

Sociale Samenhang in Beeld, het SSB nu en straks.





Sociale Samenhang in Beeld, het SSB nu en straks.

Symposium 2005

Inhoud

Voorwoord <i>Gosse van der Veen</i>	7
Doen wat je belooft? <i>Bart F.M. Bakker</i>	9
Over onderzoek naar criminaliteit en de waarde van het Sociaal Statistisch Bestand: een analyse in 8 punten. <i>Frans L. Leeuw</i>	21
Reactie op: 'Over onderzoek naar criminaliteit en de waarde van het Sociaal Statistisch Bestand: een analyse in 8 punten'. <i>Henk G. van de Bunt</i>	35
Het SSB en het zicht op de integratie van minderheden. Of: waarom surveygegevens onontbeerlijk blijven. <i>Jaco Dagevos</i>	41
De Ongekende Multiculturele Samenleving Nader Gekend. Reactie op: 'Het SSB en het zicht op de integratie van minderheden. Of: waarom surveygegevens onontbeerlijk blijven'. <i>Godfried Engbersen</i>	53
Arbeidsmarktdynamiek en vergrijzing. <i>Paul van Leeuwe, Dirk Scheele, Jules Theeuwes en Tijl Woortman</i>	63
Knap maar Karig. Naar een wetenschappelijk en beleidsmatig betekenisvol Sociaal Statistisch Bestand. Reactie op: 'Arbeidsmarktdynamiek en vergrijzing'. <i>Peter Ester en Marcel Kerkhofs</i>	87
De prijs van migratie. Selectieve verhuisstromen van de vier grote steden. <i>Jan Latten, Bas Hamers en Han Nicolaas</i>	95
In balans? Grote steden op weg naar een nieuw evenwicht. Reactie op: 'De prijs van migratie. Selectieve verhuisstromen van de vier grote steden'. <i>Jeroen Slot</i>	117

SSB tabellen in StatLine. <i>Kea Tijdens</i>	129
Reactie op: 'SSB tabellen in StatLine'. <i>Koos Arts</i>	135
Opleidingsniveau uit registers: nieuwe bronnen, maar nog niet compleet. <i>Bart F.M. Bakker, Anne Marthe Bouman en Leo van Toor</i>	141
Reactie op: 'Opleidingsniveau uit registers: nieuwe bronnen, maar nog niet compleet'. <i>Maarten H.J. Wolbers</i>	163
Speuren naar ouders. Het ouder-kind-bestand, onmisbaar voor statistieken over intergenerationeel gedrag. <i>Kees Prins</i>	169
Het ouder-kindbestand van het SSB. Reactie op: 'Speuren naar ouders. Het ouder-kind-bestand, onmisbaar voor statistieken over intergenerationeel gedrag'. <i>Clara H. Mulder</i>	183
Hoe toegankelijk is het Sociaal Statistisch Bestand (SSB)? <i>José Gouweleeuw</i>	189
De DANS om het SSB. Reactie op: 'Hoe toegankelijk is het Sociaal Statistisch Bestand (SSB)?' <i>Peter Doorn</i>	203
Paneldiscussie	212
Curriculum vitae - dagvoorzitter, sprekers, referenten en deelnemers paneldiscussie	218
Afkortingslijst SSB symposium 2005	223
Lijst van publicaties over het Sociaal Statistisch Bestand en de Virtuele volkstelling	226
Deelnemers SSB-symposium, Naturalis, 1 december 2005	239
Dankwoord	241



Voorwoord

Voor u ligt de verzameling van opstellen die zijn gebaseerd op de lezingen en referaten die op 1 december 2005 zijn gehouden tijdens het symposium in Naturalis te Leiden, over het Sociaal Statistisch Bestand (SSB). Niet zonder trots kan ik met dit SSB aantonen dat de nieuwe onderzoeksstrategie die het CBS sinds twee jaar is gestart, een succes genoemd mag worden. De nieuwe aanpak voorziet in een breed gebruik van de data van registraties, in de plaats van vele enquêtes. De beoogde voordelen zijn evident: het verdwijnen van de vele vragenlijsten heeft een forse administratieve lastenverlichting voor bedrijven en instellingen bewerkstelligd. Minstens zo belangrijk is de statistische en wetenschappelijke winst, omdat de veelal integrale bestanden over verschillende, maar aanverwante onderwerpen nu op persoonsniveau kunnen worden gekoppeld. Hierdoor kan het CBS voor het eerst samenhangende uitkomsten op alle terreinen van maatschappelijk belang opstellen en daarmee, zoals de titel van deze publicatie aangeeft, de sociale samenhang in beeld brengen.

Het SSB, dat na een aanloop, vijf jaar geleden formeel van start is gegaan, heeft nu al zijn bestaanrecht aangetoond, terwijl het duidelijk is dat we pas aan het begin staan van wat een uniek, rijk, dynamisch en vrijwel onuitputtelijke, samenhangende personendatabase moet worden.

Het succes van het SSB mag ook blijken uit de vele prominente woordvoerders en deelnemers, die zich overwegend positief over de toepassingen van dit project uitlieten, waarvoor ik hen bij deze gaarne wil bedanken. Natuurlijk zijn er ten opzichte van dit 'relatief jonge' bestand nog vele wensen. Ook die treft u aan in dit werk, en als CBS zullen wij met de vele behoeften en opmerkingen rekening proberen te houden.

Ik wens u veel leesplezier met deze bundel en hoop dat u nog vaak en nuttig gebruik zult maken van het SSB, nu en straks.

Drs. G. van der Veen

Directeur-Generaal van de Statistiek





Bart F.M. Bakker (CBS)

Doen wat je belooft?

In november 2003 is op de conferentie 'De virtuele Volkstelling en het Sociaal Statistisch Bestand' door het CBS een aantal toezeggingen gedaan aan de deelnemers. Aan de hand van een schets van de ontwikkelingen die hebben plaatsgevonden in het SSB wordt beoordeeld of gerealiseerd is wat twee jaar geleden is toegezegd. De conclusie is dat het CBS zijn toezeggingen grotendeels maar niet geheel is nagekomen. Tijdens de conferentie worden de deelnemers uitgenodigd om een oordeel te geven over de kwaliteit die het CBS heeft gerealiseerd en om suggesties te doen over de richting van het SSB in de komende jaren.

1 Inleiding

In november 2003 organiseerde het SISWO samen met het CBS een conferentie over de virtuele Volkstelling en het Sociaal Statistisch Bestand (SSB). Aanleiding was dat het CBS eind 2003 als één van de eerste lidstaten van de Europese Unie de gegevens over de zeventiende Nederlandse Volkstelling aan het Europese statistiekbureau Eurostat verstrekt had. Dat was mogelijk omdat gebruik gemaakt werd van registers die al beschikbaar waren voor andere dan statistische doeleinden en deze te combineren met enquêtes. De basis van deze Volkstelling werd gevormd door het SSB, waarin registers en enquêtegegevens aan elkaar worden gekoppeld en waarin veel meer gegevens voorkomen dan gebruikt worden voor de volkstelling.

In de conferentie stond de wijze centraal waarop het CBS voor 2001 een volkstelling heeft samengesteld, welke mogelijkheden het SSB nog meer te bieden heeft en welke gegevens sociale wetenschappers graag aan het SSB zouden willen ontleen. Het CBS heeft daarbij een aantal toezeggingen gedaan (Bakker en Putman, 2003, blz. 125-126):

- De uitkomsten uit het SSB worden gepubliceerd in twaalf grote StatLine-publicaties. StatLine is het elektronische publicatiemedium van het CBS.
- De inhoud van het SSB was in 2003 beperkt tot de arbeidsmarkt¹. Dit zou uitgebreid moeten worden tot andere domeinen. Uiterst cruciaal zijn huishoudens en onderwijs, maar ook familiale relaties zijn van belang.
- Het werd wenselijk geacht dat uit de virtuele Volkstelling bestanden ter beschikking worden gesteld die internationaal vergelijkbaar zijn.
- De toegang tot de gegevens in het SSB zou geoptimaliseerd moeten worden. De enquêtegegevens die in de virtuele volkstelling zijn gebruikt verrijkt met de registerinformatie zouden ter beschikking gesteld moeten worden via *remote execution*. Aangezien dit echter beperkt is tot arbeidsmarktgegevens zouden ook andere enquêtes dan de Enquête beroepsbevolking daarvoor gebruikt moeten worden (zorg, leren, rusten).
- De deelnemers spraken de wens uit om meer gebruikersdagen te organiseren. Door gebruikers mee te laten praten kan het CBS zijn producten beter afstemmen op de wensen van gebruikers.
- Het CBS zou beleid moeten formuleren voor het gebruik van oude gegevens. Wat kan er bijvoorbeeld beschikbaar worden gesteld uit oudere Volks- of Arbeidskrachtentellingen?

1) Het begrip arbeidsmarkt wordt in deze paper telkens in een brede context gebruikt. Het omvat eveneens de sociale zekerheid (ook als die niet arbeidsgerelateerd is).

- Er zou sprake moeten zijn van een goed versiebeheer van de bestanden die aan derden ter beschikking worden gesteld, om te voorkomen dat er allerlei verschillende versies in omloop raken. Analyses van verschillende sociale wetenschappers worden daarmee controleerbaar, omdat gebruik gemaakt wordt van dezelfde data.
- Er zouden goede metadata moeten worden geleverd bij het SSB, zowel in de publicaties over gegevens uit het SSB als bij de microdata die uit het SSB ter beschikking worden gesteld.

De vraag is in hoeverre het CBS aan de bovenstaande wensen van gebruikers tegemoet is gekomen. Om deze vraag te beantwoorden worden eerst de ontwikkelingen van het SSB sinds 2003 geschetst. In de afsluitende paragraaf wordt expliciet op de bovenstaande punten ingegaan.

2 De ontwikkelingen in het SSB sinds 2003

2.1 De situatie in 2003

Om de ontwikkelingen in het SSB sinds 2003 te kunnen schetsen is het goed om de situatie van 2003 nog even terug te halen. Op dat moment bestond het SSB uit geïntegreerde microdata over de arbeidsmarkt voor de jaargangen 1999-2001 (Bakker en Arts, 2003). Naast standen konden ook de transities op de arbeidsmarkt worden vastgesteld die door personen worden gemaakt. Verder maakten ook de demografische gegevens deel uit van het toenmalige SSB.

In 2003 waren nog maar weinig publicaties uit het SSB samengesteld. Afgezien van een aantal artikelen in een bundel waarin de methode werd beschreven (Al & Bakker, 2000) en een speciaal nummer van de Sociaal-economische maandstatistiek (Arts en Hoogteijling, 2002) waarin de eerste uitkomsten werden gepubliceerd werd er nog weinig gebruik gemaakt van het SSB. Er waren nog geen uitkomsten op StatLine beschikbaar. Wat er was gepubliceerd was wat inhoud betreft beperkt tot de arbeidsmarkt.

2.2 De kern en de satellieten

De omvang van het SSB is sinds 2003 enorm toegenomen. Naast demografische en arbeidsmarktgegevens zijn gegevens opgenomen over onderwijs, criminaliteit, allochtonen, huishoudens, etc. Om hierin enige structuur aan te brengen en de beheerbaarheid van de gegevens te garanderen worden de zogenaamde SSB-kern en de SSB-satellieten onderschei-

den. De SSB-kern bevat de demografische en sociaal-economische kerninformatie. De SSB-satellieten worden ontwikkeld voor de beschrijving van verschillende onderwerpen die buiten de arbeidsmarkt liggen.

2.3 De SSB-kern

De SSB-kern bevat demografische gegevens zoals geslacht, geboortedatum, burgerlijke staat, geboorteland en woonplaats van alle personen die vanaf 1995 enig moment tot de bevolking van Nederland behoorden. Alle wijzigingen hierin, zoals verhuizingen en echtscheidingen, zijn ook in de kern opgenomen. Het huishouden is afgebakend en getypeerd op drie peilmomenten (primo en ultimo van het jaar en ultimo derde kwartaal). Deze demografische gegevens worden ontleend aan de GBA. Verder zijn de belangrijkste sociaal-economische gegevens in de kern opgenomen. Daaronder vallen de arbeidsmarktpositie, het ontvangen van uitkeringen en de lonen, winsten, en uitkerings- en studiefinancieringsbedragen. Deze gegevens worden ontleend aan integrale registraties van de belastingdienst van werknemers, zelfstandigen en uitkeringsgerechtigden, registraties van het Uitvoeringsinstituut Werknemersverzekeringen (UWV) die naast registraties van werknemers ook de werkloosheids- en arbeidsongeschiktheidsuitkeringen verzorgt en registraties van de studiefinanciering van de Informatie Beheergroep. Daarnaast is een (grotendeels elektronische) bedrijfsenquête opgenomen waarin iets meer dan de helft van het aantal werknemers is opgenomen: de Enquête Werkgelegenheid en Lonen.

De SSB-kern bestaat niet uit een bestand, zoals de naam wel doet vermoeden, maar uit een aantal koppelbare bestanden: een waarin de banen van werknemers in Nederland worden beschreven, een waarin alle zelfstandigen zijn opgenomen, een bestand met overige arbeid in Nederland (voornamelijk freelancers en alfahulpen), een bestand met daarin mensen die in Nederland wonen maar in het buitenland werken, zes verschillende uitkeringenbestanden waaronder uitkeringen die uit het buitenland worden genoten en een personenbestand waarin behalve de eerder besproken demografische kenmerken ook de sociaal-economische kenmerken over personen zijn geaggregeerd. Dat betekent bijvoorbeeld dat voor mensen die op een moment twee banen hebben daarvan op basis van het hoogste loon de belangrijkste wordt opgenomen.

Verder wordt voor ieder verslagjaar een bestand gemaakt waaruit de StatLine-publicaties worden samengesteld. In tegenstelling tot de bovenstaande bestanden die informatie over het gehele verslagjaar bevatten, is hiervoor een peildatum gekozen: ultimo derde kwartaal van ieder verslagjaar. Deze peildatum is gekozen, omdat het op dat moment relatief rustig is op de arbeidsmarkt en er ook relatief weinig veranderingen in administraties plaats vinden. Voor de belangrijkste variabelen zijn ook de mutaties ten opzichte van een jaar eer-

der opgenomen. Daarmee worden stromen in beeld gebracht geoperationaliseerd als de transitie tussen twee peilmomenten. Op dit moment zijn gegevens voor de verslagjaren 1999-2003 beschikbaar. In juni 2006 volgt het verslagjaar 2004. De reden dat pas anderhalf jaar na afloop van het verslagjaar een SSB-kern beschikbaar is, is dat de gegevens over zelfstandigen pas 16 maanden na afloop van het verslagjaar worden geleverd. De enige bron die voldoende informatie bevat (zoals de winsten) zijn de aangiftebestanden van de fiscus. De vertraging komt omdat zelfstandigen op vrij grote schaal uitstel voor de aangifte van de inkomstenbelasting aanvragen en krijgen. Bovendien is dan pas tweederde van de zelfstandigen in dat bestand opgenomen, zodat we de overige eenderde moeten bijschatten.

Als onderdeel van de kern wordt inmiddels een opleidingsregister gebouwd met als doel om voor zoveel mogelijk mensen het opleidingsniveau vast te stellen, zodat dit gebruikt kan worden als achtergrondvariabele in allerlei onderzoek. De stand van zaken wordt elders in deze bundel gepresenteerd. Verder worden in 2006 de inkomensgegevens verbeterd door ook andere inkomensbronnen als lonen, winsten en uitkeringsbedragen daarin op te nemen. Daaronder vallen onder meer inkomen uit vermogen en huursubsidie. Bovendien worden de gegevens over huishoudens verbeterd.

3 De satellieten van het Sociaal Statistisch Bestand

De kern van het SSB bevat ook verschillende koppelingsvariabelen, zodat andere informatie voor de statistische beschrijving van specifieke onderwerpen daaraan kan worden gekoppeld. In de eerste plaats kunnen gegevens van personen worden gekoppeld. Zo zijn ook persoonsenquêtes in het SSB opgenomen. De Enquête Beroepsbevolking zorgt bijvoorbeeld voor aanvullende informatie over het beroep dat mensen hebben, zodat onderzoek naar loonverschillen gecorrigeerd voor beroep mogelijk is (Van Cruchten, et al., 2002; Schulte Nordholt, 2003). In de tweede plaats kan informatie over adressen worden gekoppeld. Zo kan bijvoorbeeld de relatie worden gelegd tussen kenmerken van mensen die in een huur- of in een koopwoning wonen. En *last but not least* kunnen gegevens van bedrijven aan de kern van het SSB worden gekoppeld.

Deze koppelmogelijkheden worden benut om voor bepaalde onderwerpen een zogenaamde SSB-satelliet maken. Dit proces verloopt in drie stappen. Eerst wordt alle informatie uit de SSB-kern die nodig is voor de beschrijving van het onderwerp gekopieerd of afgeleid. Dit is onder meer nodig om te voorkomen dat er verschillende resultaten wor-

den gepubliceerd over hetzelfde verschijnsel. In de tweede stap wordt dan de voor het onderwerp unieke informatie toegevoegd. Dan kan het gaan om registers en enquêtes. In de derde stap worden de gewichten van de enquêtes aangepast volgens de methode van consistent en herhaald wegen (Houbiers, 2004). Er zijn inmiddels satellieten voor de beschrijving van veiligheid, de integratie van allochtonen, gezondheid, ruimtelijke en sociale mobiliteit, onderwijs en voorlopers van de SSB-satelliet Sociale dynamiek en arbeidsmarkt.

In de SSB-satelliet Veiligheid zijn gegevens van de processen verbaal opgenomen die opge maakt zijn door de politie. Daarbij gaat het om alle processen verbaal waarin sprake is van een misdrijf. Door deze informatie te combineren met de demografische en sociaal-economische gegevens uit de SSB-kern kunnen daderprofielen worden opgesteld. Een eerste analyse van deze gegevens heeft een rapport opgeleverd over verschillen in verdachte zijn van allochtonen en autochtonen. Omdat de gegevens integraal zijn, worden de allochtonen onderscheiden naar generatie en een groot aantal verschillende herkomstlanden (Blom, *et al.*, 2005). De toepassingsmogelijkheden zijn uiteraard veel groter. Voor het komende jaar zijn afspraken gemaakt om ook slachtoffergegevens op te nemen en de gegevens over de rechtbankstrafzaken van het OM.

In de SSB-satelliet Integratie zijn indicatoren opgenomen waarmee de integratie van verschillende allochtone groepen kan worden gemonitord. Deze informatie is interessant vanwege het feit dat verschillende cohorten nieuwkomers vergeleken kunnen worden in de mate waarin zij sociaal-economisch integreren. Daarover is gepubliceerd in de zogenaamde Integratiekaart (Van Rijn, *et al.*, 2004; Bijl, *et al.*, 2005). Voor 2006 staat verdere uitbreiding op het programma, waarbij het streven is om alle belangrijke informatie over allochtonen bij elkaar te brengen en voor zover nodig te integreren. Daaronder vallen ook de enquêtes die het SCP samen met het CBS uitvoert (Leefsituatie van Allochtonen in de Steden en Survey Integratie Minderheden). In 2006 worden slechts enkele belangrijke stappen in die richting gezet. Vooralsnog is de satelliet beperkt tot de indicatoren die behoren tot de Integratiekaart.

De SSB-satelliet Gezondheid, ook wel Gezondheid Statistisch Bestand genoemd, is los van de rest van het SSB ontwikkeld. Er zijn gegevens in opgenomen over ziekenhuisopnames en doodsoorzaken. De komende periode worden die gegevens verder uitgebreid. Inmiddels zijn afspraken gemaakt om de SSB-satelliet gezondheid op dezelfde manier te produceren als de overige satellieten. Dat zal in de loop van 2006 en 2007 worden uitgevoerd.

De SSB-satelliet Ruimtelijke en sociale mobiliteit bevat informatie over verhuizingen, en demografische en sociaal-economische transitie. Van de woning zelf is alleen of het een huur- of een koopwoning betreft en de WOZ-waarde opgenomen. Er is onderzoek mee verricht naar de verandering van de samenstelling van de grote steden naar demografische en sociaal-economische kenmerken (Wittebrood, Latten & Nicolaas, 2005; Latten, Nicolaas & Wittebrood, 2005). Dit kan echter ook worden verrijkt door naar wijken en buurten te kijken. Verder worden in 2006 deze gegevens benut voor onderzoek naar de realisatie van verhuisplannen. Deze verhuisplannen zijn gevraagd in het Woningbehoefteonderzoek (WBO) 1998 en 2002. Door de WBO's in de satelliet op te nemen, kan de realisatie van die verhuisplannen worden vastgesteld.

De SSB-satelliet Sociale dynamiek en arbeidsmarkt bestaat nu nog uit een aantal voorlopers. De eerste voorloper is het zogenaamde immigrantenpanel. Daarbij worden de immigranten van een bepaald jaar getypeerd naar het migratiemotief. Zij worden bovendien gevolgd voor wat betreft de positie op de arbeidsmarkt (Sprangers *et al.*, 2005). Daarnaast is echter informatie opgenomen over de krimp en groei van bedrijven waarmee onderzoek gedaan kan worden naar het verband met de individuele sociale mobiliteit. Zo kan worden nagegaan wat de individuele gevolgen zijn van massaontslag. Hoe snel vinden mensen weer werk? De eerste vrucht van dit soort analyses vinden we in deze bundel (Van Leeuwe *et al.*, 2006).

In 2006 wordt ook de SSB-satelliet Onderwijs ontwikkeld. Daarin wordt met behulp van de onderwijsnummerbestanden de schoolloopbanen beschreven en deze worden in verband gebracht met kenmerken van het milieu van herkomst en de arbeidsmarktpositie van schoolverlaters.

De eerste satelliet was misschien wel de virtuele volkstelling (CBS, 2003; Schulte Nordholt, Hartgers en Gircour, 2004), maar gezien het Europese karakter hiervan wordt deze informatie niet jaarlijks samengesteld, wat wel het geval is met de overige satellieten. De hier genoemde satellieten worden allemaal nog verder ontwikkeld. Daarbij wordt steeds beoordeeld of met de inhoud van de satellieten de belangrijkste vragen van statistiekgebruikers kunnen worden beantwoord. Om deze gebruikersbehoefte goed in kaart te brengen wordt ook dikwijls samengewerkt met andere instellingen die op het betreffende terrein over veel expertise beschikken. Tevens wordt regelmatig overlegd met belangrijke gebruikers zoals ministeries en planbureaus.

4 Doen wat je belooft: de gebruikers aan het woord

De toezeggingen die het CBS in 2003 heeft gedaan en de realisatie daarvan zijn in het onderstaande schema samengevat.

Schema 1. Toezeggingen 2003 en realisaties 2003-2006

<i>Toezeggingen</i>	<i>Gerealiseerd</i>
Inhoudelijke verbreding	Huishoudens en familierelaties al opgenomen, SSB-satellieten, opleidingsniveau in 2006
Publiceren 12 grote StatLine-tabellen	Uit SSB 13 grote tabellen gepubliceerd en bijdragen aan andere StatLine-publicaties
Bestanden uit virtuele VT ter beschikking stellen	Via IPUMS-project gerealiseerd
Ontwikkelen remote access	Pilot wordt uitgevoerd en bij slagen daarvan naar verwachting in 2006 verder beschikbaar gesteld
Aandacht voor beveiliging oudere gegevens	Beveiliging kan niet minder streng, maar wel VT-1960 en VT-1971 beschikbaar gesteld
Meer gebruikersdagen	1 december 2005 was de volgende

De inhoud van het SSB was in 2003 beperkt tot de arbeidsmarkt. Dit zou uitgebreid moeten worden tot andere domeinen. Uiterst cruciaal zijn huishoudens en onderwijs, maar ook familiale relaties zijn van belang. De afbakening en typering van huishoudens is inmiddels in de SSB-kern opgenomen en wordt in 2006 verder verbeterd. Het opleidingsniveau wordt in 2006 weliswaar niet integraal maar wel voor een grote steekproef in de SSB-kern opgenomen. In de bijdrage van Bakker, Bouman en Van Toor wordt daar dieper op ingegaan. Familiale relaties worden ook in de SSB-kern opgenomen: de relatie tussen ouders en kinderen wordt vastgesteld voor zover de informatie in de voor ons beschikbare registers dat toelaat. Kees Prins gaat in zijn bijdrage in op het ouder-kind-bestand.

