaten met deze van de EU_SILC zou meer inzicht moeten verschaffen in de mate waarin de EU_SILC resultaten benaderd kunnen worden op basis van de hier gesimuleerde gegevens.

Voor de leeftijdsverdeling, gepresenteerd in Tabel 31 zijn de resultaten niet beschikbaar op basis van de EU_SILC. Een andere leeftijdsindeling op basis van de EU_SILC is wel beschikbaar en zal daarom in de toekomst ook geproduceerd worden als output van de EVAL module.

⁷⁷ Zie Eurostat (2011).

Voor wat de regionale indeling betreft, zijn wel direct vergelijkbare gegevens beschikbaar voor Vlaanderen en Wallonië.⁷⁸ Op basis van de MIMOSIS steekproef wordt het armoederisicopercentage voor Vlaanderen lichtjes overschat in vergelijking met de EU_SILC (i.e. 10.8% op basis van MIMOSIS en 10.1% op basis van de EU_SILC). Het armoederisicopercentage voor Wallonië wordt onderschat (17.3% op basis van MIMOSIS en 18.4% op basis van de EU_SILC). Voor het voor Brussel geschatte percentage van 26.6% hebben we, zoals gezegd, geen referentiepunt.

5 BESLUIT

Doorheen de tekst werd een simulatiemodel gepresenteerd dat gebouwd is op een steekproef van administratieve gegevens. Door een model te bouwen op een steekproef van administratieve gegevens, situeert het model zich tussen minstens twee alternatieve opties, met name een model dat voornamelijk enquêtegegevens als input gebruikt en het werken met de volledige administratieve populatie.

Ten opzichte van werken met de volledige bevolking, zoals zou kunnen met het Datawarehouse Arbeidsmarkt en Sociale Bescherming heeft deze steekproefaankpak het nadeel dat:

- Met minder observaties gewerkt wordt dan er in werkelijkheid beschikbaar zijn,
- Niet alle beschikbare variabelen gebruikt kunnen worden omdat ze niet allemaal in een data-aanvraag kunnen opgenomen worden en
- De meeste socio-economische toestanden enkel gemeten worden op één bepaald moment in de tijd.

Ten opzichte van het werk met enquêtes heeft de gekozen aanpak het nadeel dat niet alle variabelen die opgenomen zijn in enquêtes, en die relevant zijn voor de beoordeling van socio-economische positie van individuen of gezinnen, beschikbaar zijn in administratieve bestanden.

De gevolgde aanpak heeft, ten opzichte van enquêtes, dan weer het voordeel dat:

- de steekproefomvang veel groter kan gemaakt worden dan deze van enquêtes en
- dat de gebruikte concepten veel meer aansluiten bij de administratieve situatie die de rechten, die in het model gereconstrueerd worden, gegenereerd heeft.

Ten opzichte van het werken met de volledige populatie heeft de gevolgde aanpak ook het voordeel dat de rekentijden aanzienlijk beperkt worden zonder dat dit, in geval van een steekproef die representatief is voor het gemodelleerde fenomeen, ten koste zou mogen gaan van de betrouwbaarheid van de verworven inzichten.

Doorheen de tekst werd daarom aandacht besteed aan het aftoetsen van de representativiteit van de steekproef voor de verschillende bestudeerde en gemodelleerde deeldomeinen, i.e. a) werkloosheidsuitkeringen, b) uitkeringen ziekte en invaliditeit, arbeidsongevallen en beroepsziekten, c) pensioenen, d) gezinsuitkeringen, e) uitkeringen

⁷⁸ Zie ADSEI (2011c).

sociale bijstand, f) bijdragen en inhoudingen en g) personenbelastingen. Door bij de modellering van ieder van deze deeldomeinen een onderscheid te maken tussen manipuleerbare parameters, exogene en endogene variabelen is het mogelijk om, op basis van de bestaande versie van het model, reeds een aantal simulaties voor deze domeinen uit te voeren.

Er werd echter bij de bespreking ook de aandacht op gevestigd dat voor een aantal domeinen een bijkomende inspanning noodzakelijk is om de simulatiemogelijkheden van de beschikbare data beter te exploiteren. Dit geldt in het bijzonder voor de modules die ziekte uitkeringen en bijdragen en inhoudingen behandelen.

