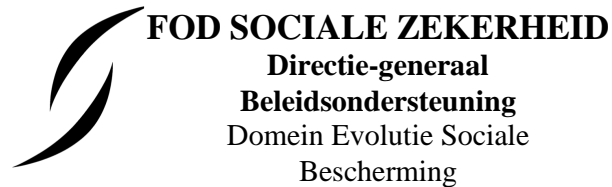


tel.: +322528 63 19
fax: +32 (0)2 528 69 67
vragen naar: Guy Van Camp
e-mail: Guy.VanCamp@minsoc.fed.be



uw brief van
uw kenmerk

ons kenmerk
datum 16 september 2010

bijlage(n)

Betreft: 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

1 BESCHRIJVING VAN DE ONDERZOEKSINSTELLING DIE DE AANVRAAG INDIENT

De data-aanvraag wordt ingediend door de Federale Overheidsdienst Sociale Zekerheid (FOD SZ) DG Beleidsondersteuning, Kruidtuinlaan 50, B-1000 BRUSSEL. Voor dit project wordt de FOD Sociale Zekerheid vertegenwoordigd door Koen Vleminckx, tel 02/528.64.34 (koen.vleminckx@minsoc.fed.be).

2 BESCHRIJVING OPZET VAN HET WETENSCHAPPELIJK ONDERZOEK

Op vraag van de FOD SZ (domein Beleidsondersteunend Onderzoek) werd in het kader van een AGORA project een microsimulatiemodel voor sociaal beleid in België ontwikkeld door drie universitaire teams. Dit model, MIMOSIS (MICrosimulation MOdel for Belgian Social Insurance Systems) genaamd, werd begin 2006 opgeleverd aan de FOD SZ.

Het microsimulatiemodel MIMOSIS maakt het mogelijk om hervormingen in een aantal beleidsdomeinen van de sociale zekerheid te simuleren op basis van geanonimiseerde maar individuele administratieve gegevens afkomstig uit het Datawarehouse Arbeidsmarkt en Sociale Bescherming (DWH AM&SB) van de Kruispuntbank van de Sociale Zekerheid. MIMOSIS richt zich op 6 beleidsdomeinen: sociale zekerheidsbijdragen, pensioenen, werkloosheidsuitkeringen, inkomensgaranties, ziekte- en invaliditeitsuitkeringen en familiale uitkeringen. Voor elke hervorming

C:\My\teksten\MIMS2009\Aanvraag\DataAanvraagUpdateMimosisV6_FullSample.doc

Finance Tower
Kruidtuinlaan 50, bus 135
1000 BRUSSEL
Contactpunt: +32 (0)2 528 64 34
E-mail : dgstrat@minsoc.fed.be
<http://socialsecurity.fgov.be>

.be

wordt via een geïntegreerde belastingmodule tevens het effect berekend op de verschuldigde personenbelasting. Door het gebruik van microgegevens laat het model een zeer gedetailleerde impactanalyse toe van de budgettaire en verdelingseffecten van wijzigingen in een aantal domeinen van de sociale zekerheid. De uitgevoerde simulaties leveren een beeld op van de bevolkingsgroepen die het meest door de bestudeerde maatregel worden getroffen (in positieve of negatieve zin) en hen die er het minst door worden getroffen. Dit laat toe door beleidsmakers voorgestelde hervormingen te toetsen aan de door hen vooropgestelde objectieven, zowel in termen van budgettaire effecten als in termen van de welvaart van de beoogde bevolkingsgroepen.¹

De huidige versie van MIMOSIS maakt voornamelijk gebruik van gegevens van het uitgaven- en inkomstenjaar 2001. Een aantal van de door MIMOSIS gebruikte databronnen veranderde sinds 2001 qua structuur. Verder zijn een aantal bronnen pas beschikbaar in het DWH AM&SB voor meer recente uitgaven- en inkomensjaren dan 2001. Met het oog op het onderhoud van het model, is het dan ook aangewezen om een aangepaste versie van MIMOSIS te ontwikkelen met gegevens voor een meer recent uitgaven- en inkomensjaar.

De voorliggende data-aanvraag wordt opgesteld om de onderliggende steekproef van MIMOSIS te kunnen updaten. Daarnaast is het de bedoeling om te bestuderen of de constructie van het equivalent beschikbaar inkomen, die wordt uitgewerkt op het niveau van de MIMOSIS-steekproef, geïmplementeerd kan worden op het niveau van het DWH AM&SB in zijn geheel. We willen in het bijzonder focussen op het equivalent beschikbaar inkomen omdat dit concept een prominente plaats inneemt bij armoedemeting en armoedemonitoring. Het equivalent beschikbaar inkomen wordt in eerste instantie bepaald als de som van de beschikbare inkomens van alle leden van een sociologisch gezin, i.e. de som van alle inkomstenbronnen min de daarop betaalde bijdragen en belastingen. Deze som wordt vervolgens gedeeld door een equivalentieschaal die de samenstelling van het gezin op een gewogen manier samenvat.

Toepassingen waarvoor MIMOSIS in het verleden onder andere gebruikt werd, zijn:

1. De evaluatie van een maatregel uit het Generatiepact met betrekking tot de aanpassing van de regels voor toegelaten arbeid in combinatie met een overlevingspensioen (op vraag van de toenmalige minister van Pensioenen),
2. De evaluatie van een aantal scenario's voor de opwaardering van een aantal sociale minima tot de Europese armoedenorm (op vraag van het Rekenhof),
3. De evaluatie van de impact op armoede van de aanpassing van de sociale supplementen in de kinderbijslag (op vraag van de RKW en in het kader van het Europese voorzitterschap),
4. De raming van de bijdrage en belastingvoeten die van toepassing zijn op sociale uitkeringen (op vraag van Eurostat).²

Toepassingen waarvoor MIMOSIS nog niet gebruikt werd, maar die naar de toekomst toe bijzondere aandacht zullen genieten zijn deze die betrekking hebben op personen met een handicap vermits deze uitkeringen verwerkt worden door de FOD SZ.

¹ Meer informatie over het model is beschikbaar op volgend adres: <http://socialsecurity.fgov.be/nl/> onder de rubriek Publicaties, 2008, Mimosis (http://socialsecurity.fgov.be/nl/nieuws_publicaties/publicaties/mimosis/index.htm).

² Verslagen van een aantal van deze toepassingen zijn terug te vinden op <http://socialsecurity.fgov.be/nl/nieuws-publicaties/publicaties/mimosis/toepassingen.htm> of op verzoek bij de verantwoordelijke voor de data-aanvraag.

3 BESCHRIJVING VAN DE GEVRAAGDE GEGEVENS

3.1 TOEPASSINGSGEBIED

De noodzakelijke input voor het microsimulatiemodel is een steekproef met sociale zekerheidsgegevens die representatief is voor de Belgische bevolking. Om simulaties te kunnen uitvoeren is het niet noodzakelijk dat individuen die opgenomen worden in de steekproef exact geïdentificeerd kunnen worden. Variabelen die aanleiding zouden kunnen geven tot identificatie worden daarom ook geweerd uit de aanvraag. Dit neemt echter niet weg dat de variabelen die wel worden opgenomen in de aanvraag, zoals inkomensvariabelen, best zo exact mogelijk weergegeven worden. Dit met het oog op het bekomen van zinvolle simulatieresultaten. Om de aanvraag van exacte nominale gegevens te kunnen verantwoorden, werd voor de voorbereiding van deze data-aanvraag gewerkt met een tweetrapsprocedure. In een eerste stap werden, op basis van een kleine steekproef met niet exacte nominale gegevens, SAS programma's ontwikkeld. Deze programma's genereren afgeleide variabelen die doorgaans niet als dusdanig voorkomen in het Datawarehouse Arbeidsmarkt en Sociale Bescherming, maar die wel kunnen dienen als input voor de hoger beschreven simulatiemodellen. In een tweede stap worden deze afgeleide variabelen aangevraagd voor een ruimere steekproef dan deze die in de eerste stap gehanteerd werd.

De voorliggende aanvraag heeft betrekking op de gegevens die voor de tweede stap gevraagd worden. De manier waarop de steekproef voor deze tweede stap getrokken dient te worden, wordt meer in detail beschreven in het document met de titel "Voorstel trekking steekproef update MIMOSIS", dat bij deze aanvraag bijgevoegd wordt als bijlage 1.

3.2 GEGEVENS

De gevraagde gegevens hebben voornamelijk betrekking op de te modelleren onderdelen, i.e. sociale zekerheidsbijdragen, pensioenen, werkloosheidsuitkeringen, inkomensgaranties, ziekte- en invaliditeitsuitkeringen, familiale uitkeringen en de op de gerelateerde inkomenscomponenten verschuldigde personenbelasting. Daarnaast worden ook een aantal variabelen gevraagd om de bekomen simulatieresultaten te kunnen evalueren.

In de volgende sectie geven we een meer gedetailleerde motivatie waarom de verschillende variabelen aangevraagd worden. In deze sectie geven we enkel een overzicht van de variabelen die worden aangevraagd. We delen deze aanvraag verder op naar de bron op basis waarvan de afgeleide variabelen geconstrueerd werden. Per bron geven we een aantal gegevens van de gevraagde variabelen. In een eerste kolom staat de naam van de variabele zoals die door het programma gegenereerd wordt. De tweede kolom bevat een omschrijving van het gegeven en de derde kolom een lijst van de mogelijke waarden die het gegeven zou kunnen aannemen. Per bron geven we ook aan voor welk tijdstip of welke tijdstippen we de variabelen aanvragen. Het uitgangspunt bij het opstellen van deze aanvraag is dat, op het moment dat de aanvraag wordt ingediend, 2008 het meest recente jaar is waarvoor in het DWH AM&SB gegevens van de 4 verschillende kwartalen werden opgenomen.

Vermits de voorliggende vraag betrekking heeft op de gegevens die voor de tweede stap gevraagd worden, voegen we in iedere deelsectie ook een onderdeel toe dat betrekking heeft op de programma's waarmee de gevraagde gegevens gegenereerd kunnen worden.

3.2.1 BRON ALGEMENE DIRECTIE STATISTIEK EN ECONOMISCHE INFORMATIE (ADSEI), RIJKSREGISTER EN AFGELEIDE VARIABLEN UIT HET DATAWAREHOUSE ARBEIDSMARKT EN SOCIALE BESCHERMING (DWH AM&SB)

Voor alle individuen die behoren tot de steekproef, worden volgende gegevens gevraagd zoals gekend op 1 januari &Year+1:

	Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu	Fictief nummer	Gegeven nodig om te kunnen koppelen met andere bestanden
2	NAREGNIS_REFPERS	Geanonimiseerd identificatienummer van de referentiepersoon van het gezin waartoe het individu behoort op 1 januari &Year+1	Fictief nummer	
3	NAREGNIS_HH_ID	Geanonimiseerd identificatienummer van het gezin waartoe het individu behoort op 1 januari &Year+1	Fictief nummer	
4	NAREGNIS_RND_SLCT	Aanduiding of het individu geselecteerd werd bij het trekken van de steekproef (= ja of nee)	0 neen 1 ja	Om het juiste steekproefgewicht van individuen te construeren is het van belang te weten wie getrokken werd voor de oorspronkelijke steekproef en wie er later bijgetrokken werd.
5	NAREGNIS_RELATION	Verband tussen het individu en het huishoudhoofd op 1 januari &Year+1	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	Variabele speelt cruciale rol in de module FAMREL en de definitie van belastingeenheden in de PIT module. In de FAMREL module wordt het sociologisch gezin ingedeeld in kleinere familiekeren. Om deze familiekeren te kunnen aanmaken moet de variabele Relatiecode geobserveerd worden in zijn beschreven detail.

6	NAREGNIS_HEADOFHH	Aanduiding of het individu op 1 januari &Year+1 gezinshoofd is van het gezin waartoe het individu behoort	0 neen 1 ja	
8	NAREGNIS_DATBIRTH	Geboortedatum van het individu in jaar en maand	mm-JJJJ	Geboortedatum is zowel van belang voor de berekening van bepaalde voordelen, zoals kinderbijslag als voor de evaluatie van maatregelen in functie van leeftijd.
9	NAREGNIS_SEX	Geslacht van het individu op 1 januari &Year+1	1 man 2 vrouw	Variabele is van belang bij het evalueren van bepaalde maatregelen, i.e. om te bekijken of een maatregel een verschillend impact-effect naar geslacht.
10	NAREGNIS_COMMUNIT	NIS gemeentecode van de woonplaats waar het individu op 1 januari &Year+1 woonde		Variabele wordt gevraagd om in de PIT module de gemeente en agglomeratiebelasting te kunnen modelleren en om de resultaten te kunnen voorstellen voor verschillende regionale indelingen.
11	NAREGNIS_TYPEHH	Typologie van het gezin waartoe het individu behoort op 1 januari &Year+1	1 Eenpersoonshuishouden 2 Gehuwd paar zonder kinderen 3 Gehuwd paar met kinderen 4 Ongehuwd paar zonder kinderen 5 Ongehuwd paar met kinderen 6 Eenoudergezin 7 Overige huishoudens	Deze variabele werd niet gevraagd voor de vorige versie van MIMOSIS. In een update zal hij gebruikt kunnen worden voor het evalueren van simulaties.
12	NAREGNIS_COLL	Aanduiding of het gezin waartoe het individu op 1 januari &Year+1 behoort een collectief gezin is	0 neen 1 ja	Om een correcte welvaartsevaluatie te kunnen uitvoeren is het van belang te weten of een individu in een privaat of een collectief gezin woont.
13	NAREGNIS_LIPRO	Typologie van het individu op basis van eigenschappen van het gezin waartoe het individu behoort op 1 januari &Year+1	SING alleenstaande MAR0 gehuwde zonder kind MAR+ gehuwde met kind CMAR kind bij gehuwd paar UNM0 ongehuwd samenwonende zonder kind UNM+ ongehuwd samenwonenden met kind(eren) CUNM kind bij ongehuwd paar HIPA hoofd eenoudergezin	Deze variabele werd niet gevraagd voor de vorige versie van MIMOSIS. In een update zal hij gebruikt kunnen worden voor het evalueren van simulaties.

			C1PA kind in eenoudergezin NFRP andere inwonende OTHR overige personen COLL wonend in collectief huishouden	
14	NAREGNIS_CIVIL	Burgerlijke staat van het individu op 1 januari & Year+1	0 andere 1 ongehuwd 2 gehuwd 3 weduwnaar, weduwe 4 uit de echt gescheiden 5 gescheiden van tafel en bed	Variabele wordt gebruikt in de module FAMREL en bij de definitie van belastingeenheden in de PIT module. Gegeven was beschikbaar voor de vorige versie maar niet voor de huidige omdat het niet opgenomen is in de lijst met variabelen uit het Rijksregister die door ADSEI aan KSZ geleverd worden.

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma "ManipPersInfoToNewData.sas".

Voor alle individuen die behoren tot de oorspronkelijke steekproef (i.e. trekking op 1 januari 2002) worden volgende gegevens gevraagd:

	Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu	Fictief nummer	Gegeven nodig om te kunnen koppelen met andere bestanden
2	NAREGNIS_REFPERS_2002	Geanonimiseerd identificatienummer van de referentiepersoon van het gezin waartoe het individu behoort op 1 januari 2002	Fictief nummer	
3	NAREGNIS_HH_ID_2002	Geanonimiseerd identificatienummer van het gezin waartoe het individu behoort op 1 januari 2002	Fictief nummer	
4	NAREGNIS_RND_SLCT_2002	Aanduiding of het individu geselecteerd werd bij het trekken van de oorspronkelijke steekproef (= ja of nee)	0 Neen 1 Ja	Gegeven laat toe om de oorspronkelijke steekproefgewichten te reconstrueren.
5	NAREGNIS_RELATION_2002	Verband tussen het individu en het huishoudhoofd op 1 januari 2002	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	Gegeven moet toelaten om de oorspronkelijke familiekeren te vergelijken met nieuwe familiekeren.
14	NAREGNIS_CIVIL_2002	Burgerlijke staat van het individu op 1 januari 2002	0 andere 1 ongehuwd 2 gehuwd 3 weduwnaar, weduwe	Gegeven laat toe om oorspronkelijke belastingeenheden aan te maken en om schatting te maken in het geval burgerlijke staat op een later

			4 uit de echt gescheiden	
			5 gescheiden van tafel en bed	

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma "ManipRRRegOldToNewData.sas".

