

Methodiek Schoolverlatersinformatiesysteem meetjaren 2006 en 2007

Timo Huijgen
Paul van der Kolk
Esther Soudant

ROA-TR-2009/4

Researchcentrum voor Onderwijs en Arbeidsmarkt
Maastricht University
Postbus 616, 6200 MD Maastricht
T 043 3883647 F 043 3884914

secretary-roa@maastrichtuniversity.nl
www.roa.nl

Methodiek Schoolverlatersinformatiesysteem meetjaren 2006 en 2007

Timo Huijgen
Paul van der Kolk
Esther Soudant

ROA-TR-2009/4
June 2009

Researchcentrum voor Onderwijs en Arbeidsmarkt
Maastricht University
Postbus 616, 6200 MD Maastricht
T 043 3883647 F 043 3884914

secretary-roa@maastrichtuniversity.nl
www.roa.nl

Inhoud

1	Algemeen	1
	1.1 Inleiding	1
	1.2 De VO- en BVE-Monitor	2
	1.3 De HBO-Monitor	4
	1.4 De WO-Monitor	6
2	Vragenlijsten	7
	2.1 Algemeen	7
	2.2 De VO- en BVE-Monitor	8
	2.3 De HBO-Monitor	10
	2.4 De WO-Monitor	10
3	Classificaties	11
	3.1 Opleidingen	11
	3.2 Beroepen	12
	3.3 Branches	12
	3.4 Plaatsnamen	13
4	Dataverzameling en respons	13
	4.1 De VO- en BVE-Monitor	13
	4.2 De HBO-Monitor	18
	4.3 De WO-Monitor	19
5	Datacleaning	20
6	Nieuwe variabelen	24

7	Weging	27
	7.1 Weging 2006	27
	7.2 Weging 2007	28
	7.3 Controle op weegfactoren	30

1 Algemeen

1.1 Inleiding

Het Researchcentrum voor Onderwijs en Arbeidsmarkt (ROA) voert jaarlijks vier grootschalige schoolverlatersonderzoeken uit. Doel van deze onderzoeken is het verschaffen van inzicht in de relatie tussen het gevolgde onderwijs en de intrede op de arbeidsmarkt. Daarnaast kan de doorstroom binnen het onderwijs nader worden onderzocht. Bij alle vier de onderzoeken worden schoolverlaters c.q. afgestudeerden ongeveer anderhalf jaar na het behalen van het diploma geënquêteerd. De vier onderzoeken bestaan uit de VO-Monitor (VMBO, HAVO en VWO), de BVE-Monitor (BOL en BBL), de HBO-Monitor (het hoger beroepsonderwijs) en de WO-Monitor (het universitair onderwijs). Bij al deze onderzoeken gaat het om de gediplomeerde uitstroom uit het reguliere, door de overheid bekostigde onderwijs. Op basis van de resultaten wordt ieder jaar een aantal rapporten gepubliceerd waarin op landelijk niveau een overzicht wordt gegeven van de doorstroom van recent gediplomeerden naar de arbeidsmarkt en het vervolgonderwijs. De resultaten op instellingsniveau kunnen daarnaast door onderwijsinstellingen gebruikt worden als instrument binnen hun kwaliteitszorg.

In meetjaar 2007 is ook onder de ongediplomeerden van het AVO, VMBO en MBO een onderzoek verricht. Hiervoor is een aparte vragenlijst ontwikkeld. Alhoewel de opzet en technische uitvoering van dit ongediplomeerden onderzoek vergelijkbaar is met de reguliere schoolverlatersonderzoeken worden de ongediplomeerden in deze rapportage verder buiten beschouwing gelaten. De cijfers in dit rapport hebben dan ook alleen betrekking op de gediplomeerde uitstroom.

Het ROA heeft de projectleiding over deze schoolverlatersonderzoeken en is verantwoordelijk voor de instrumentontwikkeling, de wetenschappelijke analyses, rapportages en de landelijke statistiek. Dit houdt o.a. in dat het ROA de vragenlijsten samenstelt, de data controleert, opschoont en bewerkt en de landelijke rapportages verzorgt. Het veldwerk wordt verzorgd door DESAN Research

Solutions BV, die o.a. de internetscripts maakt, de vragenlijsten en brieven verstuurt, de internetfaciliteiten beschikbaar stelt, de data-invoer, het coderen en de afzonderlijke instellingsrapportages verzorgt.

In dit technical report wordt ingegaan op de (technische) uitvoering van bovengenoemde onderzoeken. Hierbij wordt allereerst een korte omschrijving van de populatie en steekproef van de verschillende monitoren gegeven (paragrafen 1.2, 1.3 en 1.4). Vervolgens wordt ingegaan op de vragenlijsten (hoofdstuk 2), de gebruikte classificaties (hoofdstuk 3), de dataverzameling en de respons (hoofdstuk 4), de gehanteerde datacleaning (hoofdstuk 5), de berekende nieuwe variabelen (hoofdstuk 6) en de toegepaste weging (hoofdstuk 7).

1.2 De VO- en BVE-Monitor

Populatie VO-Monitor

Gediplomeerde schoolverlaters van het algemeen voortgezet onderwijs (HAVO en VWO) en het voorbereidend middelbaar beroepsonderwijs (VMBO, inclusief de theoretische leerweg).

Steekproef VO-Monitor

De namen en adressen van de te benaderen schoolverlaters worden aangeleverd door de deelnemende onderwijsinstellingen. De steekproef in meetjaar 2006 was afhankelijk van het aantal deelnemende instellingen. In meetjaar 2007 is in aanvulling op de deelname door instellingen een aanvullende steekproef getrokken via de IB groep te Groningen. Dit is een naar opleiding en regio, landelijk representatieve steekproef op basis van de leerlingenbestanden van de IB groep. Er wordt bij deze IB steekproef rekening gehouden met de reeds deelnemende onderwijsinstellingen, zodat er geen personen dubbel benaderd worden.

Populatie BVE-Monitor BOL

In het kader van de BVE-Monitor worden gediplomeerde en ongediplomeerde schoolverlaters van de beroepsopleidende leerweg (BOL) benaderd.

Steekproef BVE-Monitor BOL

De BOL steekproef bestaat uit twee bronnen. Allereerst is onderwijsinstellingen gevraagd deel te nemen aan de BVE-Monitor. De deelnemende instellingen leveren vervolgens de adressen van de personen die gediplomeerd de instelling verlaten hebben. Ten tweede is, zowel in meetjaar 2006 als 2007 via de IB groep een landelijke, naar opleiding en regio, representatieve steekproef getrokken op basis van hun leerlingenbestanden. Hierbij worden niet alleen gediplomeerde maar ook ongediplomeerde leerlingen benaderd. Evenals bij de VO-Monitor wordt bij deze steekproef rekening gehouden met de reeds deelnemende onderwijsinstellingen, zodat er geen personen dubbel worden benaderd.

Populatie BVE-Monitor BBL

In het kader van de BVE-Monitor worden gediplomeerde en ongediplomeerde schoolverlaters van de beroepsbegeleidende leerweg (BBL) benaderd. Leerlingen die het diploma van een bepaald BBL niveau behaald hebben en binnen de instelling verder studeren in een studie met een hoger niveau worden niet benaderd.

Steekproef BVE-Monitor BBL

Net als de BOL steekproef worden onderwijsinstellingen gevraagd deel te nemen aan de reguliere BVE-Monitor. Evenals bij de BOL zijn zowel in meetjaar 2006 als in meetjaar 2007 aanvullende steekproeven via de IB groep getrokken.

In tabel 1.1 staan de aantallen benaderde gediplomeerde schoolverlaters en de totale populatie van de afgelopen twee meetjaren, verdeeld naar onderwijssoort. Het meetjaar 2006 heeft betrekking op scholieren die in het studiejaar 2004/2005

hun diploma behaald hebben. Meetjaar 2007 heeft de gediplomeerden van schooljaar 2005/2006 als populatie.