6 REFERENTIES⁷⁹

- [1.] ADSEI (2011a), Loop van de bevolking, *webpublicatie*, Brussel: Federale OverheidsDienst Economie, Algemene Directie Statistiek en Informatie.
(<http://statbel.fgov.be/nl/statistieken/cijfers/bevolking/loop/>)
- [2.] ADSEI (2011b), Fiscale Inkomens, *webpublicatie*, Brussel: Federale OverheidsDienst Economie, Algemene Directie Statistiek en Informatie.
(http://statbel.fgov.be/nl/statistieken/cijfers/arbeid_leven/fisc/)
- [3.] ADSEI (2011c), Inkomen en levensomstandigheden, *webpublicatie*, Brussel: Federale OverheidsDienst Economie, Algemene Directie Statistiek en Informatie.
(http://statbel.fgov.be/nl/statistieken/cijfers/arbeid_leven/inkomens/)
- [4.] APS (2011), Huishoudens en familiekeren (ID: 1329), *webpublicatie*, Brussel: Studiedienst van de Vlaamse Regering. (<http://aps.vlaanderen.be/sqmi/langereeksen/1329.htm>)
- [5.] Atkinson, A., Rainwater, L. and Smeeding, T. (1995), *Income distribution in OECD Countries: Evidence from the Luxembourg Income Study*, Parijs: OECD.
- [6.] Bay, C., Maréchal, C., Perelman, S. and Van Camp, G. (2007a), MIMOSIS: MICrosimulation MOdel for Belgian Social Insurance Systems, Modelling rules for the module on social security contributions, *webpublicatie*, Brussel: Federale Overheidsdienst Sociale Zekerheid, Directie-generaal Beleidsondersteuning, Domein Evolutie van de Sociale Bescherming.
(http://socialsecurity.fgov.be/NL/nieuws_publicaties/publicaties/mimosis/wpcontrib.pdf)
- [7.] Bay, C., Perelman, S. and Van Camp, G. (2007b), MIMOSIS: MICrosimulation MOdel for Belgian Social Insurance Systems, Modelling rules for the module on family allowances, *webpublicatie*, Brussel: Federale Overheidsdienst Sociale Zekerheid, Directie-generaal Beleidsondersteuning, Domein Evolutie van de Sociale Bescherming.
(http://socialsecurity.fgov.be/NL/nieuws_publicaties/publicaties/mimosis/wpfamal.pdf)
- [8.] Berghman, J., Curvers G., Palmans S. en Peeters H. (2007), De Belgische gepensioneerden in kaart gebracht, Deel 1: De wettelijke pensioenbescherming, *Working paper Sociale Zekerheid Nr. 6*, Brussel: Federale Overheidsdienst Sociale Zekerheid, Directie-generaal Beleidsondersteuning, Domein Evolutie van de Sociale Bescherming.
(http://www.socialsecurity.fgov.be/docs/nl/publicaties/workingpapers/WP_2007-6.pdf)
- [9.] Cowell, F. and Jenkins, S. (2000), Estimating Welfare Indices: Household Weights and Sample Design, *Discussion Paper No. DARP 48*, London: London School of Economics.
(http://eprints.lse.ac.uk/2160/Estimating_Welfare_Indices_Household_Weights_and_Sample_Design.pdf)
- [10.] Decoster, A. en Valenduc, C. (eds.) (2011), *Belastingen en fiscaal beleid in België*, Leuven: Acco, pp. 291.
- [11.] Decoster, A. (2010), Berekend Vlaanderen: workshop FLEMOSI – SBO-project 90044, *webpublicatie*, Flemosi-consortium.
(<http://www.flemosi.be/uploads/100/FLEMOSI%20WORKSHOP%202011-01-21%20André.pdf>)
- [12.] Decoster, A, De Swertdt, K, Orsini, K, and Van Camp, G. (2007a), MIMOSIS: MICrosimulation MOdel for Belgian Social Insurance Systems, Modelling rules for the module on personal income taxes, *webpublicatie*, Brussel: Federale Overheidsdienst Sociale Zekerheid, Directie-generaal Beleidsondersteuning, Domein Evolutie van de Sociale Bescherming.
(http://socialsecurity.fgov.be/NL/nieuws_publicaties/publicaties/mimosis/wppit.pdf)
- [13.] Decoster, A, De Swertdt, K, Orsini, K, and Van Camp, G. (2007b), MIMOSIS: MICrosimulation MOdel for Belgian Social Insurance Systems, Modelling rules for household composition and family relationships, *webpublicatie*, Brussel: Federale Overheidsdienst Sociale Zekerheid, Directie-generaal Beleidsondersteuning, Domein Evolutie van de Sociale Bescherming.
(http://socialsecurity.fgov.be/NL/nieuws_publicaties/publicaties/mimosis/wpfamrel.pdf)

⁷⁹ Niet alle geciteerde bronnen zijn momenteel publiek beschikbaar. Ze zullen op termijn beschikbaar gesteld worden op een herwerkte versie van de website "socialsecurity.fgov.be".