3.2.2 BRON FONDS VOOR ARBEIDSONGEVALLEN (FAO)

Voor alle individuen die behoren tot de omschreven steekproef, worden volgende gegevens gevraagd. Het gaat om gegevens op jaarbasis die geconstrueerd werden door de jaargegevens samen te voegen met de gegevens van het laatste ongeval dat geobserveerd wordt in één van de 4 kwartaal-bestanden van &Year.

1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu	Fictief nummer	Gegeven nodig om te kunnen koppelen met andere bestanden
	TAUXIPRES_&Year	Graad van permanente arbeidsongeschiktheid (geregistreerd in &Year)	Waarden zoals in DWH AM&SB	Gegeven nodig om de uitkering te kunnen herberekenen
	TAUXATRES_&Year	Graad voor hulp van derden (geregistreerd in &Year)	Waarden zoals in DWH AM&SB	Gegeven nodig om de uitkering te kunnen herberekenen
	TAUXIT_&Year	Graad van tijdelijke arbeidsongeschiktheid (geregistreerd in &Year)	Waarden zoals in DWH AM&SB.	Gegeven nodig om de uitkering te kunnen herberekenen
	DDEBUT_&Year	Begin van de periode van ongeschiktheid (geregistreerd in &Year)	mm-JJJJ	Gegeven laat toe de ontvangen uitkering toe te wijzen aan een gegeven jaar
	DFIN_&Year	Einde van de periode van ongeschiktheid (geregistreerd in &Year)	mm-JJJJ	Gegeven laat toe de ontvangen uitkering toe te wijzen aan een gegeven jaar
	DUREEITT_&Year	Aantal dagen tijdelijke arbeidsongeschiktheid: volledige afwezigheid (geregistreerd in &Year)	Waarden zoals in DWH AM&SB	Gegeven nodig om de uitkering te kunnen herberekenen
	DUREEITP_&Year	Aantal dagen tijdelijke arbeidsongeschiktheid: gedeeltelijke afwezigheid (geregistreerd in &Year)	Waarden zoals in DWH AM&SB	Gegeven nodig om de uitkering te kunnen herberekenen
	MONTPERTSAL_&Year	Bedrag van het verloren loon (geregistreerd in &Year)	Bedrag in klassen van 10 euro	Gegeven nodig om de uitkering te kunnen herberekenen
	SALBASE_&Year	Voorgesteld loon dat dient als basis voor de berekening van de uitkering (geregistreerd in &Year)	Bedrag in klassen van 10 euro	Gegeven nodig om de uitkering te kunnen herberekenen
	MONTITT_&Year	Bedrag voor tijdelijke arbeidsongeschiktheid: volledige afwezigheid (geregistreerd in &Year)	Bedrag in klassen van 10 euro	Gegeven laat toe om uitbetaald bedrag als basis te nemen voor herberekening
	MONTITP_&Year	Bedrag voor tijdelijke arbeidsongeschiktheid: gedeeltelijke afwezigheid (geregistreerd in &Year)	Bedrag in klassen van 10 euro	Gegeven laat toe om uitbetaald bedrag als basis te nemen voor herberekening
	CATPROFVICT_&Year	Beroeps categorie waartoe het individu op het moment van het ongeval behoort (geregistreerd in &Year)	Waarden zoals in DWH AM&SB	Gegeven laat toe om de berekende uitkering toe te wijzen aan de juiste sector (gegeven moet toelaten om onderscheid tussen arbeiders en bedienden te maken)

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma "ManipFAOToNewData.sas".

3.2.3 BRON FONDS VOOR BEROEPSZIEKTEN (FBZ)

Voor alle individuen die behoren tot de omschreven steekproef, worden optelbare bedragen opgeteld per jaar en gecombineerd met de karakteristieken van de lijn met het hoogste bedrag. De vroegste begindatum en de laatste einddatum van de erkenningsperiode en de betalingsperiode worden weerhouden. Volgende gegevens worden gevraagd voor het jaar &Year:

	Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu	Fictief nummer	Gegeven nodig om te kunnen koppelen met andere bestanden
	FBZ_PERCENT_&Year	Percentage arbeidsongeschiktheid als gevolg van beroepsziekte (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt in de eerste plaats gevraagd om de regels die van toepassing zijn voor de bepaling van de uitkering, te kunnen modeleren. Daarnaast zou het gegeven kunnen gebruikt worden om een domeinoverschrijdende indicator van verminderde arbeidsmarktcapaciteit af te leiden (zie module nota SICK)
	FBZ_SRTUIT_&Year	Soort uitkering beroepsziekte (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de regels die van toepassing zijn voor de bepaling van de uitkering, te kunnen modeleren. (zie module nota SICK)
	FBZ_BASLOON_&Year	Basisloon gebruikt voor bepaling beroepsziekte (&Year)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de regels die van toepassing zijn voor de bepaling van de uitkering, te kunnen modeleren. (zie module nota SICK)
	FBZ_SRTPER_&Year	Soort periode beroepsziekte (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het geobserveerde bedrag om te kunnen zetten naar een correct maandbedrag. (zie module nota SICK)
	FBZ_BEGDBET_&Year	Begindatum betaling beroepsziekte (&Year)	mm-JJJJ	Gegeven wordt gevraagd om het geobserveerde bedrag om te kunnen zetten naar een correct maandbedrag. (zie module nota SICK)
	FBZ_ENDDBET_&Year	Einddatum betaling beroepsziekte (&Year)	mm-JJJJ	Gegeven wordt gevraagd om het geobserveerde bedrag om te kunnen zetten naar een correct maandbedrag. (zie module nota SICK)

	FBZ_BEDRAG_&Year	Betaald bedrag beroepsziekte (&Year)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd het beschikbaar inkomen mee te bepalen. (zie module nota SICK)
--	------------------	--------------------------------------	-------------------------------	---

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma "ManipFBZToNewData.sas".

3.2.4 BRON NATIONAAL INTERMUTUALISTISCH COLLEGE (NIC)

Voor alle lijnen met gelijke info voor karakteristieken per individu worden gegevens voor aantal dagen en bedragen opgeteld op jaarbasis en wordt de vroegste begindatum en de laatste einddatum weerhouden. Volgende gegevens worden gevraagd voor het jaar & Year:

	Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie				
1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu	Fictief nummer	Gegeven nodig om te kunnen koppelen met andere bestanden				
2	NIC_SRTDAGEN_&Year	Soort dagen (PAO, zwangerschap, ...) (&Year)	1 Incapacité primaire – Primaire arbeidsongeschiktheid Ecartement du travail : grossesse - Werkverwijdering wegens zwangerschap 3 5 Accouchement - Bevallingsrust Ecartement du travail : allaitement - Werkverwijdering wegens borstvoeding 6 Ecartement du travail : maladie contagieuse - Werkverwijdering wegens besmettelijke ziekten 7 8 Paternité - Vaderschap Paternité: nouveau - Nieuw vaderschapsverlof 9 10 Adoption - Adoptieverlof Zonder uitkering (periode andere code « niet vergoed » met inbegrip gewaarborgd loon) 00	Gegeven wordt gevraagd om de uitkering te kunnen berekenen. De rekenregels variëren met het soort uitkering..				
3	NIC_STELSEL_&Year	Stelsel waarin het recht geklasseerd wordt (&Year)	<table border="1"> <tr> <td>1</td> <td>Algemene regeling</td> </tr> <tr> <td>2</td> <td>Zelfstandige regeling</td> </tr> </table>	1	Algemene regeling	2	Zelfstandige regeling	Gegeven wordt gevraagd om het stelsel te kunnen bepalen waaronder de verzekerde valt. De rekenregels voor de bepaling van de uitkering hangen af van het stelsel.
1	Algemene regeling							
2	Zelfstandige regeling							
4	NIC_DAGENAO_&Year	Aantal dagen arbeidsongeschiktheid (&Year)	Aantal dagen in Arbeidsongeschiktheid	Gegeven wordt gevraagd om de uitkering te kunnen berekenen. De uitkering hangt namelijk af van het aantal dagen ongeschiktheid.				
8	NIC_BEGDATUM_&Year	Begindatum ziekte (&Year)	mm-JJJJ	Gegeven wordt gevraagd als back-up voor het gegeven aantal_dagen_ao. In de huidige versie van MIMOSIS wordt aantal dagen afgeleid op basis van de begin en einddatum van				

				arbeidsongeschiktheid.
9	NIC_ENDDATUM_&Year	Einddatum ziekte (&Year)	mm-JJJ	Gegeven wordt gevraagd als back-up voor het gegeven aantal_dagen_ao. In de huidige versie van MIMOSIS wordt aantal dagen afgeleid op basis van de begin en einddatum van arbeidsongeschiktheid.
10	NIC_BEDRAG_&Year	Betaald bedrag aan uitkeringen (&Year)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om het beschikbaar inkomen mee te bepalen.

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma "ManipNICToNewData.sas".

3.2.5 BRON PROGRAMMATORISCHE FEDERALE OVERHEIDSDIENST MAATSCHAPPELIJKE INTEGRATIE (POD MI)

Voor alle lijnen met gelijke info voor karakteristieken per individu worden gegevens voor optelbare variabelen opgeteld op jaarbasis en wordt de vroegste begindatum en de laatste einddatum weerhouden. Volgende gegevens worden gevraagd voor het jaar &Year:

	Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu	Fictief nummer	Gegeven nodig om te kunnen koppelen met andere bestanden
3	PODMI_BEDRAG_&Year	Bedrag betaling recht POD MI (&Year)	Bedrag in klassen van 10 euro	Het bedrag wordt gevraagd omdat het een ondergrens zou moeten zijn van wat aan een individu wordt uitbetaald.
5	PODMI_BUDART_&Year	Percentage & omschrijving van de terugbetaling door de Staat aan de OCMW's (begrotingsartikel) (&Year)	Waarden zoals in DWH AM&SB	Gegeven laat toe om het onderscheid te maken tussen verschillende soorten uitkeringen: leefloon, equivalent leefloon, activering, tewerkstelling en individuele integratiemaatregelen. In combinatie met variabelen typamt_id en legislation bepaal je type hulp.
7	PODMI_LEGIS_&Year	Wetgeving van toepassing voor bepaling recht (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de betaalde uitkering te kunnen herberekenen.
8	PODMI_CATID_&Year	Gezinscategorie voor bepaling recht POD MI (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de betaalde uitkering te kunnen herberekenen.
9	PODMI_CATCNTID_&Year	Gezinscategorie tot dewelke het individu gerekend wordt voor telling recht POD MI (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de betaalde uitkering te kunnen toewijzen.
10	PODMI_STSID_&Year	Statuut van begunstigde van de financiële steun equivalent aan het leefloon (&Year)	Waarden zoals in DWH AM&SB	Gegeven laat toe om een onderscheid te maken tussen socio economische toestand van individuen op het moment van betaling
11	PODMI_THP_&Year	Soort maatschappelijke hulpverlening toegekend krachtens de wet van 2 april 1965 (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om maatschappelijke hulp te onderscheiden van andere uitkeringen.

12	PODMI_BEGDATUM_&Year	Begindatum betaling recht POD MI (&Year)	mm-JJJJ	Gegeven laat toe om de uitkering toe te wijzen aan een bepaald jaar
13	PODMI_ENDDATUM_&Year	Einddatum betaling recht POD MI (&Year)	mm-JJJJ	Gegeven laat toe om de uitkering toe te wijzen aan een bepaald jaar

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma "ManipPODMIToNewData.sas".

3.2.6 BRON RIJKSINSTITUUT VOOR ZIEKTE EN INVALIDITEIT (RIZIV)

Voor alle individuen in de steekproef worden gegevens uit het betalingsbestand gescheiden naar stelsel en enkel de meest voorkomende betalingscode wordt weerhouden. Andere nominale gegevens zoals het betalingsbedrag en de vergoede dagen worden opgeteld over de andere lijnen. Van de gegevens uit het basisbestand wordt enkel de lijn met de meest recente ziektedatum weerhouden. Volgende gegevens worden gevraagd voor het jaar &Year:

	Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu	Fictief nummer	Gegeven nodig om te kunnen koppelen met andere bestanden
	RIZIVINA_BETALING_INVALID_&Year	Betalingscode (persoon in invaliditeit) (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de uitkering waar een individu recht op heeft te typeren (zie module nota SICK).
	RIZIVINA_STELSEL_INVALID_&Year	Code die type verzekering weergeeft (persoon in invaliditeit) (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het stelsel te bepalen waaronder een individu valt.
	RIZIVINA_AANVANGS_INVALID_&Year	Aanvangsdatum van de betalingsperiode (persoon in invaliditeit) (&Year)	mm-JJJJ	Gegeven wordt gevraagd om af te bakenen of het begin van de betaling in jaar van berekening ligt.
	RIZIVINA_EINDBETA_INVALID_&Year	Einddatum van de betalingsperiode (persoon in invaliditeit) (&Year)	mm-JJJJ	Gegeven wordt gevraagd om af te bakenen of het einde van de betaling in jaar van berekening ligt.
	RIZIVINA_BEDRAG_INVALID_&Year	Uitgekeerd bedrag (persoon in invaliditeit) (&Year)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om het beschikbaar inkomen mee te bepalen.
	RIZIVINA_VERGOEDE_INVALID_&Year	Vergoede dagen (persoon in invaliditeit) (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om te vergelijken met het aantal dagen binnen het jaar dat in aanmerking komen voor betaling (zie variabelen Aanvangsdatum_betalingsperiode en Einddatum_betalingsperiode)
	RIZIVINA_BETALING_MAT_&Year	Betalingscode (zwangerschapsverlof gedurende invaliditeit) (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de uitkering waar een individu recht op heeft te typeren (zie module nota SICK).

RIZIVINA_STELSEL_INVALID_&Year	Code die type verzekering weergeeft (zwangerschapsverlof gedurende invaliditeit) (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het stelsel te bepalen waaronder een individu valt.
RIZIVINA_AANVANGS_MAT_&Year	Aanvangsdatum van de betalingsperiode (zwangerschapsverlof gedurende invaliditeit) (&Year)	mm-JJJJ	Gegeven wordt gevraagd om af te bakenen of het begin van de betaling in jaar van berekening ligt.
RIZIVINA_EINDBETA_MAT_&Year	Einddatum van de betalingsperiode (zwangerschapsverlof gedurende invaliditeit) (&Year)	mm-JJJJ	Gegeven wordt gevraagd om af te bakenen of het einde van de betaling in jaar van berekening ligt.
RIZIVINA_BEDRAG_MAT_&Year	Uitgekeerd bedrag (zwangerschapsverlof gedurende invaliditeit) (&Year)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om het beschikbaar inkomen mee te bepalen.
RIZIVINA_VERGOEDE_MAT_&Year	Vergoede dagen (zwangerschapsverlof gedurende invaliditeit) (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om te vergelijken met het aantal dagen binnen het jaar dat in aanmerking komen voor betaling (zie variabelen Aanvangsdatum_betalingsperiode en Einddatum_betalingsperiode)
B_ZIEKTE_&Year	Begindatum primaire arbeidsongeschiktheid (&Year)	mm-JJJJ	Gegeven wordt gevraagd om het aantal weken dat er reeds een uitkering voor ziekte en invaliditeit wordt uitbetaald, te berekenen (zie module nota SICK).
CUMUL_&Year	Code die aangeeft of persoon die uitkering ontvangt wegens invaliditeit ook een uitkering ontvangt naar aanleiding van arbeidsongeval, beroepsziekte of ongeval (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om na te gaan of er gelijktijdig ook nog een anderen uitkering ontvangen wordt
AANWIJZ_&Year	Gegevens met betrekking tot de gezinslast en de arbeidsmarktpositie van het individu op het moment van de aanvraag tot erkenning invaliditeit (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de gezinslast van de invalide te bepalen (zie module nota SICK).