Tabel 1.1

Populatie en benaderd aantal schoolverlaters VO en BVE

Onderwijssoort	Meetjaar 2006		Meetjaar 2007	
	Benaderd	Populatie	Benaderd	Populatie
AVO	3.406	55.917	5.156	67.479
HAVO	1.802	36.709	2.909	38.152
VWO	1.604	19.208	2.247	29.327
VMBO	5.953	94.902	10.604	98.246
Theoretische leerweg	1.235	38.376	3.421	41.274
Landbouw	2.472	5.860	2.047	5.768
Techniek	859	16.444	1.855	14.906
Economie	765	16.629	1.436	17.580
Gezondheidszorg	622	17.593	1.845	18.718
BOL¹	21.367	79.900	20.299	75.713
Landbouw	--	3.081	876	2.858
Techniek	--	16.082	4.328	16.228
Economie	--	28.039	8.397	27.564
Gezondheidszorg	--	18.876	3.094	15.126
Gedrag en Maatschappij	--	13.822	3.604	13.937
BBL²	2.811	--	7.813	44.929
Landbouw	1.035	--	678	2.875
Techniek	768	--	2.981	16.870
Economie	531	--	2.038	10.517
Gezondheidszorg	309	--	1.525	11.481
Gedrag en Maatschappij	168	--	591	3.186

1.3 De HBO-Monitor

In samenwerking met de HBO raad voert het ROA al sinds 1990 een arbeidsmarktmonitor uit onder afgestudeerden van het hoger beroepsonderwijs.

-
1. Meetjaar 2006: voor de non-respondenten was de gevolgde opleiding niet altijd bekend. De benaderde aantallen kunnen dus voor BOL niet per sector worden weergegeven. Meetjaar 2007: de benaderde aantallen zijn exclusief de additionele steekproef die in opdracht van de Gemeente Rotterdam is uitgevoerd.
 2. In meetjaar 2006 was het niet mogelijk de BBL te wegen naar landelijke populatiecijfers.

Steekproef HBO-Monitor

De benodigde adresgegevens van de gediplomeerden zijn afkomstig uit de bestanden van de hogescholen. De deelname van de hogescholen aan de HBO-Monitor gebeurt op basis van vrijwilligheid en is al vele jaren zeer hoog. De afgelopen jaren lag het percentage benaderde afgestudeerden rond de 80 procent. Voor de studiejaren geldt dat de afstudeerdatum tussen 1 september en 31 augustus dient te liggen.

Populatie HBO-Monitor

In het gehele hoger onderwijs is sinds 2002 de 'bachelor-master' structuur ingevoerd. In het HBO zijn hierbij de bachelor opleidingen de equivalenten van de reguliere HBO opleidingen uit de oude structuur. De master opleidingen kunnen vergeleken worden met de vroegere HBO 2^e fase opleidingen (post initieel onderwijs). De populatie van de HBO-Monitor bestaat uit de afgestudeerden van de HBO bachelor opleidingen (en de nog resterende uitstroom van de vergelijkbare oude opleidingsvarianten). Daarnaast behoren de (nieuwe) master opleidingen tot de populatie voorzover ze in het Centraal Register Opleidingen Hoger Onderwijs (CROHO) als zodanig erkend worden.

Om de kloof te dichten tussen het middelbaar en het hoger beroepsonderwijs worden sinds enkele jaren zogenaamde 'Associate Degree' opleidingen aangeboden. Dit zijn (korte) HBO opleidingen die door HBO instellingen gegeven worden waarvan het diploma iets onder het niveau van de reguliere HBO bachelor variant ligt. Ook deze Associate Degree opleidingen behoren tot de populatie van de HBO-Monitor. Bij zowel de master- als de Associate Degree opleidingen gaat het echter voornamelijk om een relatief klein aantal opleidingen.

In tabel 1.2 staan de aantallen benaderde personen en de totale populatie van de afgelopen twee meetjaren. Het meetjaar 2006 heeft betrekking op afgestudeerden van het studiejaar 2004/2005; meetjaar 2007 op afgestudeerden van het studiejaar 2005/2006. In de beide meetjaren bestaat de steekproef nog voor een deel uit

uitstroom van de oude structuur opleidingen en voor een deel uit uitstroom uit nieuwe, bachelor opleidingen. In de komende jaren zal de uitstroom uit de eerstgenoemde groep langzaam verdwijnen.

Tabel 1.2

Populatie en benaderd aantal afgestudeerden HBO

Onderwijssoort	Meetjaar 2006		Meetjaar 2007	
	Benaderd	Populatie	Benaderd	Populatie
HBO	48.155	58.509	48.865	62.618
Landbouw	1.800	1.758	1.463	1.757
Onderwijs	8.936	10.845	10.421	10.404
Techniek	8.128	9.927	7.315	9.609
Economie	15.228	18.879	16.098	17.529
Gezondheidszorg	4.624	5.694	4.260	5.433
Gedrag en Maatschappij	6.355	8.355	6.373	8.032
Taal en cultuur	3.084	3.051	2.935	2.910

1.4 De WO-Monitor

Populatie WO-Monitor

De populatie van de WO-Monitor bestaat uit twee groepen. Allereerst behoren de afgestudeerden van de master opleidingen tot de populatie. Deze masters zijn te vergelijken met de uitstroom uit de vroegere initiële WO opleidingen. Ten tweede behoort de uitstroom uit de bachelor opleidingen tot de populatie voor zover die niet zijn doorgestroomd naar een master opleiding. De afstudeerdatum dient bij de WO-Monitor tussen 1 oktober en 30 september te liggen.

Steekproef WO-Monitor

De WO-Monitor vindt in VSNU verband plaats onder de afgestudeerden van alle 13 Nederlandse universiteiten. Hierbij neemt het ROA voor 10 universiteiten de uitvoering voor haar rekening. De Radboud Universiteit Nijmegen, de Universiteit Twente en de Rijksuniversiteit Groningen dragen zelf zorg voor de uitvoering. De benodigde adresgegevens zijn afkomstig uit de bestanden van de universiteiten. Evenals in het hoger beroepsonderwijs is men in het universitair onderwijs overgegaan tot de invoering van de bachelor-master opleidingstructuur. Bachelors

waarvan bekend is dat ze zijn doorgestroomd naar een master opleiding worden niet in de steekproef opgenomen. Deze master opleiding kan daarbij aan de eigen universiteit of aan een andere universiteit gevolgd worden. Reden hiervoor is dat de verder studerende bachelors toch al benaderd worden indien ze afstuderen aan een master opleiding. Daarnaast hebben ze vaak op het moment van enquêteren, ongeveer anderhalf jaar na hun afstuderen, ook de master (vervolg)opleiding met succes afgerond. Omdat studentenadministraties van universiteiten meestal niet in staat zijn om de vertrekkende bachelor studenten te blijven ‘volgen’, komt het in de praktijk er op neer dat alle bachelors die niet aan de eigen universiteit met een master studie verder gaan in de steekproef terecht komen. Dus ongeacht of men gaat werken of aan een andere universiteit een master gaat volgen.

De populatiecijfers van het WO staan in tabel 1.3. Het meetjaar 2006 heeft betrekking op afgestudeerden van het academisch jaar 2004/2005; meetjaar 2007 op afgestudeerden van het academisch jaar 2005/2006.

Tabel 1.3

Populatie en benaderd aantal afgestudeerden WO

Onderwijssoort	Meetjaar 2006 Benaderd/populatie	Meetjaar 2007 Benaderd/populatie
WO	28.889	30.404
Landbouw	1.010	961
Techniek	3.220	3.189
Economie	5.580	5.733
Gezondheidszorg	3.876	3.652
Gedrag en Maatschappij	6.405	7.446
Taal en cultuur	3.578	3.953
Recht en Openbare Orde	3.355	3.398
Natuurwetenschappen	1.874	2.072

2 Vragenlijsten

2.1 Algemeen

De ROA schoolverlatersonderzoeken worden sinds het uitvoeringsjaar 1996 volgens een geïntegreerd model uitgevoerd. Hierbij is sprake van één meetmoment en wordt er zoveel mogelijk gestreefd naar een basis set kernvragen. Ondanks deze

set kernvragen bestaan er meerdere vragenlijsten. Hiermee wordt recht gedaan aan de verschillen tussen de diverse onderwijsniveaus. Zo heeft ieder onderwijsniveau zijn eigen vragenlijst en zijn er aparte varianten voor het kunstonderwijs in het HBO.