- [14.] Decoster, A, De Swerdt, K, Orsini, K, Lefèbvre, M., Maréchal, C., Paszukiewicz, A., Perelman, S., Rombaut, K., Verbist, G. and Van Camp, G. (2008), Valorisation of the microsimulation model for social security MIMOSIS, *Final Report Project AG/01/116*, Brussel: Federaal Wetenschapsbeleid.
(http://www.belspo.be/belspo/home/pub/pub_ostc/agora/raghh116_en.pdf)
- [15.] DOCDWH (2011), Documentatie met betrekking tot het datawarehouse, *webpublicatie*, Brussel: Kruispuntbank van de Sociale Zekerheid
(http://www.ksz.fgov.be/nl/bcss/page/content/websites/belgium/statistics/statistics_01/statistics_01_05.html)
- [16.] European Commission (2003), Commission regulation (EC) No 1980/2003 of 21 October 2003 implementing Regulation (EC) No 1177/2003 of the European Parliament and of the Council concerning Community statistics on income and living conditions (EU_SILC) as regards definitions and updated definitions, Official Journal of the European Union. (<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:298:0001:0022:EN:PDF>)
- [17.] Eurostat (2011), Income, Social Inclusion and Living Conditions, *webpublicatie*, Luxemburg: Eurostat.
(http://epp.eurostat.ec.europa.eu/portal/page/portal/income_social_inclusion_living_conditions/data/database)
- [18.] FOD Sociale Zekerheid (2011a), *Vademecum van de financiële en statistische gegevens over de sociale bescherming in België: periode 2005-2011*, Brussel: Federale Overheidsdienst Sociale Zekerheid, Directie-generaal Sociaal Beleid, Domein Kerngegevens.
(http://socialsecurity.fgov.be/NL/nieuws_publicaties/publicaties/vademecum/2011/vademecum_2011_nl.pdf)
- [19.] FOD Sociale Zekerheid (2011b), Programma's voor omzetting gegevens Datawarehouse Arbeidsmarkt en Sociale Bescherming naar bestanden aangevraagd in data-aanvraag, *mimeo*, Brussel: Federale Overheidsdienst Sociale Zekerheid, Directie-generaal Beleidsondersteuning, Domein Evolutie van de Sociale Bescherming.
- [20.] FOD Sociale Zekerheid (2011c), Programma's voor omzetting gegevens uit data-aanvraag naar input file MIMOSIS, *mimeo*, Brussel: Federale Overheidsdienst Sociale Zekerheid, Directie-generaal Beleidsondersteuning, Domein Evolutie van de Sociale Bescherming.
- [21.] FOD Sociale Zekerheid (2011d), Broncode MIMOSIS: Steekproef januari 2009, gegevens 2008, *mimeo*, Brussel: Federale Overheidsdienst Sociale Zekerheid, Directie-generaal Beleidsondersteuning, Domein Evolutie van de Sociale Bescherming.
- [22.] FOD Sociale Zekerheid (2010a), *Vademecum van de financiële en statistische gegevens over de sociale bescherming in België: periode 2004-2010*, Brussel: Federale Overheidsdienst Sociale Zekerheid, Directie-generaal Sociaal Beleid, Domein Kerngegevens.
(http://socialsecurity.fgov.be/NL/nieuws_publicaties/publicaties/vademecum/2010/vademecum_2010_nl.pdf)
- [23.] FOD Sociale Zekerheid (2010b), *Jaaroverzicht 2009 in cijfers*, Brussel: Federale Overheidsdienst Sociale Zekerheid, Directie-generaal Personen met een Handicap.
(http://handicap.fgov.be/docs/vouwfolder_2009_overzicht_cijfers_nl.pdf)
- [24.] Knapen, H. (2011), Constructie inkomensnotie Datawarehouse AM&SB, *tussentijds verslag project DOCDWH*, Leuven: K.U. Leuven, Centrum voor Sociologisch Onderzoek.
- [25.] KSZ (2011), "Overzicht belangrijkste cijfers basistoepassingen 2008 – verbeterde versie.xls: populatie pensioentrekken", bijlage bij mail met subject Beheersgroep datawarehouse arbeidsmarkt en sociale bescherming – nieuw overzicht van de belangrijkste cijfers een aangepast verslag vergadering en time Vrijdag 26 augustus 2011.
- [26.] Maréchal, C., Paszukiewicz, A., Perelman, S. and Van Camp, G. (2007), MIMOSIS: MICrosimulation MOdel for Belgian Social Insurance Systems, Modelling rules for module on welfare adaptations of pensions,
(http://socialsecurity.fgov.be/NL/nieuws_publicaties/publicaties/mimosis/wppenswelf.pdf)
- [27.] Ministerie van Financiën (2011), *Fiscaal Memento 2011*, Brussel: Ministerie van Financiën, Studie en Documentatiedienst.
(http://docfin.fgov.be/intersalgnl/thema/publicaties/memento/pdf/FM2011_V01_volledig.pdf)