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma "ManipRIZIVToNewData.sas".

3.2.7 BRON RIJKSDIENST VOOR KINDERBIJSLAG VOOR WERKNEMERS (RKW)

Voor alle individuen in de steekproef worden de gegevens geordend op het niveau van het rechtgevend kind. Per kind wordt een begindatum en een einddatum van betaling bepaald in combinatie met een rechthebbende en een bijslagtrekkende. Volgende gegevens worden gevraagd voor het jaar &Year:

	Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu	Fictief nummer	Gegeven nodig om te kunnen koppelen met andere bestanden
	BEGINRKW_&Year	Begindatum betaling kinderbijslag voor kind in een ander stelsel dan dat van de zelfstandigen (&Year)	mm-JJJJ	Gegeven wordt gevraagd om te bepalen hoeveel maanden in een bepaald jaar in aanmerking genomen moeten worden.
	EINDRKW_&Year	Einddatum betaling kinderbijslag voor kind in een ander stelsel dan dat van de zelfstandigen (&Year)	mm-JJJJ	Gegeven wordt gevraagd om te bepalen hoeveel maanden in een bepaald jaar in aanmerking genomen moeten worden.
	RH_RKW_&Year	Geanonimiseerd identificatienummer van het individu dat recht op kinderbijslag opent in een ander stelsel dan dat van de zelfstandigen (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de kinderbijslaggegevens te koppelen aan andere bronnen (zie module nota FAMAL)
	BT_RKW_&Year	Geanonimiseerd identificatienummer van het individu dat kinderbijslag ontvangt in een ander stelsel dan dat van de zelfstandigen (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de kinderbijslaggegevens te koppelen aan andere bronnen (zie module nota FAMAL)
	RG_RKW_&Year	Geanonimiseerd identificatienummer van het kind dat recht op kinderbijslag genereert in een ander stelsel dan dat van de zelfstandigen (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de kinderbijslaggegevens te koppelen aan andere bronnen (zie module nota FAMAL)
	RKW_GWB_&Year	Individueel ontvangt kinderbijslag in stelsel van gewaarborgde kinderbijslag (&Year.)	Indien Bureau = "09915" of Bureau = "09918" dan waarde 1, anders 0	Gegeven gevraagd om te bepalen of individu recht heeft op gewaarborgde kinderbijslag of niet

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma "ManipRKWToNewData.sas".

3.2.8 BRON RIJKSINSTITUUT VOOR DE SOCIALE VERZEKERING DER ZELFSTANDIGEN (RSVZ)

Voor alle individuen in de steekproef worden de gegevens samengevoegd tot één lijn per individu indien er meerdere lijnen zouden voorkomen. Volgende gegevens worden gevraagd voor ieder van de 4 kwartalen van &Year:

	Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu	Fictief nummer	Gegeven nodig om te kunnen koppelen met andere bestanden
	BEROEPC_&Year	Beroepscode bij RSVZ (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om resultaten later te kunnen analyseren naar beroeps categorie en om eventueel ontbrekende inkomensgegevens aan te vullen.
	NACEC_&Year	Activiteit van een zelfstandige weergegeven onder de vorm van een NACE code (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om resultaten later te kunnen analyseren naar activiteitssector van de zelfstandige.
	BIJDCAT_&Year	Code van de bijdragecategorie bij RSVZ (in hoofdberoep, bijberoep, na pensioenleeftijd, ...) (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gebruikt om de bijdragecategorie te bepalen (zie module nota CONTRIB)
	HOEDRSVZ_&Year	Code van hoedanigheid bij RSVZ (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gebruikt om de bijdragecategorie te bepalen (zie module nota CONTRIB)
	BEGAANSC_&Year	Aanvangsdatum van de aansluiting bij de RSVZ (&Year)	mm-JJJJ	Gegeven wordt gebruikt om te bepalen onder welk bijdrageregime een individu valt (starter of niet) en om het aantal aangesloten jaren te kennen
	SCHAANSC_&Year	Datum van het einde van de aansluiting bij de RSVZ (&Year)	mm-JJJJ	Gegeven wordt gevraagd om te toetsen of er een einde van de activiteiten voorzien wordt in een gegeven jaar

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma "ManipRSVZToNewData.sas".

Voor alle individuen in de steekproef worden de gegevens geordend op het niveau van het rechtgevend kind. Per kind wordt een begindatum en een einddatum van betaling bepaald, één rechthebbende en één bijslagtrekkende. Volgende gegevens worden gevraagd voor ieder van de 4 kwartalen van &Year:

Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu	Fictief nummer	Gegeven nodig om te kunnen koppelen met andere bestanden
BT_HOED_&Year	Hoedanigheid van de persoon die kinderbijslag ontvangt in het stelsel der zelfstandigen	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de hoedanigheid van het individu in het kinderbijslagstelsel van de zelfstandigen te bepalen (zie module nota FAMAL)
RH_RSVZ_&Year	Geanonimiseerd identificatienummer van het individu dat recht op kinderbijslag opent in het stelsel der zelfstandigen (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de kinderbijslaggegevens te koppelen aan andere bronnen (zie module nota FAMAL)
BT_RSVZ_&Year	Geanonimiseerd identificatienummer van het individu dat kinderbijslag ontvangt in het stelsel der zelfstandigen (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de kinderbijslaggegevens te koppelen aan andere bronnen (zie module nota FAMAL)
RG_RSVZ_&Year	Geanonimiseerd identificatienummer van het individu dat kinderbijslag genereert in het stelsel der zelfstandigen (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de kinderbijslaggegevens te koppelen aan andere bronnen (zie module nota FAMAL)
QUART_RSVZKB_&Year.	Aanduiding van de kwartalen van het jaar &Year. waarin de gegeven combinatie geobserveerd wordt in RSVZ kinderbijslagbestanden		Gegeven gevraagd om betalingskwartalen te kennen

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma "ManipRSVZ_KBToNewData.sas".

Voor alle individuen die behoren tot de omschreven steekproef, worden volgende gegevens gevraagd zoals gekend op het einde van ieder tellingsjaar vanaf &Year-3 tot en met &Year.

	Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu	Fictief nummer	Gegeven nodig om te kunnen koppelen met andere bestanden
	jaar_x	Jaar waarin het netto bedrijfsinkomen gegenereerd werd	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het geobserveerde jaarinkomen te kunnen infleren naar het niveau van het bijdragejaar waarvoor berekeningen worden uitgevoerd
	inkomens_x	Netto bedrijfsinkomsten van het jaar geregistreerd in de variabele jaar	Bedrag in klassen van 10 euro	De gegevens worden gevraagd voor het bijdragejaar zelf (i.e. &Year) als daaraan voorafgaande jaren. Dit laatste om, in de mate van het mogelijke, pensioenrechten van zelfstandigen te kunnen berekenen.

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma “ManipRSVZIncToNewData.sas”.

3.2.9 BRON RIJKSDIENST VOOR SOCIALE ZEKERHEID (RSZ)

Voor alle individuen die behoren tot de omschreven steekproef, worden per tewerkstellingslijn volgende gegevens uit het basisbestand, het bestand statbasef, statbasem en statbaseplus samengevoegd voor het 4^{de} kwartaal (&Q=4) van &Year.

	Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu		Gegeven nodig om te kunnen koppelen met andere bestanden
	secemp_&Year&Q	sector (1-privé 2-openbare)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om bijdragen aan de juiste sector te kunnen toewijzen en gegevens met externe bronnen te kunnen vergelijken (zie module nota CONTRIB)
	nr_cp_&Year&Q	nummer paritair comité	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om te weten welk soort structurele deductie precies mag toegepast worden (zie module nota CONTRIB)
	codtra_&Year&Q	werknemerscode	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het statuut van de werknemer, met het oog op de bepaling van de bijdragen, te bepalen (arbeider, bediende, contractueel of niet)
	clatra_&Year&Q	werknemersklasse	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het statuut van de werknemer, met het oog op de bepaling van de bijdragen, te bepalen (arbeider, bediende, contractueel of niet)
	clatr2_&Year&Q	speciale klasse werknemer	Waarden zoals in DWH AM&SB	Gegeven wordt gebruikt om te zien of een werknemer een gesubsidieerd contractueel is (zie module nota CONTRIB)
	jrsrem_f_&Year&Q	aantal normale bezoldigde fulltimedagen	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het gewerkt aantal uren te kunnen berekenen
	jrsrem_p_&Year&Q	aantal bezoldigde parttimedagen	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het gewerkt aantal uren te kunnen berekenen
	jrspre_&Year&Q	aantal bezoldigde dagen vooropzeg	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om

				pensioenrechten te bepalen bij het ontbreken van info in andere bronnen
	jrsvac_&Year&Q	aantal bezoldigde vakantiedagen	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om pensioenrechten te bepalen bij het ontbreken van info in andere bronnen
	jrsasm_&Year&Q	aantal bezoldigde gelijkgestelde dagen	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om pensioenrechten te bepalen bij het ontbreken van info in andere bronnen en om eventuele uitkeringen ziekte en invaliditeit te bepalen
	codasm_&Year&Q	hoofdcode van de gelijkgestelde dagen	Waarden zoals in DWH AM&SB	Code wordt gevraagd om na te gaan of bepaalde gelijkgestelde dagen in aanmerking komen voor de bepaling van een uitkering voor ziekte en invaliditeit (zie module nota SICK)
	sal100_&Year&Q	loon 100%	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd omdat het een component is van de berekeningsbasis voor de bepaling van de bijdragen (zie module nota CONTRIB)
	primes_&Year&Q	premies	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd omdat het een component is van de berekeningsbasis voor de bepaling van de bijdragen (zie module nota CONTRIB)
	preavi_&Year&Q	vooropzeg	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd omdat het een component is van de berekeningsbasis voor de bepaling van de bijdragen (zie module nota CONTRIB)
	salatt_&Year&Q	wachtloon	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd omdat het een component is van de berekeningsbasis voor de bepaling van de bijdragen (zie module nota CONTRIB)
	salfor_&Year&Q	forfaitair loon	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd omdat het een component is van de berekeningsbasis voor de

				bepaling van de bijdragen (zie module nota CONTRIB)
	saljrs_&Year&Q	berekend dagloon	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om simulaties in functie van gewerkt aantal uren maal verdiend loon te kunnen uitvoeren
	t_prest_&Year&Q	type prestatie	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om aan te duiden of iemand al dan niet voltijds werkt
	tauxpt_&Year&Q	percentage parttime	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om afgeleide rechten, verbonden aan deeltijds werk, te kunnen schatten en eventueel te vergelijken met gegevens inzake tijdskrediet en loopbaanonderbreking
	tauxsa_&Year&Q	percentage zonder gelijkgestelde dagen	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om werkbonus te reduceren in functie van de gepresteerde arbeidstijd
	tauxaa_&Year&Q	percentage met gelijkgestelde dagen	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om werkbonus te reduceren in functie van de gepresteerde arbeidstijd
	t_dnrp_&Year&Q	type dnrp	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de tewerkstellingslijn te identificeren (combinatie fictif, t_dnrp, r_exclus en pp_rrnmc)
	r_exclus_&Year&Q	reden voor uitsluiting dnrp	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de tewerkstellingslijn te identificeren (combinatie fictif, t_dnrp, r_exclus en pp_rrnmc)
	pp_rrnmc_&Year&Q	hoofdprestatie van de combinatie rrn/stamnummer/kengetal	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de tewerkstellingslijn te identificeren (combinatie fictif, t_dnrp, r_exclus en pp_rrnmc)
	codred_&Year&Q	verminderingcode	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om te identificeren op welke vermindering iemand recht heeft (zie module nota

				CONTRIB)
	mntded_&Year&Q	gecumuleerd bedrag van de vermindering	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de gesimuleerde vermindering mee te vergelijken en om het beschikbaar inkomen te berekenen.
	cotpat_&Year&Q	bedrag werkgeversbijdrage	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de gesimuleerde werkgeversbijdrage mee te vergelijken en om de loonkost (i.e. bruto loon + werkgeversbijdrage + bijzondere bijdrage) te berekenen
	cotper_&Year&Q	bedrag persoonlijke bijdrage	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de gesimuleerde persoonlijke bijdrage mee te vergelijken en om het belastbaar inkomen (i.e. bruto inkomen - persoonlijke bijdrage) te berekenen
	cotspe_&Year&Q	bedrag bijzondere bijdragen	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de loonkost (i.e. bruto loon + werkgeversbijdrage + bijzondere bijdrage) te berekenen
	maatman_&Year&Q	Maatman	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de bijdragebasis voor de werkbonus te kunnen bepalen.
	DeductionCode_&Year&Q	Code van de bijdragevermindering	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om bijdragevermindering te kunnen herberekenen
	DeductionCalculationBasis_&Year&Q	Basis waarop de bijdragevermindering wordt berekend	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om bijdragevermindering te kunnen herberekenen
	DeductionAmount_&Year&Q	Bedrag van de toegestane bijdragevermindering	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de herberekende bijdragevermindering mee te vergelijken
	mpecvac_&Year&Q	geaggregeerd dubbel vakantiegeld	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om te controleren wat de afwijking is tussen het herberekende dubbel

				vakantiegeld en het geobserveerde aggregaat
	mcotdpv_&Year&Q	Bijdrage op geaggregeerd dubbel vakantiegeld	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om te controleren wat de afwijking is tussen de herberekende bijdrage op het dubbel vakantiegeld en het geobserveerde aggregaat
	CATTRA_&Year&Q	Werkgeverskengetal	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om bijdragen te kunnen herberekenen
	InsEtab_&Year&Q.	Arrondissement van de lokale vestigingseenheid	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om positie woon-werkverkeer te kunnen inschatten
	CONTRIBUTIONCODE_&Year &I met i= 1 tot 4.	Bijdragecode van bijdragen van de werkgever aan winstparticipaties, bedrijfsvoertuigen en extralegale pensioenen (in kwartaal &i van &Year)	Codes 861, 862 en 851	Gegeven geeft indicatie van betaling bijdragen tweede pijler
	CONTRIBUTIONCALCULATIONBASIS_&Year&i met i= 1 tot 4.	Berekeningsbasis van bijdragen van de werkgever aan winstparticipaties, bedrijfsvoertuigen en extralegale pensioenen (in kwartaal &i van &Year): gegeven is aggregaat op niveau van de werkgever	Waarden zoals in DWH AM&SB	Gegeven geeft indicatie van betaling bijdragen tweede pijler
	CONTRIBUTIONAMOUNT_&Year&i met i= 1 tot 4.	Bedrag van bijdragen van de werkgever aan winstparticipaties, bedrijfsvoertuigen en extralegale pensioenen (in kwartaal &i van &Year) : gegeven is aggregaat op niveau van de werkgever	Waarden zoals in DWH AM&SB	Gegeven geeft indicatie van betaling bijdragen tweede pijler

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma “ManipRSZQuartBasisToNewData.sas”.