In onderstaand hoofdstuk wordt ingegaan op de veranderingen die in de vragenlijsten in de meetjaren 2006 en 2007 zijn doorgevoerd. Hierbij wordt alleen gekeken naar veranderingen die structureel van aard zijn en behoren tot de set van basisvragen.

2.2 De VO- en BVE-Monitor

De VO-Monitor bestaat uit een VMBO vragenlijst en een HAVO/VWO vragenlijst. Voor beide vragenlijsten geldt dat er in 2006 en 2007 slechts een beperkt aantal vragen veranderd is. In het VMBO wordt vanaf 2006 niet meer gevraagd naar de totale werkloosheidsduur sinds het behalen van het diploma maar naar de werkloosheidsduur tot de eerste baan. De vraag over de kwaliteit van de informatie met betrekking tot verschillende aspecten van de studie- en beroepskeuze is komen te vervallen. Hiervoor in de plaats is een wat bredere vraagstelling aan de lijst toegevoegd die aan de hand van een aantal stellingen de mening van de schoolverlaters peilt over de gevolgde VMBO opleiding.

In 2007 is de vraag over het al dan niet leidinggeven in de huidige functie komen te vervallen omdat die voor startende VMBO-ers op de arbeidsmarkt zelden relevant is. In datzelfde meetjaar is een vraag naar de oordelen over de gevolgde opleiding ook op de VMBO lijst toegevoegd. Deze vraag stond ook op een groot aantal andere vragenlijsten.

Met ingang van meetjaar 2006 zijn op de HAVO/VWO vragenlijst de vragen met betrekking tot de examencijfers per gevolgd vak komen te vervallen. Deze vragen werden in het verleden gebruikt om het gevolgde vakkenpakket vast te stellen maar zijn met de invoering van de vaste studieprofielen minder relevant geworden.

Daarnaast zijn de vragen over de studie- en beroepskeuze sterk ingekort en zijn de vragen naar het oordeel van de gediplomeerden over de gevolgde opleiding gewijzigd en in overeenstemming met de andere vragenlijsten van de VO en BVE-Monitor gebracht. Dit laatste geldt eveneens voor de vragen over het voortijdig verlaten van de vervolgopleiding. Evenals in het VMBO wordt vanaf 2006 niet meer gevraagd naar de totale werkloosheidsduur sinds het behalen van het diploma maar naar de werkloosheidsduur tot de eerste baan.

In meetjaar 2007 is een vraag toegevoegd die betrekking heeft op de bekendheid van de zogenaamde Associate Degree opleidingen in het HBO. Omdat met name de HAVO uitstroom vervolgstudies in het HBO kiest kan aan de hand van de vragenlijst gekeken worden in hoeverre deze nieuwe Associate Degree opleidingen bij de doelgroep bekend waren op het moment van de studiekeuze en in hoeverre men overwogen heeft een dergelijke opleiding te gaan volgen.

In zowel de BOL als de BBL vragenlijst zijn in 2006 vragen toegevoegd over het voortijdig verlaten van de gevolgde opleiding en het eventueel voortijdig verlaten van de gekozen vervolgopleiding. Evenals in de VMBO als de HAVO/VWO vragenlijst zijn ook in de BOL en BBL lijsten de vragen opgenomen over de oordeelsvorming van de schoolverlaters met betrekking tot hun gevolgde opleiding.

In meetjaar 2007 zijn de BBL en de BOL vragenlijsten geïntegreerd tot één vragenlijst. De vragen over het voortijdig verlaten van de gevolgde opleiding, de eerder behaalde diploma's en de vraag over de totale intredewerkloosheid zijn komen te vervallen. De Associate Degree opleidingen in het HBO zijn met name opgezet om de kloof tussen het MBO en het HBO te verkleinen. Evenals in de HAVO/VWO vragenlijst is, vanwege de doorstroom van met name de BOL naar het HBO, het blok vragen toegevoegd over de bekendheid van de Associate Degree opleidingen.

In verband met het uitgebreide onderzoek in meetjaar 2007 onder ongediplomeerde schoolverlaters is in de vragenlijsten van de VO- en BVE-

Monitor een aantal nieuwe vragen toegevoegd, met name over het vervolgtraject in het onderwijs. Hierdoor is het mogelijk om op een aantal belangrijke aspecten een vergelijking te maken tussen de gediplomeerde en de ongediplomeerde uitstroom.

2.3 De HBO-Monitor

Met ingang van meetjaar 2006 wordt er weliswaar nog naar het soort buitenlandervaring van de afgestudeerden gevraagd maar niet meer naar het aantal maanden buitenlandervaring. Daarnaast is er in de jaren 2006 en 2007 aan 10% van de totale steekproefpopulatie van de sector HBO economie en aan 10% van de technische ICT opleidingen een extra set vragen voorgelegd met daarin vragen over hun domeinspecifieke competenties. In de HBO sector taal en cultuur is integraal sprake van domeinspecifieke vragensets. Ook de vragen naar het oordeel over de gevolgde opleiding zijn in 2006 enigszins aangepast. In verband met het toenemende belang van mobiliteitsvraagstukken zijn vanaf 2006 vragen toegevoegd over de woonplaatsen van de afgestudeerden tijdens verschillende fasen van hun studie/leven.

2.4 De WO-Monitor

De (kern-)vragenlijst die gebruikt wordt in de WO-Monitor wordt in VSNU verband door alle 13 universiteiten gezamenlijk opgesteld en jaarlijks aangepast. De kernvragen van de HBO-Monitor en de WO-Monitor zijn al sinds een aantal jaren vrijwel identiek. Dit omdat de benodigde kernindicatoren, voor zowel landelijke statistieken als onderwijsinstellingen zelf, in het hoger onderwijs voor een groot deel gelijk zijn. Dit komt mede door de in 2002 ingevoerde bachelor-master structuur binnen het gehele hoger onderwijs. De verschillen tussen de HBO vragenlijst en de WO vragenlijst zijn gering voor zover het de kernvragen betreft. Dit in verband met de vergelijkbaarheid binnen het gehele hoger onderwijs.

Evenals in het HBO zijn in het WO de vragen naar het oordeel over de gevolgde opleiding in 2006 enigszins aangepast, en zijn vanaf 2006 vragen toegevoegd over

de woonplaatsen van de afgestudeerden tijdens verschillende fasen van hun studie/leven.

3 Classificaties

Voor een aantal open vragen in de enquête worden de antwoorden volgens een standaardclassificatie gecodeerd. Het gaat hierbij om kenmerken als opleiding, beroepen, bedrijven c.q. branches en plaatsnamen.

3.1 Opleidingen

Voor het coderen van opleidingen wordt gebruik gemaakt van een door het ROA ontwikkelde opleidingsclassificatie. Omdat de bestaande nationale en internationale classificaties niet voldoen aan de eisen met betrekking tot mate van specificiteit en praktische bruikbaarheid is er voor gekozen een eigen indeling te hanteren. Deze classificatie is koppelbaar aan de nationaal veel gehanteerde Standaard Onderwijs Indeling (SOI) van het CBS, het Centraal Register Beroepsopleidingen (CREBO), het Centraal Register Opleidingen Hoger Onderwijs (CROHO), en de onderwijsclassificatie International Standard Classification of Education (ISCED).