- [28.] Molenberghs, G. en Gaens, E. (1994), Gezondheidsenquête door middel van interview: Statistische beschouwing, *Werkdocument Limburgs Universitair Centrum*, Diepenbeek: Universitaire Campus.
- [29.] Mutsaerts, A. (2011), Constructie Wstatus en Reconstructie CIMIRe op basis van het DWH AM&SB, *tussentijds verslag project SIPEBE*, Leuven: K.U. Leuven, Centrum voor Sociologisch Onderzoek.
- [30.] POD Maatschappelijke Integratie (2011), Studies, publicaties en statistieken: statistieken, webpublicatie, Brussel: Programmatorische OverheidsDienst Maatschappelijke Integratie. (<http://www.mi-is.be/be-nl/studies-publicaties-en-statistieken/statistieken>)
- [31.] Put, J. (2008), *Praktijkboek sociale zekerheid voor de onderneming en de sociale adviseur 2008*, Mechelen: Wolters Kluwer.
- [32.] Rijksregister (2011), Voostelling van het Rijksregister, webpublicatie, Brussel: Federale OverheidsDienst Binnenlandse Zaken. (<http://www.ibz.rm.fgov.be/index.php?id=2461&L=1>)
- [33.] RKW (2011), *Activiteitenverslag 2009*, Brussel: Rijksdienst voor Kinderbijslag voor Werknemers. (<http://www.rkw.be/NL/Documentation/Publication/AnnualReports/JV2009.pdf>)
- [34.] Rombaut, K., Van Camp, G. and Verbist, G. (2007a), *MIMOSIS: MICrosimulation MOdel for Belgian Social Insurance Systems, Modelling rules for the Unemployment Benefits Module*, (http://socialsecurity.fgov.be/NL/nieuws_publicaties/publicaties/mimosis/wpunem.pdf)
- [35.] Rombaut, K., Van Camp, G. and Verbist, G. (2007b), *MIMOSIS: MICrosimulation MOdel for Belgian Social Insurance Systems, Modelling rules for the Sickness and Disability benefits Module*, (http://socialsecurity.fgov.be/NL/nieuws_publicaties/publicaties/mimosis/wpsick.pdf)
- [36.] RSZ (2011), Verminderen van de bijdragen voor sociale zekerheid: aantal betrokken werknemers en budgettaire impact, *webpublicatie*, Brussel: Rijksdienst voor Sociale Zekerheid. (<http://www.rsz.fgov.be/nl/content/statistics/publications/reductions.html>)
- [37.] RVP (2011), Herwaardering van het loon voor de pensioenberekening, *webpublicatie*, Brussel: Rijksdienst voor Pensioenen. (<http://www.onprvp.fgov.be/NL/profes/calculation/career/wages/revaluation/Pages/default.aspx#b>)
- [38.] SAS (2011), SAS/STAT 9.2 User's Guide, *webpublicatie*. (<http://support.sas.com/documentation/cdl/en/statugkde/61796/PDF/default/statugkde.pdf>)
- [39.] Vleminckx, K. (2011), Aanvraag van gecodeerde sociale gegevens van persoonlijke aard voor het updaten van het microsimulatiemodel MIMOSIS en het bestuderen van de mogelijkheid tot armoedemeting met administratieve gegevens uit het DWH AM&SB, *mimeo*, Brussel: Federale Overheidsdienst Sociale Zekerheid, Directie-generaal Beleidsondersteuning, Domein Evolutie van de Sociale Bescherming.
- [40.] Vleminckx, K. (2008), Aanvraag van gecodeerde individuele sociale zekerheidsgegevens voor wetenschappelijk onderzoek in de projecten "Microsimulatiemodel Sociale Zekerheid" (fase II – uitbouw, validering en exploitatie), *mimeo*, Brussel: Federale Overheidsdienst Sociale Zekerheid, Directie-generaal Beleidsondersteuning, Domein Evolutie van de Sociale Bescherming.

APPENDIX 1: BESCHRIJVING BEPALING STEEKPROEFGEWICHTEN

In deze sectie beschrijven we hoe steekproefgewichten geproduceerd kunnen worden, gegeven de gevolgde methode van steekproeftrekking. In de methode van steekproeftrekking kunnen twee belangrijke stappen onderscheiden worden, met name de aanleg van een lijst van 100.000 toevallig geselecteerde individuen en vervolgens de aanvulling van deze lijst met alle gezinsleden die behoorden tot het private gezin van het random getrokken individu.

Voor de uiteenzetting introduceren we volgende symbolen:

NI : Het aantal individuen in de populatie (i.e. de omvang van het steekproefkader)

ni : het aantal individuen in de steekproef

g_j : Het aantal gezinsleden in het gezin waartoe individu j behoort

P_j : de kans dat individu j voorkomt in de populatie

p_j : de kans dat individu j geselecteerd wordt voor de steekproef

In geval van eenvoudige random steekproeftrekking is de kans dat een individu voorkomt in het steekproefkader gelijk aan de kans dat het individu geselecteerd wordt voor de steekproef. De wegingsfactor w_j , die in dat geval aan ieder individu j in de steekproef wordt toegekend, kan bepaald worden als:

$$w_j = \frac{NI}{ni} \quad (1).$$

Omdat wij in de tweede stap de random steekproef van individuen verder aanvullen met de andere gezinsleden is de kans dat een individu uit de populatie terecht komt in de steekproef niet meer gelijk voor alle individuen. De wegingsfactor in vergelijking (1) moet in dit geval aangepast worden met de verhouding $\frac{P_j}{p_j}$, i.e. de kans dat het individu j in de populatie zit gedeeld door de kans dat het individu j in de steekproef zit.⁸⁰ Het aangepaste gewicht w_j wordt in algemene termen gegeven in vergelijking (2):

$$w_j = \frac{NI}{ni} \cdot \frac{P_j}{p_j} \quad (2).$$

De kans P_j , dat een individu j geobserveerd wordt in de populatie is voor alle personen gelijk, i.e. $\frac{1}{NI}$. De kans p_j , dat een individu j geobserveerd wordt in de steekproef is ongelijk verdeeld en hangt af van de gezinsomvang van het gezin waartoe j behoort. Ieder individu kan namelijk zelf geselecteerd worden in de eerste stap of via één van zijn