Voor alle individuen die behoren tot de omschreven steekproef, worden per individu een aantal gegevens geaggregeerd op jaarbasis. Volgende gegevens worden gevraagd voor ieder jaar vanaf 2002 tot en met &Year (in onderstaande tabel is $x = 2002$ tot en met &Year)

	Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu		Gegeven nodig om te kunnen koppelen met andere bestanden
	(Sal100_x + Salatt_x + Primes_x + Preavi_x)	Het bedrag van de bezoldigingen op jaarbasis (som voor jaar = x)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om individuen te kunnen klasseren naar arbeidsmarktverleden (in het bijzonder voor de afleiding van het laatst verdiende inkomen) (zie GrossIncomeHour.doc)
	GeschatDubbelVakantieGeld_x	Dubbel vakantiegeld geschat rekening houdend met arbeidsmarktstatuut van het individu (som voor jaar = x)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om individuen te kunnen klasseren naar arbeidsmarktverleden (in het bijzonder voor de afleiding van het laatst verdiende inkomen) (zie GrossIncomeHour.doc)
	GeschatEnkelVakantieGeld_x	Geschat enkel vakantiegeld van arbeiders (som voor jaar = x)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om individuen te kunnen klasseren naar arbeidsmarktverleden (in het bijzonder voor de afleiding van het laatst verdiende inkomen) (zie GrossIncomeHour.doc)
	heurpt_x	Aantal uren deeltijds (som voor jaar = x)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om individuen te kunnen klasseren naar arbeidsmarktverleden (in het bijzonder voor de afleiding van het laatst verdiende inkomen) (zie GrossIncomeHour.doc)
	jrsasm_x	Aantal gelijkgestelde dagen (som voor jaar = x)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om individuen te kunnen klasseren naar arbeidsmarktverleden (in het bijzonder voor de afleiding van het laatst verdiende inkomen) (zie GrossIncomeHour.doc)

jrsrem_f_x	Aantal bezoldigde dagen: full time (som voor jaar = x)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om individuen te kunnen klasseren naar arbeidsmarktverleden (in het bijzonder voor de afleiding van het laatst verdiende inkomen) (zie GrossIncomeHour.doc)
jrsrem_p_x	Aantal bezoldigde dagen: part time (som voor jaar = x)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om individuen te kunnen klasseren naar arbeidsmarktverleden (in het bijzonder voor de afleiding van het laatst verdiende inkomen) (zie GrossIncomeHour.doc)
saljrs_x	Het gemiddeld dagloon (som voor jaar = x)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om individuen te kunnen klasseren naar arbeidsmarktverleden (in het bijzonder voor de afleiding van het laatst verdiende inkomen) (zie GrossIncomeHour.doc)

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma "ManipRSZYear_Before2003ToNewData.sas" voor jaren voorafgaand aan 2003 en met het programma "ManipRSZYear_From2003ToNewData.sas" vanaf 2003.

Voor alle individuen die behoren tot de omschreven steekproef, worden per individu een aantal gegevens gevraagd die een samenvatting vormen van de arbeidsmarktstatus van 1997 tot en met &Year.

	Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie								
1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu		Gegeven nodig om te kunnen koppelen met andere bestanden								
2	LABMSTAT_PRIV	Meest voorkomende arbeidsmarktstatus op de private arbeidsmarkt in de periode 1997 tot en met &Year	<table border="1"> <tr> <td>1</td> <td>Arbeider</td> </tr> <tr> <td>2</td> <td>Bediende</td> </tr> </table>	1	Arbeider	2	Bediende	Gegeven wordt gevraagd om arbeidsmarkt verleden van individu te kunnen inschatten.				
1	Arbeider											
2	Bediende											
3	LABMSTAT_PUB	Meest voorkomende arbeidsmarktstatus op de publieke arbeidsmarkt in de periode 1997 tot en met &Year	<table border="1"> <tr> <td>1</td> <td>Arbeider contractueel</td> </tr> <tr> <td>2</td> <td>Bediende contractueel</td> </tr> <tr> <td>3</td> <td>Arbeider statutair</td> </tr> <tr> <td>4</td> <td>Bediende statutair</td> </tr> </table>	1	Arbeider contractueel	2	Bediende contractueel	3	Arbeider statutair	4	Bediende statutair	Gegeven wordt gevraagd om arbeidsmarkt verleden van individu te kunnen inschatten.
1	Arbeider contractueel											
2	Bediende contractueel											
3	Arbeider statutair											
4	Bediende statutair											
4	Work_FP	Voornameelijk actief als voltijds werker of als deeltijds werker	<table border="1"> <tr> <td>0</td> <td>Totaal aantal voltijds gewerkte dagen is kleiner dan aantal deeltijds gewerkte dagen</td> </tr> <tr> <td>1</td> <td>Totaal aantal voltijds gewerkte dagen is groter dan of gelijk aan aantal deeltijds gewerkte dagen</td> </tr> </table>	0	Totaal aantal voltijds gewerkte dagen is kleiner dan aantal deeltijds gewerkte dagen	1	Totaal aantal voltijds gewerkte dagen is groter dan of gelijk aan aantal deeltijds gewerkte dagen	Gegeven wordt gevraagd om arbeidsmarkt verleden van individu te kunnen inschatten.				
0	Totaal aantal voltijds gewerkte dagen is kleiner dan aantal deeltijds gewerkte dagen											
1	Totaal aantal voltijds gewerkte dagen is groter dan of gelijk aan aantal deeltijds gewerkte dagen											

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt door achtereenvolgens de programma's "CreateLabmstatRSZPPO", "CreateLabmstatRSZ" en "CreateLabmstatCombin" te laten lopen.

3.2.10 BRON RIJKSDIENST VOOR SOCIALE ZEKERHEID VAN DE PROVINCIALE EN PLAATSELIJKE OVERHEIDSDIENSTEN (RSZPPO)

Voor alle individuen die behoren tot de omschreven steekproef, worden per tewerkstellingslijn volgende gegevens uit het basisbestand, het bestand statbasef en cuppo samengevoegd voor het 4^{de} kwartaal (&Q=4) van &Year.

	Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu		Gegeven nodig om te kunnen koppelen met andere bestanden
2	cattra_&Year&Q	Werkgeverscategorie (code voor vakantieregime niet-vastbenoemden, code voor pensioenstelsel vastbenoemden).	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het statuut van de werknemer, met het oog op de bepaling van de bijdragen, te bepalen (arbeider, bediende, contractueel of niet). Zou de vroegere variabele wcatcd moeten vervangen.
4	codimp_&Year&Q	Code van de dimensieklasse van de werkgever	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de bijdragen te kunnen bepalen. Deze hangen namelijk af van de omvang van de werkgever (zie module nota CONTRIB).
5	codtra_&Year&Q	Met het oog op een correcte inning van de socialezekerheidsbijdragen maakt deze variabele een onderscheid tussen handarbeiders, hoofdarbeiders, gesubsidieerde contractuelen, vastbenoemden, ...	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het statuut van de werknemer, met het oog op de bepaling van de bijdragen, te bepalen (arbeider, bediende, contractueel of niet). Zou de vroegere variabele wcatcd moeten vervangen.
6	clatraApl_&Year&Q	Groepering van de werknemerscodes in een aantal werknemersklassen.	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het statuut van de werknemer, met het oog op de bepaling van de bijdragen, te bepalen (arbeider, bediende, contractueel of niet). Zou de vroegere variabele wcatcd moeten vervangen.
7	clatr2Apl_&Year&Q	Gedetailleerde groepering van de werknemerscodes.	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het statuut van de werknemer, met het oog op de bepaling van de bijdragen, te bepalen (arbeider, bediende, contractueel of niet). Zou de vroegere variabele wcatcd moeten vervangen.
10	days_pc1_&Year&Q	Aantal gepresteerde dagen met prestatiecodes 1, 2 en 301:	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het gewerkt aantal uren te kunnen berekenen, indien nodig het

				aantal dagen waarvoor vergoeding ziekte betaald wordt en bij het ontbreken van informatie in andere bronnen om iets af te leiden over eventueel opgebouwde pensioenrechten.
11	hrs_pc1_&Year&Q	Aantal gepresteerde uren met prestatiecodes 1, 2 en 301:	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het gewerkt aantal uren te kunnen berekenen, indien nodig het aantal dagen waarvoor vergoeding ziekte betaald wordt en bij het ontbreken van informatie in andere bronnen om iets af te leiden over eventueel opgebouwde pensioenrechten.
12	days_pc10_&Year&Q	Aantal gepresteerde dagen met prestatiecodes 10, 11 en 13:	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het gewerkt aantal uren te kunnen berekenen, indien nodig het aantal dagen waarvoor vergoeding ziekte betaald wordt en bij het ontbreken van informatie in andere bronnen om iets af te leiden over eventueel opgebouwde pensioenrechten.
13	hrs_pc10_&Year&Q	Aantal gepresteerde uren met prestatiecodes 10, 11 en 13:	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het gewerkt aantal uren te kunnen berekenen, indien nodig het aantal dagen waarvoor vergoeding ziekte betaald wordt en bij het ontbreken van informatie in andere bronnen om iets af te leiden over eventueel opgebouwde pensioenrechten.
14	days_pc21_&Year&Q	Aantal gepresteerde dagen met prestatiecodes 21, 22, 24 en 25:	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het gewerkt aantal uren te kunnen berekenen, indien nodig het aantal dagen waarvoor vergoeding ziekte betaald wordt en bij het ontbreken van informatie in andere bronnen om iets af te leiden over eventueel opgebouwde pensioenrechten.
15	hrs_pc21_&Year&Q	Aantal gepresteerde uren met prestatiecodes 21, 22, 24 en 25:	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het gewerkt aantal uren te kunnen berekenen, indien nodig het aantal dagen waarvoor

				vergoeding ziekte betaald wordt en bij het ontbreken van informatie in andere bronnen om iets af te leiden over eventueel opgebouwde pensioenrechten.
26	t_prest_&Year&Q	Code (voltijds, deeltijds, speciaal) die informatie verstrekt over het arbeidsregime op de laatste dag van het kwartaal	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om aan te duiden of iemand al dan niet voltijds werkt
27	apltauxpt_&Year&Q	Percentage deeltijds	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om afgeleide rechten, verbonden aan deeltijds werk, te kunnen schatten en eventueel te vergelijken met gegevens inzake tijdskrediet en loopbaanonderbreking
29	eqtsp_paid_&Year&Q	Voltijds equivalent exclusief gelijkgestelde dagen	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de werkbonus te kunnen bepalen
30	eqtsp_similar_&Year&Q	Voltijds equivalent exclusief betaalde dagen	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de werkbonus te kunnen bepalen
31	tauxaa_&Year&Q	Voltijds equivalent inclusief gelijkgestelde dagen	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de werkbonus te kunnen bepalen
33	r_exclus_&Year&Q	Code die aanduidt of een record moet worden meegeteld of niet.	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om joblijn te identificeren
34	MeanWorkingHours_&Year&Q	Aantal uren dat de betrokken werknemer wekelijks dient te presteren volgens het arbeidscontract	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de bijdragebasis voor de werkbonus te kunnen bepalen.
35	RefMeanWorkingHours_&Year&Q	Aantal wekelijks te presteren uren door de maatman.	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de bijdragebasis voor de werkbonus te kunnen bepalen.
36	regime_&Year&Q	Aantal dagen per week van het arbeidsstelsel van de werknemer.	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om gegevens te kunnen omzetten naar equivalent van 6 dagenweek.
47	sal100_&Year&Q	Bedrag van de gewone bezoldiging in het kwartaal.	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de berekeningsbasis voor de sociale zekerheidsbijdragen te kennen
48	salcot_&Year&Q	Bedrag van de bijdrageplichtige loonmassa inzake sociale zekerheid in het kwartaal	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de berekeningsbasis voor de sociale zekerheidsbijdragen te kennen
49	primes_&Year&Q	Het bedrag van de bezoldigingen in het kwartaal die niet rechtstreeks verbonden zijn met de prestaties gedurende een bepaald kwartaal.	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de berekeningsbasis voor de sociale zekerheidsbijdragen te kennen
50	vehsoc_&Year&Q	Voordeel betreffende het individueel en persoonlijk gebruik van een voertuig ter beschikking gesteld door een werkgever	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om het belastbaar inkomen te kunnen bepalen

53	preavi_&Year&Q	Het bedrag van de verbrekingsvergoeding aangegeven bij de RSZPPO.	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de berekeningsbasis voor de sociale zekerheidsbijdragen te kennen
54	saljrs_&Year&Q	Gemiddeld dagloon	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om simulaties in functie van gewerkt aantal uren maal verdiend loon te kunnen uitvoeren
56	cotpat_&Year&Q	Patronale bijdragen Sociale Zekerheid	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de gesimuleerde bijdrage mee te vergelijken.
57	cotper_&Year&Q	Persoonlijke bijdragen Sociale Zekerheid	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de gesimuleerde bijdrage mee te vergelijken.
58	mntdedwr_&Year&Q	Bedrag bijdragevermindering werkgever	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de gesimuleerde bijdragevermindering mee te vergelijken.
59	mntdedocc_&Year&Q	Bedrag bijdragevermindering werknemer	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de gesimuleerde bijdragevermindering mee te vergelijken.
70	Maribel_&Year&Q	Code die aangeeft of de werknemer aangeworven werd in het kader van de sociale maribel.	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om bijdragevermindering te kunnen berekenen
71	STremAmt_&Year&Q	Loon student	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de berekeningsbasis voor de sociale zekerheidsbijdragen te kennen
72	STndays_&Year&Q	Aantal gepresteerde dagen door de student in het kwartaal	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de berekeningsbasis voor de sociale zekerheidsbijdragen te kennen
73	STctrAmt_&Year&Q	Bijdrage werknemer-student	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de gesimuleerde bijdrage mee te vergelijken
	matric_&Year&Q	gehercodeerde waarde van het inschrijvingsnummer van de werkgever	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om gegevens aan andere deelbestanden te kunnen koppelen
	maatman_&Year&Q	maatman	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de bijdragebasis voor de werkbonus te kunnen bepalen.
	t_dnr_&Year&Q	type dnr	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om joblijn te identificeren.
	rembase_&Year&Q	Basisloon	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de berekeningsbasis voor de sociale zekerheidsbijdragen te kennen

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma "ManipRSZPPOQuartBasisToNewData.sas".

Voor alle individuen die behoren tot de omschreven steekproef, worden per individu een aantal gegevens geaggregeerd op jaarbasis. Volgende gegevens worden gevraagd voor ieder jaar vanaf 2002 tot en met &Year (in onderstaande tabel is $x = 2002$ tot en met &Year)

	Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu		Gegeven nodig om te kunnen koppelen met andere bestanden
2	Sal100_x + Primes_x + Preavi_x	Het bedrag van de 'gewone' bezoldigingen op jaarbasis (som voor jaar = x)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om individuen te kunnen klasseren naar arbeidsmarktverleden (in het bijzonder voor de afleiding van het laatst verdiende inkomen) (zie GrossIncomeHour.doc)
7	GeschatDubbelVakantieGeld_x	Dubbel vakantiegeld geschat rekening houdend met arbeidsmarktstatuut van het individu (som voor jaar = x)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om individuen te kunnen klasseren naar arbeidsmarktverleden (in het bijzonder voor de afleiding van het laatst verdiende inkomen) (zie GrossIncomeHour.doc)
9	RemunDay_x	Vergoede dagen (som voor jaar = x)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om individuen te kunnen klasseren naar arbeidsmarktverleden (in het bijzonder voor de afleiding van het laatst verdiende inkomen) (zie GrossIncomeHour.doc)
10	AssimDay_x	Gelijkgestelde dagen waarvoor vergoeding betaald wordt door werkgever (som voor jaar = x)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om individuen te kunnen klasseren naar arbeidsmarktverleden (in het bijzonder voor de afleiding van het laatst verdiende inkomen) (zie GrossIncomeHour.doc)
12	RemunHour_x	Vergoede uren (som voor jaar = x)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om individuen te kunnen klasseren naar arbeidsmarktverleden (in het bijzonder voor de afleiding van het laatst verdiende inkomen) (zie GrossIncomeHour.doc)
13	AssimHour_x	Gelijkgestelde uren waarvoor vergoeding betaald wordt door werkgever (som voor jaar = x)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om individuen te kunnen klasseren naar arbeidsmarktverleden (in

				het bijzonder voor de afleiding van het laatst verdiende inkomen) (zie GrossIncomeHour.doc)
15	saljrs_x	Het gemiddeld dagloon (som voor jaar = x)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om individuen te kunnen klasseren naar arbeidsmarktverleden (in het bijzonder voor de afleiding van het laatst verdiende inkomen) (zie GrossIncomeHour.doc)

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma "ManipRSZPPOYear_From2005ToNewData.sas" voor jaren vanaf 2005 en met "ManipRSZPPOYear_Before2005ToNewData.sas" voor jaren voorafgaand aan 2005.