In de ROA classificatie wordt onderscheid gemaakt naar onderwijstype, onderwijsvorm (regulier/duaal), fase (korte (middel-) lange opleidingen), sector, opleiding en eventueel de richting binnen de opleiding. De classificatie wordt jaarlijks bijgewerkt en wordt gebaseerd op studiegidsen, bestaande administratieve opleidingsindelingen zoals CREBO en CROHO en informatie van de deelnemende onderwijsinstellingen zelf.

Zoals eerder vermeld is in 2002 de Bachelor-Master structuur ingevoerd in het hoger onderwijs. Het onderscheid tussen opleidingen uit de oude structuur en de nieuwe BaMA opleidingen is voor onderzoeksdoeleinden van groot belang. Daarom zijn er in de opleidingsclassificatie voor de BaMa opleidingen ook nieuwe

opleidingscodes gecreëerd. In het HBO zijn dit met name bachelor varianten aangezien deze vergelijkbaar zijn met de reguliere, oude HBO opleidingen. HBO master opleidingen behoren vooralsnog slechts in beperkte mate tot de populatie. De master HBO opleidingen zijn dus slechts incidenteel in de classificatie terug te vinden. Voor het WO zijn zowel de bachelor als de master varianten in de classificatie opgenomen.

3.2 Beroepen

De beroepen worden gecodeerd aan de hand van de Standaard Beroepenclassificatie van het CBS (SBC-92). De SBC-92 is een classificatie van beroepen op grond van niveau en richting van de benodigde bekwaamheden om het beroep uit te oefenen in combinatie met de (maximaal 3) belangrijkste werksoorten. Er worden 1211 beroepen onderscheiden, ingedeeld in 43 beroepsklassen en 121 beroepsgroepen. De SBC is koppelbaar aan de internationaal veel gehanteerde ISCO indeling.

Omdat sommige beroepen binnen de SBC voor met name rapportagedoeleinden niet specifiek genoeg zijn is er door het ROA in incidentele gevallen een nog verdere onderverdeling gemaakt. Hierdoor zijn in de ROA schoolverlatersbestanden de beroepencodes 6 cijferig en niet 5 cijferig zoals in de CBS classificatie. In de ROA beroepencode-variant is als 6^e digit aan alle SBC codes een nul toegevoegd zodat er, daar waar nodig, toch ruimte is voor een nadere specificatie. Zo is bijvoorbeeld de SBC code 93208 ‘Advocaat, notaris, rechter, officier van justitie, juridisch adviseur’ in de schoolverlatersdata niet alleen uitgebreid naar de 6 digit code 932080 maar ook nader gedifferentieerd naar 932081 ‘Advocaat’, 932082 ‘Officier van justitie’ enz.

3.3 Branches

Voor het coderen van bedrijven en branches is gebruik gemaakt van de Standaard Bedrijfsindeling (SBI-1993) van het CBS. De SBI is een systematische hiërarchische indeling van economische activiteiten die zes niveaus kent. Evenals

de SBC is de SBI voor de schoolverlatersonderzoeken uitgebreid van een 5 cijferige code naar een 6 digit code. De SBI sluit aan bij de door Eurostat vastgestelde NACE indeling.

3.4 Plaatsnamen

Plaatsnamen worden gecodeerd aan de hand van gemeentecodes uit de meest recente versie van het Plaatsnamenregister welke afkomstig is van CENDRIS (een onderdeel van TNT Post).

4 Dataverzameling en respons

De wijze van dataverzameling gebeurde in 2006 en 2007 veelal schriftelijk of via Internet. In tabel 4.1 staat een schematisch overzicht van het veldwerk per onderwijssoort voor de uitvoeringsjaren 2006 en 2007.

4.1 De VO- en BVE-Monitor

AVO en VMBO

De adressen voor het benaderen van schoolverlaters van het AVO en het VMBO zijn enerzijds afkomstig van de deelnemende onderwijsinstellingen en anderzijds van de IB groep. Voor de deelnemende onderwijsinstellingen gebeurt de verzending van de aanbiedingsbrieven door DESAN of door de onderwijsinstellingen zelf. Bij centrale verzending leveren de scholen het adressenbestand aan DESAN. In meetjaar 2006 was er bij zowel het AVO als het VMBO nog sprake van decentrale verzending, dat wil zeggen verzending door de scholen zelf. Bij de AVO steekproef ging het om 22% van alle uitgezette vragenlijsten, en bij het VMBO om 14%. In 2007 is voor het eerst binnen het AVO en het VMBO een groep schoolverlaters benaderd via de IB groep. Ook de integraal deelnemende onderwijsinstellingen is de mogelijkheid geboden om dit via de IB groep te doen. In 2007 zijn er dan ook geen AVO of VMBO instellingen geweest die zelf de verzending van de vragenlijsten nog verzorgd hebben.

Tabel 4.1

Wijze van dataverzameling

	AVO		VMBO		BOL		BBL		HBO		WO	
	2006*	2007*	2006	2007	2006	2007	2006	2007	2006	2007	2006	2007
Verzending via:												
Onderwijs-instelling	22%		14%		5%		23%		1%	1%	22%	21%
DESAN	78%	48%	86%	45%	22%	39%	66%	21%	99%	99%	78%	79%
IB Groep		52%		55%	73%	61%	11%	79%				
Eerste mailing**												
	post	post	post	post	post	post	post	post	email+	email+	email+	email+
Vragenlijst bijgevoegd?	nee	nee	nee	nee	nee	nee	nee	nee	nee	nee	nee	nee
Eerste rappel												
Vragenlijst bijgevoegd?	nee	nee	ja	ja	ja	ja	ja	ja	nee	nee	nee	nee
Tweede rappel												
Vragenlijst bijgevoegd?	nee	nee	nee	nee	nee	nee	nee	nee	nee	nee	nee	nee
Telefonische benadering	nee	nee	nee	nee	nee	nee	nee	nee	ja	ja	ja	ja

* = dataverzameling alleen via internet

** = HBO en WO: ingeval onderwijsinstellingen e-mailadressen hebben aangeleverd via email anders reguliere post

In beide meetjaren zijn zowel voor het AVO als voor het VMBO alle eerste aanbiedingsbrieven met de reguliere post verstuurd. Aangezien de mogelijkheid bestond om via internet de enquête in te vullen, was bij de eerste mailing nog geen vragenlijst toegevoegd. De papieren vragenlijst is voor het VMBO met het eerste rappel verstuurd. Vanwege de traditioneel hoge (internet-) responscijfers onder de AVO schoolverlaters is bij de AVO steekproef geen schriftelijke lijst meegestuurd. Dit betekent dat deze groep alleen de mogelijkheid gekregen heeft om via internet de vragenlijst in te vullen. Voor AVO en VMBO scholen waar ook na het eerste rappel nog sprake was van een lage respons is nog een tweede rappelzending verstuurd.

In tabel 4.2 staan per onderwijssector de wijze van enquêteren en de responscijfers vermeld. In meetjaar 2006 lag de respons onder de AVO schoolverlaters op 45%; voor het VMBO was dit 41%. Voor het AVO was dit vergelijkbaar met meetjaar 2005, en voor het VMBO betekende dit een daling van 2%-punten. In 2007 was de respons onder de AVO en VMBO schoolverlaters gelijk namelijk 36%. Voor het AVO betekende dit een flinke daling ten opzichte van 2006. Het VMBO kende weliswaar ook een daling in de respons maar deze was minder groot dan bij het AVO. Deze daling heeft alles te maken met het verschil in de opbouw van de steekproeven tussen de genoemde meetjaren. In tegenstelling tot 2006 bestond in 2007 het overgrote deel van zowel de AVO als de VMBO steekproef uit de IB-steekproef. In 2007 is een groot deel van de aangeschrevenen benaderd via de IB groep en waren ook de NAW gegevens afkomstig van de IB groep. In 2007 zijn er zelfs geen decentrale, door de scholen zelf verzorgde, verzendingen gedaan (zie ook tabel 4.1). In het laatste meetjaar was weliswaar de steekproef groter maar is verhoudingsgewijs een veel groter deel niet meer uit naam van de individuele onderwijsinstelling benaderd. In het verleden is gebleken dat een benadering uit naam van een school een hogere respons oplevert dan een algemene benadering. Daarnaast hadden scholen die in het verleden zelf voor de verzending van de vragenlijsten zorgden hun administraties vrij goed op orde waardoor er weinig vervuiling in de adressenbestanden van deze scholen was en deze instellingen dus meestal een relatief hoge respons hadden.