In 2003 hebben we aangegeven dat de uitkomsten uit het SSB gepubliceerd zouden worden in twaalf grote StatLine-publicaties. StatLine is het elektronische publicatiemedium van het CBS. Inmiddels zijn 13 StatLine-tabellen gepubliceerd. Kea Tijdens geeft in haar bijdrage een oordeel over de kwaliteit van deze informatie.

Naast de uitbreiding van de kern is het SSB ook flink uitgebreid met een aantal satellieten die ieder een specifiek thema beschrijven: veiligheid, allochtonen, ruimtelijke en sociale mobiliteit, onderwijs, sociale dynamiek, etc. Het oordeel over de satelliet Veiligheid wordt geveld in de bijdrage van Frans Leeuw, de directeur van het WODC, de instelling

waarmee het CBS heeft samengewerkt om de satelliet te ontwikkelen. De satelliet Integratie wordt onder de loep genomen door Jaco Dagevos van het SCP, een deskundige op dit terrein. Jan Latten van het CBS zal enkele uitkomsten presenteren uit de satelliet Ruimtelijke en sociale mobiliteit. Of het CBS daarmee de goede kant op gaat is beoordeeld door Jeroen Slot, hoofd onderzoek van de Dienst Onderzoek en Statistiek van de gemeente Amsterdam.

Het werd wenselijk geacht dat uit de virtuele Volkstelling bestanden ter beschikking worden gesteld die internationaal vergelijkbaar zijn. Inmiddels heeft het CBS aan de Universiteit van Minnesota steekproeven geleverd uit de Volkstellingen van 1960 en 1971 en de Virtuele Volkstelling van 2001. Deze bestanden zullen ook via DANS worden aangeboden aan Nederlandse onderzoekers. José Gouweleeuw gaat daar in haar bijdrage op in.

De toegang tot de gegevens in het SSB zouden geoptimaliseerd moeten worden. De enquêtegegevens die in de virtuele volkstelling zijn gebruikt verrijkt met de registerinformatie zouden ter beschikking gesteld moeten worden via *remote execution*. Aangezien dit echter beperkt is tot arbeidsmarktgegevens zouden ook andere enquêtes dan de Enquête beroepsbevolking daarvoor gebruikt moeten worden (zorg, leren, rusten). Op dit moment wordt een pilot uitgevoerd om microdata via *remote access* ter beschikking te stellen, een meer gebruikersvriendelijke manier van onsluiten van microdata dan *remote execution*. *Remote access* houdt in dat een speciale PC op de werkplek van de onderzoeker wordt geplaatst waarmee de microdata kunnen worden geanalyseerd. De beveiliging is onder meer geregeld via het afsluiten van een contract en het verifiëren van de vingerafdruk van de personen die gebruik maken van deze service. Ook hierop gaat José Gouweleeuw in.

De deelnemers spraken de wens uit om meer gebruikersdagen te organiseren. Door gebruikers mee te laten praten kan het CBS zijn producten beter afstemmen op de wensen van gebruikers. Sinds 2003 zijn enkele gebruikersdagen georganiseerd voor specifieke onderwerpen. Zo zijn in 2004 een conferentie over de uitkomsten van de virtuele volkstelling en een conferentie over sociale dynamiek georganiseerd. Het SSB stond in die conferenties niet centraal, maar speelde wel een belangrijke rol op de achtergrond. De bijeenkomst op 1 december was weer een echte gebruikersdag van het SSB.

Het CBS zou beleid moeten formuleren voor het gebruik van oude gegevens. Wat kan er bijvoorbeeld beschikbaar worden gesteld uit oudere Volks- of Arbeidskrachtentellingen? Dit is besproken binnen het CBS en er is besloten om de beveiliging van oudere gegevens niet minder streng te maken. Wel zijn steekproeven uit oudere Volkstellingen van 1960 en 1971 ter beschikking gesteld.

Er zou sprake moeten zijn van een goed versiebeheer van de bestanden die aan derden ter beschikking worden gesteld, om te voorkomen dat er allerlei verschillende versies in omloop raken. Analyses van verschillende sociale wetenschappers worden daarmee controleerbaar, omdat gebruik gemaakt wordt van dezelfde data. Dit is een uiterst relevant punt dat nog verder moet worden uitgewerkt voor het SSB. Gezien de snelle ontwikkeling van het SSB is versiebeheer cruciaal. Omdat tot op heden de analyse van bestanden voornamelijk in samenwerkingsverbanden met externe partijen is gebeurd, is het versiebeheer vrij gemakkelijk te realiseren. Dit wordt echter anders wanneer de gegevens ook via *remote access* ter beschikking worden gesteld en er op grotere schaal op wordt geanalyseerd.

Er zouden goede metadata moeten worden geleverd bij het SSB, zowel in de publicaties over gegevens uit het SSB als bij de microdata die uit het SSB ter beschikking worden gesteld. Ook hiervoor moeten nog de nodige stappen worden gezet. De metadata beperken zich tot op heden tot codeboeken en rechte tellingen. Deze documentatie wordt echter uitgebreid op het moment dat de gegevens ter beschikking worden gesteld via *remote access*.

Het geheel overziend heeft het CBS voor een deel gedaan wat het heeft beloofd. Er wordt inmiddels ruimschoots gepubliceerd over uitkomsten uit het SSB, er zijn enkele cruciale variabelen zoals type huishouden toegevoegd en de relaties tussen ouders en kinderen kunnen worden gelegd, en er heeft inhoudelijke verbreding plaats gevonden door de ontwikkeling van SSB-satellieten. Ook zijn er gebruikersdagen georganiseerd. Versiebeheer en metadata moeten de komende periode nog meer aandacht krijgen.

Referenties

- Al, P. & B.F.M. Bakker (eds.), 2000, Re-engineering Social Statistics by micro-integration of different sources. The-manummer Netherlands Official Statistics, jrg. 15, nr. summer.
- Arts, C.H. & E.M.J. Hoogteijling, 2002, Het Sociaal Statistisch Bestand 1998 en 1999, In: Sociaal-economische maandstatistiek, december 2002, blz. 13-21.
- Bakker, B.F.M., & C.H. Arts, 2003, Dynamiek op de arbeidsmarkt; gegevens over stromen uit het Sociaal Statistisch Bestand. In: B.F.M. Bakker & L. Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS) blz. 59-70.
- Bakker, B.F.M. & L. Putman (red.), 2003, De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS).

- Bijl R.V., A. Zorlu, A.S. van Rijn, R.P.W. Jennissen, M. Blom, 2005, Integratiekaart 2005; de maatschappelijke integratie van migranten in de tijd gevolgd: trend- en cohortanalyses. WODC-Cahier 2005-16 (Den Haag/Voorburg: WODC/CBS).
- Blom, M., J. Oudhof, R.V. Bijl & B.F.M. Bakker (red.), 2005, Verdacht van criminaliteit. Allochtonen en autochtonen nader bekeken. WODC-cahier 2005-2 (Den Haag/Voorburg: WODC/CBS).
- CBS, 2003, <http://www.cbs.nl/nl-nl/menu/themas/dossiers/volkstellingen/cijfers/incidenteel/maatwerk/2003-volkstelling-excel.htm>.
- Cruchten, J.M.J. van, M.I. Hartgers en E. Schulte Nordholt, 2002, Loon- naar opleidings- en beroepsniveau: het loonstrucuuronderzoek 1998, In: Sociaal-economische maandstatistiek, jrg. 19, december 2002, blz. 72-81.
- Houbiers, M. 2004, Towards a Social Statistical Database and Unified Estimates at Statistics Netherlands, In: Journal of Official Statistics, Vol. 20, nr. 1, blz. 55-75.
- Latten, J., H. Nicolaas & K. Wittebrood, 2006, Concentratie allochtonen toegenomen, In: Bevolkingstrends, jrg. 53, nr. 3, blz. 90-95.
- Leeuwe, P. van, Theeuwes, J., D. Scheele, J. Theeuwes & T. Woortman, 2006, Arbeidsmarktdynamiek en vergrijzing (deze publicatie).
- Rijn, A.S. van, A. Zorlu, R.V. Bijl & B.F.M. Bakker, 2004, De ontwikkeling van een Integratiekaart WODC-Cahier 2004-9 (Den Haag/Voorburg: WODC/CBS).
- Schulte Nordholt, E., M. Hartgers & R. Gircour (red.), 2004, The Dutch Virtual Census of 2001, Analysis and Methodology (Voorburg/Heerlen: Statistics Netherlands).
- Sprangers, A.H., H. Nicolaas, A. Zorlu en J. Hartog, 2005, Vluchtelingen en gezinsmigranten vinden moeilijker werk, In: Sociaal-economische trends, jrg. 2005, nr. 4, blz. 29-37.
- Wittebrood, K., J. Latten & H. Nicolaas, 2005, Wonen, leefbaarheid en veiligheid, In: SCP, WODC, CBS, Jaarrapport Integratie 2005, blz. 132-147.





Frans L. Leeuw* (WODC & UM)

Over onderzoek naar criminaliteit en de waarde van het Sociaal Statistisch Bestand: een analyse in 8 punten.

Deze bijdrage gaat in op de betekenis van SSB- en andere gegevens voor onderzoek naar criminaliteit. De betekenis van het SSB en zijn satellieten wordt ten eerste geplaatst tegen de achtergrond van verschillende typen dataverzameling binnen de criminologie.

Vervolgens wordt een aantal voordelen van het werken met het SSB uiteengezet. Pluspunten betreffen onder andere de betere kartering van waar voor de criminologie relevant gedrag zich maatschappelijk voordoet, die met 'all sources analysis' mogelijk is, de longitudinaliteit van de gegevens en het belang ervan voor evaluatieonderzoek. Ook wordt gewezen op het belang dat het SSB vooral over gegevens over gedrag(suitkomsten) gaat en niet slechts attitudes, percepties, ervaringen en belevenissen weergeeft.

Echter, ook een aantal beperkingen, resp. uitdagingen worden genoemd. Zij betreffen o.a. de dekking door het SSB van lieden die zich niet willen 'laten' 'registreren', de relatie tussen theorie-gericht criminologisch onderzoek en registerdata en de noodzaak andere typen data (zoals: waarnemingen-ter-plekke en sociaal kapitaal gegevens) te linken met het SSB. Afgerond wordt met een aantal wensen.

* Dank aan Martine Blom, André van der Laan, Rob Bijl, Stefan Bogaerts, Frank Willemsen (WODC) en de anonieme CBS-reviewer voor eerder commentaar en suggesties.

Punt een

De tijd dat in de criminologie nauwelijks (grootschalig) empirisch onderzoek gebeurde is –gelukkig— al langere tijd voorbij. Mede door de uitbouw van universitair criminologisch onderzoek, de *Onderzoekschool Maatschappelijke Veiligheid* en door instituten als het WODC en het NSCR beschikt de Nederlandse samenleving over behoorlijk wat empirisch onderzoek op dit terrein. Weliswaar niet als criminologisch instituut, maar ook het CBS is actief op dit terrein. In het –recente—verleden bijvoorbeeld door de afname van de POLS (‘permanente leefsituatie survey’) maar ook omdat het ontwikkelen van de ‘SSB-Veilighedsatelliet’ de komende jaren prioriteit heeft en omdat het CBS de ‘Veiligheidsmonitor’ thans uitvoert.

Ik schets in het navolgende eerst kort waar we, naar mijn oordeel en als we de *aard en wijze van gegevensverzameling* centraal stellen, thans zo’n beetje staan in het onderzoek naar criminaliteit. Ik doe daarbij géén poging enige historiciteit te realiseren. Ook laat ik gemakshalve het *kwalitatieve onderzoek* buiten beschouwing; op deze conferentie waar statistieken centraal staan, is dat niet ongepast.

Bovendien zijn de voorbeelden die ik geef uit het *hier en nu*. Dus onmiddellijk zullen sommigen onder U roepen dat ik dingen vergeet. Correct, maar ingecalculeerd.

Daarna ga ik in op wat de rol en betekenis van het SSB kan zijn voor de criminologie en haar onderzoeksinstituten. Tenslotte schets ik een aantal wensen met betrekking tot het SSB voor de toekomst.

Punt twee

Dat er veel meer empirisch onderzoek plaatsvindt naar determinanten en gevolgen van criminaliteit mag waar zijn, tegelijkertijd moet geconstateerd worden dat omvangrijk, *op populatieniveau gericht kwantitatief empirisch onderzoek* nog niet erg vaak in de sociale wetenschappen /criminologie te vinden is.

De *Recidivemonitor* van het WODC is een voorbeeld van een langlopend onderzoek van alle (rechts-)personen die met de Nederlandse justitie in aanraking zijn gekomen, inclusief het overzicht van de strafzaken waarin zij als verdachte centraal stonden (Wartna et al, 2005). Maar dit onderzoek richt zich slechts op een deel van de (brede) strafrechtsketen:

veroordeling, executie van de straf en recidivisme en het richt zich bovendien op slechts een beperkt aantal (onafhankelijke) variabelen. Wel krijgen we op deze manier beter zicht op de uitstroomresultaten van strafrechtelijke interventies en op het verloop van criminele carrières, zowel bij jeugdige als bij volwassen daders. De metingen in het kader van de Recidivemonitor verlopen steeds op dezelfde manier. De uitkomsten van het onderzoek zijn hierdoor onderling vergelijkbaar. Sommige metingen worden op gezette tijden herhaald, zodat van de belangrijkste dadergroepen ook het verdere verloop van hun strafrechtelijke carrière op de voet kan worden gevolgd.

Een ander voorbeeld is de *Integratiekaart* van het WODC, die samen met het CBS ontwikkeld is. Daarin wordt, op grond van combinaties van twee SSB-satellieten (de SSB-satelliet Allochtonen en de SSB-satelliet Veiligheid met daarin o.m. HKS-gegevens)—onder andere—naar de relatie tussen bepaalde achtergrondvariabelen en criminaliteit gekeken. Maar overigens vooral naar een tiental andere theoretisch, met integratie samenhangende variabelen (onderwijsniveau, interetnisch huwen, arbeidsmarktparticipatie enz.). Ook hier worden op populatieniveau trends in kaart gebracht en worden, voor cohorten vanaf 1999, ontwikkelingen van (herkomst)groeperingen (aanzienlijk meer dan alleen de vier klassieke) gevolgd (vergelijk Bijl et al., 2005).

Punt drie

Veel bekender dan populatieonderzoeken zijn *de steekproefonderzoeken*. Zij hebben betrekking op tal van vormen van criminaliteit. Recente Nederlandse voorbeelden zijn de volgende:

- De (jaarlijkse) steekproef *Monitor Criminaliteit Bedrijfsleven (MCB)*. De MCB is in 2002 voor het eerst in zijn volle breedte uitgevoerd en is in 2004 in iets gewijzigde opzet herhaald. Doel van deze Monitor is inzicht te geven in de aard en omvang van de criminaliteit tegen bedrijven in Nederland en in de preventieve maatregelen die bedrijven nemen tegen criminaliteit, en in de ontwikkeling daarvan door de tijd.
- De *Veiligheidsmonitor* (CBS, WODC, Justitie & BZK) die de vervanger van de *Politie-monitor Bevolking, POLS en enquête Leefbaarheid en veiligheid* wordt.

- *Surveys van (zelfgerapporteerd) gedrag van jongeren* is een andere vorm van grootschalig empirisch onderzoek in de Nederlandse criminologie (vergelijk bijvoorbeeld van der Laan, 2005). Dit soort studies heeft een vrij lange traditie, net zoals dat het geval is met onderzoek onder *steekproeven van slachtoffers*. Ook het SCP is actief op dit terrein als ook het nieuwe, aan de Universiteit van Tilburg verbonden Intervict.
- Studies waarin *netwerken* van criminaliteit in de woonomgeving of op scholen (Weerman et al, 2005)² onderzocht worden, kunnen hier aan worden toegevoegd.

Vaak overheerst de vragenlijst in deze onderzoeken als waarnemingsinstrument. Maar er zijn ook uitzonderingen zoals onderzoek van Bernasco (2004)³ naar verplaatsingscriminaliteit waarbij gegevens, waaronder foto's, gemaakt en verzameld worden over de mate van 'vervuiling' en 'verloedering' van wijken en waarbij gegevens over wijkkwaliteit in relatie worden gebracht met crimineel gedrag. Ook in buitenlands onderzoek waarin de *broken-windows theorie* centraal staat, gebeurt dat (vgl. Sampson & Raudenbush, 1999)⁴. En met soms onverwachte gevolgen⁵.

Punt vier

Eigenlijk is met dit laatste voorbeeld de brug geslagen naar het derde 'pakket' aan *empirisch criminologisch onderzoeken, namelijk die waarin kwantitatieve en kwalitatieve vormen van dataverzameling en – analyse gecombineerd worden*. Dat soort onderzoeken is eveneens in vrij ruime mate aanwezig. Recente Nederlandse voorbeelden zijn onderzoeken naar:

- schattingen van illegalen in Nederland (Engbersen et al, 2002);,
- de ervaringen met het pilot project rondom twee man op één cel (Moors et al., 2004),
- aard en omvang van corruptie in bestuurlijk Nederland (Huberts en Nelen, 2005),
- een samen met de VU/EUR door het WODC in uitvoering zijnd onderzoek naar 'criminele carrières in de georganiseerde misdaad', waarbij kwalitatieve informatie over ongeveer 1000 'zware jongens' gecodeerd wordt, teneinde daarop kwantitatieve analyses mogelijk te maken,
- het door Rovers en de Vries Robbé (2005) uitgevoerde onderzoek naar interne criminaliteit (in de logistieke sector). Naast zo'n 140 interviews met bedrijven, is ook in kaart gebracht hoe de afhandeling van aangiften door politie en justitie plaatsvindt,

3) Bernasco, W., Lezing gegeven ter gelegenheid van een NSCR/WODC seminar over het belang van netwerkonderzoek voor de criminologie. (19 april 2004). Vgl. ook: Bernasco, W. & Luykx, F. (2002). De ruimtelijke spreiding van woninginbraak - een

analyse van Haagse buurten. Tijdschrift voor Criminologie, 44, 3, 231-246.

4) Sampson, Robert J. and Steve Raudenbush. 1999. Systematic Social Observation of Public Spaces: A

New Look at Disorder in Urban Neighborhoods . American Journal of Sociology 105: 603-651.

5) In dit onderzoek waarbij duizenden videobeelden gemaakt werden en in relatie gebracht werden met

- een laatste voorbeeld is het onderzoek “determinanten van subjectieve veiligheid” dat door het SCP is uitgevoerd. Daar wordt op basis van een op postcode gerealiseerde koppeling van bestaande bestanden (Politiemonitor, Woning Behoeftte Onderzoek, POLS, e.d.) met geografische informatie het vraagstuk van de subjectieve veiligheid geanalyseerd.

Punt vijf

Voor de geïnteresseerde lezer van wat in Nederland aan onderzoek wordt verricht, verwijs ik naar het *elk jaar* door het WODC gemaakte overzicht *wie wat in Nederland* aan criminologisch onderzoek ‘doet’. Het betreft het ‘special issue’ van *Justitiële Verkenningen*.

Punt zes

Ik heb *deze wereld* (in zeer globale termen) geschetst, omdat het een van de *contexten* is waarbinnen de relevantie van het sinds enige tijd door het CBS in ontwikkeling genomen SSB gezien kan worden.

Kom ik thans tot de vraag *wat de pluspunten van het SSB zijn als men onderzoek wil doen naar criminaliteit in Nederland?* Het door het WODC en het CBS gepubliceerde rapport *Verdacht van criminaliteit* (Blom et al., 2005) is een (eerste) voorbeeld van wat met de koppeling van het SSB en het HKS tot stand te brengen is. Ik verwijs naar dit rapport voor meer informatie (Blom et al., 2005).

criminaliteitsgegevens bleek dat 'there is a much smaller connection than commonly believed between a neighborhood's appearance and its crime rate. ...factors other than neighborhood orderliness hold the key to understanding predatory crime.

One such factor is "collective efficacy," the capacity of neighbors to work together to strengthen their community. Another major factor is concentrated poverty'. Uit: William Harms, Study conducted in Chicago neighborhoods calls 'broken-windows'

theory into question, in: The University of Chicago Chronicle, Jan. 6, 2000, Vol. 19 No. 7.

Pluspunt 1

De ervaringen opgedaan in het WODC-CBS onderzoek van Blom et al (2005) maken duidelijk dat koppeling van SSB gegevens aan de HKS-dataset in kaart kunnen brengen waar, dwz bij welke groeperingen met welke kenmerken, het verdacht zijn van criminaliteit zich voordoet. Het koppelen van HKS gegevens aan, het zij toegegeven, een nog maar beperkt aantal relevante criminogene achtergrondvariabelen, zoals:

- *schooluitval* (via ingewikkelde koppeling van meerdere bestanden),
- *sociaal-economische en sociaal-demografische achtergrond* en
- *herkomst (1^{ste} én 2^{de} generatieopsplitsing)*

geeft ook een wat *meer theoretisch* inzicht.

Op deze manier wordt niet alleen beschikt over enkele persoonskenmerken en gegevens van sociaal-demografische en economische aard, maar is ook kennis beschikbaar over kenmerken van de sociale omgeving, waar verdachten uit afkomstig zijn.

Dat het in het SSB om *geregistreerd gedrag* gaat en niet om *antwoorden op vragen over attitudes, opvattingen, percepties of gedrag(sintenties)*, resp. over in de ogen van de respondenten gepleegd gedrag, is een *fort* van deze werkwijze. De in de registers voorkomende gedragsgegevens kunnen immers niet beïnvloed zijn door eventuele sociaal wenselijke antwoordtendenties op vragen. Er wordt de ‘respondent’ namelijk niets gevraagd. Antwoordtendenties zoals het over- of onderdrijven, liegen of ‘zo maar iets roepen’ kunnen de kwaliteit van de gegevens derhalve niet bedreigen. Levitt en Hubner (2005) hebben fraaie voorbeelden gegeven van het verschil tussen het meten van gerapporteerd gedrag via vragenlijsten en het gebruiken (en analyseren) van (register)bestanden waarin gedrag(suitkomsten) zijn opgenomen. De subtitel van hun boek heet niet voor niets ‘the dark side of everything’. Het is overigens verstandig nog eens in herinnering te roepen wat in de jaren 70 al door Bem (1972)⁶ werd beargumenteerd: ‘individuals come to know their own attitudes, emotions and other internal states partially by inferring them from their own overt behaviour’⁷.

Pluspunt 2

Omdat het SSB inmiddels, dwz na de studie van Blom et al. (2005) *op jaarbasis* gekoppeld is aan HKS én het SSB verder uitgebreid wordt, zijn in de nabije toekomst ook *longitudinale analyses, waaronder cohort studies* mogelijk. Immers, doordat er op micro/individueel niveau gekoppeld wordt, kunnen *cohorten* gevolgd worden. Dat is cruciaal voor vele beleidsvragen; criminele carrières in beeld brengen, integratiecarrières in beeld brengen; combinaties van carrières: bijvoorbeeld hoe doen mensen het qua arbeidsmarktpositie en zien we linken met hun criminele gedrag? Als meer Justitiële bestanden gekoppeld worden, behoren duuranalyses ook tot de mogelijkheden. Daarmee kunnen vra-

6) Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*, (Vol. 6, pp. 1-62). New York: Academic Press.

7) Op de bijna overloze aantallen attitude-gedrag-sintentie-onderzoeken in de traditie van het model van het beredeneerd gedrag ga ik niet in. Zie voor een fundamentele analyse daarvan de nog steeds lezenswaardige dissertatie van Frits Tazelaar (1980)

Mentale incongruenties - sociale restricties - gedrag. Een onderzoek naar beroeps participatie van gehuwde vrouwelijke academici. Utrecht University.

gen beantwoord worden over hoe lang iemand in kringen van Justitie verblijft en of dat verschilt tussen verschillende bevolkingsgroepen/tijdstippen/regio's enz.

Pluspunt 3

Omdat vrijwel alle registers in het SSB integraal zijn, kan meer en betrouwbaarder informatie worden samengesteld over kleine regionale eenheden en kleine groepen in de samenleving.