⁸⁰ Zie Cowell, F. and Jenkins, S. (2000), p. 11.

gezinsleden in de tweede stap. De selectiekans van een individu j kan als volgt geschreven

worden $\frac{g_j}{\sum_{i=1}^M g_i}$. Het gewicht w_j kan in dit geval bepaald worden als:

$$w_j = \frac{\sum_{i=1}^M g_i}{ni \cdot g_j} \quad (3).$$

APPENDIX 2: GEWOGEN RESULTATEN STEEKPROEF NAAR REGIO, LEEFTIJD EN GESLACHT

Tabel 33: Gewogen aantal observaties op individueel niveau naar regio en leeftijd (mannen)⁸¹

		0-4 JAAR	5-9 JAAR	10-14 JAAR	15-19 JAAR	20-24 JAAR	25-29 JAAR	30-34 JAAR	35-39 JAAR	40-44 JAAR	45-49 JAAR
Gewogen aantal op basis van steekproef											
	België	316,089	301,683	309,986	333,564	326,344	346,023	346,948	382,905	406,619	408,775
	Regio Brussel Hoofdstad	40,224	32,539	29,578	29,669	33,947	44,661	47,037	45,015	40,495	35,102
	Regio Vlaanderen	172,438	164,267	174,577	188,554	182,132	196,081	192,433	215,687	240,219	245,510
	Regio Wallonië	103,427	104,877	105,831	115,341	110,265	105,281	107,478	122,203	125,905	128,163
Werkelijk aantal op basis van bevolkingsst atistieken 2009											
	België	288,971	293,651	311,591	330,478	320,003	344,636	328,241	363,808	401,351	409,016
	Regio Brussel Hoofdstad	38,530	32,626	29,884	29,167	32,488	38,906	39,027	38,833	38,034	35,564
	Regio Vlaanderen	153,874	158,023	173,986	189,444	175,099	198,596	186,108	206,833	238,479	245,965
	Regio Wallonië	96,567	103,002	107,721	111,867	112,416	107,134	103,106	118,142	124,838	127,487
Procentuele afwijking gewogen aantal en werkelijk aantal											
	België	-9.4	-2.7	0.5	-0.9	-2.0	-0.4	-5.7	-5.2	-1.3	0.1
	Regio Brussel Hoofdstad	-4.4	0.3	1.0	-1.7	-4.5	-14.8	-20.5	-15.9	-6.5	1.3
	Regio Vlaanderen	-12.1	-4.0	-0.3	0.5	-4.0	1.3	-3.4	-4.3	-0.7	0.2
	Regio Wallonië	-7.1	-1.8	1.8	-3.1	1.9	1.7	-4.2	-3.4	-0.9	-0.5

⁸¹ De gegevens in kolom werkelijke verdeling zijn gebaseerd op gegevens van ADSEI, zie ADSEI (2011a): Bevolking naar geslacht, leeftijdsgroepen en leeftijdsklassen.

Tabel 33 (vervolg): Gewogen aantal observaties op individueel niveau naar regio en leeftijd (mannen)⁸²

		50-54 JAAR	55-59 JAAR	60-64 JAAR	65-69 JAAR	70-74 JAAR	75-79 JAAR	80-84 JAAR	85-89 JAAR	90-94 JAAR	95 JAAR EN MEER
Gewogen aantal op basis van steekproef											
	België	377,220	339,739	301,676	219,183	206,333	171,250	109,706	52,235	9,952	2,421
	Regio Brussel Hoofdstad	30,277	26,502	22,450	16,530	14,577	12,357	8,910	4,994	1,072	314
	Regio Vlaanderen	226,024	200,878	181,908	138,067	131,955	107,309	66,975	31,606	6,067	1,482
	Regio Wallonië	120,919	112,359	97,318	64,586	59,801	51,584	33,821	15,635	2,813	625
Werkelijk aantal op basis van bevolkingsst atistieken 2009											
	België	377,759	346,092	318,169	227,248	219,340	174,977	118,318	57,826	10,691	2,516
	Regio Brussel Hoofdstad	29,960	25,434	27,840	18,416	15,527	15,314	10,207	5,771	1,343	231
	Regio Vlaanderen	229,505	204,173	190,240	139,337	139,872	106,914	70,823	35,733	6,529	1,751
	Regio Wallonië	118,294	116,485	100,089	69,495	63,941	52,749	37,288	16,322	2,819	534
Procentuele afwijking gewogen aantal en werkelijk aantal											
	België	0.1	1.8	5.2	3.5	5.9	2.1	7.3	9.7	6.9	3.8
	Regio Brussel Hoofdstad	-1.1	-4.2	19.4	10.2	6.1	19.3	12.7	13.5	20.2	-35.9
	Regio Vlaanderen	1.5	1.6	4.4	0.9	5.7	-0.4	5.4	11.5	7.1	15.4
	Regio Wallonië	-2.2	3.5	2.8	7.1	6.5	2.2	9.3	4.2	0.2	-17.0

⁸² De gegevens in kolom werkelijke verdeling zijn gebaseerd op gegevens van ADSEI, zie ADSEI (2011a): Bevolking naar geslacht, leeftijdsgroepen en leeftijdsklassen.