Tabel 4.2Respons en wijze van enquêteren VO- en BVE-Monitor³

Onderwijs- soort	2006				2007			
	Respons	Enquêtemethode			Respons	Enquêtemethode		
		Internet	Schrifte- lijk	Tele- fonisch		Internet	Schrifte- lijk	Tele- fonisch
%	%	%	%	%	%	%	%	
AVO	45	100			36	100		
HAVO	40	100			34	100		
VWO	50	100			38	100		
VMBO	41	80	20		36	79	21	
VMBO TL	44	88	12		40	78	22	
Landbouw	38	78	22		24	98	2	
Techniek	32	76	24		32	75	25	
Economie	41	85	15		36	73	27	
Gezondheidszorg	54	75	25		46	77	23	
BOL niveau 1/2	27	87	13		21	74	26	
Landbouw	35	80	20		25	67	33	
Techniek	25	91	9		19	78	22	
Economie	32	86	14		19	75	25	
Gezondheidszorg	22	85	15		26	68	32	
BOL niveau 3/4	40	89	11		32	78	22	
Landbouw	39	84	16		44	75	25	
Techniek	39	91	9		30	83	17	
Economie	46	92	8		26	80	20	
Gezondheidszorg	38	84	16		37	74	26	
Gedrag en maatsch.	35	89	11		35	75	25	
BBL niveau 1/2	-	79	21		24	71	29	
Landbouw	-	65	35		26	72	28	
Techniek	-	93	7		21	71	29	
Economie	-	84	16		24	75	25	
Gezondheidszorg	-	86	14		37	67	33	
BBL niveau 3/4	-	87	13		32	73	27	
Landbouw	-	82	18		31	76	24	
Techniek	-	94	6		29	73	27	
Economie	-	90	10		27	76	24	
Gezondheidszorg	-	86	14		41	71	29	
Gedrag en maatschappij	-	84	16		36	74	26	

3. Bij de landelijke BOL steekproef door de IB groep is voor het jaar 2006 het niveau en de richting van de gevolgde opleiding niet bekend. Daardoor zijn de responscijfers voor BOL in dat jaar exclusief de IB Groep steekproef.

Dit alles betekent dat door de verandering van de steekproefopzet de responscijfers van het AVO en het VMBO van 2007 niet te vergelijken zijn met die van voorgaande jaren. De meting van 2008 moet uitwijzen in hoeverre de responscijfers zich voor deze onderwijssoorten stabiliseren. Ten opzichte van de andere onderwijssoorten betekent dit dat het AVO nu lager scoort dan het HBO en WO, terwijl het AVO voorheen altijd de hoogste responscijfers kende.

BOL en BBL

Voor zowel de BOL als de BBL geldt dat de adressen voor een deel van de deelnemende scholen afkomstig zijn en voor een deel van de IB groep. Voor de BOL steekproef werd in 2006 73% en in 2007 61% van de schoolverlaters via de IB groep benaderd. In het BBL is er pas vanaf 2007 sprake van een grote landelijke IB steekproef. In beide onderwijssoorten was er alleen in meetjaar 2006 nog sprake van verzending door enkele scholen zelf. In 2007 gebeurde de gehele verzending centraal via DESAN of via de IB groep.

Voor beide onderwijssoorten zijn alle mailings in beide meetjaren per reguliere post verzonden en is bij de tweede zending, de eerste rappelbrief, een schriftelijke versie van de vragenlijst bijgevoegd. Deze schriftelijke vragenlijst heeft er in 2006 in geresulteerd dat op BBL niveau 1/2 de respons voor één vijfde bestaat uit schriftelijke respons en de rest uit internetrespons. Op BBL niveau 3/4 was dit 13%. Op zowel BOL niveau 1/2 als 3/4 koos iets meer dan één tiende van de respondenten voor de schriftelijke variant. Voor meetjaar 2007 liggen deze percentages over de gehele linie hoger hetgeen opmerkelijk is. In de afgelopen jaren heeft er wat betreft invulmethode een verschuiving van schriftelijk naar internet plaatsgevonden. Deze trend lijkt met de 2007 cijfers een halt te zijn toegeroepen en zelfs enigszins teruggedraaid te zijn. De cijfers van 2007 wijzen er in ieder geval op dat de schriftelijke lijst nog altijd een belangrijk alternatief voor de online vragenlijsten is.

Vergelijkbaar met het VMBO en het AVO zijn ook voor de BOL de responscijfers in 2007 gedaald ten opzichte van 2006. Ook hier wordt dit veroorzaakt door het grotere aandeel dat de IB steekproef heeft in de totale steekproef. Ten opzichte van 2005 is de respons voor het BOL in 2006 weinig veranderd. Door de daling zijn de responscijfers voor 2007 uitgekomen op 21% voor BOL niveaus 1/2 en 32% voor BOL niveaus 3/4. Voor de BBL gelden in 2007 vergelijkbare responspercentages (zie tabel 4.2).

4.2 De HBO-Monitor

In tabel 4.3 staan enkele cijfers over de respons en de wijze van enquêteren onder de afgestudeerden van het hoger beroepsonderwijs, de HBO-Monitor van de jaren 2006 en 2007. In het verleden is gebleken dat een groot deel van de respondenten de vragenlijst bij voorkeur via internet invult. Evenals in de vorige twee meetjaren heeft de HBO-Monitor dan ook plaatsgevonden via een internet vragenlijst.

Tabel 4.3

Respons en wijze van enquêteren HBO-Monitor

Onderwijs- soort	Respons %	2006 Enquêtemethode		Respons %	2007 Enquêtemethode	
		Internet %	Tele- fonisch %		Internet %	Tele- fonisch %
		HBO	42		99	1
Landbouw	44	99	1	44	99	1
Onderwijs	42	99	1	37	98	2
Techniek	45	99	1	40	99	1
Economie	39	99	1	35	99	1
Gezondheidszorg	46	99	1	43	99	1
Gedrag en maatschappij	44	98	2	42	99	1
Taal en cultuur	40	100		39	100	

Alle adressen (postadressen en indien bekend ook email adressen) die nodig zijn voor het benaderen van HBO afgestudeerden zijn afkomstig van de deelnemende hogescholen. Voor vrijwel alle vragenlijsten geldt dat ze door DESAN centraal verstuurd worden. Omdat de eerste mailing bij voorkeur per email gaat, worden de hogescholen verzocht om behalve de postadressen ook zoveel mogelijk email adressen aan te leveren. In de gevallen waar email adressen niet beschikbaar of niet

bruikbaar zijn, wordt een brief met de gewone post verstuurd. In beide meetjaren zijn na de eerste (elektronische) mailing nog twee rappels met de reguliere post verstuurd. Omdat de vragenlijst enkel nog via internet in te vullen is, zijn er in beide jaren geen schriftelijke versies van de vragenlijsten verstuurd. Aangezien er bij de intake een responsgarantie gegeven wordt aan de hogescholen wordt er door DESAN in sommige gevallen nog een telefonische belronde gehouden. Hierbij wordt in eerste instantie gevraagd om de internetvragenlijst alsnog in te vullen. Daarna gaat men pas over tot het afnemen van een telefonische enquête. Zoals uit tabel 4.3 blijkt, gaat het hierbij om slechts een zeer klein deel van de respondenten.