Pluspunt 4

Wie voor de criminologie het belang van *all-sources analysis* onderschrijft, heeft veel aan onderzoek waarin SSB-gegevens met gegevens uit justitiële en politionele statistieken verbonden worden.

Wie de nadruk legt op *verklarend onderzoek*, kan met behulp van het SSB *preciezere verklaringsvragen* stellen *dan voorheen*, bijv. hoe het komt dat bepaalde bevolkingsgroepen een aanzienlijke hogere kans hebben verdacht van criminaliteit te zijn dan andere. Doordat gecontroleerd kan worden voor voor de hand liggende versturende, intermediaire variabelen zoals leeftijd(sverdeling), sekse enz., kunnen deze vragen tegen *een preciezere achtergrondkennis* worden gesteld dan wanneer men niet over SSB-/SSB-satellieten beschikt.

Wie *verklaringsvragen wil beantwoorden*, zal nog *niet veel* aan het SSB hebben (zie ook een van de beperkingen hieronder)

Pluspunt 5

Wie evaluatieonderzoek doet naar de effecten van justitiële interventies, kan ook zijn voordeel met het SSB doen. Dit kan langs twee manieren. *De eerste manier* is om discussies over hoe de 'werkelijkheid' er *voor* de interventie uit zag, te beslechten. Niet onbekend is het verschijnsel dat wanneer een evaluatie een nameting heeft maar geen voormeting, discussies ontstaan over waar de nametingresultaten mee vergeleken moeten worden. Alhoewel dat soort discussies voorkomen moet worden door een expliciete voor- en nameting te hebben (en bij voorkeur ook aan de andere eisen van het (quasi-)experimenteel design te voldoen), kan het aanroepen van *contextspecifieke SSB-gegevens* wel (enigszins) helpen. De werkelijkheid van voor de interventie kan daardoor namelijk 'objectiever' in kaart gebracht worden.

De tweede manier is door voorafgaand aan de beleidsinterventie preciezer te segmenteren *wie* met het beleid geconfronteerd wordt en *wie niet*. Of door beter zicht te krijgen op de groepen/mensen waar verdiepend onderzoek naar gedaan moet worden. Door op grond van uitsplitsing binnen het SSB krachtigere inzichten te hebben over contextspecifieke kenmerken van groepen waar –als voorbeeld– (jeugd)criminaliteit zich voordoet, kan in een verdiepend (survey)onderzoek met meer precisie daar de aandacht op gericht worden⁸.

8) Het nieuwe Woning Behoeft Onderzoek (WOON) van VROM werkt op deze manier. Daar worden gegevens over inkomen maar ook over energetische waarde van de woning, kadastragegevens en tal van andere achteraf verzamelde data

gekoppeld. Hiermee wordt de interviewtijd aanzienlijk terug gebracht en neemt de betrouwbaarheid van de data toe.

Punt zeven

Ik kom nu toe aan de vraag *wat de beperkingen van, of anders geformuleerd: de uitdagingen voor het SSB, resp. haar satellieten zijn, als men onderzoek wil doen naar criminaliteit in Nederland?*

Beperking/uitdaging 1

De SSB-satelliet *criminaliteit* omvat momenteel alleen nog de HKS van de KLPD: het gaat om verdachten, en niet om veroordeelden. Wezenlijker: het gaat om *geregistreeerde criminaliteit*, dus om degenen die tegen de lamp gelopen zijn. Andere beperkingen zijn dat HALT afdoeningen er niet in zitten. Het gaat om ca. 20.000 afdoeningen op jaarbasis, waarvan ruim 11.400 betrekking hebben op een misdrijf; omdat in het huidige HKS-SSB gegevens beschikbaar zijn over ca. 180.000 verdachten van misdrijven, gaat het om een niet geringe gemiste kans. Gegevens van bijzondere opsporingsdiensten (FIOD en inspecties) zitten er evenmin in. Organisatiecriminaliteit is dus grotendeels ‘missing’. Ook dat is jammer.

Beperking/uitdaging 2

Illegalen en niet-geregistreeerden zitten er –vermoedelijk—niet in, dit vanwege de koppeling met het GBA. Een schatting van Hoogteijling (2002) is dat er in 2001 tussen de 46.500 en 115.600 illegalen in Nederland woonden. In hoeverre zwerfjongeren in het GBA staan geregistreerd is ook onduidelijk (Algemene Rekenkamer, 2004). Overigens wordt ook wel gewag gemaakt van hogere absentiecijfers. Daar komt bij dat de dekking van het GBA ook door andere redenen niet 100% is. Juist (echte) ‘slechterikken’ zullen hun best doen *buiten de GBA* registratie te blijven. Dit stelt voor criminologisch onderzoek een *bijzonder probleem*. Wat overigens wel helpt is om met behulp van specialistische schattingstechnieken én het combineren van gegevensbestanden zoals HKS aan GBA, toch zicht te krijgen is op dit type vraagstelling (zie ook Sikkel et al., 2006).

Beperking/uitdaging 3

Aan het gebruik van registergegevens kleven voor- en nadelen. Weliswaar zijn registers in het algemeen integraal, treden er daarom geen steekproeffouten op en is er geen probleem met selectieve non-respons, maar de kwaliteit van de administratieve informatie is niet altijd voldoende. Daarnaast zijn lang niet alle theoretisch gewenste variabelen op basis van registerinformatie te operationaliseren. Overigens worden juist door het (uit)bouwen van het SSB aangeleverde registers grondig geschoond.

Beperking/uitdaging 4

Het is niet of zeer beperkt mogelijk *theory-driven* te werk te gaan. Gegevens over criminogene factoren zijn maar zeer beperkt beschikbaar in de registers. De variabelen in het SSB zijn *proxies* voor variabelen die indirect gerelateerd zijn aan criminaliteit. Informatie over directe risico- en protectieve factoren zoals (in)adequate opvoedingsstijlen, (delinquente) vrienden of (crimineel) schoolklimaat is niet voorhanden.

Anders gezegd: de in SSB opgenomen registraties zijn – noodzakelijkerwijs – vaak *tame-lijk plat*. De registraties waren ook *niet* voor sociaal-wetenschappelijk onderzoek gericht op verklaringen, ingericht. Daarom is het een belangrijke uitdaging meer werk te maken van het *koppelen van survey-gegevens aan SSB-registers* (bijvoorbeeld onderzoek naar slachtoffers, naar (zelf-gerapporteerde) jeugddelinquentie enz.). Gelukkig is het CBS daarmee wel reeds gestart.

Beperking/uitdaging 5

Once you're in, you're never out, was en is de kritiek op het werken met gecombineerde (grote) databestanden. *Big brother is watching you* is een topic dat een vergelijkbaar gevoel geeft. Hoe te voorkomen dat de privacy in het geding is en dat mensen met een aanvangs'fout' in een bestand komen en er vervolgens nooit meer van 'af' komen, is een vijfde minpunt/uitdaging, zij het van een andere aard.

Punt acht

Wat zijn *verbeteringswensen* vanuit criminologisch onderzoek met betrekking tot het SSB?

- Zorgen voor meer en relevantere *achtergrondvariabelen* waardoor krachtiger, en meer *theorie-gestuurde analyses* mogelijk zijn. Uiteraard vindt de selectie van die variabelen plaats op grond van (theoretische) literatuur, maar waar mogelijk zou de band tussen dit (theorie-gestuurde) werk en het SSB verder verstevigd moeten worden.
- Koppeling met *buurtkenmerken* zou verkend moeten worden. Uit tal van sociologische studies is bekend hoe belangrijk (gebrek aan) sociaal kapitaal is voor (deviant) gedrag. De eerder geciteerde studie van Sampson en Raudenbusch laat zien dat fysieke kenmerken (verloedering) en sociale kenmerken van een buurt ('collective efficacy') van invloed zijn op het niveau van criminaliteit in een buurt. Wat kan én zou het CBS aan de beschikbaarstelling van dit soort gegevens vis-a-vis het SSB kunnen doen? Meer samenwerking met het NWO/KNAW DANS instituut? Nog beter is het als het CBS werk gaat maken van *sociaal-kapitaal statistieken*: wie investeert (sociaal) in wie; wie deelt met wie, wie ruilt met wie, wie bouwt op wie voort ('positief sociaal kapitaal') maar ook: wie wordt slachtoffer van wie en wie is dader van welk slachtoffer ('zuur sociaal kapitaal'). Omdat veel delicten samen worden gepleegd, is het ook belangrijk voor onderzoekers de *samenhang tussen personen* te kennen (samenplegen / netwerkonderzoek). Ook hier kan –gelukkig– geconstateerd worden dat het CBS deze wens onderschrijft. Als op enigszins afzienbare termijn survey data en HKS-gegevens opgenomen zijn in het SSB, zal ook meer werk van buurt-analyses mogelijk zijn.
- *Halt-gegevens* opnemen in de bestanden, waarbij met name belangrijk is na te gaan wat na de Halt-interventie met de jongeren is gebeurd en in hoeverre ze de Halt-verwijzing volbrengen. Te overwegen is ook gegevens over de betrokken 12-minners (*de Stop-maatregel*) toe te voegen.
- Nagaan wat de mogelijkheden zijn het *ESB* (het Economisch Statistisch Bestand) en het *SSB te koppelen vanuit het gezichtspunt van* onderzoek naar organisatiecriminaliteit, bevelen wij eveneens aan.

- Te komen tot een *gearticuleerd beleid* door het CBS ten aanzien van de vraag *welke bestanden wel en welke niet gekoppeld worden*, is de voorlaatste oproep. Daar zou meer dan tot nu toe de scientific, i.e. criminological community bij betrokken moeten worden. Een voorbeeld uit de WODC praktijk: is het duidelijk wat het CBS zal gaan besluiten ten aanzien van gegevens over *inburgering*. Voor een project als de Integratiekaart, is zo'n besluit wezenlijk, maar onduidelijk is wat en waarom besloten gaat worden.
- De laatste wens is het meer werk te maken van het koppelen van HKS en SSB bestanden andere justitie-databestanden. Het WODC is thans doende deze in een pilot project te koppelen en pro's en con's op een rijtje te zetten. Denk aan:
 - *OMDATA* – OMDATA bevat gegevens over de afdoening van strafzaken in eerste aanleg.
 - *Onderzoeks- en Beleidsdatabase Justitiële Documentatie (OBJD)* – De OBJD is een geanonimiseerde database zonder verjaringsregels waarin de complete justitiële voorgeschiedenis van personen wordt geregistreerd, inclusief de zaken die in hoger beroep zijn behandeld en de daarbij behorende onherroepelijke uitspraken en afdoeningen.
 - *TenUitvoerleggingsProgramma gevangeniswezen en JJI's (TULP/GW/JJI)* – Dit is het registratiesysteem van de sector gevangeniswezen van de Dienst Justitiële Inrichtingen (DJI), resp. de justitiële jeugdinrichtingen.

Ook gegevens betreffende de reclassering en het *Landelijk Alcohol en Drugs Informatie-Systeem (LADIS)* zouden in aanmerking kunnen komen. Dit laatste is een bestand dat niet binnen Justitie aanwezig is. Het bevat gegevens van de reguliere (GGZ) verslavingszorg en de Stichting Verslavingsreclassering GGZ Nederland (SVG); het betreft enkel de ambulante⁹ verslavingszorg. LADIS bevat informatie over de soort verslaving en behandeling van hulpzoekers in de reguliere verslavingszorg én verslaafden die behandeld worden via de verslavingsreclassering.

9) Ambulante zorg betekent dat de patiënt zich voor behandeling verplaatst naar de zorgverstreker. Klinieken waar een cliënt overnacht vallen hier dus niet onder.

Tot slot

Koppeling van gegevensbestanden (voor wetenschappelijk onderzoek) vereist grote precisie. Koppelingsfouten kunnen desastreuze gevolgen hebben. Voor wetenschap, beleid en samenleving. Precisie kost tijd en tijd is schaars. Het wetenschappelijk onderzoek zal derhalve moeten leren leven met enerzijds een steeds informatiever, vollediger en 'rijker' worden van het SSB (en haar companen), maar anderzijds ook begrip moeten hebben voor strenge veiligheidseisen en zekere beschikbaarheidsproblemen.

Referenties

- Algemene Rekenkamer, Opvang zwerfjongeren, nr. 28 265, vergaderjaar 2001-2002, 2004.
- Bijl, R. et al. De Integratiekaart 2005. WODC/CBS, Den Haag, 2005.
- Blom, M. et al, Verdacht van criminaliteit, Wodc/CBS, Den Haag, 2005.
- Engbersen, Godfried et al. Illegale vreemdelingen in Nederland: omvang, overkomst, verblijf en uitzetting, Risbo, Rotterdam, 2002.
- Hoogteijling, E., Illegalen in Nederland, Maandstatistiek van de Bevolking, CBS, 2002.
- Huberts, L. en Nelen, H. Corruptie in het Nederlands openbaar bestuur, omvang, aard en afdoening. VU, Amsterdam, 2005.
- Laan, A. van der, M. Blom & G. Huijbregst, Monitor Jeugd Terecht 2005, WODC.
- Levitt, S.D and J. Dubner, Freakonomics : a rogue economist explores the hidden side of everything, New York, William Moore, 2005.
- Moors, J.A. et al, Kiezen voor delen? Evaluatie van de eerste fase van de invoering van meerpersoonscegebruik, IVA/WODC, Boom Pers, 2004.
- Rovers, G.B. & de Vries Robbé, E., Interne criminaliteit in de logistieke sector, WODC-Boom Pers, 2005.
- Sikkel, D., P.G.M. van der Heijden en G. van Gils, Methoden voor omvangschattingen van verborgen populaties, met name illegalen, WODC, sixtat, Universiteit Utrecht FSW, WODC-Boom Pers, 2006.
- Wartna, B. et al, Door na de gevangenis; Een cijfermatig overzicht van de strafrechtelijke recidive onder ex-gedetineerden, WODC, 2005.
- Wartna, B., M.Blom & N.Tollenaar, De WODC-Recidivemonitor, WODC, 2004.





Henk G. van de Bunt (EUR/VU)

Reactie op: Over onderzoek
naar criminaliteit en de waarde van
het Sociaal Statistisch Bestand:
een analyse in 8 punten.

1 Inleiding

In zijn inleiding behandelt Frans Leeuw een aantal voorbeelden van grootschalig empirisch onderzoek in de criminologie op basis van grote gegevensbestanden. Hij wijst op een aantal interessante databestanden van politie (HKS) en justitie (JDS). Het HKS registreert de dader- en daadgegevens van alle door de politie aangehouden verdachten. Het JDS registreert de afdoeningen van de justitiecontacten van personen die in Nederland werden verdacht van een strafbaar feit. Leeuw gaat vervolgens specifieker in op enkele plus- en minpunten van het werken met SSB gegevens in het kader van criminologisch onderzoek. Ik ben het grotendeels met zijn beoordeling eens. In deze korte reactie zal ik enkele aanvullende opmerkingen maken en eindigen met een punt van aandacht.

2 De mogelijkheden van HKS en JDS

HKS en JDS bieden een schat aan informatie over de (geregistreerde) criminaliteit in Nederland, over het functioneren van de strafrechtelijke instellingen en het effect van hun optreden. Wie enige jaren terug gaat in het JDS, kan bekijken of afgestrafte daders weer opnieuw in het HKS of het JDS terugkeren. Dit is precies wat het WODC in de afgelopen jaren heeft gedaan. Zo is bekeken hoeveel personen die in de periode 1996-1999 uit de gevangenis zijn gekomen, binnen 6 jaar weer opnieuw tegen een strafrechtelijke veroordeling opliepen. Hieruit komt naar voren dat de meeste ex-gedetineerden in deze periode zelfs meerdere keren opnieuw met justitie in aanraking zijn gekomen. De recidive van deze groep ex-gedetineerden is ongeveer 4 keer zo hoog als onder de daders in het algemeen (Wartna et al. 2005).

Er zijn tal van mogelijkheden om de beide bestanden nog verder te exploreren. Binnen het JDS zou specifiekere gekeken kunnen worden naar recidivekansen per delicts- of leeftijdscategorie. Het onderzoek naar recidivepatronen verkeert nog in een pril stadium. In mindere mate geldt dit ook voor het HKS-onderzoek. Onderzoek naar het HKS kan in kaart brengen wie 'veelplegers' zijn en waar zij wonen. Bovendien bevat het HKS veel gegevens over aangiften van misdrijven, waardoor er informatie kan worden gegenereerd over de spreiding van misdrijven over tijd, plaats en tijdstip van plegen (c.q. ontdekking). Dit zijn gegevens die voor het stellen van prioriteiten en het bepalen van de politie-inzet uitermate belangrijk zijn.

Vertraagd door de politieke correctheid uit de jaren negentig, is pas in de afgelopen jaren een begin gemaakt is met het systematisch gebruik van deze belangrijke bestanden. Er is nog steeds sprake van onderbenutting van de bestaande mogelijkheden. Inmiddels dienen zich nog weer nieuwe mogelijkheden aan, namelijk de koppeling van politie- en justitiebestanden aan gegevens uit het SSB.

3 HKS/JDS en SSB

De koppeling van politie- en justitiebestanden aan het SSB of SSB-satellieten kent naar mijn mening twee belangrijke voordelen.

In de *eerste* plaats kunnen dankzij dergelijke koppelingen gebreken of tekortkomingen in de huidige politie- en justitieregistraties worden opgevangen. Een belangrijke tekortkoming bijvoorbeeld van HKS en JDS is dat de etnische afkomst van verdachten/daders onvoldoende duidelijk wordt vermeld. Alleen de geboorteplaats van de betrokkene wordt vermeld, niet de herkomst van de ouders. Dit betekent dat het criminaliteitsprobleem van de tweede generatie allochtonen niet op basis van uitsluitend HKS in kaart gebracht kan worden. In de vele rapportages over het aandeel van allochtonen in de geregistreerde criminaliteit kon tot voor kort slechts een beeld gegeven worden van het aandeel van de eerste generatie allochtonen. Maar dankzij de koppeling die recentelijk – op individueel niveau - is aangebracht tussen het HKS en SSB is het gelukt van alle betrokken HKS-verdachten de etnische afkomst vast te stellen. In 2005 verscheen hierover een belangwekkende rapportage (Blom et al. 2005). Voor het eerst is het aandeel van de tweede generatie allochtonen in de geregistreerde criminaliteit zichtbaar gemaakt. Van de in totaal 163.000 HKS-verdachten (2002) bleken 22.000 tot deze tweede generatie gerekend te kunnen worden. Met inbegrip van de eerste generatie (39.000) is 37,5% van alle in Nederland woonachtige geregistreerde verdachten van een misdrijf van allochtone herkomst.

Dit type koppeling biedt talloze andere mogelijkheden die nog beter benut kunnen worden. Wellicht kan ook JDS gekoppeld worden aan het SSB om in het kader van recidive het aandeel van de tweede generatie zichtbaar te kunnen maken. Meer in het algemeen gesproken kan de koppeling eraan bijdragen dat - voor politie en justitie - relevante *beschrijvingsvragen* beter beantwoord kunnen worden. Immers er kunnen met HKS/JDS koppelingen worden gemaakt met het SSB en diens satellieten die informatie bevatten over onder meer huwelijkse status, opleiding (schooluitval, diploma), arbeidspositie (werkend/uitkering) en woonplek en verhuisgedrag.

In de *tweede* plaats kan de koppeling *theoriegericht* onderzoek stimuleren. In de criminologie wordt veel aandacht besteed aan de vraag of crimineel gedrag al dan niet onderdeel uitmaakt van een algemener, afwijkend gedragspatroon. Het is het probleem van de ‘generality of deviancy’. Volgens Gottfredson en Hirschi hebben misdrijfplegers een lage mate van zelfcontrole. Die eigenschap speelt ze in alle fasen van hun leven op allerlei gebieden parten. Misdrijfplegers hebben, als gevolg van hun lage zelfbeheersing, meestal veel conflicten in de persoonlijke sfeer (echtscheidingen, regelmatige arbeidsconflicten) en maken ook op straat brokken (verkeersongevallen) en leven erg ongezond door onmatig eten, roken en drinken. Op grond van deze intrigerende theorie omtrent de oorzaken van crimineel gedrag zou het aangewezen zijn de mogelijkheden te verkennen van het koppelen van een daderpopulatie aan bijvoorbeeld de GBA, WW/ABW of POLS.

Er is nog een geheel andere theoretische en beleidsmatig relevante invalshoek, en dat betreft de mobiliteit in wijken en de sociale spreiding binnen wijken. Veel lokaal veiligheidsbeleid is er op gericht wijken ‘leefbaar’ te houden. De assumptie is dat de leefbaarheid en de sociale controle achteruit gaan wanneer wijken een eenzijdige ‘zwarte’ samenstelling krijgen. Recent is de zogenaamde Rotterdam-wet aanvaard die gemeentebesturen om die reden de bevoegdheid geeft om een spreidingsbeleid te voeren in daarvoor in aanmerking komende wijken. Koppeling van HKS aan SSB kan vermoedelijk meer licht werpen op een eventueel bestaand verband. Wonen in ‘zwarte’ wijken relatief meer misdrijfplegers dan in gemêleerd samengestelde wijken? Is er een verband tussen onveiligheidsgevoel en verhuismobiliteit?

4 Onbetrouwbaarheid van mensen en... bestanden

Het is, kortom, aantrekkelijk om registers aan elkaar te koppelen. Frans Leeuw voegt nog een ander pluspunt toe, en dat is dat het in het SSB vooral om geregistreerd gedrag gaat en niet om percepties of attitudes. Hij noemt dit zelfs een *fort* van deze registers. Op dit punt deel ik zijn analyse niet. Het argument gaat er aan voorbij dat ook de ogenschijnlijk zo objectieve gedragsregistraties manco’s bezitten. Een deel van de misdrijfplegers blijft namelijk onzichtbaar: zij komen direct uit het buitenland, en zijn niet bekend in de Nederlandse SSB registers. Of zij verblijven al langere tijd in Nederland, maar maken gebruik van valse persoons- of verblijfsdocumenten. Zij staan dan wel geregistreerd in registers, maar de gegevens zijn onbetrouwbaar. Een selectieve uitval derhalve van gegevens over misdrijfplegers in de SSB registers. Uiteraard geldt dat ook voor de politie- en justitie-regis-

ters. Maar daar is sprake van een selectieve ‘overrapportage’. Het is een bekend gegeven in de criminologie dat bij het gebruik van politiebesteden rekening gehouden moet worden met de selectiviteit in de pakkans en oplossing van misdrijven. Straatcriminelen lopen een veel grotere kans opgepakt te worden dan de goed opgeleide, blanke fraudeur. Deze onbetrouwbaarheid van politiebesteden is altijd al een belangrijk punt van zorg geweest. Maar de kwaliteit van de bestanden dreigt de laatste jaren verder achteruit te gaan doordat het optreden van de politie steeds meer beïnvloed wordt door de lokale politiek. Er worden prioriteiten in de opsporing gesteld en over aantallen geregistreerde en opgehelderde misdrijven worden tevoren afspraken gemaakt. Naarmate politiekorpsen zich meer gaan vastleggen op prestatieafspraken, zal de neiging bestaan om het doen van met name aangiften met een lage ophelderingskans te ontmoedigen. Ook deze ontwikkeling draagt bij aan de verdere selectieve vervuiling van de bestanden. De politieregisters zijn in de huidige ‘afrekencultuur’ geen rustig bezit meer.

Concluderend, het is uitermate belangrijk om politie- en justitiebestanden te koppelen aan het SSB of SSB-satellieten. Maar één waarschuwing is op zijn plaats: de mogelijkheden die het koppelen biedt, kunnen alleen optimaal worden benut als de politie- en justitiebestanden van goede kwaliteit zijn.

Referenties

Blom, M., J. Oudhof, R.V. Bijl en B.F.M. Bakker (2005), Verdacht van criminaliteit; allochtonen en autochtonen nader bekeken. WODC/CBS: Den Haag.

Gottfredson, M.R. & T. Hirschi (1990) *A General Theory of Crime*, Stanford: Stanford U.P.

Wartna, B.S.J., N. Tollenaar en M. Blom (2005), *Recidive 1997*. WODC: Den Haag.





Jaco Dagevos (SCP)

Het SSB en het zicht op de integratie van minderheden. Of: waarom surveygegevens onontbeerlijk blijven.