3.2.11 BRON RIJKSDIENST VOOR ARBEIDSVOORZIENING (RVA)

Voor alle individuen die behoren tot de omschreven steekproef, wordt op basis van de hoogte van het uibetaalde bedrag, het type werkloosheid en de vergoedingscategorie, één lijn geselecteerd per jaar. Volgende gegevens gevraagd voor het jaar &Year:

	Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu		Gegeven nodig om te kunnen koppelen met andere bestanden
2	FICHE7_&Year	Typering werkloosheid (statuut persoon bij RVA) (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om type uitkering te identificeren
3	DAGEN_&Year	Aantal dagen waarvoor een werkloosheidsuitkering wordt ontvangen (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de uitkering te simuleren
4	DGNMND_&Year	Bedrag van de dagvergoeding die aan de werkloze toegekend wordt (&Year)	Bedrag in klassen van 2 euro	Gegeven wordt gevraagd om de gesimuleerde daguitkering mee te vergelijken of indien informatie ontbreekt om een bedrag te berekenen
5	BEDRAGEN_&Year	bedrag van de werkloosheidsuitkering ontvangen tijdens het jaar (&Year)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om het beschikbaar inkomen te bepalen in geval van ontbrekende gegevens
6	REDENLO_&Year	Reden van loopbaanonderbreking (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om type uitkering die berekend moet worden te identificeren (zie module nota UNEM)
9	DUUR_&Year	Werkloosheidsduur (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om regels bepaling werkloosheidsuitkering opnieuw toe te passen
10	URENPWA_&Year	Aantal uren gewerkt in het kader van PWA gedurende het jaar (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de PWA uitkering te simuleren
11	FICHE61_&Year	Vergoedingscategorie van de werkloze (categorie van RVA) (&Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om het gesimuleerde gezinstype mee te vergelijken

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma “ManipRVAToNewData.sas”.

3.2.12 BRON RIJKSDIENST VOOR PENSIOENEN (RVP), PENSIOENKADASTER

De betalingsgegevens uit deze bestanden worden samengevoegd tot een aggregaat op jaarbasis per code uit soort_pensioen en voor ieder gegeven wordt de eerst geobserveerde betalingsdatum geselecteerd. Voor ieder individu uit de steekproef worden volgende gegevens gevraagd voor &Year:

	Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu		Gegeven nodig om te kunnen koppelen met andere bestanden
	RPWE_&Year	rustpensioen werknemer (betalingsjaar = &Year)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
	FIRPCAP_&Year	voordelen vrije verzekerden (rust) (kapitaal) (betalingsjaar = &Year)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
	FISPCAP_&Year	voordelen vrije verzekerden (overleving) (kapitaal) (betalingsjaar = &Year)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
	RPSE_&Year	rustpensioen zelfstandige (betalingsjaar = &Year)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
	RPCIV_&Year	rustpensioen overheid (betalingsjaar = &Year)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
	RPFOR_&Year	buitenlands rustpensioen (betalingsjaar = &Year)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
	RPDOSZ_&Year	rustpensioen DOSZ (betalingsjaar = &Year)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
	SPWE_&Year	overlevingspensioen werknemer (betalingsjaar = &Year)	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)

				PENSWELF)
SPSE_&Year	overlevingspensioen zelfstandige (betalingsjaar = &Year)	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
SPCIV_&Year	overlevingspensioen overheid (betalingsjaar = &Year)	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
SPFOR_&Year	buitenlands overlevingspensioen (betalingsjaar = &Year)	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
SPDOSZ_&Year	overlevingspensioen DOSZ (betalingsjaar = &Year)	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
CAPFIRST_&Year	kapitalen eerste pijler (betalingsjaar = &Year)	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
IGO_&Year	GIB/IGO (betalingsjaar = &Year)	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
DISAB_&Year	tegemoetkoming mindervaliden (betalingsjaar = &Year)	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
RPRENT_SPILPRIV_&Year	tweedepijler(rust)pensioen, rente-uitkering door private instelling (betalingsjaar = &Year)	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
RPRENT_SPILPUB_&Year	tweedepijler(rust)pensioen, rente-uitkering door publieke instelling (betalingsjaar = &Year)	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
SPRENT_SPILPRIV_&Year	tweedepijler(overlevings)pensioen, rente-uitkering door private instelling (betalingsjaar = &Year)	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota

				PENSWELF)
SPRENT_SPILPUB_&Year	tweedepijler(overlevings)pensioen, rente-uitkering door publieke instelling (betalingsjaar = &Year)	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
RPCAP_SPILPRIV_&Year	tweedepijler(rust)pensioen, kapitaaluitkering door private instelling (betalingsjaar = &Year)	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
RPCAP_SPILPUB_&Year	tweedepijler(rust)pensioen, kapitaaluitkering door publieke instelling (betalingsjaar = &Year)	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
SPCAP_SPILPRIV_&Year	tweedepijler(overlevings)pensioen, kapitaaluitkering door private instelling (betalingsjaar = &Year)	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
SPCAP_SPILPUB_&Year	tweedepijler(overlevings)pensioen, kapitaaluitkering door publieke instelling (betalingsjaar = &Year)	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
DIRPENS_&Year	afgeleide pensioenen (betalingsjaar = &Year)	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
OTHPENS_&Year	overige pensioenbedragen (betalingsjaar = &Year)	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen simuleren (zie module nota PENSWELF)
RPWE_PDATE_&Year	startdatum betaling rustpensioen werknemer (betalingsjaar = &Year)	jjjjmm		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
FIRPCAP_PDATE_&Year	startdatum betaling voordelen vrije verzekerden (rust) (kapitaal) (betalingsjaar = &Year)	jjjjmm		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
FISPCAP_PDATE_&Year	startdatum betaling voordelen vrije verzekerden (overleving) (kapitaal) (betalingsjaar = &Year)	jjjjmm		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen

				uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
	RPSE_PDATE_&Year	startdatum betaling rustpensioen zelfstandige (betalingsjaar = &Year)	jjjjmm	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
	RPCIV_PDATE_&Year	startdatum betaling rustpensioen overheid (betalingsjaar = &Year)	jjjjmm	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
	RPFOR_PDATE_&Year	startdatum betaling buitenlands rustpensioen (betalingsjaar = &Year)	jjjjmm	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
	RPDOSZ_PDATE_&Year	startdatum betaling rustpensioen DOSZ (betalingsjaar = &Year)	jjjjmm	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
	SPWE_PDATE_&Year	startdatum betaling overlevingspensioen werknemer (betalingsjaar = &Year)	jjjjmm	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
	SPSE_PDATE_&Year	startdatum betaling overlevingspensioen zelfstandige (betalingsjaar = &Year)	jjjjmm	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
	SPCIV_PDATE_&Year	startdatum betaling overlevingspensioen overheid (betalingsjaar = &Year)	jjjjmm	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
	SPFOR_PDATE_&Year	startdatum betaling buitenlands overlevingspensioen (betalingsjaar = &Year)	jjjjmm	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
	SPDOSZ_PDATE_&Year	startdatum betaling overlevingspensioen DOSZ (betalingsjaar = &Year)	jjjjmm	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de

				datum van ingang van het recht (zie module nota PENSWELF)
CAPFIRST_PDATE_&Year	startdatum betaling kapitalen eerste pijler (betalingsjaar = &Year)	jjjjmm		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
IGO_PDATE_&Year	startdatum betaling GIB/IGO (betalingsjaar = &Year)	jjjjmm		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
DISAB_PDATE_&Year	startdatum betaling tegemoetkoming mindervaliden (betalingsjaar = &Year)	jjjjmm		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
RPRENT_SPILPRIV_PDATE_&Year	startdatum betaling tweedepijler(rust)pensioen, rente-uitkering door private instelling (beta	jjjjmm		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
RPRENT_SPILPUB_PDATE_&Year	startdatum betaling tweedepijler(rust)pensioen, rente-uitkering door publieke instelling (betalingsjaar = &Year)	jjjjmm		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
SPRENT_SPILPRIV_PDATE_&Year	startdatum betaling tweedepijler(overlevings)pensioen, rente-uitkering door private instelling (betalingsjaar = &Year)	jjjjmm		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
SPRENT_SPILPUB_PDATE_&Year	startdatum betaling tweedepijler(overlevings)pensioen, rente-uitkering door publieke instelling (betalingsjaar = &Year)	jjjjmm		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
RPCAP_SPILPRIV_PDATE_&Year	startdatum betaling tweedepijler(rust)pensioen, kapitaaluitkering door private instelling (betalingsjaar = &Year)	jjjjmm		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
RPCAP_SPILPUB_PDATE_&Year	startdatum betaling tweedepijler(rust)pensioen, kapitaaluitkering door publieke instelling (betalingsjaar = &Year)	jjjjmm		Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht

				(zie module nota PENSWELF)
	SPCAP_SPILPRIV_PDATE_&Year	startdatum betaling tweedepijler(overlevings)pensioen, kapitaaluitkering door private instelling (betalingsjaar = &Year)	jjjjmm	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
	SPCAP_SPILPUB_PDATE_&Year	startdatum betaling tweedepijler(overlevings)pensioen, kapitaaluitkering door publieke instelling (betalingsjaar = &Year)	jjjjmm	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
	DIRPENS_PDATE_&Year	startdatum betaling afgeleide pensioenen (betalingsjaar = &Year)	jjjjmm	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
	OTHPENS_PDATE_&Year	startdatum betaling overige pensioenbedragen (betalingsjaar = &Year)	jjjjmm	Gegeven wordt gevraagd om welvaartsaanpassingen te kunnen uitvoeren in functie van de datum van ingang van het recht (zie module nota PENSWELF)
	code_gezinslast_&Year	Code gezinslast (geregistreerd in &Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om gezinssituatie met betrekking tot pensioenregeling te kunnen beoordelen.
	code_echtgenote_&Year	Code echtgenote ten laste (geregistreerd in &Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om gezinssituatie met betrekking tot pensioenregeling te kunnen beoordelen.
	aantal_kind_&Year	Aantal kinderen ten laste (geregistreerd in &Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om gezinssituatie met betrekking tot pensioenregeling te kunnen beoordelen.
	aantal_andere_&Year	Aantal andere personen ten laste (geregistreerd in &Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om gezinssituatie met betrekking tot pensioenregeling te kunnen beoordelen.
	Code_alleenstaande_&Year	Code alleenstaande (geregistreerd in &Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om gezinssituatie met betrekking tot pensioenregeling te kunnen beoordelen.

	begindatum_pensioen_c_&Year	begindatum_pensioen_c (geregistreerd in &Year)	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om simulaties in functie van datum van opruststelling te kunnen uitvoeren.
--	-----------------------------	--	----------------------------	---

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma "ManipRVPToNewData.sas".

3.2.13 BRON GEGEVENS DG PERSONEN MET EEN HANDICAP VAN DE FOD SOCIALE ZEKERHEID

Voor alle individuen die behoren tot de omschreven steekproef, worden voor ieder van de 4 kwartalen van &Year volgende gegevens gevraagd:

	Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
1	GECODEERD_INSZ	Geanonimiseerd identificatienummer van het individu		Gegeven nodig om te kunnen koppelen met andere bestanden
	Type_rec	Het type record geeft aan uit welk bestand een record afkomstig is.	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om aanduiding te hebben van de herkomst van het gegeven (bestand met medische gegevens, rechten of betalingsgegevens)
	Code_leg	De code wetgeving geeft aan op welke wetgeving het betreffende record van toepassing is.	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om van toepassing zijnde wetgeving te kennen
	Date_start_recon	Begindatum van de medische erkenningsprocedure van de handicap	mm-JJJJ	Gegeven wordt gevraagd om periode van medische erkenning af te bakenen
	Date_end_recon	Einddatum van de medische erkenningsprocedure van de handicap	mm-JJJJ	Gegeven wordt gevraagd om periode van medische erkenning af te bakenen
	Memb_inf_50perc	Erkenning van de handicap 50% onderste ledematen	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd als indicator van de eigenschappen van het individu.
	Cecite_comp	Erkenning van de handicap volledige blindheid	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd als indicator van de eigenschappen van het individu.
	Amp_memb_sup	Erkenning van de handicap amputatie van de bovenste ledematen	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd als indicator van de eigenschappen van het individu.
	Paral_memb_sup	Erkenning van de handicap verlamming van de bovenste ledematen	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd als indicator van de eigenschappen van het individu.
	incapacite	Percentage van de ongeschiktheid van het kind	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de oude bijslag te kunnen benaderen
	Pillier_un	Het aantal punten dat het kind scoort voor de gevolgen van de aandoening op het vlak van de lichamelijke of geestelijke ongeschiktheid.	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de uitkering nieuwe wetgeving te kunnen reconstrueren.
	Pillier_deux	Het aantal punten dat het kind scoort voor de gevolgen van de aandoening op het vlak van de activiteit en de participatie.	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de uitkering nieuwe wetgeving te kunnen reconstrueren.
	Pillier_trois	Het aantal punten dat het kind scoort voor de gevolgen van de aandoening voor de familiale omgeving	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de uitkering nieuwe wetgeving te

				kunnen reconstrueren.
Arr	Deze variabele geeft aan of er een erkenning is van de vermindering van het verdienvermogen	Waarden zoals in DWH AM&SB		Gegeven wordt gevraagd om het onderscheid te kunnen maken met IT uitkering
Critere_un	Het aantal punten dat de volwassene scoort voor de vermindering zelfredzaamheid, functie verplaatsingsmogelijkheden	Waarden zoals in DWH AM&SB		Gegeven wordt gevraagd om uitkering nieuwe wet te kunnen bepalen
Critere_deux	Het aantal punten dat de volwassene scoort voor de vermindering zelfredzaamheid, functie voedsel nuttigen of bereiden	Waarden zoals in DWH AM&SB		Gegeven wordt gevraagd om uitkering nieuwe wet te kunnen bepalen
Critere_trois	Het aantal punten dat de volwassene scoort voor de vermindering zelfredzaamheid, functie persoonlijke hygiëne en zichzelf kleden	Waarden zoals in DWH AM&SB		Gegeven wordt gevraagd om uitkering nieuwe wet te kunnen bepalen
Critere_quatre	Het aantal punten dat de volwassene scoort voor de vermindering zelfredzaamheid, functie woning onderhouden en huishoudelijk werk verrichten	Waarden zoals in DWH AM&SB		Gegeven wordt gevraagd om uitkering nieuwe wet te kunnen bepalen
Critere_cinq	Het aantal punten dat de volwassene scoort voor de vermindering zelfredzaamheid, functie leven zonder toezicht	Waarden zoals in DWH AM&SB		Gegeven wordt gevraagd om uitkering nieuwe wet te kunnen bepalen
Critere_six	Het aantal punten dat de volwassene scoort voor de vermindering zelfredzaamheid, functie communicatie en sociaal contact	Waarden zoals in DWH AM&SB		Gegeven wordt gevraagd om uitkering nieuwe wet te kunnen bepalen
Date_deces	Overlijdensdatum van de betrokkene	mm-JJJJ		Gegeven wordt gevraagd om periodes te kunnen afbakenen.
montant	Theoretisch bedrag voor de betalingsperiode	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om het een indicatie geeft van het recht
Montant_emis	Werkelijk bedrag betaald gedurende de betalingsperiode	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd om het gezinsinkomen te kunnen reconstrueren
Class_stat	Gegeven classificeert de toegekende tegemoetkoming op een unieke wijze	Waarden zoals in DWH AM&SB		Gegeven wordt gevraagd omdat het informatie bevat over veel onderdelen die een rol spelen bij het bepalen van de uitkering
Per_start_pay	Begindatum betalingsperiode waarop de betaling betrekking heeft	mm-JJJJ		Gegeven wordt gevraagd om de betalingsperiode te kunnen afbakenen.
Per_end_pay	Einddatum betalingsperiode waarop de betaling betrekking heeft	mm-JJJJ		Gegeven wordt gevraagd om de betalingsperiode te kunnen afbakenen.
Mnt_mens_sim	Totaal gesimuleerd maanbedrag	Bedrag in klassen van 10 euro		Gegeven wordt gevraagd omdat het een indicatie geeft van de hoogte van het recht

Mnt_mens_AL_sim	Gesimuleerd maanbedrag IT	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd omdat het een indicatie geeft van de hoogte van het recht
periode	Maand ten opzichte van dewelke geïndexeerd moet worden	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om link te kunnen leggen tussen recht en feitelijk betaald bedrag.
Date_start	Begindatum recht	mm-JJJJ	Gegeven wordt gevraagd om periode van het recht af te bakenen.
Date_end	Einddatum recht	mm-JJJJ	Gegeven wordt gevraagd om periode van het recht af te bakenen.
Date_decision	Datum beslissing tot eventuele herziening van het recht	mm-JJJJ	Gegeven laat toe om de lijn van het laatst van toepassing zijnde recht te selecteren.
Date_decision_corr	Gewijzigde datum van de beslissing	mm-JJJJ	Gegeven laat toe om de lijn van het laatst van toepassing zijnde recht te selecteren.
Gecodeerd_insz_part	Geanonimiseerd identificatienummer van de partner van de rechthebbende	Waarden zoals in DWH AM&SB	Gegeven wordt gevraagd om de partner van de rechthebbende te kunnen identificeren. De definitie van partnerschap wijkt bij personen met een handicap af van deze in het Rijksregister.
Date_start	Startdatum partnerschap met partner gekend onder gecodeerd_insz_part	mm-JJJJ	Gegeven wordt gevraagd om periode van partnerschap te kunnen afbakenen.
Date_end	Einddatum partnerschap met partner gekend onder gecodeerd_insz_part	mm-JJJJ	Gegeven wordt gevraagd om periode van partnerschap te kunnen afbakenen.