Sinds meetjaar 2007 is ook de respons onder de HBO afgestudeerden onder de 40% gedoken. In het verleden was de respons onder deze groep afgestudeerden altijd relatief hoog ten opzichte van bijvoorbeeld het MBO. Van een respons van 45% in meetjaar 2005 is de deelname gedaald naar zo'n 38% in het laatste meetjaar. Met name onder de afgestudeerden van de sector economie is de respons laag te noemen. Om een eind te maken aan deze dalende trend is het voornemen om in meetjaar 2008 toch weer een schriftelijke vragenlijst mee te sturen.

4.3 De WO-Monitor⁴

Evenals de HBO-Monitor heeft de WO-Monitor de afgelopen jaren alleen via een internetenquête plaatsgevonden (zie tabel 4.4).

4. Het ROA draagt zorg voor de uitvoering van de WO-Monitor voor 10 van de 13 universiteiten die in landelijk verband meedoen met de monitor. Paragraaf 4.3 heeft betrekking op de dataverzameling van alleen deze 10 instellingen.

Tabel 4.4

Respons en wijze van enquêteren WO-Monitor

Onderwijssoort	Respons %	2006 Enquêtemethode		Respons %	2007 Enquêtemethode	
		Internet %	Telefonisch %		Internet %	Telefonisch %
WO	44	100		40	99	1
Landbouw	50	100		52	99	1
Onderwijs	42	100		43	99	1
Techniek	55	100		54	99	1
Economie	35	100		33	100	
Gezondheidszorg	46	100		45	100	
Gedrag en maatschappij	46	100		43	99	1
Taal en cultuur	41	100		39	99	1
Recht en opb. Orde	36	100		33	99	1
Natuurwetenschappen	44	99	1	39	99	1

De email gegevens en de postadressen zijn door alle 10 universiteiten aan DESAN geleverd, en de verzending heeft dus in beide meetjaren volledig centraal plaatsgevonden. In zowel meetjaar 2006 als meetjaar 2007 is de eerste mailing voor zover mogelijk in elektronische vorm verstuurd. Daarna zijn nog twee schriftelijke rappels verstuurd per reguliere post. Vanwege de aan universiteiten gegeven responsgarantie is er in een aantal sectoren en voor een aantal universiteiten nog een telefonische belronde geweest. De responscijfers zijn iets hoger dan die in het HBO. In 2006 was de gemiddelde respons onder WO afgestudeerden 44%; in 2007 was dit gedaald naar 40%. Met name afgestudeerden uit de sectoren economie en recht en openbare orde reageren minder vaak. De sector techniek en de sector landbouw doen het daarentegen redelijk goed wat responscijfers betreft.

5 Datacleaning

Na de dataverzameling voegt DESAN de verschillende bestanden samen en worden er enkele controles uitgevoerd alvorens de data naar ROA voor verdere datacleaning worden gestuurd. Deze controles hebben betrekking op: de aanwezigheid van dubbele cases in de data (indien de respondent bijvoorbeeld abusievelijk zowel de schriftelijke - als de internet vragenlijst gebruikt) en gaten in

de responsadministratie (zijn er groepen of scholen overgeslagen?) en de gecodeerde opleidingsvariabele. Daarnaast worden voor de HBO vragenlijsten de vraag over voltijd/deeltijd/duale studievariant aangevuld met administratiegegevens indien de vraag niet is ingevuld door de respondent. Aan de hand van de gecodeerde landenvariabelen berekent DESAN tenslotte nog enkele variabelen die betrekking hebben op etniciteit en landenindelingen (bijvoorbeeld EU versus niet EU).

De datacleaning en dataverwerking die het ROA uitvoert na ontvangst van de bestanden van DESAN bestaat uit drie hoofdonderdelen:

- controle op variabelen en verwijzingsstructuren, dit wordt in de rest van dit hoofdstuk besproken;
- het aanmaken van samengestelde, zogenaamde ‘nieuwe’ variabelen, dit wordt in hoofdstuk 6 nader toegelicht;
- het wege van de data; dit wordt in hoofdstuk 7 besproken.

De datacleaning van het ROA richt zich met name op het controleren van de vragenlijststructuur met betrekking tot verwijzingsvragen en inconsistenties tussen antwoorden van respondenten. Deze inconsistenties zijn onvermijdelijk omdat er altijd respondenten zijn die zich vergissen, vragen verkeerd begrijpen of vragen niet invullen.

Algemene variabelencontrole

Allereerst wordt gekeken of de variabelen de goede naam, het goede label, de goede value labels en het goede format hebben. Daarnaast kunnen er variabelen niet in de aangeleverde bestanden zitten of kunnen er extra variabelen in aanwezig zijn. Deze controle wordt gedaan met behulp van een variabelen database die bij de opmaak van de vragenlijsten door het ROA bijgewerkt wordt. Deze variabelen database dient als basis voor de data-entry programma's en internetscripts, en is ook in de verdere datacleaning een belangrijke informatiebron.

Controle basisvariabelen

Een aantal basisvariabelen (zoals gevolgde opleiding, diplomabezit en afstudeerjaar) wordt gecontroleerd op validiteit. Naar aanleiding hiervan worden enkele eerder door DESAN aangemaakte hulpvariabelen over de afstudeerperiode en bruikbaarheid (de zogenaamde afboekcodes) van cases bijgewerkt.

Verwijzingsvragen

Iedere vragenlijst wordt gecontroleerd aan de hand van de ingebouwde verwijzingsstructuren. Vragen die men niet had hoeven invullen worden leeggemaakt. Doordat meer en meer gebruik gemaakt wordt van internetscripts worden heel veel verwijzingen automatisch doorgevoerd waardoor onterecht ingevoerde antwoorden steeds minder voorkomen. Aan de hand van de verwijzingsstructuur wordt een onderscheid gemaakt tussen 'system' missing waarden en 'user' missing waarden. In beide gevallen is de variabele niet gevuld voor de desbetreffende case. System missing houdt in dat de variabele (vraag) niet relevant is. Zo hoeven personen zonder werk vragen over de huidige baan niet in te vullen. User missing (in het bestand altijd de waarde -9) houdt in dat de respondent de vraag wel had moeten invullen maar dit niet heeft gedaan. Zo zal bijvoorbeeld niet iedereen die betaald werk heeft zijn of haar inkomen willen invullen.

Voor het cleanen van verwijzingsvragen zijn vaste regels die voor alle vragenlijsten gehanteerd worden. Indien de respondent een antwoord heeft gegeven waar een verwijzing op van toepassing is, hangt de cleaning van de verwijzingsvraag af van het aantal vervolgvragen dat niet ingevuld hoefde te worden. Als er drie of minder vragen zijn, wordt de verwijzingsvraag zelf gecorrigeerd indien al deze vragen (toch) zijn ingevuld. Zijn er vier of meer vragen die volgens de verwijzing niet hadden hoeven te worden ingevuld dan wordt de verwijzingsvraag pas gecorrigeerd indien minstens 3 van deze vragen (toch) een geldig antwoord hebben. Indien de respondent volgens zijn antwoord niet onder de verwijzing valt en de volgende vraag gewoon moet invullen, worden deze op -9 (user missing) gezet indien hij de

vragen (toch) niet ingevuld heeft. De verwijzingsvraag zelf wordt in dit laatste geval niet gecorrigeerd.

Inhoudelijke variabelencontrole

Naast de controle op format, naam en labels wordt gekeken of de variabelen in de databestanden gevuld zijn met toegestane, valide waarden. Bij gesloten vragen moeten de waarden van variabelen overeenkomen met de antwoordcategorieën die op de vragenlijst staan. Daarnaast wordt bij numerieke variabelen gecontroleerd op extreme en niet valide waarden. Extreme waarden worden tijdens de cleaning terug gezet naar user-missing (-9). Bij de vraag over de maatschappelijke positie (belangrijkste bezigheid op moment van enquêteren) worden de ingevulde antwoorden bij de 'anders'-categorie waar mogelijk teruggebracht naar een bestaande voorgecodeerde categorie.