Deze bijdrage gaat in op de betekenis van SSB- en surveygegevens voor het inzicht in de integratie van minderheden. De stelling wordt verdedigd dat voor een adequaat beeld van de integratie van minderheden surveygegevens noodzakelijk zijn en blijven. Aan de registergegevens kleven, net als aan de surveygegevens, nadelen. De ene bron is niet zonder meer superieur aan de andere. Registergegevens schieten vooral tekort op het vlak van de sociaal-culturele integratie. Het is wenselijk dat in de nabije toekomst SSB- en surveygegevens meer met elkaar in verband worden gebracht. Dit kan door middel van verrijking van surveybestanden met gegevens uit het SSB. Tevens valt winst te boeken door bij de formulering en beantwoording van onderzoeksvragen beide bronnen in onderlinge samenhang te benutten.

1 Inleiding

Inzicht in de integratie van minderheden in Nederland is sinds enige decennia in hoge mate gebaseerd op informatie afkomstig uit survey-onderzoeken en registratiegegevens. Met het operationeel worden van het SSB is het gebruik van registratiegegevens verder toegenomen, ook op het terrein van etnische minderheden. Een mooi moment voor een pas op de plaats. Deze bijdrage zet surveygegevens tegenover gegevens uit het SSB. De interesse gaat uit naar de vraag hoe de kwaliteit van deze gegevens zich tot elkaar verhouden en welke inzichten SSB- en surveygegevens opleveren? En hieruit voortvloeiend: betekent de ontwikkeling van het SSB dat op minderheden gerichte surveys kunnen worden afgebouwd, of is daarvan geen sprake?

In de volgende paragraaf worden eerst de sterke kanten van het SSB geïnventariseerd, gevolgd door een overzicht van de merites van (minderheden)surveys. Paragraaf 4 bevat de conclusies van deze vergelijking samen en doet enkele voorstellen voor de optimalisering van het gebruik van survey- en registratiegegevens in het minderhedenonderzoek van de komende jaren.

2 De sterke punten van het SSB

2.1 Integrale tellingen

Registratiegegevens kenmerken zich doorgaans door grote aantallen. Dat maakt deze gegevens aantrekkelijk. Op het terrein van het minderhedenonderzoek betekent dit dat ook de kleine allochtone groepen in beeld komen, die in surveys doorgaans buiten beschouwing moeten blijven. Dit wordt wel als duidelijk voordeel boven surveygegevens aangemerkt (vgl. Van Rijn et al., 2004). Grote aantallen maken nauwkeurige detaillering mogelijk, zowel naar achtergrondkenmerken als naar ruimtelijk schaalniveau. Daar kunnen surveys niet tegenop.

Het is aan de andere kant wel zo dat in Nederland ook survey-onderzoek is uitgevoerd onder bijvoorbeeld de Somaliërs. Dit is met 22.000 personen een kleine groep, hetgeen laat zien dat ook kleine minderheden via een survey te bereiken zijn.

SSB maakt gebruik van bestaande registratiegegevens, geen enquêtedruk

Eén van de redenen om te investeren in het SSB ligt in de kosten en moeite die gepaard gaan met de uitvoering van surveys. Dit argument geldt in het bijzonder de surveys gericht op

10) Dit laat onverlet dat de betrekkelijk lage responscijfers vooral worden veroorzaakt door het feit dat allochtonen vaak grote stadsbewoners zijn. Indien hiervoor wordt gecorrigeerd, blijkt dat de verschillen in nonrespons tussen autochtonen en

allochtonen nog maar weinig van elkaar afwijken (Schmeets en Van der Bie 2005). Ook in het LAS-onderzoek vinden we geen dramatische verschillen in respons tussen allochtonen en autochtonen. In het LAS-onderzoek kijkt de respons van autochtonen

nauwelijks af van die van Turken en Antillianen. De respons van autochtonen is wel hoger dan bij de Marokkanen en Surinamers. Een definitieve uitspraak is nog niet te geven. Dan zou een met Schmeets en Van der Bie vergelijkbare analyse

minderheden. Deze zijn peperduur, vooral vanwege de noodzaak dat de dataverzameling face-to-face moet worden uitgevoerd, bij voorkeur door tweetalige enquêteurs. De kosten worden verder opgejaagd door het betrekkelijk hoge aantal benaderingen dat nodig is om de gewenste aantallen minderheden te bereiken.¹⁰ Door de naar verhouding hoge nonrespons neemt het risico van een selectieve onderzoeksgroep toe, bijvoorbeeld doordat beter geïntegreerden sterker zijn vertegenwoordigd en degenen die de Nederlandse samenleving minder goed gezind zijn deelname aan een survey vaker weigeren.

Verder rijzen er bij surveys altijd vragen rondom betrouwbaarheid en validiteit, en dit geldt in het bijzonder voor surveys onder minderheden. Taalproblemen en culturele verschillen kunnen de beantwoording van vragen beïnvloeden. Het construeren van valide meetinstrumenten die rekening houden met culturele verschillen staat in Nederland nog in de kinderschoenen. Interview- en vraagstellingseffecten spelen in minderhedensurveys een belangrijke rol, al is het gewicht nog steeds niet helemaal duidelijk (zie voor een overzicht van deze discussie Houtkoop-Steenstra en Veenman, 2002).

Problemen met betrekking tot nonrespons en de daaruit mogelijk voortvloeiende selectiviteit van de onderzoeksgroep spelen bij registratiegegevens meestal niet of nauwelijks parten. In beginsel gaat het om een integrale telling van personen (die als inwoner van een gemeente, als werknemer of als uitkeringsontvanger geboekstaafd staan). Of individuen in een registratiebestand terechtkomen, is doorgaans geen kwestie van een keus. Wie maandelijks zijn salaris wil ontvangen, zal als werknemer geregistreerd staan. Voor personen die afhankelijk zijn van een uitkering geldt hetzelfde: men dient zich te registreren. Vraagstukken rondom selectiviteit treffen registratiegegevens in aanzienlijk mindere mate. Dit wil overigens niet zeggen dat registratiegegevens zonder problemen zijn, maar daarover zo dadelijk meer.

2.2 Nieuwe mogelijkheden voor longitudinaal onderzoek

Om ontwikkelingen in de individuele levensloop (b.v. overgang op de arbeidsmarkt, veranderingen in inkomensbronnen) vast te stellen, voorzien surveys via panelonderzoek of retrospectieve vragen informatie. Hoewel veelgebruikt, zijn de nadelen niet te veronachtzamen. Panels worden gekenmerkt door, vaak aanzienlijke, paneluitval. Vaak is die uitval selectief. Retrospectieve vragen hebben als nadeel dat geheugeneffecten de kwaliteit van de gegevens beïnvloeden. Bovendien zijn er grenzen aan het aantal retrospectieve vragen dat aan respondenten kan worden voorgelegd.

In dit licht biedt het SSB een belangrijke stap vooruit. Het door het CBS opgebouwde Immigrantenpanel levert interessante informatie over de positieverwerving van cohorten nieuwkomers. Onder andere onderscheiden naar jaar van komst naar Nederland, het

moeten worden uitgevoerd, die rekening houdt met verschillen tussen allochtone groepen en autochtonen in o.a. urbanisatiegraad, opleidingsniveau en arbeidsmarktpositie.

migratiemotief en etnische groep wordt nagegaan welk deel een betaalde baan, een uitkering of geen inkomen heeft. Doordat deze gegevens jaarlijks beschikbaar zijn, leveren analyses op het Immigrantenpanel belangrijke inzichten op over het tempo waarin groepen zich op de arbeidsmarkt manifesteren (Sprangers et al., 2004a en 2004b; Bijl et al., 2005). Binnen surveys zou dergelijke informatie nauwelijks zijn te verzamelen. Men zou over een groot aantal respondenten moeten beschikken dat zich in bepaalde jaren in Nederland heeft gevestigd. Het jaar-op-jaar bevragen van de maatschappelijke positie doet een sterk beroep op het geheugen en kost veel (interview)tijd, zeker voor cohorten die al geruime tijd in Nederland zijn. En los daarvan, in Nederland is voor allochtone nieuwkomers dergelijke informatie gewoonweg niet voorhanden.

2.3 'Moeilijke' groepen en 'moeilijke' onderwerpen

Een voordeel van registratiegegevens boven surveygegevens is dat ze voor bepaalde onderwerpen gegevens bevatten die in surveys niet of niet goed zijn te meten. Inkomensmetingen in surveys zijn doorgaans problematisch; de item-nonrespons is aan de hoge kant en om tot bepaalde inkomensbegrippen te komen zijn gedetailleerde deelvragen nodig, die veel interviewtijd en hoofdbreken kosten. Informatie over criminaliteit is een ander voorbeeld. Zeker met betrekking tot het ouderschap zijn surveys niet de meest geëigende bronnen. De trefkans is gering, zowel doordat bepaalde delicten betrekkelijk weinig voorkomen als vanwege de waarschijnlijk selectieve respons van de groep ouders.¹¹ Registratiegegevens, hoewel ook niet zonder problemen, zijn dan belangrijk om de lacune te vullen.

Door koppeling van diverse registratiebestanden aan de geboortelandgegevens uit het GBA is in de afgelopen jaren informatie losgekomen over onder andere criminaliteit, de participatie in het voortgezet en hoger onderwijs, inkomens en uitkeringen en deelname van allochtonen aan gesubsidieerde arbeid.

Koppeling van de bestanden van de Kamers van Koophandel aan het GBA geeft gedetailleerd inzicht in de betekenis van het etnisch ondernemerschap, de vertegenwoordiging in diverse branches, welke overlevingskansen hun bedrijven hebben en welke verschuivingen zich tussen de eerste en tweede generatie voltrekken (zie o.m. Van den Tillaart en Doesborgh, 2004). Deze gegevens zijn een belangrijke aanvulling op informatie uit surveys en nog lang niet uitgeanalyseerd.

11) Al kunnen gerichte surveys deze bezwaren voor een belangrijk deel wegnemen, zie Wittebrood 2006.

3 Zijn minderhedensurveys achterhaald?

Hierboven zijn verschillende voordelen onderscheiden die (gekoppelde) registratiegegevens hebben boven surveygegevens. Wat betekent dit voor deze surveys? Kunnen we in de komende jaren toe met steeds minder surveygegevens? Het antwoord is ontkennend. Om adequaat de integratie van minderheden te blijven volgen, zijn surveys onontbeerlijk, ook nu steeds meer informatie aan het SSB valt te ontlenen. Want surveys hebben zo hun eigen merites.

3.1 Survey is flexibeler

In de kern komen de nadelen van registratiegegevens erop neer dat het registratiegegevens zijn. Dit betekent onder meer dat je voor je informatie afhankelijk bent van wat in registraties aanwezig is. Dit hoeven niet altijd de vraagstukken te zijn waar beleid of wetenschap behoefte aan heeft. Een periodiek gehouden survey is aanmerkelijk flexibeler. Vraagstukken die (meer) aandacht behoeven, kunnen in de dataverzameling worden ingebracht. In de laatste jaren is bijvoorbeeld de discussie over integratie steeds meer verschoven van structurele naar sociaal-culturele integratie. Een survey kan hierop inspelen, zoals ook is gebeurd in de verschillende edities van de SPVA die inhoudelijk wezenlijk van elkaar verschillen, registratiegegevens laten zich veel moeilijker aanpassen. Het in 2004 en 2005 uitgevoerde LAS-onderzoek vormt tot op zekere hoogte weer een aanvulling op de SPVA-onderzoeken, doordat op meer uitgebreide wijze is ingegaan op onderwerpen die betrekking hebben op burgerschap, zoals participatie in verenigingen, politieke participatie en cultuurdeelname.

3.2 Actuele gegevens

Gegevens uit het SSB lopen vaak twee jaar achter. Binnen het universitaire onderzoek wordt hier meestal niet zwaar aan getild, maar het beleid vraagt – en dit geldt zeker voor het beleidsterrein van integratie – om actuele cijfers. Hoewel het een race tegen de klok is, leveren surveys kort na afronding van het veldwerk al gegevens op. Met de juiste inzet van middelen en menskracht kan de integratie van minderheden met zeer recente cijfers worden gevolgd.

Registratiegegevens lossen niet alle problemen rondom meetfouten en paneluitval op. Voorstanders van registratiegegevens wijzen op meetfouten in surveygegevens. Dit is, zoals hiervoor al toegelicht, een belangrijk probleem. Maar registratiegegevens zijn evenmin vrij van problemen. Er zitten fouten in registratiebestanden, bijvoorbeeld doordat mutaties niet worden doorgegeven. Dit speelt onder andere de longitudinale vergelijking

van de integratie van nieuwkomers parten (zie b.v. Sprangers et al., 2004a). Personen die niet langer woonachtig zijn in de gemeente waar ze volgens het GBA behoren te wonen, worden, zoals dat heet, administratief afgevoerd. Het kan zijn dat personen zijn geëmigreerd, zonder dat dit is gemeld of dat ze in Nederland verblijven zonder zich te hebben ingeschreven. Het is aannemelijk dat dit laatste het geval is bij een behoorlijk deel van de nieuwkomers. Bij de vergelijking in de tijd raak je deze personen dus kwijt. Afgaande op de bevindingen van Van Rijn et al. (2004) gaat het om forse aandelen.

Registratiegegevens dekken verder maar een deel van de werkelijkheid. In criminologisch onderzoek bijvoorbeeld is dit een belangrijk punt van discussie (zie Wittebrood, 2006 voor een overzicht). Kern van de zaak is hier dat slechts een deel van de criminaliteit in het vizier van instanties komt, en in dat verband wordt geregistreerd. Een zeer groot deel van de delicten en slachtoffers onttrekt zich aan de waarneming van politie of O.M. en komen derhalve niet in de registraties terecht. Registratiegegevens beschrijven de geregistreeerde werkelijkheid en hetgeen daarbuiten valt moet via andere onderzoekstechnieken, waaronder surveys, worden opgespoord.

3.3 Geen aansluiting op formele definities

Een ander nadeel van registratiegegevens is dat ze soms op net een andere wijze het verschijnsel meten dan volgens formele definities eigenlijk zou moeten.

Gegevens over de arbeidsmarkt zijn hier een voorbeeld van. Via de EBB meten we de netto-arbeidsparticipatie af aan personen die minstens 12 uur per week werken. In de registratiegegevens is echter het aantal gewerkte uren niet bekend. Je krijgt dus niet de cijfers die je volgens de formele definities eigenlijk zou willen hebben. En misschien nog bezwaarlijker: er worden uiteenlopende gegevens geproduceerd over min of meer hetzelfde verschijnsel. Voor gebruikers van gegevens kan dit tot verwarring leiden. Voor voorstanders van de 1-cijfer gedachte – voor elk verschijnsel wordt één cijfer geproduceerd – is het een gruwel. Een ander voorbeeld betreft de wijze van afbakening van kleine banen. In de EBB gebeurt dit, conform de formele definitie, op grond van het aantal gewerkte uren, in sommige analyses op basis van het SSB worden kleine banen geoperationaliseerd als personen die niet meer verdienen dan eenderde van het minimumloon.

Voor de bepaling van de werkloosheid bieden registratiegegevens nog minder soelaas. De formele werkloosheidsdefinitie stelt onder meer eisen wat betreft het zoekgedrag en beschikbaarheid van niet-werkenden. Deze informatie is niet aan registraties te ontleen. En om nog even op het terrein van arbeid te blijven; we missen nogal wat gegevens, zoals het beroepsniveau van werkenden, het aantal uren dat wordt gewerkt (voltijd- en deeltijdwerk), de aard van het dienstverband (vast/flexibel) en meer subjectieve opvattingen over de werkloosheids- en arbeidsmarktpositie. Bij dit laatste kan men onder andere denken aan de inschatting van allochtone werkzoekenden om werk te vinden en aan de opvatting van

werkenden of hun opleidingsniveau aansluit op het niveau van hun baan. Informatie over opleidingsniveau, arbeidsmarktpositie en inkomen en percepties op deze terreinen is wel beschikbaar in surveys. Een niet te veronachtzamen voordeel boven registratiegegevens is bovendien dat tevens informatie kan worden verzameld over de in het land van herkomst verworven opleidingskwalificaties en bereikte arbeidsmarktpositie.

3.4 SSB dekt slechts een klein deel van het integratievraagstuk

Naast de zojuist genoemde beperkingen, vormt een belangrijk probleem dat op het terrein van het integratievraagstuk de reikwijdte van het SSB begrensd is. De achilleshiel van het SSB ligt vooral in het ontbreken van informatie over de sociaal-culturele integratie.

Registratiegegevens bevatten geen cijfers over opvattingen, over de mate, aard en wenselijkheid van contacten met autochtonen, over de beheersing en het gebruik van de Nederlandse taal, over de identificatie met de eigen groep en de gerichtheid op de autochtone samenleving, over de mate van religiositeit, over ervaren discriminatie, etc. De ‘zachte’ kant van de integratie wordt nauwelijks gedekt door registratiegegevens. Een uitzondering daarop zijn gegevens over gemengde huwelijken, die als een belangrijke indicator gelden voor sociaal-culturele integratie (zie Bakker en Giesbertz, 2005). Maar alles bijeengenomen, blijft het ontbreken van inzicht in de sociaal-culturele integratie een zwak punt van het SSB, en dit wringt des te meer doordat de publieke en politieke discussie de laatste jaren sterk is opgeschoven naar vraagstukken met betrekking tot sociaal-culturele integratie.

Surveys zijn, met al hun gebreken, beter in staat om in deze informatiebehoefte te voorzien. Nog belangrijker is dat belangrijke indicatoren van structurele en sociaal-culturele integratie op het zelfde moment en bij dezelfde persoon zijn gemeten. Dit biedt mogelijkheden om de samenhang tussen beide vormen van integratie te analyseren. Dergelijke analyses wijzen op tussen groepen uiteenlopende integratiepatronen en daarmee duidelijk maken dat er in Nederland grote verschillen bestaan in het tempo en de wijze van integratie van allochtone groepen (zie b.v. Dagevos, 2005; Gijsberts en Dagevos, 2005). De uitkomsten zijn beleidsmatig van belang en sluiten aan bij actuele wetenschappelijke discussies, onder meer in relatie tot het vraagstuk van gesegmenteerde assimilatie (b.v. Portes en Zhou, 1994).

Ook andere relevante thema's in minderhedenonderzoek zijn grotendeels afhankelijk van surveyinformatie. Er bestaat in Nederland en daarbuiten inmiddels een uitgebreide onderzoekstraditie gericht op wederzijdse beeldvorming van verschillende bevolkingsgroepen (zie b.v. Gijsberts et al., 2004). Welke stereotiepe kenmerken groepen toeschrijven aan andere groepen en in hoeverre de populariteit van groepen van elkaar verschilt, kan niet anders dan via surveys worden vastgesteld. Ditzelfde geldt voor de meting van diverse aspecten van burgerschap, een concept dat in de discussie over wat integratie is, een steeds belangrijker

plaats krijgt. In hoeverre allochtonen actief zijn in vrijwilligerswerk, politiek betrokken en actief zijn of aan cultuur deelnemen kan niet aan registratiegegevens worden ontleend.

4 SSB en surveys: kansen voor een LAT-relatie?

Hierboven zijn de gegevens over de integratie van minderheden uit surveys en het SSB tegen elkaar afgewogen. Onderstaande tabel vat de plussen en minnen van beide gegevensbronnen samen.

Tabel 1

SSB: +	<i>(minderheden) surveys: -</i>
Grootschalig, in beginsel alle groepen, detaillering	Accent op grote(re) groepen
Gebruik beschikbare registraties	Enquetedruk, duur
Volgen van individuen	Panels, met selectieve uitval
Info over 'moeilijke' onderwerpen en 'moeilijke' groepen	(Selectieve) nonrespons, item-nonrespons
SSB: -	<i>(minderheden) surveys: +</i>
Inhoud afhankelijk van registraties	Inhoud flexibel aan te passen, toesnijden op voor beleid en wetenschap relevante thema's
Loopt enkele jaren achter	Actueel
Verskil met officiële definities	Aansluiting formele definities
Deelonderwerpen integratie	Breed overzicht intergratie-indicatoren

Wat betekenen deze verschillen nu voor de informatievoorziening op het terrein van de integratie van minderheden? Uit het voorgaande is al duidelijk geworden dat er zwaarwegende argumenten zijn om de integratie van minderheidsgroepen te blijven volgen met (minderheden)surveys. Met het SSB is een imposant bouwwerk opgetrokken van elkaar verbonden registratiegegevens, maar het kan surveys niet vervangen, al is dit uit oogpunt van kosten en interviewbaarheid binnen allochtone groepen misschien te betreuren.

Veel meer perspectief is er om beide gegevensbronnen op elkaar te betrekken. Dit kan op verschillende manieren. Het SSB kan dienen als steekproefkader voor surveys, hetgeen ook al effectief is gebeurd. Het SSB kan in beginsel zowel de steekproef als het uiteindelijke bestand 'verrijken' met registratiegegevens. Wanneer hierdoor kenmerken van de steekproefpersoon bekend zijn, zoals generatie, migratiemotief, jaar van vestiging, kan het veldwerk veel gericht worden uitgevoerd (b.v. matching op geslacht van respondent en

enquêteur, gerichte inzet van tweetalige enquêteur bij 1e generatie respondent). Dit heeft waarschijnlijk gunstige gevolgen voor de respons en het verloop van het gesprek. Tevens kan het uiteindelijke surveybestand aangevuld worden met registratiegegevens, en dan vooral voor onderwerpen die moeilijk in surveys zijn te meten. Al blijft het een nadeel dat sommige registratiegegevens enkele jaren achterlopen. Hoe dit ook zij, een verdergaande fysieke integratie van bronnen is mogelijk.

Op dit moment is bij het CBS de vormgeving van een zogenoemde Allochtonensatelliet in voorbereiding, waarin zoveel mogelijk relevante bronnen aan elkaar worden geknoopt.

Tot op heden worden in publicaties (tegenstrijdige) bevindingen uit surveys en het SSB nauwelijks met elkaar in verband gebracht. Dat is jammer. Uitkomsten van het SSB leiden tot nieuwe vragen, die mogelijk door surveys kunnen worden beantwoord. Uit het SSB blijkt bijvoorbeeld dat het integratietempo van bepaalde cohorten nieuwkomers van elkaar verschilt. Thans beschikbare surveygegevens kunnen factoren aandragen die meer duidelijkheid bieden over de achtergronden van deze verschillen. Het omgekeerde komt natuurlijk ook voor: uitkomsten uit survey-onderzoek kunnen aanleiding zijn om registratiegegevens nader te verkennen. Hier is tot op heden weinig van terecht gekomen. Het gezamenlijk formuleren van thema's die gebruik maken van beide typen databronnen zou dan een belangrijke stap zijn.

Referenties

- Bakker, B. en P. Giesbertz (2005). 'Heeft Cupido een maat(je)? Over de integratie van allochtonen op de arbeidsmarkt'. In: *Bevolkingstrends* (2), p. 65-74.
- Bijl, R.V., A. Zorlu, A.S. van Rijn, R.P.W. Jennissen, M. Blom (2005). *Integratiekaart 2005. De maatschappelijke integratie van migranten in de tijd gevolgd: trend- en cohortanalyses*. Den Haag: WODC/CBS.
- Dagevos J., m.m.v. J. Iedema en R. Schellingerhout (2005). 'Gescheiden werelden? De etnische signatuur van vrijetijdscontacten van minderheden'. *Sociologie* (1) 1, p. 52-69.
- Gijsberts, M. en J. Dagevos (2005). *Uit elkaars buurt. De invloed van etnische concentratie op integratie en beeldvorming*. Den Haag: Sociaal en Cultureel Planbureau (SCP-publicatie 2005/13).
- Gijsberts, M., L. Hagendoorn en P. Scheepers (red.) (2004). *Nationalism and Exclusion of Migrants. Cross-National Comparisons*. Aldershot: Ashgate.
- Houtkoop-Steenstra, H. en J. Veenman (2002). *Interviewen in de multiculturele samenleving. Problemen en oplossingen*. Assen: Van Gorcum.
- Portes A. en M. Zhou (1994). 'Should immigrants assimilate?' In: *Public Interest* (18), p. 116-125.
- Rijn, A.S. van, A. Zorlu, R.V. Bijl & B.F.M. Bakker (2004). *De ontwikkeling van een integratiekaart*. Den Haag: WODC/CBS.
- Schmeets, H. en R. van der Bie (2005). *Enquêteonderzoek onder allochtonen. Problemen en oplossingen. Voorburg/Heerlen: CBS*.
- Sprangers, A., A. Zorlu, J. Hartog en H. Nicolaas (2004a) 'Immigranten op de arbeidsmarkt'. In: *Bevolkingstrends* (2), p. 45-53.
- Sprangers A., A. Zorlu, J. Hartog en H. Nicolaas (2004b). 'Werk en uitkering van Turkse en Marokkaanse immigranten'. In: *Bevolkingstrends* (3), p. 90-96.
- Tillaart, H. van den en J. Doesborgh (2004). *Demografie etnisch ondernemerschap in de detailhandel*. Nijmegen: ITS.
- Wittebrood, K. (2006). *Slachtoffers van criminaliteit in Nederland. Omvang, achtergronden en beleid*. Den Haag: Sociaal en Cultureel Planbureau.