Tabel 34: Gewogen aantal observaties op individueel niveau naar regio en leeftijd (vrouwen)⁸³

		0-4 JAAR	5-9 JAAR	10-14 JAAR	15-19 JAAR	20-24 JAAR	25-29 JAAR	30-34 JAAR	35-39 JAAR	40-44 JAAR	45-49 JAAR
Gewogen aantal op basis van steekproef											
	België	301,448	289,318	296,352	320,482	324,396	344,963	342,202	373,829	394,900	402,081
	Regio Brussel Hoofdstad	38,498	31,473	28,084	29,146	37,916	49,069	46,334	41,571	36,836	34,406
	Regio Vlaanderen	164,444	157,490	167,109	180,576	178,634	192,199	189,040	211,129	232,243	238,548
	Regio Wallonië	98,506	100,355	101,159	110,760	107,846	103,695	106,828	121,129	125,821	129,127
Werkelijk aantal op basis van bevolkingsst atistieken 2009											
	België	270,697	281,696	299,980	322,151	313,158	333,551	326,266	370,588	393,257	409,547
	Regio Brussel Hoofdstad	36,383	30,992	29,447	30,205	33,125	41,043	40,253	38,853	36,714	35,410
	Regio Vlaanderen	146,578	150,920	169,917	180,361	174,380	189,936	182,113	211,521	235,028	243,728
	Regio Wallonië	87,736	99,784	100,616	111,585	105,653	102,572	103,900	120,214	121,515	130,409
Procentuele afwijking gewogen aantal en werkelijk aantal											
	België	-11.4	-2.7	1.2	0.5	-3.6	-3.4	-4.9	-0.9	-0.4	1.8
	Regio Brussel Hoofdstad	-5.8	-1.6	4.6	3.5	-14.5	-19.6	-15.1	-7.0	-0.3	2.8
	Regio Vlaanderen	-12.2	-4.4	1.7	-0.1	-2.4	-1.2	-3.8	0.2	1.2	2.1
	Regio Wallonië	-12.3	-0.6	-0.5	0.7	-2.1	-1.1	-2.8	-0.8	-3.5	1.0

⁸³ De gegevens in kolom werkelijke verdeling zijn gebaseerd op gegevens van ADSEI, zie ADSEI (2011a): Bevolking naar geslacht, leeftijdsgroepen en leeftijdsklassen.

Tabel 34 (vervolg): Gewogen aantal observaties op individueel niveau naar regio en leeftijd (vrouwen)⁸⁴

		50-54 JAAR	55-59 JAAR	60-64 JAAR	65-69 JAAR	70-74 JAAR	75-79 JAAR	80-84 JAAR	85-89 JAAR	90-94 JAAR	95 JAAR EN MEER
Gewogen aantal op basis van steekproef											
	België	375,940	340,669	312,151	241,082	246,275	235,946	184,717	113,968	31,675	12,035
	Regio Brussel Hoofdstad	31,661	28,836	25,597	19,948	19,341	19,277	16,891	12,028	3,829	1,541
	Regio Vlaanderen	220,609	196,478	183,263	147,278	151,193	140,758	105,786	63,391	17,699	6,841
	Regio Wallonië	123,670	115,355	103,291	73,856	75,741	75,911	62,040	38,549	10,147	3,653
Werkelijk aantal op basis van bevolkingsst atistieken 2009											
	België	381,372	349,580	324,956	252,175	260,767	250,312	199,774	121,011	34,319	13,243
	Regio Brussel Hoofdstad	36,134	33,529	27,783	25,531	23,795	24,078	21,750	13,460	5,137	1,805
	Regio Vlaanderen	217,944	198,685	190,744	150,802	159,747	148,347	113,128	67,109	19,204	7,404
	Regio Wallonië	127,294	117,366	106,429	75,842	77,225	77,887	64,896	40,442	9,978	4,034
Procentuele afwijking gewogen aantal en werkelijk aantal											
	België	1.4	2.5	3.9	4.4	5.6	5.7	7.5	5.8	7.7	9.1
	Regio Brussel Hoofdstad	12.4	14.0	7.9	21.9	18.7	19.9	22.3	10.6	25.5	14.6
	Regio Vlaanderen	-1.2	1.1	3.9	2.3	5.4	5.1	6.5	5.5	7.8	7.6
	Regio Wallonië	2.8	1.7	2.9	2.6	1.9	2.5	4.4	4.7	-1.7	9.4

⁸⁴ De gegevens in kolom werkelijke verdeling zijn gebaseerd op gegevens van ADSEI, zie ADSEI (2011a): Bevolking naar geslacht, leeftijdsgroepen en leeftijdsklassen.