Meerkeuzevragen worden in de data vastgehouden als dichotome variabelen, waarbij iedere antwoordcategorie een aparte variabele is. Hierbij staat de waarde 0 voor 'niet genoemd' en de waarde 1 voor 'wel genoemd'. De gehanteerde datacleaning houdt bij deze vragen het volgende in: indien er geen enkele variabele gevuld is, krijgen alle variabelen van deze vraag de waarde -9 (user-missing). Als er minimaal 1 variabele gevuld is met de waarde 1 (wel genoemd) krijgen alle niet gekozen variabelen van deze vraag de waarde 0 en worden dus aangemerkt als 'niet genoemd'.

Bij sommige vragen dient de respondent aanvullende informatie (numerieke informatie of tekst) te geven indien er voor een bepaalde antwoordcategorie gekozen is. Zo wordt bijvoorbeeld naar het aantal personen waar men leiding aan geeft gevraagd indien er 'ja' aangegeven is op de vraag of men leiding geeft. Dit soort vragen worden gecontroleerd en gecorrigeerd op onderlinge consistentie.

Arbeidsuren en inkomen

In de vragenlijsten van het hoger onderwijs wordt onderscheid gemaakt tussen het aantal arbeidsuren in de reguliere werkweek en het aantal uren uit overwerk. Als het aantal uren van de reguliere werkweek ingevuld is en het aantal uren overwerk niet, wordt het aantal uren overwerk op 0 gezet. Dit wordt gedaan omdat sommige respondenten geneigd zijn voor hen niet relevante vragen niet in te vullen. Als het aantal uren in de reguliere werkweek onbekend is en het aantal uren overwerk wel ingevuld is, wordt het aantal uren in de reguliere werkweek gevuld met het aantal uren overwerk. Het aantal uren overwerk wordt vervolgens op 0 gezet. Eenzelfde procedure wordt gehanteerd voor de vraag over arbeidsuren in eventuele nevenfuncties. Bij de drie inkomensvragen in het hoger onderwijs worden eveneens, indien minstens één van de inkomens is ingevuld, de andere variabelen gehercodeerd van missing naar 0.

6 Nieuwe variabelen

Na het proces van datacleaning wordt een aantal nieuwe variabelen aangemaakt. Nieuwe variabelen zijn in het bestand te herkennen aan de 'n_' notatie waarmee de variabelenaam begint. Deze variabelen worden gemaakt om het gebruiksgemak van de bestanden te vergroten en te zorgen voor continuïteit en vergelijkbaarheid tussen zowel de verschillende vragenlijsten als de verschillende meetjaren. Een nieuwe variabele kan om twee redenen worden gemaakt. Allereerst om (kleine) verschillen in vraagstelling tussen de onderlinge vragenlijsten op te heffen. Vanwege de verschillen in onderzoekspopulaties is bijvoorbeeld de vraagstelling met betrekking tot het volgen van vervolgonderwijs niet overal hetzelfde. Om echter toch aan te kunnen geven welke respondenten naar vervolgonderwijs zijn gegaan, wordt hiervoor één indicator berekend. Daarnaast zijn er nieuwe variabelen die niet op één maar op meerdere vragen, en dus meerdere variabelen van een vragenlijst, gebaseerd zijn. Hierbij gaat het met name om variabelen die betrekking hebben op

werkloosheid, beroepsbevolking en dergelijke. In de rest van dit hoofdstuk worden enkele van de belangrijkste nieuwe variabelen kort toegelicht.

Baankenmerken

- huidige functie aantal arbeidsuren (n_hau): in verband met extreme waarden wordt 0.5% van de bovengrens van deze variabele gehercodeerd naar -9 (user missing).
- bruto uurloon huidige functie (n_hbiu): de variabele is gebaseerd op het inkomen en de arbeidsuren van de huidige functie exclusief overwerk. In verband met extreme waarden wordt 0.5% van de ondergrens en 0.5% van de bovengrens op missing (-9) gezet.
- volledige of deeltijd betrekking (n_hau33): personen die meer dan 32 uur werken in een normale week, dus zonder overuren en uren van eventuele nevenfuncties, hebben een voltijd betrekking.
- dienstverband (n_hdvb2): dienstverband huidige functie (zoals 'loondienst' en 'zelfstandige') .
- soort aanstelling (n_hasd): deze indicator geeft aan of de huidige aanstelling een tijdelijk of vast karakter heeft.
- functie op eigen niveau (n_hnwg): de variabele geeft aan of de respondent een functie heeft die minimaal op zijn eigen niveau is of niet. Deze variabele wordt bepaald door het door de werkgever vereiste opleidingsniveau te vergelijken met het niveau van de gevolgde opleiding van de afgestudeerde.
- functie binnen eigen richting (n_horwg): indicator die aangeeft of iemand in zijn eigen/verwante richting werkt of in een andere (of geen specifieke) richting.

Arbeidsmarktkenmerken

- beroepsbevolking (n_berbv1, n_berbv2): iemand die 12 uur of meer werkt behoort tot de werkzame beroepsbevolking; iemand die niet of minder dan 12 uur werkt maar wel op zoek is naar (meer) betaald werk behoort tot de werkloze beroepsbevolking. Respondenten die aangegeven hebben zichzelf als student te zien worden niet meegenomen bij de berekening van de n_berbv2 variabele.

Deze correctie is noodzakelijk vanwege de specifieke samenstelling van de onderzoekspopulatie. Van de populatie studeert een aanzienlijk deel verder, en hebben velen naast hun studie een bijbaantje van een aantal uren in de week. Het niet uitsluiten van deze groep zou leiden tot een vertekend beeld van de werkzame en werkloze beroepsbevolking. De variabele *n_berbv2* wordt dan ook standaard in rapportages gebruikt.

- aanbieden op de arbeidsmarkt (*n_aanb1*, *n_aanb2*): de variabele geeft aan of iemand zich wel of niet aanbiedt op de arbeidsmarkt. Degenen die zich aanbieden behoren tot de beroepsbevolking (werkend of werkloos). Ook hier worden de 'studenten' bij de tweede variant van de variabele (*n_aanb2*) buiten beschouwing gelaten.

Gevolgde opleiding

- sector van gevolgde opleiding (*n_sector*): indeling van de gevolgde opleidingen naar onderwijssector.
- sector en niveau van gevolgde opleiding (*n_gokop2*): indeling van de gevolgde opleidingen naar onderwijssector en onderwijsniveau. Het aantal sectoren kan per onderwijsniveau verschillen.

Vervolgonderwijs

- sector en niveau van de vervolgopleiding (*n_vokop2*): indeling van de gevolgde vervolgopleidingen naar onderwijssector en onderwijsniveau. Het aantal sectoren kan per onderwijsniveau verschillen.
- niveau vervolgonderwijs (*n_nivvvo*): deze variabele geeft aan of de vervolgopleiding van een hoger niveau is dan de gevolgde, verlaten opleiding.

De hier genoemde nieuwe variabelen vormen slechts een klein deel van het totale aantal nieuwe variabelen. Een volledig overzicht van de nieuwe variabelen is te vinden in het bij de bestanden behorende documentatiesysteem.

7 Weging

7.1 Weging 2006

Voor de weging worden alleen de gediplomeerde respondenten geselecteerd. De grootte en de spreiding van de steekproef wordt bepaald door de deelname van de onderwijsinstellingen. Hierdoor ontstaat er een ongelijke spreiding naar opleiding, regio, duale-/voltijd-/ deeltijdopleidingen en geslacht. Als gevolg van de beperkte deelname van onderwijsinstellingen voor BBL, wordt BBL niet gewogen. Daar tegenover staat dat alle instellingen van het WO (uitzondering is de Open Universiteit) deelnemen aan het onderzoek waardoor er voor het WO geen ongelijke spreiding verkregen wordt en er dus niet gecorrigeerd hoeft te worden.