Godfried Engbersen (EUR)

De ongekende multiculturele samenleving nader gekend.*

* Reactie op: 'Het SSB en het zicht op de integratie van minderheden.
Of: waarom surveygegevens onontbeerlijk blijven'.

1 Inleiding

Het Sociaal Statistisch Bestand (SSB) is gelet op de thema van het symposium bedoeld om de sociale samenhang in Nederland beter in beeld te krijgen. In mijn reactie op het betoog van Jaco Dagevos zal ik het acroniem SSB als leidraad van mijn betoog nemen. Ik zal drie vragen centraal stellen in mijn reactie. De eerste vraag is in hoeverre het SSB meer kennis zal opleveren over de *sociale positie* van migrantengroepen in de Nederlandse samenleving (de eerste S). De tweede vraag is in hoeverre het SSB een diepgaand inzicht zal bieden in de *sociale samenhang van de multiculturele samenleving* (de tweede S) en de derde vraag is in hoeverre het SSB tot een scherper *beeld* komt van ongekende aspecten van de Nederlandse multiculturele samenleving (de laatste B).

In de beantwoording van deze vragen zal blijken dat ik het op alle punten eens ben met Jaco Dagevos. Ik zal zijn betoog op een aantal punten aanvullen. Bijvoorbeeld als het gaat om de noodzaak tot aanvullend survey-onderzoek. Daarnaast wil ik ook nog eens aandringen op een verbreding van het Sociaal Statistisch Bestand.

2 Kennis over de sociale positie van migrantengroepen

Een belangrijk winstpunt van het SSB is dat het gebaseerd is op de registratiegegevens waardoor problemen van non-respons en selectieve respons worden voorkomen. Exacte gegevens zijn moeilijk te geven, maar de response die in landelijk dekkend *face-to-face* onderzoek in de vijftiger jaren nog meer dan 90% bedroeg, is via circa 70% in de jaren zeventig teruggelopen naar circa 40%. Voor telefoonenquête en schriftelijke enquête gelden soortgelijke of nog lagere cijfers. Vooral laag opgeleide en langdurig werkloze migranten zijn daardoor slecht bereikbaar. Ik herinner me nog heel goed mijn promotieonderzoek naar de sociale gevolgen van langdurige werkloosheid in de Rotterdamse wijk 'Het Nieuwe Westen'. Nagenoeg niemand reageerde op mijn brieven, deuren bleven gesloten en toen ik mijn respondenten probeerde te bellen bleek 45% van de langdurig werklozen in mijn onderzoekswijk een geheim telefoonnummer te bezitten. Later heb ik dit aspect sociologisch geïdentificeerd als een indicator van selectieve sociale afsluiting dat vooral voorkomt bij kwetsbare maatschappelijke groepen (Engbersen, 2006 [1990]) Het fenomeen van 'gehei-

me nummers' speelt overigens nog steeds in achterstandbuurten blijktens een recente publicatie van de Groningse sociologen Van Goor, Jansma en Veenstra (2005). Processen van selectieve sociale afsluiting zijn een groot probleem in survey-onderzoek onder migranten. Het SSB ondervangt dit probleem. Daardoor kan in een aantal cruciale aspecten van de maatschappelijke positie van migranten (bijvoorbeeld hun arbeid-, inkomens- en huisvestingspositie) een goed inzicht worden verkregen.

Een tweede voordeel is dat meer informatie beschikbaar komt over andere migrantengroepen dan de klassieke groepen zoals Turken, Marokkanen, Surinamers en Antillianen. In 2002 kende Nederland zo'n 56 migrantengroepen met meer dan 4000 personen. In tabel 1 wordt een overzicht gegeven van de allochtone samenstelling van de Nederlandse bevolking. Daaruit kan worden afgeleid dat er een groeiend aantal migrantengroepen is met meer dan 15.000 mensen. Het is van groot wetenschappelijk belang om inzicht te krijgen in het integratieproces bij deze nieuwe migrantengroepen. Het SSB kan daarvoor cruciale informatie bieden.

Een serieus probleem is echter dat het SSB nog geen opleidingsgegevens biedt. We kunnen we veel te weten komen over de inkomens-, arbeidsmarkt- en hun huisvestingspositie en zelfs van hun mogelijke betrokkenheid bij criminaliteit. Dit is zeer relevant voor inzicht in de integratie van migrantengroepen, maar tot op heden ontbreekt inzicht in het opleidingsniveau (het human of cultural capital van migranten). Ik heb begrepen dat voor deze omissie een oplossing in het verschiet ligt. Maar het ontbreken van opleidingsniveau is in onze meritocratische samenleving een belangrijke lacune. Het onderzoek van Dagevos (2006) en anderen laat immers zien hoe bepalend het opleidingsniveau is voor de maatschappelijke kansen van migranten.

Tabel 1. Allochtone bevolking in Nederland naar etnische herkomst 1995-2004, 2-jaars gemiddelden*

	als % van de allochtone bevolking						
	absolute aantallen * 1000						
	95-96	97-98	99-00	01-02	03-04	95-96	03-04
Totaal	15530	15707	15926	16149	16282		
Autochtone bevolking	13004	13047	13103	13147	13176		
Allochtone bevolking	2527	2660	2823	3002	3105		
Waarvan							
Westerse landen	1330	1348	1377	1411	1422	52.7	45.8
Waarvan							
Duitsland	410	405	400	395	388	16.2	12.5
Engeland	66	67	71	75	76	2.6	2.5
België	111	112	113	113	113	4.4	3.6
Polen	26	28	30	33	38	1.0	1.2
Vml. Joegoslavië	58	62	69	75	76	2.3	2.5
Vml. Sovjet Unie	14	19	26	37	43	0.6	1.4
Verenigde Staten	23	25	27	29	30	0.9	1.0
Indonesië	411	407	405	402	397	16.3	12.8
Niet-Westerse landen	1196	1312	1446	1590	1684	47.3	54.2
Waarvan							
Turkije	276	295	314	336	355	10.9	11.4
Marokko	229	247	267	290	311	9.1	10.0
Somalië	22	27	29	28	23	0.9	0.8
Zuid Afrika	10	11	13	15	15	0.4	0.5
Ghana	13	14	16	18	19	0.5	0.6
Kaapverdië	17	18	18	19	20	0.7	0.6
Egypte	12	13	15	17	18	0.5	0.6
Suriname	283	294	306	318	327	11.2	10.5
Ned. Antillen en Aruba	88	96	112	127	131	3.5	4.2
Irak	14	26	36	42	43	0.5	1.4
Afghanistan	6	14	24	33	37	0.3	1.2
China	24	27	31	37	43	1.0	1.4
Iran	18	21	24	27	29	0.7	0.9
Vietnam	13	14	15	16	18	0.5	0.6
Pakistan	14	15	16	18	18	0.6	0.6
Hongkong	17	17	18	18	18	0.7	0.6

*Selectie 15.000+

Bron: CBS (www.cbs.nl)

3 Inzicht in de sociale samenhang van de multiculturele samenleving

Ik kom op mijn tweede S, namelijk die van de sociale samenhang. Wat is de betekenis van het SSB voor dit aspect van de migrantenthematiek?

Ten eerste is het van belang om op te merken dat het SSB koppelingen mogelijk maakt tussen compositievariabelen van migrantengroepen (in termen van arbeidsmarktpositie, leeftijd, inkomen, gezondheid) en contextvariabelen (zoals kenmerken van buurten). Dergelijke micro-macro-verbindingen zijn uiterst relevant. Het levert, onder andere, nieuwe inzichten op over de mogelijke maatschappelijke effecten van de ruimtelijke concentratie en segregatie van migrantengroepen (armoede, criminaliteit en gezondheid) (Latten, 2005)

Maar is het genoeg? Niet helemaal. En hier komt de noodzaak naar voren van additioneel survey-onderzoek dat respondenten ondervraagt over andere dan sociaal-economische aspecten. Ik sluit hierbij aan bij collega Dagevos. Voor het doorgronden van de maatschappelijke positie is het van belang om ook naar opvattingen en attitudes te vragen. Naar sociaal-psychologische aspecten als vertrouwen in politiek, openbaar bestuur, buurten en medeburgers, of naar het bestaan van negatieve gevoelens van ressentiment, desidentificatie en wrok (Burgers en Engbersen, 1994; De Swaan, 1999). De antwoorden op dergelijke vragen geven inzicht in de mate waarin autochtone en allochtone groepen zich thuis voelen in de Nederlandse samenleving en zich begrepen voelen door politiek en openbaar bestuur. Antwoorden op dergelijke vragen geven ook inzicht in het mogelijk ontstaan van processen van radicalisering bij Islamitische jongeren. Ook is het van belang om nader inzicht te verkrijgen in ervaringen met discriminatie. Voor dit soort type onderzoek is dus aanvullend survey- en etnografisch onderzoek noodzakelijk. Er is veel kritiek uitgeoefend op de sociale wetenschappen omdat zij de Fortuyn-revolutie niet hebben zien aankomen. Er zijn derhalve gevoelige onderzoeksantennes nodig om onderhuidse gevoelens van onvrede op het spoor te komen (Boutellier, 2006).

Een tweede aspect van de maatschappelijke samenhang betreft de aard van de sociale netwerken waarin migranten zijn ingebed. Daarbij is niet alleen de vraag relevant naar de mate van contacten tussen autochtonen en allochtonen, maar ook de vraag naar de kwaliteit van sociale netwerken. Op dit moment heeft onderzoek naar 'sociaal kapitaal' een enorme vlucht genomen in de sociale wetenschappen. Sociaal kapitaal heeft betrekking op de sociale netwerken waarin mensen zijn ingebed en de hulpbronnen (informatie, geld, steun) die zij daaruit weten te destilleren (Portes, 1998 en Putnam, 2000). Naast cultureel

kapitaal is sociaal kapitaal een uiterst belangrijke ‘sociale muntsoort’ voor sociale integratie en stijging. Om het sociale kapitaal van individuele migranten en migrantengroepen te kennen, is echter aanvullend survey-onderzoek nodig. Dergelijke data zijn niet via registratiebestanden te krijgen.

Dat geldt ook voor een laatste aspect dat nader aandacht behoeft, namelijk het feit dat migranten deel uitmaken van netwerken die landgrensoverschrijdend zijn (transnationalisme). Het SSB is in de eerste plaats een *nationaal bestand*. Inzicht in grensoverschrijdende activiteiten en identificaties (in welke mate identificeren migranten zich met personen uit het land van herkomst of de wijdere dispora) is ook relevant voor het begrijpen van hedendaagse processen van integratie. Veel migranten, ook de tweede generatie, kennen transnationale bindingen. Dat blijkt bijvoorbeeld uit de onderstaand tabel waarin de resultaten zijn verwerkt van een verkennend onderzoek onder zes verschillende migrantengroepen (Snel et al. 2006). Uit dat onderzoek blijkt dat veel migranten grensoverschrijdende sociaal-culturele activiteiten ontwikkelen, naast (alledaagse) economische en politieke activiteiten. Een belangrijke vraag is in hoeverre patronen van transnationalisme een belemmering dan wel een extra voordeel zijn voor succesvolle integratie. In de internationale literatuur keren beide standpunten terug (zie Snel et al. 2006). Nader empirisch onderzoek moet uitsluitsel bieden.

Tabel 2. Overzicht van transnationale activiteiten naar migrantengroep (in % van de totale groep)

	Marokko (N=50)		Nederlandse Antillen (N=50)		Irak (N=50)		Voormalig Joegoslavië (N=50)		Japan (N=50)		USA (N=50)	
<i>Alledaagse economische activiteiten</i>												
Maakt geld over naar familie in herkomstland	40	16	72	62	2	4						
Goederen sturen naar familie in herkomstland	28	14	4	54	12	14						
Bezit van huis in herkomstland	16	8	14	46	24	14						
Bijdragen aan goede doelen in herkomstland	12	2	30	26	0	26						
Totaal (betrokken bij minstens een activiteit)	54	28	78	82	32	42						
<i>Professionele economische activiteiten</i>												
Investeren in ondernemingen in herkomstland	2	2	6	0	10	24						
Handel drijven met herkomstland	4	2	0	0	8	2						
Bezoek aan herkomstland voor business	0	8	0	2	14	32						
Totaal (betrokken bij minstens een activiteit)	4	10	6	2	22	50						
<i>Politieke activiteiten</i>												
Leest kranten uit herkomstland	10	58	62	70	54	66						
Op de hoogte van politiek in herkomstland	70	56	80	76	64	96						
Lid van politieke partij in herkomstland	2	6	24	0	6	20						
Participeert in demonstraties m.b.t. herkomstland	4	6	52	58	4	6						
Totaal (betrokken bij minstens een activiteit)	72	76	88	94	82	100						
<i>Sociaal-culturele activiteiten in herkomstland</i>												
Bezoekt familie en vrienden in herkomstland	90	78	36	92	72	86						
Frequent contact met familie in herkomstland	72	82	94	92	92	98						
Lid van maatschappelijke organisaties in herkomstland	2	4	22	0	6	16						
Totaal (betrokken bij minstens een activiteit)	94	96	94	96	96	100						
<i>Sociaal-culturele activiteiten in Nederland</i>												
Lid van organisaties m.b.t. herkomstland	16	16	62	8	18	16						
Bezoekt bijeenkomsten met voornamelijk landgenoten	50	60	62	60	38	16						
Bezoekt culturele activiteiten	56	44	28	34	66	56						
Totaal (betrokken bij minstens een activiteit)	86	90	96	84	88	68						

De cijfers in deze tabel geven aan welk percentage respondenten een bepaalde activiteit rapporteert.

Bron: Snel, Engbersen en Leerkes 2006

4 Een scherper beeld van de multiculturele samenleving?

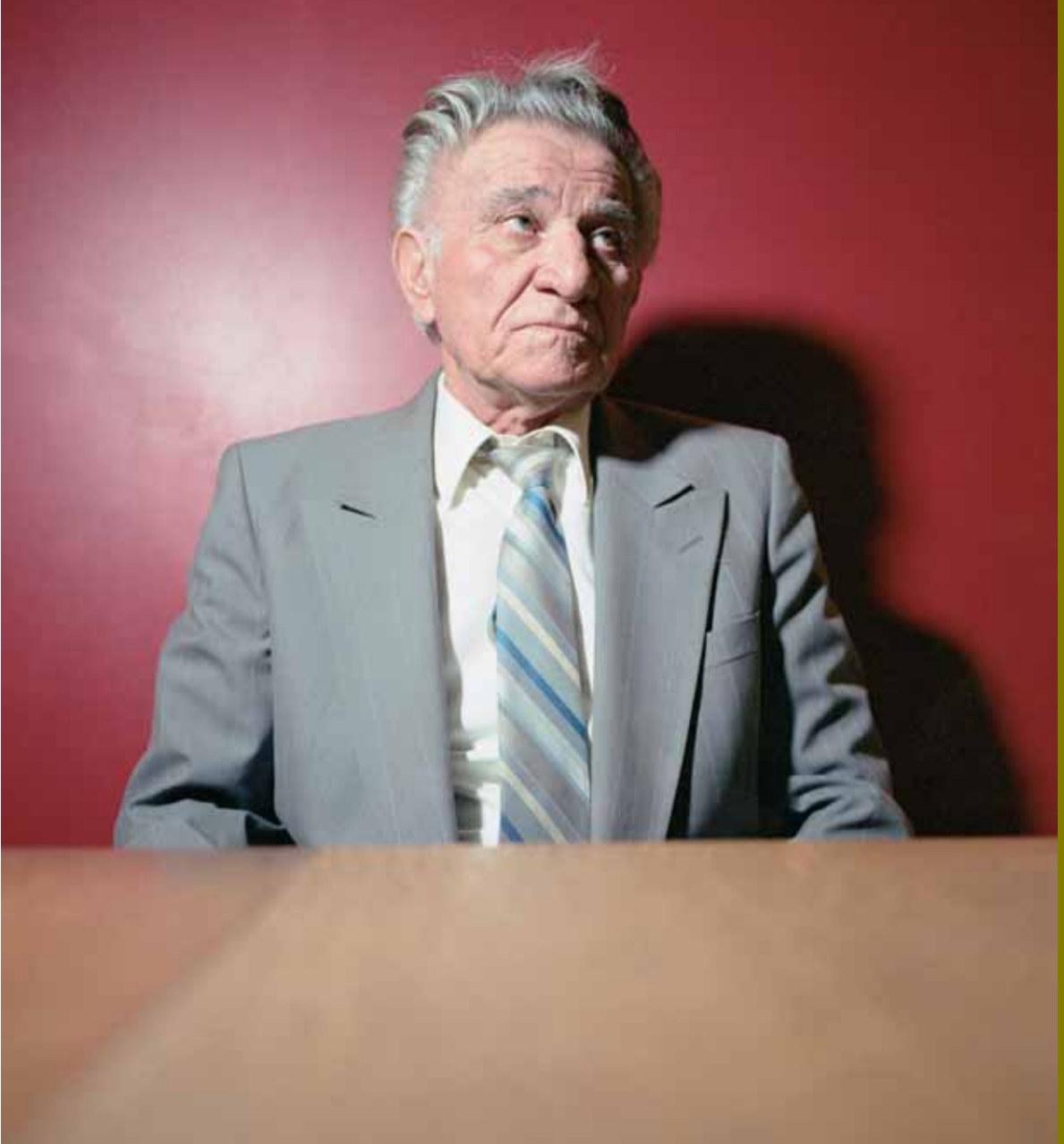
Mijn stelling is dat het SSB diverse aspecten van de multiculturele samenleving scherper kan belichten door optimaal gebruik te maken van registratiegegevens en die adequaat te koppelen. Een duidelijk voorbeeld daarvan is onderzoek naar de betrokkenheid van migrantengroepen (eerste en tweede generatie) bij criminaliteit. De relatie tussen internationale migratie, integratieprocessen en criminaliteit is een klassiek thema in de sociologie (denk aan de publicaties van de beruchte Chicago School: Thomas and Znaniecki, *The Polish Peasant in Europe and America*, 1918-1920 of Trasher, *The Gang*, 1927). Het is een goede zaak dat het taboe op een onderwerp als deze is opgeheven in Nederland. Wel is het zaak om dit vraagstuk op een gedistantieerde manier te behandelen en niet te vervallen in sterk normatieve, moraliserende analyses.

Ik zou echter willen dat het SSB nog een stapje verder gaat door ook meer systematisch onderzoek te doen naar een relatief nieuw aspect van de migrantenproblematiek, en wel het vraagstuk van ongedocumenteerde of illegale migranten. Met mijn onderzoeksgroep heb ik de afgelopen jaren veel onderzoek gedaan naar *De Ongekende Stad* (Burgers en Engbersen, 1999) en naar de *Verborgene Stad* (Burgers en Engbersen, 2001). Dat onderzoek is deels gebaseerd op etnografisch onderzoek onder irreguliere of illegale migranten, en deels gebaseerd op nadere analyses van diverse registratiegegevens, in het bijzonder het VAS (Vreemdelingen Administratie Systeem), het HKS, het GBA en cijfers van het IND. Door optimaal gebruik te maken van deze registratiegegevens was het mogelijk om inzicht te krijgen in de samenstelling van de illegalenpopulatie, de ruimtelijke spreiding van illegalen in Nederland en in de Randstad en in de verbanden die er bestaan tussen asielmigratie, irreguliere migratie en criminaliteit. Ook is een schatting gemaakt van de omvang van irreguliere migranten in Nederland. De ironie is derhalve dat door optimaal gebruik te maken van bepaalde registratiesystemen inzicht ontstaat in de zogenaamde 'ongedocumenteerde' werkelijkheid. In dit verband zou men het SSB - en het gebruik van andere registratiebestanden, inclusief relevante koppelingen - als het antwoord van het CBS kunnen beschouwen op de these van 'De Ongekende Samenleving' (DOS) door de politiek-filosoof Van Gunsteren. Middels het SSB kan DOS beter worden gekend.

Referenties

- Boutellier, H. (2006), 'Valkuilen voor sociologen: kanttekeningen vanuit beleidsonderzoek'. In: Godfried Engbersen en Jos de Haan (red.) *Balans en toekomst van de sociologie*. Amsterdam: Amsterdam University Press, pp. 284-287.
- Burgers, J. en G. Engbersen (1994), 'De wrok van Spangen'. In: *Psychologie en Maatschappij*, Jaargang 18, nummer 4, 1994, pp. 292-302.
- Burgers, J. en G. Engbersen (1999), *De ongekende stad I: Illegale vreemdelingen in Rotterdam*. Amsterdam: Boom.
- Dagevos, J. (2006), *Hoge jeugdwerkloosheid onder etnische minderheden*. Den Haag: Sociaal en Cultureel Planbureau.
- Engbersen, Godfried (2006[1990]), *Publieke bijstandsgeheimen. Het ontstaan van een onderklasse in Nederland*. Amsterdam: Amsterdam University Press.
- Engbersen, Godfried en Jack Burgers (2001), *De verborgen stad. De zeven gezichten van Rotterdam*. Amsterdam: Amsterdam University Press.
- Goor, Henk van, Folbert Jansma en René Veenstra (2005), 'Differences in undercoverage and nonresponse between neighborhoods in a telephone survey'. *Psychological Reports*, 96, 867-878.
- Latten, Jan (2005), *Zwanger van segregatie*. Amsterdam: AUP (oratie).
- Portes, A. (1998), 'Social capital: its origins and applications in modern sociology'. *Annual Review of Sociology*, 24, pp. 1-24.
- Putnam, R. D. (2000), *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon & Schuster.
- Snel, Erik, Godfried Engbersen en Arjen Leerkes, (2006), 'Transnational involvement and social integration'. *Global Networks*, 6 (3): 285-308.
- Swaan, A. de (1999), 'Uitdijende kringen van desidentificatie: gedachten over Rwanda'. In: A. de Swaan, *De draagbare De Swaan*, Amsterdam: Prometheus, pp. 227-245.

1982 van de dynamische
mogelijke van de toekomst?





Jules Theeuwes (UvA/WRR), Paul van Leeuwe (WRR),
Dirk Scheele (WRR) & Tijn Woortman (WRR)

Arbeidsmarktdynamiek en vergrijzing.

Met behulp van de gegevens van het Sociaal Statistisch Bestand (SSB) brengen we de dynamiek van banen en van personen op de Nederlandse arbeidsmarkt tussen 1999 en 2002 in kaart. We schenken daarbij ook aandacht aan de mobiliteit van de oudere werknemer. Uit onze analyse komt naar voren dat de Nederlandse arbeidsmarkt wat betreft dynamiek vergelijkbaar is met andere landen. Wel zijn er indicaties dat de 'last van de veranderingen' onevenredig bij bepaalde groepen op de arbeidsmarkt is geconcentreerd.

1 Inleiding en Probleemschets

Bij de WRR werken we momenteel aan een rapport over arbeidsdynamiek. Daarbij gebruiken we gegevens van het Sociaal Statistisch Bestand (SSB) om ontwikkelingen in kaart te brengen en hypothesen te toetsen of beter gezegd om te kijken of bepaalde a-priori veronderstellingen niet worden tegengesproken door de data. We geven in dit onderdeel een korte motivatie waarom we een rapport schrijven over arbeidsdynamiek. Vervolgens presenteren we in onderdeel 2 een korte beschrijving van het SSB van het CBS en van de data die wij daaruit gebruiken. Op basis van de SSB gegevens kan zowel naar de dynamiek van banen worden gekeken (onderdeel 3) als naar de dynamiek van personen (onderdeel 4). In beide onderdelen bespreken we een aantal werkhypothesen en lichten toe hoe we daarbij het SSB materiaal gebruiken. Onderdeel 5 bevat conclusies over de bewegelijkheid van de arbeidsmarkt en het gebruik van SSB data voor het analyseren van arbeidsmarktdynamiek.