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma "ManipDGPMHToNewData.sas".

3.2.14 BRON COMPTE INDIVIDUEL DE PENSION (ASBL CIMIRE)

Voor alle individuen die behoren tot de oorspronkelijke steekproef (i.e. trekking 2001) worden voor alle aansluitingsjaren per jaar volgende gegevens gevraagd:

Record T22: Gegevens over de loopbaan

Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
Loopbaanjaar	Jaar waarin prestaties geleverd werden	YYYY	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
Loopbaancode	Code die type loopbaan van het individu weer geeft	Mogelijke codes opgenomen in Appendix 4 van de aanvraag	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
Vergoeding	Reële jaarlijkse vergoeding ontvangen door het individu voor de geleverde prestaties	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
Vergoede dagen	Aantal effectief gepresteerde dagen	Aantal dagen	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
Gelijkgestelde dagen	Aantal gelijkgestelde dagen	Aantal dagen	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
Aantal uren maatman	Aantal uren gepresteerde door een individu dat voltijds werkt met dezelfde loopbaancode (maatman)	Aantal uren	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
Aantal uren deeltijds	Aantal uren gepresteerde als deeltijdse arbeid	Aantal uren	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden

Record T23: Gegevens over loopbaancodes 200 tot 247

Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
Loopbaancode	Code die type loopbaan van het individu weer geeft	Mogelijke codes opgenomen in Appendix 4 van de aanvraag	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRe te vermijden
Begindatum	Begindatum gelijkgestelde periode	mm-JJJJ	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRe te vermijden
Einddatum	Einddatum gelijkgesteld periode	mm-JJJJ	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRe te vermijden
Percentage	Percentage invaliditeit individu	Percentage	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRe te vermijden

Record T24: Gegevens over loopbaancodes 248 tot 262

Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
Loopbaancode	Code die type loopbaan van het individu weer geeft	Mogelijke codes opgenomen in Appendix 4 van de aanvraag	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
Eerste datum	Begindatum gelijkgestelde periode	mm-JJJJ	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
Tweede datum	Einddatum gelijkgestelde periode of begindatum van de periode met het gewijzigd percentage van beroepsziekte	mm-JJJJ	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
Derde datum	Einddatum gelijkgestelde periode of datum van regularisatie van de gelijkgestelde periode	mm-JJJJ	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
Percentage	Percentage invaliditeit individu	Percentage	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
Basissalaris	Salaris op basis waarvan de schadeloosstelling berekend wordt	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden

Record T25: Bijkomende informatie in geval van meerdere loopbaancodes per jaar

Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
Loopbaanjaar	Jaar waarin prestaties geleverd werden	YYYY	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
Loopbaancode	Code die type loopbaan van het individu weer geeft	Mogelijke codes opgenomen in Appendix 4 van de aanvraag	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
Vergoeding	Reële jaarlijkse vergoeding ontvangen door het individu voor de geleverde prestaties	Bedrag in klassen van 10 euro	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
Vergoede dagen	Aantal effectief gepresteerde dagen	Aantal dagen	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden

Gelijkgestelde dagen	Aantal gelijkgestelde dagen	Aantal dagen	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
Aantal uren maatman	Aantal uren gepresteerd door een individu dat voltijds werkt met dezelfde loopbaancode (maatman)	Aantal uren	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
Aantal uren deeltijds	Aantal uren gepresteerd als deeltijdse arbeid	Aantal uren	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden

Record T26: Informatie ivm inkomensgarantieuitkering

Variabele	Omschrijving variabele	Mogelijke waarden	Motivatie
Loopbaancode	Code die type loopbaan van het individu weer geeft	Mogelijke codes opgenomen in Appendix 4 van de aanvraag	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
Loopbaanjaar	Jaar waarop de geregistreerde gegevens slaan	YYYY	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
maand (alloc_garantie)	Aanduiding per maand van het jaar of er inkomensgarantieuitkering was of niet	Variabele op 12 posities voor alle maanden van het jaar. 1 positie per maand, met aanduiding 1 of 0 per maand. 1 = met inkomensgarantieuitkering; 0 = zonder	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
eerste datum	begindatum statuut behoud van rechten	mm-JJJJ	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
tweede datum	begindatum deeltijdse arbeid bij werkgever	mm-JJJJ	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
derde datum	einddatum contract deeltijdse arbeid indien gekend door RVA	mm-JJJJ	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden
code (indic_assuj)	Aanduiding of het RSZ nummer gekend is of niet	0 = RSZnr. niet gekend; = 1 RSZnr gekend	Gegeven wordt gevraagd om de loopbaanhistoriek van het individu te kennen en om een nieuwe consultatie van CIMIRE te vermijden

Opmerkingen bij aanmaak van deze gegevens:

1. De gevraagde gegevens worden aangemaakt met het programma "ManipCIMIREToNewData.sas".

3.3 MOTIVATIE

De voorliggende aanvraag heeft tot doel om de dataset die gebruikt wordt als input voor het statisch microsimulatiemodel MIMOSIS up te daten.

De bronnen die gevraagd worden voor het jaar &Year, worden in eerste instantie gevraagd voor de modellering in MIMOSIS. Er kunnen drie grote blokken in de gevraagde gegevens onderscheiden worden:

- variabelen die toelaten om de regelgeving van een bepaald beleidsdomein te modelleren,
- variabelen die toelaten om het berekende resultaat te evalueren door bedragen, berekend met het model, te vergelijken met een geregistreerd bedrag,
- variabelen die toelaten om de gesimuleerde resultaten te evalueren.

Met betrekking tot dit laatste punt zijn de gegevens die toelaten om het concept “equivalent beschikbaar inkomen” te reconstrueren van bijzonder belang. Dit omdat bij simulaties die tot nog toe met MIMOSIS werden uitgevoerd, bijzondere aandacht besteed werd aan de impact van de gesimuleerde maatregel op personen onder de armoedegrens (vastgelegd op 60% van de mediaan van het equivalent beschikbaar inkomen).

Naast de gegevens van het jaar &Year zijn een aantal gegevens van voorafgaande jaren, ook van belang voor gebruik in MIMOSIS. Ten eerste omdat het laatst verdiende arbeidsinkomen in bepaalde beleidsdomeinen een centrale rol vervult. Ten tweede omdat het van belang kan zijn bij de evaluatie van bepaalde maatregelen, groepen te onderscheiden in functie van hun arbeidsmarktverleden.

4 FREQUENTIE

De aanvraag voor de betreffende gegevens is eenmalig. Het is wel de bedoeling om op basis van de gevraagde gegevens een meer gestileerde set van variabelen af te leiden die naar de toekomst toe op regelmatige tijdstippen geüpdate zou moeten worden, dit teneinde het microsimulatiemodel up to date te houden.

5 TERMIJN

De opgesomde gegevens worden gevraagd voor een termijn van 3 jaar. In deze periode zou het mogelijk moeten zijn om zowel de nodige bewerkingen op de basisdata uit te voeren alsook om het programma dat de gemanipuleerde basisdata verwerkt, aan te passen.

6 DERDEN

De gegevens worden aangevraagd door de FOD Sociale Zekerheid. De FOD Sociale Zekerheid zal de gevraagde gegevens enkel ter beschikking stellen van personen die onder de verantwoordelijkheid van de FOD Sociale Zekerheid werken en op PC's die door deze FOD beveiligd worden.

7 TIMING

Aangezien de gevraagde variabelen, op één uitzondering na, afgeleid kunnen worden van de gegevens die reeds beschikbaar zijn in de huidige versie van het DWH AM&SB wordt voorgesteld om alle gegevens met één enkele levering te bezorgen. Indien het toch wenselijk zou zijn om de levering in stukken te verdelen, wordt voorgesteld om de gegevens van het jaar &Year eerst te leveren en in tweede instantie deze van alle voorgaande jaren.

8 VOORSTEL VAN ANALYSE

De gegevens die hier gevraagd worden hebben betrekking op de tweede fase van de aanvraag. In de eerste fase werden, op basis van een kleine steekproef met variabelen uit DWH AM&SB, een aantal programma's geschreven die variabelen genereren die verder verwerkt worden door de verschillende modellen. In de tweede fase van de aanvraag worden de bestaande programma's gelanceerd op een grotere steekproef, om op die manier de gevraagde afgeleide variabelen te genereren voor een steekproef van voldoende omvang.

Op één uitzondering na, zijn alle geconstrueerde variabelen produceerbaar met gegevens die reeds deel uitmaken van het DWH AM&SB. Wel worden voor nagenoeg iedere bron nog een aantal opmerkingen gemaakt behandeld moeten worden alvorens een definitie versie van het afgeleide bestand kan aangemaakt worden. Volgende stappen lijken daarom noodzakelijk:

- Bepaling van de eenheden in de steekproef.
- Aanpassing van de basisdataset op basis waarvan de afgeleide variabelen geconstrueerd worden en indien nodig aanpassing van het programma dat de data moet genereren (zie opmerkingen per bron).
- Productie van de gevraagde variabelen per bron.

De steekproefomvang van de in deze aanvraag opgesomde gegevens wordt beperkt in vergelijking met de steekproef waarop het model finaal zal draaien. De ontwikkeling van het model zal immers in twee grote stappen verlopen. In een eerste stap zullen de gegevens uit deze aanvraag gebruikt worden om programma's te schrijven waarmee gegevens op een dusdanige wijze worden samengevoegd dat ze geschikt zijn voor input in het model. In een tweede fase zullen met deze programma's afgeleide variabelen worden aangemaakt voor een steekproef die ruimer is in omvang, maar die voornamelijk bestaat uit variabelen die een samenvoeging of afleiding zijn van de variabelen die in de voorliggende data-aanvraag worden opgesomd.

9 BIJLAGEN

9.1 BIJLAGE 1: VOORSTEL TREKKING STEEKPROEF UPDATE MIMOSIS

9.1.1 INLEIDING

De onderliggende data van het model zijn afkomstig uit het Datawarehouse Arbeidsmarkt en Sociale Bescherming (DWH AM&SB) van de Kruispuntbank van de Sociale Zekerheid (KSZ). De gegevens die voor de huidige versie gebruikt worden, zijn deze van het bijdragen en

uitkeringsjaar 2001. Het is de bedoeling om in een geüpdate versie van het model, gegevens uit het DWH AM&SB van een meer recent bijdragen en uitkeringsjaar te gebruiken. In deze nota bespreken we hoe de steekproef voor een update van de onderliggende dataset precies dient getrokken te worden.³

Voor de verdere bespreking van de steekproefprocedure is het niet van belang om te weten voor welk jaar de steekproef uiteindelijk zal getrokken worden. We geven hier wel mee dat een update van de onderliggende data bij voorkeur zou uitgevoerd moeten worden met gegevens voor het bijdragen en uitkeringsjaar 2008, omdat dit op het moment van de aanvraag het meest recente volledig beschikbare jaar in het DWH AM&SB is. Er wordt dus verondersteld dat de string &Year in het voorgaande, vervangen kan worden door 2008.

9.1.2 STEEKPROEFONTWERP

9.1.2.1 Het steekproefkader en de observatieeenheid

Het steekproefkader is de set van alle observatieeenheden die een kans verschillend van nul hebben om getrokken te worden.⁴ Alvorens we het steekproefmechanisme zelf beschrijven, is het van belang de beperking van ons steekproefkader onder ogen te zien. Het mini rijksregister (mini RN) vormt het vertrekpunt voor de constructie van onze steekproef met administratieve gegevens.⁵ Het mini RN omvat, in principe, identificatiegegevens van twee groepen van individuen:

1. Individuen die nog in leven zijn, die opgenomen zijn in het Rijksregister en die een verblijfsadres hebben in een Belgische gemeente, i.e. individuen die zijn ingeschreven in een bevolkingsregister van een Belgische gemeente.
2. Individuen die zijn opgenomen in het wachtregister, i.e. kandidaat politiek vluchtelingen.⁶

In de groep omschreven onder punt 1 hierboven zitten dus voornamelijk personen die fysiek ingezet zijn in België. Belgen die in het buitenland verblijven kunnen deel uitmaken van de onder punt 1 omschreven 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.⁷ Normaal is deze periode van een tijdelijke missie beperkt tot maximaal 12 maanden, maar er zijn uitzonderingen mogelijk op deze regel.

³ Aangezien het de bedoeling is de onderliggende data van het model te updaten, baseren we ons voor het steekproefplan in deze nota op de procedure die gehanteerd werd voor trekking van de steekproef van het oorspronkelijke model (zie Project AG/01/086 (2003)).

⁴ Zie Molenberghs en Gaens (1994), p. 11.

⁵ Het miniRN is een bestand met Rijksregisternummers beheerd door de KSZ. Het miniRN bevat niet alle identificatiegegevens die wettelijk geregistreerd worden en dus opgenomen zijn in het Rijksregister. Van de wettelijk geregistreerd gegevens bevat het miniRN enkel het geslacht, de geboortedatum en de verblijfplaats.

⁶ Voor de latere constructie van steekproefgewichten is het van belang aan te stippen dat de bevolkingsstatistieken die door ADSEI gepubliceerd worden enkel individuen omvatten die een verblijfsadres in België hebben en die opgenomen zijn in het Rijksregister. Individuen die zijn opgenomen in het wachtregister, worden dus niet meegeteld in de bevolkingsstatistieken die door ADSEI gepubliceerd worden.

⁷ Belgen die tijdelijk in het buitenland verblijven zijn niet verplicht dit te melden aan de gemeente waar ze voor hun vertrek verblijven.