APPENDIX 3: GEWOGEN RESULTATEN STEEKPROEF NAAR BURGERLIJKE STAAT, LEEFTIJD EN GESLACHT

Tabel 35: Gewogen aantal observaties op individueel niveau naar burgerlijke staat en leeftijd voor België (mannen)⁸⁵

		0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49
		JAAR	JAAR	JAAR	JAAR	JAAR	JAAR	JAAR	JAAR	JAAR	JAAR
Gewogen aantal op basis van steekproef											
Ongehuwd		316,089	301,683	309,986	333,358	314,373	271,116	189,461	142,398	101,565	72,178
Gehuwd		0	0	0	202	11,553	70,454	141,801	203,327	246,029	264,903
Weduwstaat		0	0	0	0	0	55	215	615	1,382	2,641
Gescheiden		0	0	0	0	413	4,398	15,471	36,565	57,643	69,053
Werkelijk aantal op basis van bevolkingsstatistieken 2009											
Ongehuwd		288,972	293,650	311,590	330,443	314,492	292,118	204,782	166,226	139,034	110,538
Gehuwd		0	0	0	36	5,511	51,180	114,932	175,996	225,849	249,720
Weduwstaat		0	0	0	0	0	0	0	196	995	1,368
Gescheiden		0	0	0	0	0	1,337	8,528	21,391	35,473	47,390
Procentuele afwijking gewogen aantal en werkelijk aantal											
Ongehuwd		-9.4	-2.7	0.5	-0.9	0.0	7.2	7.5	14.3	26.9	34.7
Gehuwd		0.0	0.0	0.0	-461.1	-109.6	-37.7	-23.4	-15.5	-8.9	-6.1
Weduwstaat		0.0	0.0	0.0	0.0	0.0	0.0	0.0	-213.8	-38.9	-93.1
Gescheiden		0.0	0.0	0.0	0.0	0.0	-228.9	-81.4	-70.9	-62.5	-45.7

⁸⁵ De gegevens in kolom werkelijke verdeling zijn gebaseerd op gegevens van ADSEI, zie ADSEI (2011a): Bevolking naar geslacht, leeftijdsgroepen en leeftijdsklassen.

Tabel 35 (vervolg): Gewogen aantal observaties op individueel niveau naar burgerlijke staat en leeftijd voor België (mannen)⁸⁶

		50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95 JAAR
		JAAR	JAAR	JAAR	JAAR	JAAR	JAAR	JAAR	JAAR	JAAR	EN MEER
Gewogen aantal op basis van steekproef											
	Ongehuwd	46,775	30,323	21,163	13,545	12,069	9,658	6,087	2,647	540	143
	Gehuwd	258,882	246,403	229,777	170,870	160,535	127,706	73,655	28,891	3,893	529
	Weduwstaat	4,665	7,618	11,122	12,636	18,794	25,317	26,362	19,412	5,333	1,725
	Gescheiden	66,898	55,395	39,614	22,132	14,935	8,569	3,602	1,285	186	16
Werkelijk aantal op basis van bevolkings- statistieken 2009											
	Ongehuwd	80,695	56,226	39,324	22,751	21,326	12,856	7,934	4,571	936	134
	Gehuwd	246,285	235,794	231,964	167,832	165,803	127,459	79,453	29,816	4,275	559
	Weduwstaat	2,087	3,765	5,086	7,955	11,640	16,480	14,855	11,681	3,751	1,395
	Gescheiden	48,693	50,307	41,795	28,710	20,572	18,181	16,077	11,757	1,729	428
Procentuele afwijking gewogen aantal en werkelijk aantal											
	Ongehuwd	42.0	46.1	46.2	40.5	43.4	24.9	23.3	42.1	42.3	-6.7
	Gehuwd	-5.1	-4.5	0.9	-1.8	3.2	-0.2	7.3	3.1	8.9	5.4
	Weduwstaat	-123.5	-102.3	-118.7	-58.8	-61.5	-53.6	-77.5	-66.2	-42.2	-23.7
	Gescheiden	-37.4	-10.1	5.2	22.9	27.4	52.9	77.6	89.1	89.2	96.3

⁸⁶ De gegevens in kolom werkelijke verdeling zijn gebaseerd op gegevens van ADSEI, zie ADSEI (2011a): Bevolking naar geslacht, leeftijdsgroepen en leeftijdsklassen.

Tabel 36: Gewogen aantal observaties op individueel niveau naar burgerlijke staat en leeftijd voor België (vrouwen)⁸⁷