Arbeidsdynamiek is een oud thema in Europa en in Nederland. De discussie over sclerose op de arbeidsmarkt dateert al vanaf het midden van de jaren zeventig toen zich in het voetspoor van de oliecrises een aantal economische crises voordeed waarbij de werkloosheid in Europa tot grote hoogte steeg. De diepere oorzaak van de economische crises werd mede gezocht in de geringe flexibiliteit van de arbeidsverhoudingen. Als gevolg van starre arbeidsmarktinstituties (zoals genereuze uitkeringen, hoge ontslagbescherming, machtige vakbonden) zouden in Europa beduidend meer werknemers dan in de Verenigde Staten langdurig buiten het arbeidsproces staan. Vooral de trage instroom van werklozen naar betaalde arbeid laat in Europa veel te wensen over. Door het aanwijzen van de starheid van de arbeidsmarktinstituties als oorzaak voor de problemen op de arbeidsmarkt werd de noodzaak van hervormingen bevestigd (zie bijvoorbeeld C. Bean (1994)). De discussie over starre arbeidsmarkten in Europa gaat in deze eeuw onverminderd verder (O. Blanchard [2005]). Het Nederlandse poldermodel heeft enkele gloriejaren gekend aan het eind van de jaren negentig maar heeft bij het begin van deze eeuw alweer zijn glans verloren. Hervormingen in de richting van meer marktconforme arbeidsverhoudingen en een grotere flexibiliteit van arbeid worden nog steeds als onontbeerlijk gezien. Deze discussie is de laatste jaren eerder in intensiteit toegenomen dan afgenomen. Daar is een aantal oorzaken voor te noemen.

De noodzaak van flexibilisering wordt de laatste tijd steeds nadrukkelijker verbonden met de door de globalisering aangewakkerde concurrentiedruk. Een groeiende groep van Aziatische landen manifesteert zich nadrukkelijk op de wereldmarkt. Recentelijk komen ook

de economische potenties van een enorm groot land als China in beeld. En ook dichterbij huis neemt de concurrentie toe. De economie van de landen in Midden en Oost Europa raakt steeds meer verweven met de economie van de landen in West Europa. Een groter wordende flexibilisering van productiestructuren brengt met zich mee dat delen van de productie uit West Europa verplaatst worden naar landen waar kostenvoordelen zijn te behalen. De signalering van dit proces zet een druk op de flexibilisering van de arbeid in West Europa. In het Europese debat wordt hier op ingespeeld. Ook de snellere technologische ontwikkeling en de introductie van nieuwe productieprocessen en producten is een mogelijke oorzaak van grotere dynamiek op de arbeidsmarkt.

Een tweetal ontwikkelingen is in Europa aanleiding om de noodzaak van de flexibilisering van de arbeid verder te benadrukken. Door de monetaire eenwording in Europa is het aanpassingsmechanisme van de wisselkoersen weggefallen. Daardoor is meer druk komen te staan op aanpassingen via de arbeidsmarkt om de spanningen tussen de productmarkten in verschillende lidstaten weg te nemen. Een andere ontwikkeling is de vergrijzing van de beroepsbevolking. Oudere werknemers worden over het algemeen als minder flexibel gezien. Zonder verdere maatregelen zou de flexibiliteit van de arbeid dan door de vergrijzing worden aangetast.

Een tweede reden waarom de vergrijzing mogelijk problematisch kan zijn heeft te maken met de bronnen van de economische groei. In het recente verleden kon Nederland zijn economische groei vooral verwezenlijken door toestroom van relatief grote cohorten schoolverlaters op de arbeidsmarkt en van (her)intredende vrouwen. In de volgende decennia daalt door de demografische veranderingen de groei van het aanbod van arbeid en daarmee droogt een bron van economische groei op. Een manier om dit te compenseren is door oudere werknemers langer productief aan het werk te houden. Een tweede manier om de productiviteitsgroei van het productiebestel te vergroten is door meer sociale innovatie in het arbeidsbestel (te denken valt aan meer investeringen in oudere werknemers, functie-wisseling en deeltijdpensioenering).

Met deze beleidszorg over dynamiek in het achterhoofd zijn de volgende vragen voor ons belangrijk. In onze analyse van de SSB data proberen we onder meer die vragen te beantwoorden.

- Zijn er aanwijzingen dat de Nederlandse arbeidsmarkt minder dynamisch is dan andere arbeidsmarkten?
- Zijn er aanwijzingen dat de dynamiek op de arbeidsmarkt recentelijk is toegenomen, onder invloed van toegenomen internationale concurrentie, technologische ontwikkeling, opkomst van China en India en andere externe schokken?

- Is de impact van de arbeidsmarktdynamiek verschillend voor de boven- en de onderkant van de arbeidsmarkt? Is er wat betreft arbeidsmarktdynamiek sprake van insiders en outsiders waarbij outsiders meer de 'last van de dynamiek' dragen dan de insiders?
- Wat is in het bijzonder de positie van de oudere werknemers in de dynamiek van de Nederlandse arbeidsmarkt?

2 Analyse van de gegevens

De gegevens die worden gepresenteerd zijn gebaseerd op cijfermateriaal afkomstig uit het Sociaal Statistisch Bestand (SSB) van het CBS. In het inleidende hoofdstuk geeft Bart Bakker een beschrijving van het SSB. Voor een uitgebreide toelichting kan daar naar verwezen worden. Hieronder volgen een aantal voor ons onderzoek saillante kenmerken van het SSB.

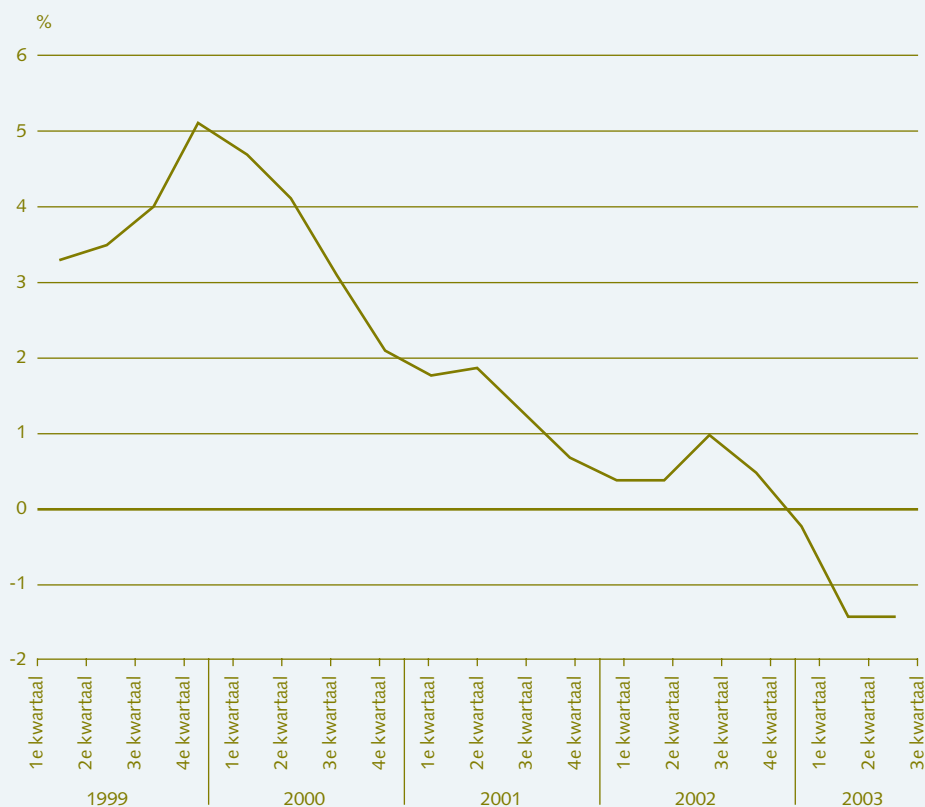
Het SSB is in belangrijke mate gebaseerd op administratieve gegevens. De backbone van het SSB is de Gemeentelijke Basis Administratie persoonsgegevens (GBA) waarin elke inwoner in Nederland is opgenomen. Vervolgens worden aan de GBA bestandsgegevens uit andere administratieve bestanden gekoppeld, zoals gegevens van de belastingadministratie, van de salarisadministratie van werkgevers, de gegevens van uitkering- en verzekerdenadministraties etc. Ook de gegevens uit enquêtes die het CBS houdt onder de bedrijven (zoals de Enquête Werkgelegenheid en Lonen) en onder personen (zoals de Enquête Beroepsbevolking (EBB)) worden aan het SSB bestand gekoppeld. De sleutel om te koppelen is daarbij meestal het SOFI nummer. Het SSB is o.m. beschikbaar als personenbestand (maar er is ook informatie voor de eenheid baan of uitkering). Van dat kenmerk maken we gebruik om de dynamiek van personen te meten.

Werkende personen zijn verbonden aan een bedrijf. Elk bedrijf kan worden geïdentificeerd met behulp van een bedrijfsidentificatie code. Op basis daarvan kunnen de data ook worden georganiseerd met de bedrijven als eenheid van observatie. De banen in die bedrijven worden bezet door de personen die er werken. Op basis van de organisatie van het SSB met de bedrijven en hun banen als focus, en de mogelijkheid om daarmee volgtijdelijke analyses te doen (door koppeling van jaargangen) kunnen we de dynamiek van banen bestuderen.

Op dit moment beschikken we over de SSB gegevens voor de jaren 1999 tot en met 2002.¹² De veranderingen in de status van personen en in de banen bij bedrijven wordt gemeten door de situatie op de peildatum in het ene jaar te vergelijken met de situatie op de peildatum in het jaar dat daar aan voorafgaat. Als peildatum in de SSB-data wordt de laatste vrijdag in september van ieder jaar gehanteerd. Dat betekent dat we de veranderingen en de dynamiek bestuderen die heeft plaatsgevonden tussen de jaren 1999-2000; 2000-2001 en 2001-2002.

Bij de analyses in dit hoofdstuk is het goed om de ontwikkeling van de conjunctuur in het achterhoofd te houden. In het begin van de periode 1999-2003 piekte de economische hoogconjunctuur waarna een omslag optrad in de groei. Dit is duidelijk te zien in de ontwikkeling van de BBP groeicijfers per kwartaal (volumemutaties) in de onderstaande grafiek.

Figuur 1. Conjunctuur (groei BBP per kwartaal)



12) We hebben recentelijk ook de beschikking over de gegevens van 2003 gekregen. Maar deze gegevens zijn voor dit hoofdstuk niet geanalyseerd.

De conjunctuurindicator piekt in het 4e kwartaal 1999 en zakt vervolgens naar een steeds lager niveau. De groei wordt zelfs negatief vanaf het 1e kwartaal 2003.

3 Dynamiek van banen

We beginnen dit onderdeel met een overzicht van de verdeling van bedrijven in groeiende, krimpende en neutrale bedrijven wat betreft werkgelegenheidsontwikkeling (paragraaf 3.1). Vervolgens geven we in paragraaf 3.2 een algemene beschrijving van de baancreatie en destructie in deze periode. Het meest in het oog springende feit is de substantiële omvang van creatie en destructie van werkgelegenheid. Hieruit blijkt dat de jaarlijkse netto groei van het aantal banen geen goed beeld geeft van de banendynamiek op de Nederlandse arbeidsmarkt. De wijze waarop we baancreatie en destructie meten sluit aan bij wat in de internationale literatuur gebruikelijk is (Davis, Haltiwanger en Schuh, 1998)¹³. In deze paragraaf zoomen we ook in op de verschillen tussen bedrijfstakken wat betreft baancreatie en destructie. We willen weten in welke sectoren de banendynamiek zich vooral afspeelt.

3.1 Groeiende, krimpende en neutrale bedrijven

In 1999 waren er iets meer dan 425 duizend bedrijven in Nederland met meer dan 7 miljoen banen.

Tabel 1. Aantal bedrijven en banen naar grootteklasse (in % van totaal)

<i>Grootte (aantal banen)</i>	<i>% Bedrijven</i>	<i>% Banen</i>
<=1	26,2%	1,5%
2<5	27,9%	4,4%
5<10	15,0%	5,8%
10<20	9,5%	7,4%
20<50	6,3%	11,0%
50<100	2,1%	8,5%
100<200	1,1%	8,5%
200-<500	0,6%	11,0%
>=500	0,4%	41,7%
Aantal	425.600	7.282.000

13) Baanstromen zijn op meerdere manieren te meten. Zo kan dat gebeuren via meting van aantallen banen bij bedrijven op twee peildata (dat is hier het geval). Een andere methode is aan de hand van gesloten of ontbonden contracten. Beide methoden

geven een onderschatting omdat meervoudige transities binnen het observatietijdvak niet worden meegenomen. Met het SSB is het wel mogelijk om alle gesloten of ontbonden arbeidscontracten binnen een jaar mee te tellen. Kee en Milot (2005) komen dan

uit op ruwweg een verdubbeling van de dynamiek, dat wil zeggen een totale turnover van circa 8 miljoen. Dat wordt vrijwel volledig verklaard door het aangaan of ontbinden van kortlopende – vaak uitzend - contracten.

Zoals uit de bovenstaande tabel blijkt is de meerderheid van de bedrijven klein van omvang. Meer dan de helft van de bedrijven telt minder dan vijf banen. Deze kleine bedrijven vertegenwoordigen echter slechts circa 6% van het totale aantal banen (zie rechterkolom). Van het aantal bedrijven is 1% als groot te kenschetsen (200 en meer banen). In deze grote bedrijven zit meer dan 52% van de banen. Er zijn dus heel veel kleine bedrijven met relatief weinig banen en heel weinig grote bedrijven met relatief veel banen.

Van jaar op jaar zijn er bedrijven waar het aantal banen groeit, krimpt, dan wel gelijk blijft. We noemen dit in de volgende tabel groeiende, krimpende en neutrale bedrijven. Tussen het ene jaar en het andere zijn er ook nieuwe bedrijven en bedrijven die worden opgeheven. In de volgende tabel bekijken we voor de jaren 1999-2000 de verandering in de werkgelegenheid in deze verschillende soorten bedrijven (als een percentage van de werkgelegenheid in 1999 – behalve dan voor nieuwe bedrijven die per definitie in 1999 nog niet bestonden)¹⁴.

Tabel 2a. Arbeidsmarktstromen 1999-2000, totale aantallen werknemers (x 1000)

	1999	Blijvers	Vertrokken	Naar andere baan	Nieuw	Van andere baan	2000
Neutraal	565,1	478,3	27,7	59,1	31,9	54,9	565,1
Groei	3587,2	2892,8	214,4	480,1	457,5	774,0	4124,2
Krimp	3052,3	2052,7	307,4	692,2	210,8	329,8	2593,3
Nieuw					65,7	107,7	173,4
Opgeheven	77,3		31,9	45,4			
Totaal	7281,9	5423,8	581,4	1276,8	765,9	1266,4	7456,0

Tabel 2b. Arbeidsstromen 1999-2000, percentages van werkgelegenheid in 1999 (*)

	1999	Blijvers	Vertrokken	Naar andere baan	Nieuw	Van andere baan	2000
	A	B	C	D	E	A-B-C+D+E	
Neutraal	100,0%	84,6%	4,9%	10,5%	5,6%	9,7%	100,0%
Groei	100,0%	80,6%	6,0%	13,4%	12,8%	21,6%	115,0%
Krimp	100,0%	67,3%	10,1%	22,7%	6,9%	10,8%	85,0%
Nieuw bedrijf	0,0%				37,9%	62,1%	100,0%
Opgeheven	100,0%		41,3%	58,7%			0,0%
Totaal	100,0%	74,5%	8,0%	17,5%	10,5%	17,4%	102,4%

(*) Met uitzondering van de nieuwe bedrijven, waarbij het percentage berekend t.o.v. de werkgelegenheid in 2000

Vertrokken: uit een baan in 1999 vertrokken naar uitkering of inactiviteit (zonder uitkering)

Naar andere baan: vertrokken naar een andere werkgever

Nieuw: binnengekomen vanuit een uitkering of inactiviteit (zonder uitkering)

Van andere baan: binnengekomen van een andere werkgever

14) De teleenheid in tabel 2a en b (en de tabellen 3a,b en 4a,b) is eigenlijk het aantal banen die werknemers nog steeds hebben (blijvers) of waar vandaan ze zijn vertrokken (van andere baan) of waar ze naar toe gekomen zijn (naar andere baan).

Omdat het gemakkelijker is om in termen van werknemers te denken gebruiken we dat ook in de kop van de tabel. Het verschil tussen het aantal werknemers dat naar een andere baan vertrekt en van een andere baan binnenkomt, ontstaat ondermeer

omdat sommige werknemers in jaar t-1 twee banen hebben, in de loop van het jaar wordt een van deze banen beëindigd en houden ze in jaar t maar een baan over.

In de bovenste helft van de tabel presenteren we de bedrijven die zowel in 1999 als in 2000 bestonden. In de neutrale bedrijven is de werkgelegenheid in 1999 en 2000 hetzelfde. Desondanks verandert de samenstelling van de werkgelegenheid. Bijna vijf procent vertrekt naar een uitkeringssituatie of naar buiten de arbeidsmarkt. Iets meer dan 10,5% vertrekt naar een andere baan. In dezelfde periode komen er 5,6% nieuwe werknemers die vanuit een uitkeringssituatie of van buiten de arbeidsmarkt komen en 9,7% die vanaf een andere baan komen. Slechts 84,6% van de werknemers zijn zowel in 1999 als in 2000 in hetzelfde bedrijf werkzaam.

In de groeiende en de krimpende bedrijven zijn de relatieve verschuivingen groter dan in neutrale bedrijven. In groeiende bedrijven is goed zichtbaar dat relatief grote percentages werknemers worden aangetrokken: 12,8% komt van een uitkering of van buiten de arbeidsmarkt en 21,6% komt van een andere werkgever. In de krimpende bedrijven stromen behoorlijk veel werknemers uit, vooral naar een andere werkgever 22,7%.

In de tweede helft staat de informatie over de nieuwe en opgeheven bedrijven. Bijna 2/3 van de werknemers die in een nieuw bedrijf worden aangenomen komen van een andere werkgever. Bij de verdwijnende bedrijven gaat bijna 60% naar een ander bedrijf.

In de laatste regel staat de totale verandering. In de cel onderaan rechts staat dat de totale werkgelegenheid in 2000 met 2,4% was gegroeid ten opzichte van de werkgelegenheid in 1999. De tabel geeft goed weer hoezeer dit saldo de resultante is van grote stromen op de arbeidsmarkt. In totaal stroomde 25,5% uit een baan (8% + 17,5%) tussen 1999 en 2000 en stroomde 27,9% in een nieuwe baan (10,5% + 17,4%). De turnover (som van instroom en uitstroom) is 53,4%. Er vindt dus vrij veel creatie en destructie plaats op de Nederlandse arbeidsmarkt. Het grootste deel daarvan (17,5% plus 17,4%) betreft baanwisselaars.

Tabel 3a. Arbeidsmarktstromen 2000-2001, totale aantallen werknemers (x 1000)

	2000	<i>Blijvers</i>	<i>Vertrokken</i>	<i>Naar andere baan</i>	<i>Nieuw</i>	<i>Van andere baan</i>	2001
Neutraal	566,5	482,6	27,2	56,7	30,7	53,2	566,5
Groei	3692,8	3029,9	208,9	454,0	443,6	747,7	4221,2
Krimp	3111,8	2080,0	325,0	706,0	216,3	338,0	2635,1
Nieuw					79,6	129,1	208,7
Opgeheven	85,0		36,3	48,7			
Totaal	7456,1	5593,3	597,4	1265,4	770,2	1268,0	7631,5

Tabel 3b. Arbeidsmarktstromen 2000-2001, percentages van werkgelegenheid in 2000

	2000	<i>Blijvers</i>	<i>Vertrokken</i>	<i>Naar andere baan</i>	<i>Nieuw</i>	<i>Van andere baan</i>	2001
Neutraal	100,0%	85,2%	4,8%	10,0%	5,4%	9,4%	100,0%
Groeiend	100,0%	82,0%	5,7%	12,3%	12,05%	20,2%	114,3%
Krimpnd	100,0%	66,8%	10,4%	22,7%	7,0%	10,9%	84,7%
Nieuw	0,0%				38,1%	61,9%	100,0%
Opgeheven	100,0%		42,7%	57,3%			0,0%
Totaal	100,0%	75,0%	8,0%	17,0%	10,3%	17,0%	102,4%

In de vergelijkbare tabel arbeidsmarktstromen voor 2000-2001 zien we min of meer dezelfde stroomgrootheden en percentages als voor de periode 1999-2000. De netto groei van de werkgelegenheid is eveneens 2,4%. Wie wat nauwkeuriger kijkt ziet toch enige tekenen van afzwakking van de conjunctuur. De groeiende bedrijven groeien iets minder hard dan in de periode daarvoor (114,3% versus 115%) en de krimpende bedrijven krimpen iets meer (84,7% versus 85% in de vorige periode). Er verdwijnen iets meer banen via de opgeheven bedrijven (85 duizend versus 77,3 duizend banen in het jaar er voor). De groei van het aantal banen in de periode 2000-2001 komt vooral uit de banen die door nieuwe bedrijven wordt gecreëerd (208,7 duizend versus 173,4 duizend banen).

Tabel 4a. Arbeidsmarktstromen 2001-2002, totale aantallen werknemers (x 1000)

	2001	<i>Blijvers</i>	<i>Vertrokken</i>	<i>Naar andere baan</i>	<i>Nieuw</i>	<i>Van andere baan</i>	2002
Neutraal	615,8	533,4	28,4	54,0	30,6	51,8	615,8
Groeiend	3529,8	2960,0	199,5	370,4	391,5	665,3	4016,8
Krimpnd	3393,2	2305,2	363,1	724,9	202,4	325,0	2832,6
Nieuw					67,1	128,1	195,1
Opgeheven	92,6		37,8	54,8			
Totaal	7631,5	5798,5	628,8	1204,2	691,6	1170,2	7660,3

Tabel 4b. Arbeidsmarktstromen 2001-2002, percentages van werkgelegenheid in 2001

	2001	<i>Blijvers</i>	<i>Vertrokken</i>	<i>Naar andere baan</i>	<i>Nieuw</i>	<i>Van andere baan</i>	2002
Neutraal	100,0%	86,6%	4,6%	8,8%	5,0%	8,4%	100,0%
Groeiend	100,0%	83,9%	5,7%	10,5%	11,1%	18,8%	113,8%
Krimpnd	100,0%	67,9%	10,7%	21,4%	6,0%	9,6%	83,5%
Nieuw	0,0%				34,4%	65,7%	100,0%
Opgeheven	100,0%		40,8%	59,2%			0,0%
Totaal	100,0%	76,0%	8,2%	15,8%	9,1%	15,3%	100,4%

In de periode 2001-2002 is er nog nauwelijks een netto groei van de werkgelegenheid (0,4%). De neergaande conjunctuur zet zich door. Bewegingen die reeds in de vorige periode zichtbaar waren zetten zich versterkt door: de groeiende bedrijven groeien minder hard, de krimpende bedrijven krimpen meer en opgeheven bedrijven veroorzaken een nog groter verlies aan banen. Toch komen er nog relatief veel banen bij in nieuwe bedrijven (195,1 duizend banen).

Onze eerste voorzichtige conclusie uit de bovenstaande analyses is dat ‘het wel meevalt’ met de arbeidsmarktdynamiek in Nederland. Op het eerste gezicht lijkt wat we in de bovenstaande tabellen zien niet te wijzen op een versterde sclerotische arbeidsmarkt.

3.2 Creatie en destructie van banen

In de internationale literatuur wordt in navolging van de pioniers Davis, Haltiwanger en Schuh (1998) creatie van banen gedefinieerd als de som van de netto banengroei in de groeiende bedrijven plus de baancreatie van nieuw opgerichte bedrijven. Destructie van banen wordt gedefinieerd als de som van de (netto) banendestructie bij krimpende bedrijven plus het verlies van banen bij opgeheven bedrijven. In de onderstaande tabel brengen we creatie en destructie volgens deze internationale definities in kaart.