Belgen die permanent in het buitenland verblijven worden dus niet opgenomen in ons steekproefkader, ook al zouden ze hun verblijfsplaats kenbaar gemaakt hebben via een consulaire register van een consulaire of diplomatieke post. De groep van Belgen in het buitenland die nog in aanraking komen met de Belgische sociale zekerheid, bijvoorbeeld omdat aan hen nog een pensioen wordt uitbetaald, zou bestudeerd kunnen worden aangezien zij voorkomen met een sociaal zekerheidsnummer in het BIS-register.⁸

Binnen dit mini RN kunnen minstens twee observatie-eenheden onderscheiden worden. Ten eerste alle personen die voorkomen met een uniek mini RN nummer, i.e. individuen. Ten tweede gezinnen. Gezinnen kunnen in deze context omschreven worden als alle individuen die hetzelfde verblijfsadres hebben.⁹ Op gezinsniveau zou men verder het onderscheid kunnen maken tussen private en collectieve gezinnen. Collectieve gezinnen omvatten samenlevingsgemeenschappen zoals bejaardentehuizen, hospitalen, gevangenis, kloosters,¹⁰

Finaal willen we beschikken over een steekproef met observaties op individueel niveau. Voor individuen die in een privaat gezin wonen dienen we ook te beschikken over informatie omtrent de samenstelling van het gezin. Een privaat gezin, zoals gedefinieerd in het Rijksregister, vormt allicht 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.¹¹ Voor de rest van deze nota gebruiken we daarom de termen private en sociologische gezinnen door elkaar. Als beslissingseenheid zijn sociologische gezinnen allicht relevant vanuit beleidsoogpunt, terwijl dit niet het geval is voor collectieve gezinnen. Daarom vragen we enkel gegevens in verband met de gezinssamenstelling voor sociologische gezinnen.

Eén mogelijkheid om de steekproef uit het mini RN te trekken, zou er in kunnen bestaan om collectieve gezinnen volledig te negeren en een steekproef samen te stellen die enkel representatief is voor private gezinnen. Met representatief bedoelen we hier dat de steekproef een afspiegeling is van de oorspronkelijke populatie waaruit de observatie-eenheden getrokken worden.¹²

Wanneer enkel een steekproef getrokken wordt van individuen uit private gezinnen, missen we echter een deel van de populatie van bejaarden, wat later in het model gevolgen zou kunnen hebben bij evaluaties van bepaalde hervormingen. Daarom wordt voorgesteld om een steekproef te trekken uit de set van individuen, geobserveerd in het mini RN, en die in leven zijn op een bepaald moment in de tijd.

Om het precieze tijdstip te bepalen waarop de populatie “gefotografeerd” wordt, houden we rekening met het soort gegevens dat later zal opgevraagd worden. De latere gegevensaanvraag omvat onder andere gegevens omtrent personenbelastingen. Op het belastingformulier vullen belastingplichtigen de gezinstoestand in van 1 januari van het jaar dat volgt op het jaar waarin het gerapporteerde inkomen verdiend werd. Optimaal zouden we in 2010 streven naar inkomensgegevens voor het jaar 2008, en zou de populatie dus best gefotografeerd worden op 1 januari 2009.

⁸ In principe zou het BIS-register enkel individuen mogen omvatten die in het buitenland verblijven en nog in aanraking komen met de Belgische sociale zekerheid. Wanneer individuen fysiek in België verblijven en in contact komen met de Belgische sociale zekerheid nog voor ze zijn opgenomen in het Rijksregister, dan zullen deze individuen in regel ook voorkomen in het BIS-register. In theorie zouden deze gevallen echter niet mogen voorkomen.

⁹ Zie Ministerie van Binnenlandse Zaken (2003).

¹⁰ Zie Administratie Planning en Statistiek (2003a).

¹¹ Zie Atkinson e.a. (1995), p. 16 voor meer info omtrent mogelijke omschrijvingen van het begrip huishouden.

¹² Zie Johnson and Kotz (1989), volume 8, p. 77-81 voor een overzicht van een aantal mogelijke interpretaties van het begrip representativiteit in de statistische literatuur.

Voor de steekproeftrekking stellen we dus voor om gebruik te maken van de gegevens uit het Rijks- en Bis-register op 1 januari van het jaar &Year+1. Het is mogelijk dat op het moment van de verwerking van deze aanvraag de gegevens uit het Rijks- en Bis-register van 1 januari 2009 nog niet aanwezig zullen zijn in het DWH AM&SB. Indien dit het geval is, wordt voorgesteld om de steekproef te trekken op 1 januari 2008 en alle andere elementen uit de aanvraag ongewijzigd te laten. In wat volgt zullen we dan ook &Year+1 gebruiken als passe-partout voor het jaar van steekproeftrekking. Idealiter is &Year+1 gelijk aan 2009. Indien dit niet mogelijk is, is &Year+1 gelijk aan 2008.

Van zodra individuen getrokken worden uit de volledige set van individuen in leven op 1 januari &Year+1, kan ook de bijhorende gezinsinformatie voor zij die in een privaat gezin wonen opgevraagd worden. Voor getrokken individuen, wonend in een privaat gezin, zouden de mini RN nummers van alle andere gezinsleden die in leven zijn op 1 januari &Year+1 moeten opgevraagd worden. Voor individuen uit een collectief gezin hoeft geen informatie opgevraagd te worden omtrent andere gezinsleden.

9.1.2.2 Het steekproefmechanisme

Voor het samenstellen van de huidige steekproef wordt de vorige steekproef (i.e. deze getrokken op 1 januari 2002) als uitgangspunt genomen. De voornaamste reden om dit te doen is dat we op die manier een nieuwe bevraging van het CIMIRE register kunnen vermijden. De hoger opgelijste gegevens worden dus in eerste instantie gevraagd voor alle individuen uit de oude steekproef die in leven zijn op 1 januari &Year+1 en, indien ze tot een privaat gezin behoren, voor de gezinsleden van de gezinnen waartoe deze individuen behoren.

Om de uitval, die tussen 1 januari 2002 en 1 januari &Year+1 heeft plaatsgevonden, op te vangen wordt een systeem van systematische steekproeftrekking voorgesteld. Deze procedure moet toegepast worden op de groep van individuen ingeschreven in het Rijksregister tussen 1 januari 2002 en 1 januari &Year+1.

In de oorspronkelijke steekproef werden 100.000 individuen getrokken via systematische steekproeftrekking. Daarom wordt voorgesteld om het verschil tussen a) 100.000 en b) het totaal aantal nog in leven zijnde individuen die in de oorspronkelijke steekproef systematisch getrokken werden, via systematische steekproeftrekking te trekken uit de groep van individuen ingeschreven in het Rijksregister tussen 1 januari 2002 en 1 januari &Year+1.

In geval van systematische steekproeftrekking wordt enkel het eerste nummer toevallig getrokken. Alle volgende nummers worden dan systematisch afgeleid vertrekkende van het eerste toevallig getrokken nummer. Bij de start van een systematische steekproeftrekking zijn de omvang van de populatie, weergegeven door N , en de omvang van de steekproef, weergegeven door n , gekend. Iedere observatie in het steekproefkader krijgt vervolgens een uniek nummer gaande van 1 tot N . De eerste observatie wordt geselecteerd als een toevallig gekozen nummer tussen 1 en x , met $x = \left\lceil \frac{N}{n} \right\rceil$.¹³ Dit toevallig gekozen nummer wordt ook de random start genoemd. Vertrekkende van het random start nummer wordt vervolgens iedere observatie uit het steekproefkader geselecteerd waarvoor het verschil tussen het ordeningsnummer en het ordeningsnummer van de voorgaande geselecteerde observatie gelijk is aan x .

¹³ In deze notatie symboliseren de haakjes dat de verhouding N gedeeld door n wordt omgezet naar het dichtstbijgelegen gehele getal. Zie Greene (1997), p. 177 voor een beschrijving van een random number generator.

De waarde N is in dit geval dus gelijk aan het totaal aantal individuen ingeschreven in het Rijksregister tussen 1 januari 2002 en 1 januari $\&Year+1$. De waarde n is gelijk aan het verschil tussen 100.000 en het totaal aantal nog in leven zijnde individuen die in de oorspronkelijke steekproef systematisch getrokken werden.

Ook voor deze extra getrokken individuen en, indien ze in een privaat gezin wonen, voor de gezinsleden van de gezinnen waartoe deze individuen behoren, worden de hoger opgelijste gegevens gevraagd.

9.1.3 OMVANG VAN DE POPULATIE EN DE STEEKPROEF

Om de omvang van 100.000 systematisch getrokken individuen te verantwoorden, maken we gebruik van de omgekeerde hypothesetoets. De normale statistische procedure bestaat er in om eerst gegevens met betrekking tot een bepaald fenomeen te verzamelen, vervolgens om met deze gegevens parameters te schatten die dit fenomeen samenvatten en tenslotte te toetsen hoe waarschijnlijk of onwaarschijnlijk de geobserveerde parameterwaarden zijn, gegeven de statistische eigenschappen van de gebruikte schatter. Aangezien de omvang van de steekproef opgenomen is in de formule van bepaalde schatters, kan men deze gang van zaken ook omkeren. Indien men een veronderstelling maakt over de waarde van een te schatten parameter in de totale populatie, kan men de vraag stellen hoeveel observaties men nodig heeft om een vooropgestelde hypothese omtrent dit fenomeen te accepteren of te verwerpen voor een vooraf vastgelegd precisieniveau. In de rest van deze sectie wordt dit idee verder uitgewerkt voor hypothesetoetsen op één variabele.¹⁴

Het fenomeen voor hetwelk we later een hypothesetoets willen uitvoeren, wordt in de populatie gevat door de variabele Y . In wat volgt gebruiken we hoofdletters om te verwijzen naar populatiefenomenen en kleine letters voor steekproefwaarden. Meer specifiek maken we gebruik van de symbolen:

Y_i om een realisatie van Y voor de i -de eenheid in de populatie aan te duiden,

y_i om een realisatie van Y voor de i -de eenheid in de steekproef aan te duiden,

$\bar{Y} = \frac{\sum_{i=1}^N Y_i}{N}$ om het gemiddelde van Y in de populatie aan te duiden,

$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$ om het gemiddelde van Y in de steekproef aan te duiden.

Bij wijze van voorbeeld veronderstellen we dat de realisaties van Y zich vertalen in binaire waarden. De populatie kan dus opgedeeld worden in twee elkaar uitsluitende klassen C_1 en C_2 . De klasse C_1 bevat bij wijze van voorbeeld enkel mannen en de klasse C_2 bevat enkel vrouwen. Verder veronderstellen we dat deze variabele Y enkel de waarden 0 en 1 kan aannemen. We kunnen Y_i en y_i dan als volgt definiëren:

¹⁴ De uiteenzetting in deze sectie is voornamelijk gebaseerd op Jolliffe (1986), Chapter 2.

$$Y_i = \begin{cases} 1 & \text{indien de observatieeenheid } i \text{ in de populatie behoort tot klasse } C_1 \\ 0 & \text{in het andere geval} \end{cases},$$

$$y_i = \begin{cases} 1 & \text{indien de observatieeenheid } i \text{ in de steekproef behoort tot klasse } C_1 \\ 0 & \text{in het andere geval} \end{cases}.$$

Gegeven deze veronderstellingen kan men het gemiddelde in de populatie en de steekproef respectievelijk als volgt definiëren:

$$\bar{Y} = \frac{\sum_{i=1}^N Y_i}{N} = P, \text{ i.e. de proportie van het aantal gevallen waarvoor de variabele } Y \text{ de waarde 1 aanneemt in de populatie en}$$

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n} = p, \text{ i.e. de proportie van het aantal gevallen waarvoor de variabele } Y \text{ de waarde 1 aanneemt in de steekproef.}$$

We veronderstellen ook dat de steekproef die we gebruiken om een hypothese te toetsen inzake het fenomeen dat wordt gevat door de variabele Y een toevallige steekproef is. In dat geval kan men aantonen dat \bar{y} een onvertekende schatter is van \bar{Y} .¹⁵ In geval van een toevallige steekproef is de formule onder (1) een onvertekende schatter van de variantie.¹⁶

$$\text{var}(\bar{y}) = \frac{N-n}{nN(N-1)} \sum_{i=1}^N (Y_i - \bar{Y})^2. \quad (1)$$

Wat in het geval van het hoger veronderstelde binaire fenomeen neerkomt op de formule onder (2).¹⁷

$$\text{var}(p) = \frac{N-n}{(N-1)} \frac{P(1-P)}{n}. \quad (2)$$

Om uit deze formule vervolgens een redelijke waarde voor de steekproefomvang af te leiden, moeten we een welbepaald precisieniveau voor een uit te voeren hypothesetoets vastleggen. Het precisieniveau drukt uit hoe ver de waarde van het fenomeen, geschat op basis van steekproefwaarden, mag afwijken van de waarde van hetzelfde fenomeen in de populatie. Dit beoogde precisieniveau wordt hier gedefinieerd als d . Gegeven d , kan het vooropgestelde precisie criterium gedefinieerd worden zoals in vergelijking (3).

$$\text{Prob}(|p-P| \geq d) \leq \alpha. \quad (3)$$

¹⁵ Zie Jolliffe (1986), p. 29-30.

¹⁶ Zie Jolliffe (1986), p. 31-32.

¹⁷ Zie Jolliffe (1986), p. 36.

Vergelijking (3) stelt dat de kans dat het verschil in absolute waarde tussen de waarde p en P groter is dan het vooropgestelde precisieniveau d , niet groter mag zijn dan α . Indien we dan veronderstellen dat de verdeling van p benaderd wordt door de normale verdeling, kunnen we vergelijking (3) als volgt vertalen:¹⁸

$$\frac{d}{\sqrt{\frac{N-n}{N-1} \frac{P(1-P)}{n}}} \geq z_\alpha. \quad (4)$$

In (4), symboliseert z_α de kritische waarde van een standaardnormale verdeling met significantieniveau α . Wanneer we (4) nu oplossen naar n , levert dit een minimale waarde op voor de steekproefomvang die een bepaald precisieniveau zou garanderen om een bepaald fenomeen te schatten. We geven vergelijking (4), herschreven naar n , in vergelijking (5).

$$n \geq \frac{z_\alpha^2 NP(1-P)}{(N-1)d^2 + z_\alpha^2 P(1-P)}. \quad (5)$$

Bij wijze van voorbeeld veronderstellen we dat $P=0,5$. Het fenomeen neemt dus de waarde 1 aan in 50% van de gevallen. Indien we nu vergelijking (5) willen gebruiken om voor een gekende populatieomvang, een zinvolle waarde voor n af te leiden, moeten we nog het precisieniveau d en de waarde voor z_α vastleggen. Er zijn geen strikte criteria om het significantieniveau α voor een hypothesetoets vast te leggen. Een veelgebruikte waarde is $\alpha=0,01$.¹⁹ Indien we veronderstellen dat we geïnteresseerd zijn in een tweezijdige hypothesetoets, i.e. $H_0: p=P$ en $H_1: p \neq P$, dan impliceert deze waarde voor α een waarde 2,576 voor z_α .²⁰ De waarde voor d dient eveneens op een eerder arbitraire manier gekozen te worden. We leggen ze hier vast op 0,005, wat een afwijking van 1% boven en onder de veronderstelde gemiddelde waarde van P impliceert.

Gebruik makend van deze veronderstellingen, impliceert dit volgende invulling van vergelijking (5), gegeven de onbekende steekproef en populatieomvang:

$$n = \frac{(2,576)^2 \cdot N \cdot 0,5 \cdot 0,5}{(N-1)(0,005)^2 + (2,576)^2 \cdot 0,5 \cdot 0,5}.$$

Het tot nog toe gehanteerde voorbeeld is op minstens twee manieren zeer specifiek. Het is ten eerste gebaseerd op de veronderstelling dat de steekproef slechts zal gebruikt worden om één enkele hypothesetoets uit te voeren. Ten tweede hebben we verondersteld dat deze test betrekking heeft op proporties.

¹⁸ Zie Jolliffe (1986), p. 93.

¹⁹ Zie Siegel and Castellan (1988), p. 9.

²⁰ Zie Siegel and Castellan (1988), p. 322.