De weging voor HAVO, VWO, VMBO en BOL vindt plaats op opleiding, regio en geslacht. Duale-/ voltijd-/ deeltijdopleidingen komen hier niet voor. Alleen voor het HBO wordt het vierde kenmerk (dual/ voltijd/ deeltijd) meegenomen, wat resulteert in een vierdimensionale weging. De weging naar opleiding vindt plaats op het niveau van de unieke Centraal Register opleidingscodes, die gebruikt worden in de Integrale Leerling Telling (ILT). Voor HAVO, VWO, VMBO en BOL is dit de CREBO code en voor HBO de CROHO code. Bij regio wordt een onderscheid gemaakt in West vs. de rest, waarbij West de provincies Noord-Holland, Zuid-Holland, Utrecht en Zeeland omvat. Voor het bepalen van de regio indeling wordt de plaats van de onderwijsinstelling als basis genomen. Dual-/ voltijd-/ deeltijdopleidingen (alleen HBO) is een administratiekenmerk welk door de deelnemende instellingen wordt opgegeven. Geslacht is een vraag op de vragenlijst.

Alhoewel de voor de weging gebruikte populatiecijfers afkomstig zijn van het CBS vormen de ILT data eigenlijk de basis voor de landelijke aantallen. De cijfers voor het wegen van de ((V)MBO, HBO) landbouw sector zijn afkomstig van (het voormalige) STOAS Onderzoek te Wageningen. Voor HBO opleidingen uit de sector kunst wordt het totale aantal uitgezette vragenlijsten gebruikt als populatiecijfer omdat de ILT data niet tot betrouwbare aantallen leiden. De

weegfactor wordt bepaald door het betreffende populatie-aantal te delen door het bijbehorende steekproefaantal. Tegelijkertijd betekent dit dat het steekproefaantal geprojecteerd wordt naar het bijbehorende populatie-aantal. Voor het WO wordt alleen de totale steekproefomvang geprojecteerd naar het populatie-aantal. Dit is geen correctiefactor. Bij wijze van spreken wordt de steekproef “opgeblazen” naar het populatie-aantal.

Doordat het aantal deelnemende instellingen met betrekking tot de BOL opleidingen (ROC's) onvoldoende was, is gebruik gemaakt van een aanvullende steekproef via de IB-Groep (Informatie Beheer Groep in Groningen). Dit is een volledig a-selecte steekproef. Hierdoor ontstaan er in de steekproef van BOL twee groepen, de deelnemende instellingen en de a-selecte aanvullende steekproef. De deelnemende instellingen hebben we conform bovenstaande procedure gewogen. De bijbehorende populatie van de deelnemende instellingen is in mindering gebracht op de totale BOL populatie. Vervolgens is de a-selecte aanvullende steekproef geprojecteerd naar het resterende populatie-aantal.

Na controle van de BOL weging bleek dat er een ongelijke verdeling ontstond in de niveaus van de BOL opleidingen. De lagere niveaus waren ondervertegenwoordigd, wat mede veroorzaakt wordt doordat scholieren die binnen een instelling dezelfde opleiding gaan volgen op een hoger niveau, minder goed responderen. De BOL niveau verdeling in de gewogen steekproef is gecorrigeerd naar de BOL niveau verdeling van de populatie.

7.2 Weging 2007

Mede door de ontstane mogelijkheden van de IB-Groep om gedetailleerde a-selecte steekproeven te trekken voor HAVO, VWO, VMBO, BOL en BBL, is er in 2007 van deze steekproeven gebruik gemaakt. Ook is er in 2007 een aparte steekproef getrokken onder ongediplomeerde schoolverlaters, waarvoor ook een aparte vragenlijst was. Voor HBO en WO is in 2007 geen aanvullende steekproef nodig, doordat het aantal deelnemende instellingen een hoge dekkinggraad geeft.

Rotterdam neemt eveneens deel aan het schoolverlatersonderzoek met een ruime steekproef van onderwijsinstellingen in Rotterdam. Hierdoor ontstaan er verschillende steekproefgroepen die voor de weging apart behandeld worden:

De HAVO/VWO steekproef bestaat uit de volgende groepen:

- Reguliere deelnemende instellingen
- Rotterdamse onderwijsinstellingen
- IB-steekproef

De VMBO, BOL en BBL steekproef bestaat uit:

- Reguliere deelnemende instellingen
- Rotterdamse onderwijsinstellingen
- Agrarische onderwijsinstellingen
- IB-steekproef

In de IB-steekproef wordt rekening gehouden met de reeds deelnemende instellingen, zodat schoolverlaters niet dubbel in de steekproef terecht komen. Aan de hand van de ILT-data wordt per onderwijssoort en per steekproefgroep de populatie bepaald. Bij VMBO wordt een onderscheid gemaakt naar sector (techniek, zorg en welzijn, economie) en binnen sector naar leerweg (Basis Beroepsgericht, Kader Beroepsgericht, Gemengde Leerweg) en tenslotte de theoretische leerweg. Bij BOL en BBL wordt afzonderlijk de populatie bepaald naar niveau (AB, BB, MKF, ZB). Voor de agrarische instellingen wordt de populatie bepaald aan de hand van de IB-steekproefbestanden voor VMBO naar sector en leerweg en voor BOL en BBL naar niveau. Zo wordt elk onderwijsniveau en elke steekproefgroep gewogen door het betreffende populatie-aantal te delen door het steekproefaantal. Tegelijkertijd wordt het steekproefaantal geprojecteerd naar het bijbehorende populatie-aantal.

De weging voor HBO is conform de wegingsprocedure van 2006 uitgevoerd. Dit houdt in een vier dimensionele weging naar opleiding (CROHO-niveau), duaal/

voltijd/ deeltijd, regio en geslacht. Het steekproefaantal van het WO is wederom geprojecteerd naar het populatie-aantal.

Nieuw in 2007 zijn de ongediplomeerden, welke zijn gewogen naar onderwijssoort HAVO, VWO, VMBO, BOL, BBL. De populatiecijfers komen van de steekproefbestanden van de IB-Groep.

7.3 Controle op weegfactoren

Een weegfactor wordt bepaald door het populatie-aantal te delen door het steekproefaantal. Hierbij kunnen twee problemen aan het licht komen:

- het steekproefaantal is groter dan het populatie-aantal;
- het populatie-aantal is 0.

In beide gevallen wordt het populatie-aantal gebruikt van het jaar ervoor. Indien dit populatie-aantal ook niet toereikend of niet bekend is, wordt het totale aantal benaderde afgestudeerden als populatie-aantal gebruikt. Aangezien echter voor de non respondenten het kenmerk geslacht onbekend is, vindt de weging plaats op de kenmerken opleiding, regio en eventueel dual/ voltijd/ deeltijd.

Na het bepalen van de weegfactoren worden deze gecontroleerd op consistentie en waarschijnlijkheid. Wanneer de populatie van het jaar ervoor minstens twee keer zo groot was, wordt uitgezocht wat hier de oorzaak van is. Tot op heden werd dit veroorzaakt door kleine aantallen of nieuwe opleidingen. Daarnaast moet de uiteindelijke weegfactor kleiner zijn dan drie maal de weegfactor van de onderwijssoort als totaal. Is dit wel het geval dan wordt één weegkenmerk buiten beschouwing gelaten en vervolgens de controleregels opnieuw bekeken. Genoemde controle vindt alleen plaats in gevallen waar het steekproefaantal kleiner is dan 50. In een aantal gevallen wordt de weegfactor 1 toegekend, namelijk wanneer de opleidingscode niet specifiek genoeg is of het kenmerk geslacht onbekend is.

Met de keuze van de a-selecte steekproeven is het wegen naar opleiding, regio en geslacht niet meer noodzakelijk. Een weging op het niveau van onderwijssoort, sector, leerweg of niveau volstaat omdat een a-selecte steekproef geen ongelijke verdeling veroorzaakt op regio en geslacht. Doordat het lagere aggregatieniveau hogere steekproef- en populatie-aantallen genereert, zijn hoge weegfactoren, die gecorrigeerd moeten worden, niet te verwachten.