Tabel 5. Creatie en destructie van banen, totale aantallen werknemers en percentages van werkgelegenheid (x 1000)

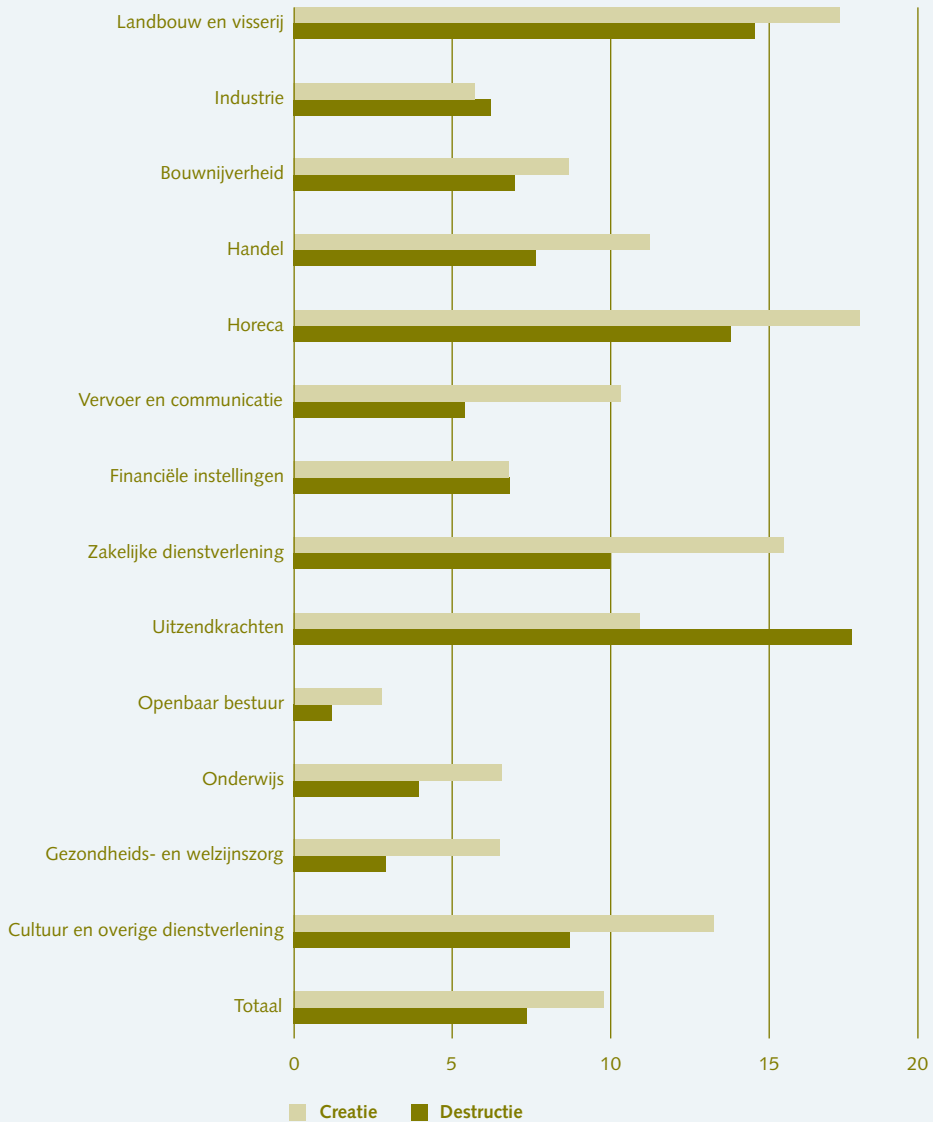
<i>Jaar</i>	<i>Banen</i>	<i>Groeiend</i>	<i>Nieuw</i>	<i>Krimpend</i>	<i>Opgeheven</i>	<i>Creatie</i>	<i>Destructie</i>	<i>Groei</i>
		(1)	(2)	(4)	(5)	(3)=(1)+(2)	(6)=(4)+(5)	(3)+(6)
1999	7282	537	173	-459	-77	710	-536	174
2000	7456	528	208	-476	-85	737	-561	175
2001	7631	487	195	-560	-92	682	-653	28
1999	100%	7,4%	2,4%	-6,3%	-1,1%	9,8%	-7,4%	2,4%
2000	100%	7,1%	2,8%	-6,4%	-1,1%	9,9%	-7,5%	2,4%
2001	100%	6,4%	2,6%	-7,3%	-1,2%	8,9%	-8,6%	0,4%

In 1999 zijn er 7.282.000 banen en in 2002 is het aantal opgelopen tot 7.660.400 banen in 2002. Dat zijn 378.400 banen erbij in drie jaar tijd. In de eerste twee periodes 1999-00 en 2000-01 groeide het aantal banen elk jaar met circa 175 duizend banen. In de periode 2001-02 is de groei al veel minder fors, nauwelijks 30 duizend banen. Die mindere groei van het aantal banen wordt vooral veroorzaakt door de bestaande bedrijven: de groeiende bedrijven groeien minder hard en vooral de krimpende bedrijven stoten veel meer banen af.

In de tabel definiëren we de kolom *creatie* als de som van de netto groei van het aantal banen in de groeiende bedrijven en banen in de nieuwe bedrijven. In de periode 1999-2000 zijn dat ruim 700 duizend extra banen of 9,8% van het aantal banen in 1999. De kolom destructie is de som van de banen die netto verdwenen zijn in de krimpende bedrijven en de weggevallen banen in de opgeheven bedrijven. Totaal is dat meer dan 500 duizend banen of 7,4%. De turnover is nu 17,2% (9,8% plus 7,4%). Dit is vergelijkbaar met wat voor andere landen wordt gevonden.

Een belangrijke vraag is in welke sectoren creatie en destructie vooral plaats vindt: zit de banendynamiek vooral in de industriële sectoren of eerder in de dienstensector? Hoe zit het bij de overheidssector en de quataire sector (zoals zorg en onderwijs)? Sectoren kunnen worden gekenschetst door de mate waarin ze technologisch geavanceerd zijn, matige of hoge productiviteitgroei kennen, bloot staan aan buitenlandse concurrentie (open sectoren) of beschermd worden omdat ze alleen voor de binnenlandse markt werken (beschermd sectoren). Hieronder geven we een overzicht van baancreatie en destructie in 13 geaggregeerde sectoren van de economie.

Figuur 2. Creatie en destructie naar sector (in % van banen in 1999)



De bovenstaande figuur illustreert het relatieve aandeel (percentage van werkgelegenheid in de desbetreffende sector) van de in de periode 1999-2000 gecreëerde of vernietigde banen.

Nationaal wordt gemiddeld 9,8% van de banen gecreëerd en 7,4% van de banen vernietigd (zie tabel 5). Sectoren die daaronder blijven voor wat betreft creatie en destructie zijn dus minder mobiele sectoren. Sectoren die daar boven uitkomen zijn zeer mobiel¹⁵. Sectoren met relatief veel creatie en destructie (beiden >10%) zijn uitzendbureaus, zakelijke dienstverlening, horeca en landbouw en visserij. Een niet zo verbazingwekkende lijst van sectoren. Sectoren waar de creatie groter is dan de destructie zijn de quartaire sectoren: cultuur, gezondheid en onderwijs (in deze sectoren werd door Paars II flink geïnvesteerd in deze jaren) en de dienstensectoren: zakelijke dienstverlening, vervoer en handel.

4 Dynamiek van personen

Bij het analyseren van de dynamiek van personen in het SSB bestand zijn we geïnteresseerd in de statusveranderingen van personen en daarbij in het bijzonder in:

- De stromen van werk naar uitkeringen en omgekeerd. Baancreatie en destructie gaat gepaard met aanname en vrijwillig of gedwongen vertrek (of beëindiging contract) van werknemers. We willen weten waar de ontslagen werknemers naar toe stromen: naar een andere baan of naar een uitkeringssituatie of naar buiten de arbeidsmarkt (studeren, inactiviteit). En we willen weten waar de werknemers die een gecreëerde baan bezetten vandaan komen: uit een andere baan of van buiten de arbeidsmarkt.
- De baanwisselaars: voor een belangrijk deel gaat creatie en destructie gepaard met het wisselen van baan (reallocatie van arbeid). Tussen welke sectoren lopen die stromen?

15) Op het oog lijkt het alsof het relatieve belang van creatie en destructie per bedrijfstak sterk gecorreleerd is. Creatie en destructie zijn of samen hoog of samen laag.

4.1 Statusveranderingen

Tabel 6. Statusverdeling 1999

	x 1000	%
Totaal	16108,7	100,0%
Werknemers	6217,9	38,6%
Zelfstandigen	620,4	3,9%
Uitkering	1257,8	7,8%
Pensioen	2557,5	15,9%
Scholier/student	3168,2	19,7%
Inactief	2286,9	14,2%

In het SSB bestand kunnen personen zich in een aantal mogelijke statussen bevinden zoals aangegeven in tabel 6 (voor 1999). De mogelijkheden zijn: werknemer, zelfstandige, uitkeringstrekker, gepensioneerd, scholier of student en inactief. Het betreft hier een uitputtende opdeling voor de totale Nederlandse bevolking (van meer dan 16 miljoen personen)¹⁶.

Personen veranderen op bepaalde momenten van status. In wat bekend staat als de stromen en toestandenbenadering van de arbeidsmarkt wordt de verandering van status gemeten als een stroom in en uit een bepaalde arbeidsmarktstatus. In de onderstaande tabel geven we deze stromen voor de periode 1999 – 2000. De stroomgegevens worden afgeleid door de status in het jaar 1999 te vergelijken met de status in 2000. Dat is qua periode vrij lang. Belangrijke stromen zoals die tussen werk en werkloosheid en vice versa hebben een veel hogere frequentie. De stromen in de tabel hieronder zijn hoogstwaarschijnlijk een onderschatting van de werkelijke stromen op de arbeidsmarkt. Desondanks zijn ze nog vrij omvangrijk.

Tabel 7. Statusveranderingen 1999-2000

Absoluut (x 1.000)	Naar (2000)							Totaal
	Van (1999)	Werknemers	Zelfstandigen	AO	WW	Pensioen	Onderwijs	
Werknemers	5998,55	38,98	25,93	85,87	62,41	50,84	218,51	6481,09
Zelfstandigen	33,46	555,39	4,33	0,85	10,24	0,35	22,99	627,61
AO	11,16	3,34	430,28	4,24	23,76	0,29	12,43	485,50
WW	26,30	2,33	1,76	61,15	6,52	0,12	17,66	115,84
Pensioen	3,60	3,77	0,85	0,78	2557,21	0,00	4,91	2571,12
Onderwijs	176,02	1,88	3,37	0,55	0,00	2999,48	34,51	3215,81
Inactief	187,84	17,81	19,14	21,58	73,87	215,69	2193,26	2729,19
Totaal	6436,93	623,50	485,66	175,02	2734,01	3266,77	2504,27	16226,16

AO: arbeidsongeschiktheidsuitkeringen
Bron: eigen berekeningen op basis van SSB

16) Personen die in meer dan een categorie vallen zijn toegedeeld onder de categorie waar de belangrijkste inkomensbron onder valt.

Relatief (obv herkomst = 100%)	Naar (2000)								
	<i>Van (1999)</i>	<i>Werknemers</i>	<i>Zelfstandigen</i>	<i>AO</i>	<i>WW</i>	<i>Pensioen</i>	<i>Onderwijs</i>	<i>Inactief</i>	<i>Totaal</i>
Werknemers		93%	1%	0%	1%	1%	1%	3%	100%
Zelfstandigen		5%	88%	1%	0%	2%	0%	4%	100%
AO		2%	1%	89%	1%	5%	0%	3%	100%
WW		23%	2%	2%	53%	6%	0%	15%	100%
Pensioen		0%	0%	0%	0%	99%	0%	0%	100%
Onderwijs		5%	0%	0%	0%	0%	93%	1%	100%
Inactief		7%	1%	1%	1%	3%	8%	80%	100%

AO: arbeidsongeschiktheidsuitkeringen

Bron: eigen berekeningen op basis van SSB

4.2 Baanwisselingen

In deze paragraaf kijken we apart naar de baanwisselaars. Dat zijn werknemers die zowel in het ene jaar als het andere jaar in de status ‘werkend’ verkeren, maar dan wel bij een andere werkgever. We beginnen met een internationale vergelijking.

Tabel 8. Baanwisselaars (in %)

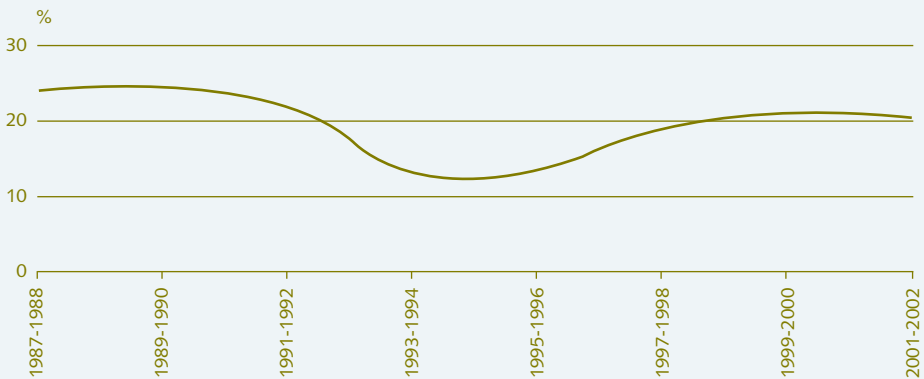
<i>Baanwisselingen</i>	<i>2001</i>
Verenigd Koninkrijk	24,0
Ierland	23,4
Spanje	21,8
Denemarken	21,0
Nederland	20,1
Finland	20,0
Portugal	13,3
Oostenrijk	11,2
België	10,6
Duitsland	9,3
Italië	8,8

Bron: Belot en Van Ours 2004

De eerste conclusie die uit deze internationaal vergelijkende tabel volgt is dat werknemer-reallocatie in Nederland nog relatief hoog is. Nederland zit in de groep met reallocatie van 20 procent en meer.

Uit de gegevens van het OSA arbeidsaanbodpanel, die ook gebruikt zijn door het SCP(2005), blijkt dat de werknemersmobiliteit een cyclisch patroon laat zien maar zeker geen tekenen vertoont van toename in dynamiek over de tijd (figuur 3).

Figuur 3. Werknemersmobiliteit

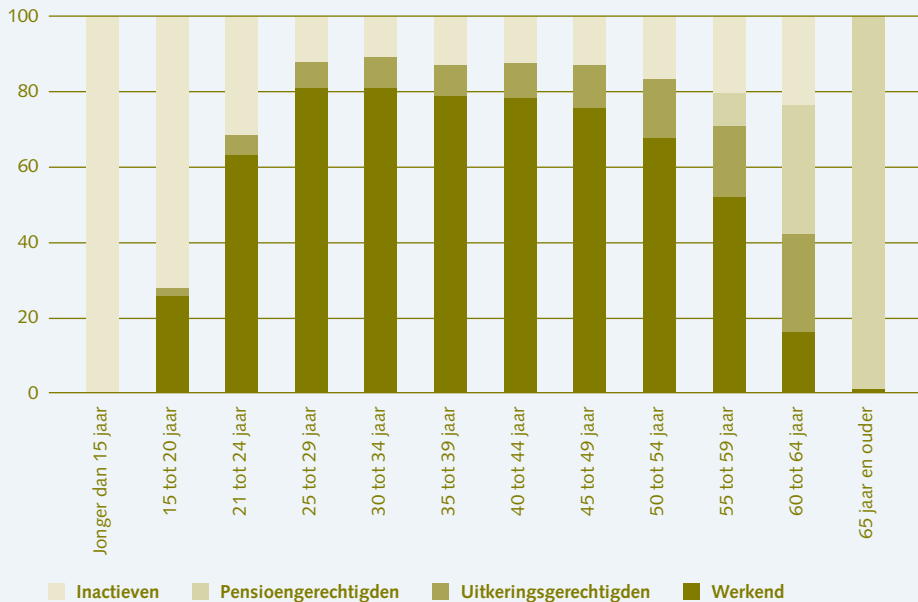


Bron: SCP (2005) en OSA (Arbeidsaanbodpanel 1988-2002)

4.3 Dynamiek van oudere werknemers

Met het oog op de aankomende vergrijzing en daarmee ook de opkomst van de oudere werknemers in het personeelsbestand van de meeste bedrijven en instellingen, is het nuttig apart te kijken naar de arbeidsmarktdynamiek van oudere werknemers. We beginnen met een verdeling over de verschillende arbeidsmarkttoestanden naar leeftijd in figuur 4.

Figuur 4. Arbeidsmarktstatus naar leeftijd, 2001



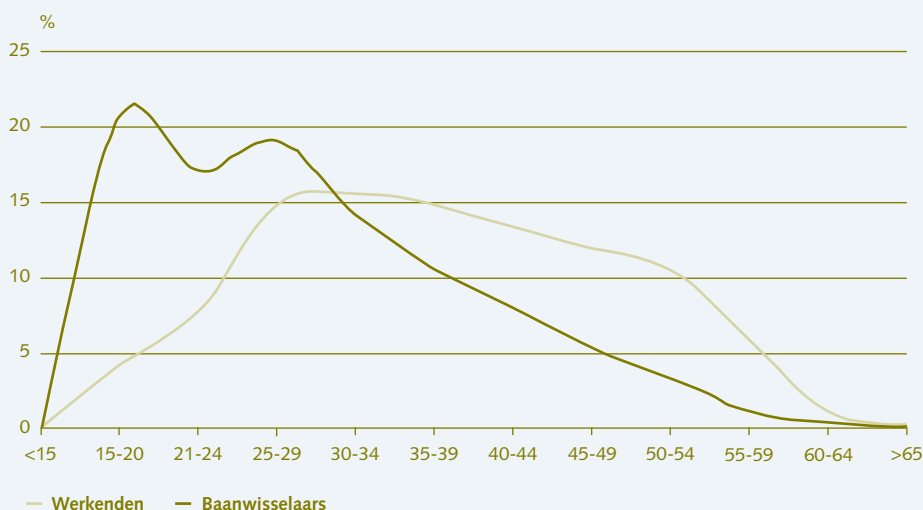
Bron: eigen berekeningen op basis van het SSB

Jongeren tot begin twintig behoren meestal –bekeken vanuit de arbeidsmarkt – tot de groep inactieven. Ze zitten op school of studeren. Halverwege de twintig treden ze toe tot de arbeidsmarkt en dan wordt ‘werkend’ de dominante status tot ongeveer de leeftijdsgroep van 55 jaar en ouder. Dan volgt een snelle toename van de toestand ‘pensioen- en uitkeringsgerechtigd’. Na leeftijd 65 heeft, op een enkele uitzondering na, iedereen een pensioen.

Figuur 5 toont het verschil tussen baanwisselaars en werkenden naar leeftijd (relatieve aandeel van het aantal werkenden en het aantal baanwisselaars per leeftijdsgroep).

Uit figuur 5 blijkt duidelijk dat baanwisseling vooral iets is voor de jongere leeftijdsgroepen en dat werken dominant is voor middelbare leeftijd. Hogere leeftijdsgroepen worden door baanwisseling noch hoog percentage werkenden (bij werkenden zakt het pas significant vanaf 50-54 jaar) gekenschetst.

Figuur 5. Verdeling van werkenden en baanwisselaars naar leeftijd (1999-2000)



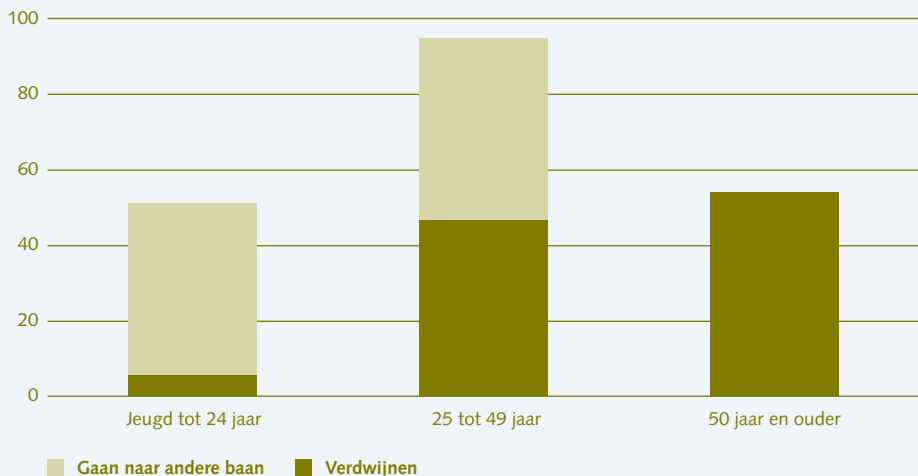
Bron: eigen berekeningen op basis van het SSB

Baancreatie en –destructie brengen met zich mee dat werknemers een baan vinden en andere werknemers hun baan verliezen. Het effect van baancreatie en –destructie pakt totaal anders uit voor de verschillende leeftijdsgroepen op de arbeidsmarkt. Figuur 6 laat de leeftijdseffecten zien van baandestructie voor jongeren (tot 24 jaar), middelbare leeftijd (24 – 49) en oudere werknemers (50+). Figuur 7 toont de leeftijdseffecten van baancreatie.

Werknemers wier baan wordt opgeheven hebben twee mogelijke bestemming: of ze stappen over naar een andere baan of ze komen in een uitkeringssituatie terecht (gemakshalve noemen we dit ‘verdwijnen van de arbeidsmarkt’). De vernietigde banen worden bezet door personen. Van die personen is – aan de hand van het SSB – de leeftijd bepaald én hun uiteindelijke arbeidsmarktstatus: andere baan of buiten de arbeidsmarkt. Van die twee bestemmingen is de (procentuele) verdeling naar leeftijd berekend. In grafiek 6 zijn vervolgens per leeftijdscategorie de twee verdelingen gecumuleerd. In grafiek 7 is eenzelfde – maar spiegelbeeldige – benadering toegepast aan de hand van toetreden tot de arbeidsmarkt dan wel komend van een andere baan.

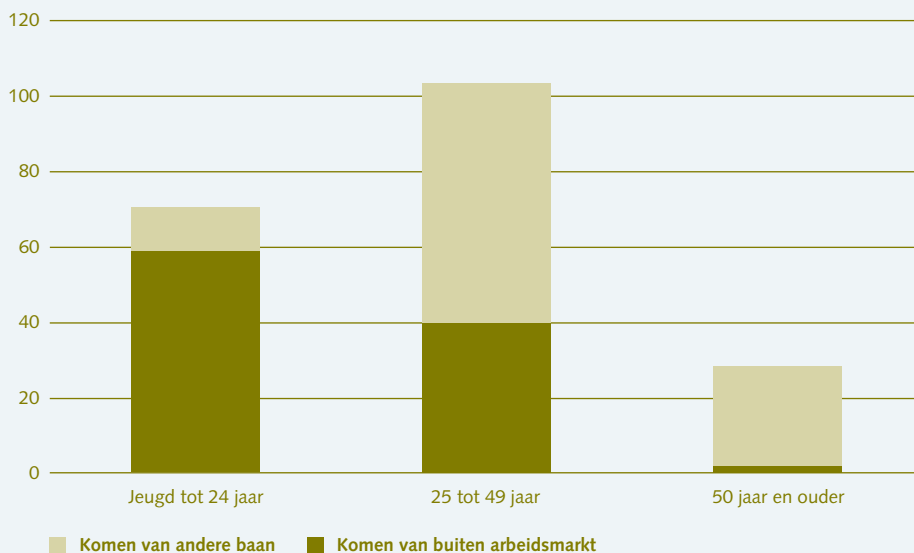
Figuur 6 is zo geconstrueerd dat de twee bestemmingen (‘andere baan’ en ‘verdwijnen van de arbeidsmarkt’) over de drie leeftijdsgroepen elk optellen tot 100%. Wat onmiddellijk opvalt, is het extreme verschil tussen de jongste en de oudste leeftijdsgroep. Jongeren die hun baan kwijtraken stappen voor de overgrote meerderheid over naar een andere baan, terwijl de oudste leeftijdsgroep grotendeels van de arbeidsmarkt verdwijnt. Voor de middelbare leeftijdsgroep zijn de beide bestemmingen meer gelijkmatig verdeeld.