Om de eerste veronderstelling af te zwakken, kan men hetzelfde principe toepassen voor verschillende hypothesetoetsen en de optimale omvang n voor al deze toetsen afzonderlijk bepalen. De uiteindelijke omvang van de steekproef zou dan een compromis moeten zijn tussen de waarden voor n die men bekomt op basis van de verschillende toetsen. Het ligt voor de hand dat een bepaalde steekproef gebruikt wordt om meerdere hypothesen te toetsen. Desondanks wordt de optimale omvang toch vaak afgeleid op basis van één enkele omgekeerde hypothesetoets.²¹

Voor wat de tweede veronderstelling betreft, ligt het voor de hand dat de door ons beoogde steekproef voor meer dan alleen maar toetsen op proporties zal gebruikt worden. Om de steekproefomvang ook te kunnen bepalen op basis van omgekeerde hypothesetoetsen op continue fenomenen, zoals inkomens, introduceren we in vergelijking (6) een meer algemene versie van vergelijking (5).

$$n \geq \frac{z_{\alpha}^2 N \text{var}(Y)}{Nd^2 + z_{\alpha}^2 \text{var}(Y)}. \quad (6).$$

In vergelijking (6), symboliseert d het absolute verschil tussen \bar{y} en \bar{Y} en $\text{var}(Y)$ de variantie van de variabele Y in de populatie, i.e. $\frac{\sum_{i=1}^N (Y_i - \bar{Y})^2}{N-1}$.²² Voor toepassing van vergelijking (6) is het dus noodzakelijk om veronderstellingen te maken over de gemiddelde waarde \bar{Y} en variantie $\text{var}(Y)$ van het fenomeen Y in de populatie, daar waar het voor vergelijking (5) volstond om enkel een veronderstelling te maken over de waarde van P (naast natuurlijk de veronderstellingen over d en α die in beide gevallen vereist zijn).

We hanteren in wat volgt vergelijking (5) of (6) om de steekproefomvang af te leiden die nodig is om fenomenen waarin we later mogelijk geïnteresseerd zullen zijn, met een vooropgesteld precisieniveau te kunnen schatten. We herinneren er daarom aan dat de beoogde steekproef uiteindelijk gebruikt zal worden om fenomenen inzake sociale zekerheidsbijdragen, pensioenen, ziekte- en invaliditeitsuitkeringen werkloosheidsuitkeringen, sociale bijstandsuitkeringen, gezinsbijslagen of personenbelastingen te beschrijven. We verzamelden daarom gegevens met betrekking tot een aantal van de later te beschrijven fenomenen en pasten er vervolgens vergelijking (5) of (6) op toe. De resultaten van deze oefening kan men terugvinden in Tabel 1.

²¹ Zie Jolliffe (1986), p. 28.

²² Zie Molenberghs en Gaens (1994), p. 106.

Tabel 1: Steekproefomvang voor hypothesetoetsen op een aantal variabelen die relevant zouden kunnen zijn in latere simulatieoefeningen met MIMOSIS

Label	\bar{Y}	d =1% of \bar{Y}	$\text{var}(Y)$	z_α	n $N = 10.000.000$
(1)	(2)	(3)	(4)	(5)	(6)
(1) Beschikbaar jaarinkomen per belastingeenheid voor het aanslagjaar 2000	17.346	173,46	279.852.800	2,576	61.341
(2) Personen met een rustpensioen in het stelsel der werknemers in 2006 uitgedrukt als een fractie van de totale bevolking	0,130	0,00130	0,113100	2,576	425.204
(3) Personen met een overlevingspensioen in het stelsel der werknemers in 2006 uitgedrukt als een fractie van de totale bevolking	0,047	0,00047	0,044791	2,576	1.185.940
(4) Werkzoekende uitkeringsgerechtigde volledig werklozen (UVW – WZ) in 2006 uitgedrukt als een fractie van de totale bevolking	0,044	0,00044	0,042064	2,576	1.260.096
(5) Oudere werklozen in 2006 uitgedrukt als een fractie van de totale bevolking	0,011	0,00011	0,010879	2,576	3.736.756
(6) Andere werklozen dan oudere en UVW – WZ in 2006 uitgedrukt als een fractie van de totale bevolking	0,012	0,00012	0,011856	2,576	3.533.140
(7) Personen met een invaliditeitsuitkering in 2006 uitgedrukt als een fractie van de totale bevolking	0,021	0,00021	0,020559	2,576	2.362.643
(8) Aantal rechthebbenden met een uitkering voor arbeidsongevallen in 2006 uitgedrukt als een fractie van de totale bevolking	0,019	0,00019	0,018639	2,576	2.551.852
(9) Personen met een inkomensvervangende en/of integratietegemoetkoming voor gehandicapten in 2006 uitgedrukt als een fractie van de totale bevolking	0,013	0,00013	0,012831	2,576	3.350.217

In kolom (2) van Tabel 1 geven we telkens de waarde die we veronderstellen het populatiegemiddelde te zijn, i.e. de waarde voor \bar{Y} .²³ In kolom (3) geven we het precisieniveau dat we voor ieder van deze fenomenen willen bereiken met de latere steekproef. Zoals blijkt uit kolom (3) varieert het absolute niveau van het beoogde precisieniveau over de variabelen, maar in relatieve termen is het telkens gelijk aan 1% van de veronderstelde gemiddelde populatiewaarde, i.e. \bar{Y} . In kolom (4) van Tabel 1 geven we de veronderstelde variantie van Y in de populatie. Alle variabelen in Tabel 1, behalve die in lijn (1), zijn aandelen in de totale populatie. Uit vergelijking (2) volgt dat de variantie van dergelijke fenomenen benaderd kan worden door de waarde $P(1-P)$.²⁴ Voor alle variabelen veronderstellen we verder dat ze afkomstig zijn uit een normale verdeling. Met een voldoende aantal observaties impliceert dit dat een t-statistiek voor dit fenomeen, een standaard normale verdeling volgt. De kritische waarde z_α , die in dat geval overeenkomt met een significantieniveau voor $\alpha = 0,01$ voor een tweezijdige hypothesetoets is 2,576. Voor toepassing van de omgekeerde hypothesetoets hebben we verder ook een idee nodig van de omvang van de totale populatie. Aangezien we de precieze omvang van het steekproefkader op dit moment niet kennen, veronderstellen we dat de totale populatie een omvang heeft van 10.000.000 observatie-eenheden. In kolom (6) van Tabel 1 geven we tenslotte de steekproefomvang die minimaal vereist is om de verschillende parameters te schatten, gegeven het vooropgestelde precisie- en significantieniveau. We maken hierbij voor het fenomeen in lijn (1) de bijkomende vereenvoudigende veronderstelling dat de resultaten, geschat voor belastingeenheden, vergelijkbaar zijn met gegevens op individueel niveau.

²³ In sectie 9.1.6 van deze nota wordt meer uitleg gegeven over de precieze oorsprong van deze veronderstelde populatiewaarden.

²⁴ De precieze oorsprong van het gemiddelde en de variantie in lijn (1) van Tabel 1, geven we eveneens in sectie 9.1.6 van deze tekst.

De resultaten in Tabel 1 suggereren dat we zeer grote steekproeven zouden moeten trekken indien we de betreffende fenomenen zouden schatten met het vooropgestelde precisie- en significantieniveau. Alhoewel het vooropgestelde significantieniveau zeer precies is, verandert de basisconclusie over een steekproef met een zeer aanzienlijke omvang niet echt als het significantieniveau zou aangepast worden naar een waarde van $\alpha = 0,1$. Bij een zelfde precisieniveau zou men in dit geval bijvoorbeeld nog steeds 177.874 observatie-eenheden in de steekproef nodig hebben om het aandeel personen met een rustpensioen te schatten.²⁵

Uit Tabel 1 blijkt ook dat voor de schatting van een gemiddeld inkomen, de omvang van de steekproef veel beperkter mag zijn, gegeven het vooropgestelde precisie- en significantieniveau (zie lijn (1) in Tabel 1). Om de steekproef beheersbaar te houden lijkt het daarom aangewezen om te kiezen voor een compromiswaarde. We stellen daarom de eerder arbitrair gekozen waarde van 100.000 toevallig getrokken individuen voor, voor de steekproef die gebruikt zal worden voor de tweede etappe (i.e. de creatie van een data-set met voornamelijk afgeleide variabelen).

9.1.4 IMPLEMENTATIE VAN HET STEEKPROEFMECHANISME

De voorgaande procedure kan samengevat worden in een aantal stappen. De uitvoering van deze stappen, zou moeten resulteren in de steekproef die we beogen.

Stap 1: Leg een lijst aan van alle individuen uit de vorige steekproef die nog in leven zijn op 1 januari &Year+1.

Stap 2: Leg een lijst aan van alle individuen uit de vorige steekproef die nog in leven zijn op 1 januari &Year+1 en die voor de vorige steekproef getrokken werden via het mechanisme van systematische steekproeftrekking. Bepaal het cijfer n als 100.000 min de omvang van de groep die in deze stap omschreven werd.

Stap 3: Leg een lijst aan van alle individuen die ingeschreven zijn in het Rijksregister tussen 1 januari 2002 en 1 januari &Year+1. Bepaal het cijfer N als de omvang van de groep die in deze stap omschreven werd.

Stap 4: Trek n individuen uit het steekproefkader omschreven in Stap 3 op systematische wijze. Het startnummer dient in dit geval bepaald te worden als een random geselecteerd nummer tussen 1 en x , met $x = \left\lceil \frac{N}{n} \right\rceil$.

Stap 5: Voeg de lijsten, aangemaakt in Stap 1 en Stap 4 samen en leg een lijst aan met unieke nummers die weergeven tot welk gezin een individu behoort. Duid in de uitgebreide lijst aan welk individu geselecteerd werd via de procedure in Stap 4.

Stap 6: Voeg aan de lijst van individuen die reeds geselecteerd werden in Stap 5 alle individuen toe die a) in leven zijn op 1 januari &Year+1, b) die behoren tot hetzelfde private gezin als de reeds geselecteerde individuen en c) die nog niet tot de lijst uit Stap 5 behoren.

Stap 7: Vraag voor ieder individu, geselecteerd in Stap 6, de gegevens, opgesomd in sectie 3.2 van de tekst op

²⁵ De waarde van z_α is 1,645 in dit geval (zie Siegel and Castellan (1988), p. 322).

9.1.5 REFERENTIES

- [1] Administratie Planning en Statistiek (2003a), <http://aps.vlaanderen.be/sgml/reeksen/2571.htm>
- [2] ADSEI (2008), http://statbel.fgov.be/figures/d21_nl.asp
- [3] Atkinson, A., Rainwater, L. and Smeeding, T. (1995), *Income distribution in OECD Countries: Evidence from the Luxembourg Income Study*, Parijs: OECD.
- [4] FOD Sociale Zekerheid (2008a), http://socialsecurity.fgov.be/NL/nieuws_publicaties/publicaties/mimosis/
- [5] FOD Sociale Zekerheid (2008b), *Vademecum van de financiële en statistische gegevens over de sociale bescherming in België: periode 2002-2008*, Brussel: Federale Overheidsdienst Sociale Zekerheid, Directie-generaal Sociaal Beleid, Domein Kerngegevens.
- [6] Greene, W. H. (1997), *Econometric Analysis*, 3rd edn., Prentice-Hall, New Jersey: Upper Saddle River.
- [7] Johnson and Kotz (1989), *Encyclopedia of Statistical Sciences*, 9 volumes+suppl., Wiley, New York.
- [8] Jolliffe, F. R. (1986), *Survey design and analysis*, Ellis Horwood Limited, Chichester.
- [9] Ministerie van Binnenlandse Zaken (2003), <http://www.rijksregister.fgov.be/>.
- [10] Molenberghs, G. en Gaens, E. (1994), Gezondheidsenquête door middel van interview: Statistische beschouwing, *Werkdocument Limburgs Universitair Centrum*, Diepenbeek: Universitaire Campus.
- [11] Project AG/01/086 (2003): "Microsimulation Model Social Security": Note 5 Description of the sample design, *mimeo*, Leuven: Centrum voor Economische Studiën, Katholieke Universiteit Leuven.
- [12] Siegel, S. and Castellan, N. J. (1988), *Nonparametric systems for the behavioural sciences*, McGraw-Hill, New York.

9.1.6 BESCHRIJVING VARIABELEN GEBRUIKT IN TABEL 1

Voor het uitvoeren van de hierboven besproken omgekeerde hypothesetoetsen, worden in Tabel 1 een aantal populatiewaarden gerapporteerd onder \bar{Y} . In Tabel 2 geven we aan uit welke bron deze gegevens gerecupereerd werden.

Tabel 2: Toelichting bij gekozen populatiewaarden die gehanteerd worden voor de omgekeerde hypothesetoetsen

Label	Totaal aantal gevallen	\bar{Y}	$\text{var}(Y)$
(1)	(2)	(3)	(4)
(1) Beschikbaar jaarinkomen per belastingeenheid voor het aanslagjaar 2000		17.346	279.852.800
(2) Personen met een rustpensioen in het stelsel der werknemers in 2006 uitgedrukt als een fractie van de totale bevolking	1.367.599 ²⁶	0,130	0,113100
(3) Personen met een overlevingspensioen in het stelsel der werknemers in 2006 uitgedrukt als een fractie van de totale bevolking	496.946 ²⁷	0,047	0,044791
(4) Werkzoekende uitkeringsgerechtigde volledig werklozen (UVW – WZ) in 2006 uitgedrukt als een fractie van de totale bevolking	460.133 ²⁸	0,044	0,042064
(5) Oudere werklozen in 2006 uitgedrukt als een fractie van de totale bevolking	116.169 ²⁹	0,011	0,010879
(6) Andere werklozen dan oudere en UVW – WZ in 2006 uitgedrukt als een fractie van de totale bevolking	125.166 ³⁰	0,012	0,011856
(7) Personen met een invaliditeitsuitkering in 2006 uitgedrukt als een fractie van de totale bevolking	215.822 ³¹	0,021	0,020559
(8) Aantal rechthebbenden met een uitkering voor arbeidsongevallen in 2006 uitgedrukt als een fractie van de totale bevolking	195.391 ³²	0,019	0,018639
(9) Personen met een inkomensvervangende en/of integratietegemoetkoming voor gehandicapten in 2006 uitgedrukt als een fractie van de totale bevolking	134.011 ³³	0,013	0,012831

Zowel het gemiddelde als de standaardafwijking voor het gegeven in lijn (1) van Tabel 2 werd geschat op basis van een steekproef van gegevens die betrekking hebben op het aanslagjaar 2000.³⁴

De gegevens voor lijnen (2) tot en met (9) in kolom (3) van Tabel 2 werden afgeleid door de aantallen in kolom (2) te delen door 10.511.382. Dit laatste cijfer is een schatting van het aantal individuen in de populatie op 1 januari 2006.³⁵ De variantie in lijnen (2) tot en met (9) in kolom (4) van Tabel 2 werden bepaald door het gegeven uit kolom (3) te vermenigvuldigen met 1 minus het gegeven uit kolom (3), i.e. de benaderende formule $P(1-P)$ voor de variantie van variabelen die enkel de waarde 0 of 1 kunnen aannemen.

²⁶ FOD Sociale Zekerheid (2008b), Tabel 4.1, p. 161, som van mannen en vrouwen.

²⁷ FOD Sociale Zekerheid (2008b), Tabel 4.1, p. 161, som van mannen en vrouwen.

²⁸ FOD Sociale Zekerheid (2008b), Tabel 8.1, p. 194.

²⁹ FOD Sociale Zekerheid (2008b), Tabel 8.1, p. 194.

³⁰ FOD Sociale Zekerheid (2008b), Tabel 8.1, p. 194.

³¹ FOD Sociale Zekerheid (2008b), Tabel 3.8, p. 153.

³² FOD Sociale Zekerheid (2008b), Tabel 6.4, p. 180.

³³ FOD Sociale Zekerheid (2008b), Tabel 2.1, p. 316.

³⁴ Zie Project AG/01/086 (2003), p. 10.

³⁵ Zie ADSEI (2008).