		0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49
		JAAR	JAAR	JAAR	JAAR	JAAR	JAAR	JAAR	JAAR	JAAR	JAAR
Gewogen aantal op basis van steekproef											
	Ongehuwd	301,448	289,318	296,352	318,311	288,328	221,915	143,612	99,930	66,904	47,020
	Gehuwd	0	0	0	2,160	34,622	113,411	172,652	221,749	253,876	267,154
	Weduwstaat	0	0	0	0	48	336	1,037	2,680	5,420	10,376
	Gescheiden	0	0	0	9	1,397	9,301	24,901	49,470	68,700	77,531
Werkelijk aantal op basis van bevolkings- statistieken 2009											
	Ongehuwd	270,697	281,696	299,980	321,429	293,379	242,146	166,809	132,749	104,562	87,519
	Gehuwd	0	0	0	721	19,636	88,001	143,627	202,843	235,881	252,330
	Weduwstaat	0	0	0	0	0	143	730	1,472	4,150	7,849
	Gescheiden	0	0	0	0	144	3,260	15,102	33,523	48,663	61,849
Procentuele afwijking gewogen aantal en werkelijk aantal											
	Ongehuwd	-11.4	-2.7	1.2	1.0	1.7	8.4	13.9	24.7	36.0	46.3
	Gehuwd	0.0	0.0	0.0	-199.6	-76.3	-28.9	-20.2	-9.3	-7.6	-5.9
	Weduwstaat	0.0	0.0	0.0	0.0	0.0	-135.0	-42.1	-82.1	-30.6	-32.2
	Gescheiden	0.0	0.0	0.0	0.0	-870.1	-185.3	-64.9	-47.6	-41.2	-25.4

⁸⁷ De gegevens in kolom werkelijke verdeling zijn gebaseerd op gegevens van ADSEI, zie ADSEI (2011a): Bevolking naar geslacht, leeftijdsgroepen en leeftijdsklassen.

Tabel 36 (vervolg): Gewogen aantal observaties op individueel niveau naar burgerlijke staat en leeftijd voor België (vrouwen)⁸⁸

		50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95 JAAR
		JAAR	JAAR	JAAR	JAAR	JAAR	JAAR	JAAR	JAAR	JAAR	EN MEER
Gewogen aantal op basis van steekproef											
	Ongehuwd	30,887	20,375	14,454	10,470	11,351	12,158	11,405	7,716	2,618	1,158
	Gehuwd	255,569	235,962	214,712	155,734	140,636	102,484	50,327	16,052	1,513	185
	Weduwstaat	17,427	26,214	39,365	49,198	75,521	108,442	115,704	86,610	26,696	10,403
	Gescheiden	72,057	58,118	43,620	25,680	18,767	12,862	7,281	3,590	848	289
Werkelijk aantal op basis van bevolkings- statistieken 2009											
	Ongehuwd	62,880	44,315	32,482	19,744	16,013	15,667	11,325	9,592	3,428	1,019
	Gehuwd	246,336	229,120	213,258	159,308	143,776	107,894	52,292	15,517	1,786	89
	Weduwstaat	17,177	21,726	36,698	48,259	77,912	108,957	123,969	90,163	27,763	11,813
	Gescheiden	54,977	54,419	42,516	24,864	23,065	17,794	12,188	5,739	1,341	321
Procentuele afwijking gewogen aantal en werkelijk aantal											
	Ongehuwd	50.9	54.0	55.5	47.0	29.1	22.4	-0.7	19.6	23.6	-13.6
	Gehuwd	-3.7	-3.0	-0.7	2.2	2.2	5.0	3.8	-3.4	15.3	-107.9
	Weduwstaat	-1.5	-20.7	-7.3	-1.9	3.1	0.5	6.7	3.9	3.8	11.9
	Gescheiden	-31.1	-6.8	-2.6	-3.3	18.6	27.7	40.3	37.4	36.8	10.0

⁸⁸ De gegevens in kolom werkelijke verdeling zijn gebaseerd op gegevens van ADSEI, zie ADSEI (2011a): Bevolking naar geslacht, leeftijdsgroepen en leeftijdsklassen.

APPENDIX 4: HERWAARDERINGSFACTOREN GESCHATTE DAGLONEN

Tabel 37: Herwaarderingsfactoren geschatte daglonen per observatiejaar⁸⁹

Prestatiejaar	Herwaarderingsfactor	Prestatiejaar	Herwaarderingsfactor
1940	6.405881	1975	2.861783
1941	6.405881	1976	2.621545
1942	6.405881	1977	2.447630
1943	6.405881	1978	2.342965
1944	6.405881	1979	2.242618
1945	6.405881	1980	2.102809
1946	6.405881	1981	1.953805
1947	6.405881	1982	1.796984
1948	6.405881	1983	1.669099
1949	6.405881	1984	1.600875
1950	6.405881	1985	1.557084
1951	6.405881	1986	1.537166
1952	6.405881	1987	1.543957
1953	6.405881	1988	1.526233
1954	6.405881	1989	1.480237
1955	6.255990	1990	1.430892
1956	6.111607	1991	1.386377
1957	5.923573	1992	1.353496
1958	5.848569	1993	1.317196
1959	5.777393	1994	1.306018
1960	5.760077	1995	1.285983
1961	5.703567	1996	1.265062
1962	5.62469	1997	1.248598
1963	5.506432	1998	1.232904
1964	5.286357	1999	1.221376
1965	5.079292	2000	1.198833
1966	4.876030	2001	1.166827
1967	4.741017	2002	1.152043
1968	4.612922	2003	1.132669
1969	4.446323	2004	1.111100
1970	4.279467	2005	1.081046
1971	4.101257	2006	1.059012
1972	3.889245	2007	1.027111
1973	3.636347	2008	1.000000
1974	3.227094		

⁸⁹ De gegevens in deze tabel zijn gebaseerd op de gegevens gepubliceerd door RVP (2011).