Figuur 6. Baandestructie en de bestemming van overbodige werknemers



Bron: eigen berekeningen op basis van het SSB

Figuur 7. Baancreatie en de afkomst van de nieuwe werknemers



Bron: eigen berekeningen op basis van het SSB

Figuur 7 is het spiegelbeeld van de vorige figuur. Wanneer de nieuwe banen worden opgevuld door jeugdige werknemers dan zijn die voor het grootste deel afkomstig van buiten de arbeidsmarkt (schoolverlaters bijvoorbeeld). Bij de oudste leeftijdsgroep komt er nauwelijks iemand van buiten de arbeidsmarkt. Oudere werknemers zijn minder dan de andere leeftijdsgroepen betrokken bij het opvullen van de nieuwe banen. In de middelbare leeftijdscategorie komen de twee afkomstmogelijkheden voor, zij het dat ‘komende van een andere baan’ wat vaker te zien is dan komende van buiten de arbeidsmarkt (bijvoorbeeld herintreders).

4.4 Mobiliteit tussen sectoren

Werknemers bouwen door ervaring en door opleiding kennis op. Hierdoor kunnen ze moeilijker veranderen; door een verandering kan de investering in kennis teniet gedaan worden. Zo kunnen werknemers moeilijk de ene sector voor de andere verruilen, als kennis gebonden is aan een sector. In hoeverre dit het geval, is de vraag. Mogelijk is veranderen van functie of van beroep moeilijker dan veranderen van sector. Er zijn veel functies die in verschillende sectoren uitgeoefend kunnen worden.

Zoals blijkt uit tabel 9 veranderen per jaar meer dan een miljoen werknemers van baan. Voor deze baanwisselaars kunnen we vervolgens nagaan of hun nieuwe baan in dezelfde sector (zelfde sbi code) is als hun oude baan, of dat hun nieuwe baan in een andere sec-

tor (en dus andere sbi code) ligt. Een zeer opvallend resultaat daarbij is dat de overgrote meerderheid van de baanwisselaars (rond de 70%) in een andere sector terecht komt. De rol van sectorspecifieke kennis blijkt niet heel groot te zijn. Een mogelijke verklaring voor dit bijzondere resultaat is dat deelmarkten eerder via beroepslijnen lopen dan via sectorlijnen. Wanneer secretaresses of accountants een nieuwe baan zoeken dan zoeken ze een nieuwe baan in hun beroep en dat hoeft niet per se in dezelfde sector te zijn. Dit empirische resultaat werpt ook een ander licht op de afspraken van werkgevers en werknemers (bijvoorbeeld in cao's) die vaak op de sector zijn afgestemd.

Tabel 9. Baanwisselingen binnen en tussen sectoren

	Aantallen (x 1000)		Percentage	
	<i>Zelfde sector</i>	<i>Andere sector</i>	<i>Zelfde sector</i>	<i>Andere sector</i>
1999-2000	338,1	791,1	30%	70%
2000-2001	334,5	784,4	30%	70%
2001-2002	332,7	712,8	32%	68%

Bron: eigen berekeningen op basis van het SSB

5 Conclusies

We komen tot de volgende conclusies wat betreft de vragen die we aan de SSB data hebben gesteld.

- Zijn er aanwijzingen dat de Nederlandse arbeidsmarkt minder dynamisch is dan andere arbeidsmarkten?

Gegeven onze empirische analyse in deze bijdrage lijkt, kunnen we stellen dat qua dynamiek de Nederlandse arbeidsmarkt vergelijkbaar is met de andere ontwikkelde landen.

- Zijn er aanwijzingen dat de dynamiek op de arbeidsmarkt recentelijk is toegenomen, onder invloed van toegenomen internationale concurrentie, technologische ontwikkeling, opkomst van China en India en andere externe schokken?

Onze conclusie is dat er geen aanwijzingen zijn in de SSB data (al dan niet gecombineerd met andere gegevens) dat er een toename van turbulentie op de arbeidsmarkt te bespeuren valt.

- Is de impact van de arbeidsmarktdynamiek verschillend voor de boven- en de onderkant van de arbeidsmarkt? Is er wat betreft arbeidsmarktdynamiek sprake van insiders en outsiders waarbij outsiders meer de ‘last van de dynamiek’ dragen dan de insiders?

Het is niet mogelijk om op basis van het bovenstaande conclusies te trekken over een ongelijke verdeling van de ‘last van de arbeidsmarktdynamiek’. De enige aanwijzing in dit verband is het groter aantal baanwisseling bij jongeren en een zekere stabiliteit bij werknemers van middelbare en oudere leeftijd. Waarbij ook nog niet duidelijk is of de baanwisseling vrijwillig of onvrijwillig is. Hier is verder onderzoek noodzakelijk, dat ook deels met de gegevens van het SSB kan gebeuren.

- Wat is in het bijzonder de positie van de oudere werknemers in de dynamiek van de Nederlandse arbeidsmarkt?

We zien een vermindering van arbeidsmarktdynamiek naarmate werknemers ouder worden. De analyses over welke leeftijdsgroepen vooral betrokken zijn bij job creatie en destructie laten zien dat creatie vooral iets is waar jongere en middelbare leeftijdsgroepen van profiteren en dat bij destructie de ouderen vooral de arbeidsmarkt verlaten. Deze resultaten zijn belangrijk voor de volgende decennia wanneer het zwaartepunt van de werkende bevolking bij de oudere werknemers komt te liggen.

Ten slotte nog een aantal opmerkingen over de voor- en nadelen van het werken met SSB data bij de analyse van de arbeidsmarktdynamiek.

Een zeer belangrijk voordeel van de SSB data is het integrale karakter van de gegevens. Plus dat de gegevens kunnen worden georganiseerd zowel vanuit het standpunt van personen (waardoor de mobiliteit en veranderingen van personen kan worden gemeten) als vanuit het standpunt van bedrijven (waardoor bedrijven en banendynamiek kan worden beschreven). Daarmee kan als het ware ‘alle’ dynamiek op de arbeidsmarkt in kaart worden gebracht. Het is dus mogelijk om met behulp van de SSB data veel dieper te graven in de dynamiek dan voorheen met dwarsdoorsnede data mogelijk was.

Een tweede belangrijk voordeel is dat personen en bedrijven in de loop van de tijd kunnen worden gevolgd. Op dit moment is de loop van de tijd weliswaar nog beperkt tot de periode 1999-2003. Die periode wordt in de toekomst steeds langer, maar dat belet niet

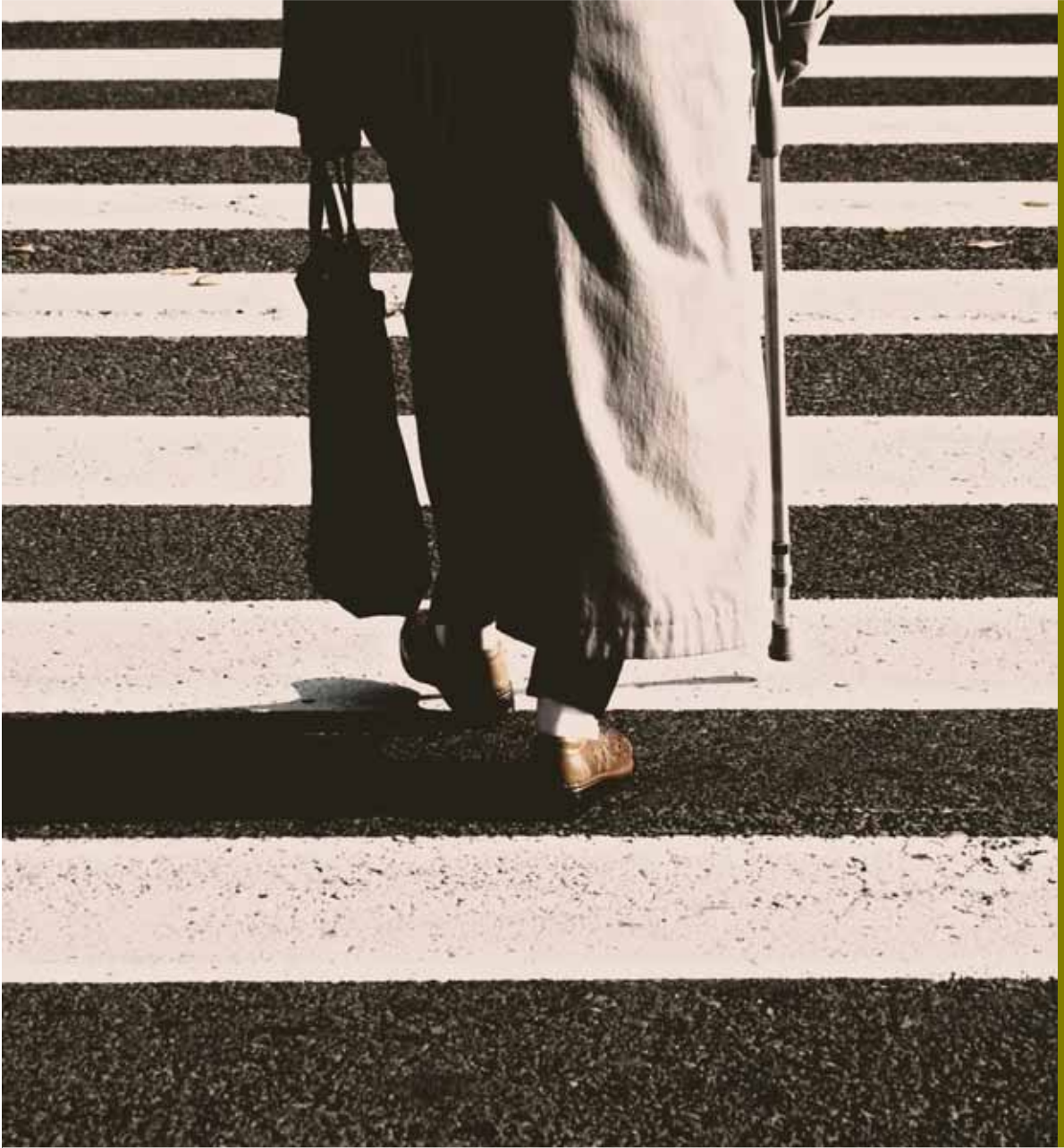
dat ook nu reeds interessante dynamische ontwikkelingen bij dezelfde persoon of bedrijf kunnen worden bestudeerd. Bijvoorbeeld de eerste arbeidsmarktervaringen van een schoolverlater kunnen worden bestudeerd door de schoolverlaters die in 1999 de arbeidsmarkt instroomden te volgen tot 2003. Ook het succes of het falen van een in 1999 opgericht bedrijf op deze manier worden gevolgd in de eerste jaren van het bestaan.

Bij onze analyse liepen we tegen twee nadelen aan (die in principe oplosbaar zijn). Een belangrijke variabele in de arbeidsmarktstudie is de hoogst behaalde opleiding. Die ontbreekt momenteel nog maar kan natuurlijk via koppelingen met bestanden waarin de opleiding is opgenomen op den duur worden meegenomen. Zie het hoofdstuk Bakker, Bouman en van Toor in deze bundel over 'Opleidingsniveau uit registers: nieuwe bronnen, maar nog niet compleet'.

Een tweede lacune waar we tegen aanliepen is dat voor sommige arbeidsmarkttransities de jaarlijkse frequentie te laag is. Denk bijvoorbeeld aan werkloosheid. De duur van de werkloosheid is voor de meeste werklozen korter dan een aantal maanden. Dus als iemand op zowel de eerste als de tweede peildatum in het werklozenbestand zit, dan is dat ofwel een langdurige werkloze (en behoort hij tot een aparte probleemgroep op de arbeidsmarkt) of het is ook mogelijk dat hij in de tussenliggende periode reeds werk heeft gevonden en ook daar weer werkloos is geworden (en daarmee behoort hij misschien tot wederom een aparte groep van draaideur werklozen). Wat we niet meten is de normale werkloosheidservaring. Dit kan verholpen worden door peildata van minder dan een jaar in te voeren. Hoe dan ook, het is onze ervaring dat het beschikbaar komen van de SSB data een belangrijke impuls kan geven aan het empirisch arbeidsmarktonderzoek en aan een beter begrip van de dynamische processen die zich in de economie afspeelen.

Referenties

- Bean, C. (1994), "European Unemployment: A Survey", *Journal of Economic Literature*, June.
- Belot, M. en J. van Ours (2004), *Arbeidsmarktinstuties en arbeidsmarktdynamiek*, preadvies ten behoeve van het WRR project 'Economische Structuur', mimeo, december.
- Blanchard, O. (2005), *European Unemployment: The evolution of facts and ideas*, NBER Working Paper 11750, November.
- Davis, S.J., J.C. Haltiwanger en S. Schuh, (1998), *Job Creation and Destruction*, MIT Press, Cambridge.
- Kee, P. en R. Milor (2005), *Arbeidsmarktdynamiek en banen*, De Nederlandse conjunctuur, aflevering 4, CBS, Voorburg en Heerlen.
- Leeuwe, P. van (2006), *Keeping Older Workers in the Workforce. An Assessment of Recent Policies in the Netherlands*, Lecture in the International Seminar for Experts "Labour Market Reform in Europe – Learning from the Best Practices, Cicero Foundation series Great Debates, www.cicerofoundation.org, Parijs, 23-24 februari 2006.
- OSA, *Het OSA-Arbeidsaanbodpanel 1985-2004*, zie website <http://www.uvt.nl/osa/data/> Universiteit van Tilburg.
- SCP (2005), *Arbeidsmobiliteit in goede banen*, publicatie nr 2005/12, juni.





Peter Ester (OSA/UvT) & Marcel Kerkhofs (OSA)

Knap maar karig.
Naar een wetenschappelijk
en beleidsmatig betekenisvol
Sociaal Statistisch Bestand.*

* Reactie op: 'Arbeidsmarktdynamiek en vergrijzing'.

1 Inleiding

Om met de deur in huis te vallen: wij zijn enthousiast over het SSB. Met de ontwikkeling van het SSB laat het CBS zien dat het in staat is om op belangrijke onderdelen van haar kernactiviteiten stevig te innoveren en de primaire processen van grootschalige dataverzameling en data-analyse in ons land een paar ferme stappen vooruit te helpen. Hulde aan het CBS derhalve. Tegelijkertijd zijn wij van mening dat het CBS het ambitieniveau (veel) verder zou moeten oprekken. In de huidige vorm is het SSB te karig: het geeft wel zicht op basale ontwikkelingen die zich op bepaalde beleidsterreinen in de Nederlandse samenleving voordoen, maar het bevat niet of nauwelijks inhoudelijke indicatoren om deze ontwikkelingen ook betekenisvol te verklaren, te ontleden en te duiden. Het SSB blijft daarmee in zijn eigen adolescentie steken en geraakt niet tot volle wasdom. Hier is evenwel een goede remedie voorhanden: investeer vooral ook in de koppeling van primair administratieve data zoals verzameld in het SSB *en* meer subjectieve of evaluatieve data zoals verzameld in tal van Nederlandse bevolkingssurveys (binnen maar vooral ook buiten het CBS). Dit naar ons idee cruciale punt laat zich goed illustreren aan de hand van een bespreking van de bijdrage van Van Leeuwe *et al.* in deze bundel over arbeidsmarktdynamiek en vergrijzing. Dit vormt de kern van het betoog van ons artikel. We sluiten dit korte referaat af met niet minder dan een pleidooi voor een ‘grote coalitie’ tussen het CBS en andere kennisinstituten in het land, met als inzet het huidige SSB aanmerkelijk te verrijken door het koppelen van administratieve data en surveygegevens.

2 Wat het SSB vermag

De twee grootste voordelen van het SSB zijn zonder twijfel het integrale en longitudinale karakter ervan. Zeker indien we de beschikking krijgen over langere tijdreeksen, zal er sprake zijn van een schat aan materiaal. Dit is primair een kwestie van tijd: kenmerk van het SSB is dat het zichzelf als het ware structureel vult en verrijkt door de opname van nieuwe peildata en nieuwe modules. Een ongekend voordeel van het SSB waarop Van Leeuwe en zijn collega's konden rekenen, is de koppeling van personen aan de bedrijven waar deze personen werkzaam zijn. Het behoeft weinig fantasie om de verstrekkende voordelen van een dergelijke koppeling – droom van iedere onderzoeker – in beeld te krijgen. Op deze wijze zijn we in staat om de dynamiek van de stromen op de arbeidsmarkt, zowel in termen van personen als banen, veel preciezer in kaart te brengen. In theorie biedt dit

prachtige analysemogelijkheden die de studie van bewegingen op de arbeidsmarkt een belangrijke opwaartse stimulans kunnen bieden. Maar zijn de SSB-data die geanalyseerd worden ook *inhoudelijk* integraal? Het antwoord op deze fundamentele vraag is naar onze mening doorslaggevend voor de uiteindelijke wetenschappelijke en beleidsmatige relevantie van het SSB.

Drie punten willen we in het kort aanroeren. Allereerst is het van belang te beseffen dat het doel van de registratie bepaalt wat we met administratieve data meten. Zo kunnen werknemers nadat hun dienstverband is beëindigd nog enige tijd in de financiële administratie van een bedrijf aanwezig blijven; of kun je als zelfstandige onder de radar van de belastingdienst doorglijden als in een jaar geen inkomen uit winst is aangegeven. Een tweede punt dat voor de studie van dynamiek relevant is, is dat het in de tot nu toe beschikbaar gekomen SSB-bestanden weliswaar mogelijk is de mensen door de tijd te volgen, maar niet continu. Veel van de informatie in deze bestanden heeft namelijk betrekking op de stand van zaken op een aantal opeenvolgende peildata. Hierdoor is het alsof je van een film slechts een reeks stilstaande beelden te zien krijgt. Zelfs als dat veel beelden zijn, blijft er de kans dat je juist de actiescènes mist. De eerlijkheid gebiedt te zeggen dat het in de opzet van het SSB in principe mogelijk is dit bezwaar weg te nemen door de frequentie van de peildata fors te vergroten. Maar of dit ook in praktische zin een begaanbare weg is, moet nog blijken. Ten derde is er veel interessants aan dynamiek en mobiliteit dat zich aan registratie onttrekt. Interne mobiliteit, creatie van nieuwe of het opheffen van bestaande functies, de redenen of motieven van een transitie. Daarmee wil niets ten nadele van het SSB gezegd zijn. De toegevoegde waarde en de kwaliteit daarvan staan buiten kijf. Maar deze punten maken wel duidelijk dat er naast een prachtbestand als het SSB altijd behoefte zal zijn aan andere gegevens, bijvoorbeeld zoals die door middel van surveys worden verzameld.

3 Dynamiek op de arbeidsmarkt

Wat het SSB allemaal wel mogelijk maakt, wordt mooi geïllustreerd door het onderzoek van Van Leeuwe et al. Deze analyse op basis van SSB-gegevens geeft een informatief en helder beeld van de dynamiek op de arbeidsmarkt, trends in arbeidsstromen, trends in groeiende, krimpende en neutrale bedrijven, trends in baancreatie en baandestructie. Een prikkelend resultaat is de zeer hoge intersectorale mobiliteit. Maar liefst 70% van de baanwisselingen is grensoverschrijdend in de zin dat de werknemer in een andere sector belandt

dan waaruit hij of zij vertrok. Heeft het sector-denken daarmee zijn langste tijd gehad? Alle reden om op dat punt verder te reageren. Denkbaar is dat dit cijfer belangrijk lager zal uitkomen als bijbanen van scholieren en studenten en uitzendwerk buiten de analyse worden gelaten.

Het hoofdstuk biedt evenzeer een beeld van de beperkingen van het SSB. Immers, met de beschikbare administratieve data kan niet goed worden beantwoord *waarom* sommige mensen van baan wisselen en anderen juist niet, *waarom* bedrijven groeien of krimpen, *waarom* baancreatie of baandestructie plaatsvindt. Het zijn deze en vergelijkbare ‘waarom-vragen’ die uiteindelijk de kern vormen van de wetenschappelijke en beleidsmatige reflectie op en analyse van de dynamiek op de arbeidsmarkt, vragen waarmee het (huidige) SSB maar moeilijk uit de voeten kan. De hiervoor benodigde informatie over attitudes, wensen, motieven en verwachtingen blijft in administratieve bestanden buiten beeld. Hoe we het ook draaien of keren: voor de beantwoording van deze wezenlijke vragen zijn surveys onder bedrijven en werknemers onontbeerlijk. Indien we willen weten waarom en hoe bedrijven en werknemers reageren op arbeidsmarktstimuli – vaak letterlijk de beweegredenen – dan zullen we hen daarnaar moeten vragen.

4 Mobiliteit van jongeren en ouderen: leeftijd, periode of cohort?

Van Leeuwe en zijn collega’s vinden dat ouderen minder flexibel zijn dan jongeren en menen dat dit in een vergrijzende samenleving en dynamische kenniseconomie “mogelijk een zorgelijke ontwikkeling” is. We denken dat dit iets complexer ligt. Immers, jongeren kunnen mobieler zijn omdat ze pas aan het begin van hun loopbaan staan, volop experimenteren met loopbaankeuzes en ook geacht worden mobieler te zijn (Bekker *et al.*, 2005; Vinken *et al.*, 2002). Bovendien mag verwacht worden dat als het aandeel ouderen op de arbeidsmarkt substantieel toeneemt, er ook meer mobiliteit onder ouderen zal zijn: immers, mobiliteit is ook een kwestie van vraag en aanbod. Minstens zo belangrijk is dat de werkende ouderen van nu of gisteren niet de werkende ouderen van morgen zullen zijn. Immers, verschillende generaties van ouderen zullen verschillen in termen van arbeids-ethos, arbeidsattitudes, opleiding, vaardigheden, loopbaanverwachtingen en levensloop-oriëntaties. Daarbij stelt zich de fundamentele vraag of deze verschillen te duiden zijn als een leeftijdseffect (werknemers worden ouderen en daarom veranderen hun denkbeelden), een periode-effect (de impact van bijvoorbeeld economische hoog- of laagconjunctuur), of

een cohort-effect (cohorten vormen generaties met een eigen blik op het domein van de arbeid). Een klassiek probleem overigens in de analyse van sociale veranderingen (vgl. Blossfeld, 1986; Hagens & Cobben, 1978; Riley, 1973). Dergelijke effecten kunnen echter alleen ontrafeld worden met gedegen longitudinaal panelsurvey-onderzoek, tijdreeksen gebaseerd op vooral administratieve data zijn daarvoor niet toereikend.

5 Tot besluit

Hoe kunnen we nu de doorstoot maken naar de door ons als wetenschappelijk en beleidsmatig superieur beschouwde koppeling van administratieve en surveydata? Zeggen dat het CBS gewoon meer moet doen dan het nu voor ogen heeft, is een voor de hand liggende maar ook wat luie optie. Wij zouden veel verder willen gaan en pleiten voor een ‘grote coalitie’ tussen het CBS als hoedster van grootschalige administratieve databestanden en de Nederlandse (academische en para-universitaire) kennisinstituten als initiators van tal van grootschalige surveystudies (vaak ook op panelbasis). Als voorbeelden van deze laatste studies noemen wij slechts de omvattende publieke-opiniesurveys van het Sociaal en Cultureel Planbureau, de langjarige arbeidsmarktpanelen van de OSA, de hoogfrequente panelstudies van CentERdata, de Netherlands Kinship Panel Study o.l.v. het NIDI, de Schoolverlatersenquête van het ROA, maar ook de Nederlandse modules van internationale surveys (zoals de Eurobarometers, het International Social Survey Project, de European Values Study, het SHARE project, het European Social Survey). Het idee is dan dat het CBS en de kennisinstituten hun databestanden inbrengen in een ambitieus kennispartnerschap, het CBS de microkoppeling tussen deze bestanden verricht en de aldus verkregen – sterk verrijkte – bestanden extern benut kunnen worden voor beleidsmatig en wetenschappelijk onderzoek, uiteraard onder strikte privacy- en geheimhoudingsvoorwaarden. DANS (Data Archiving and Networked Services), het nieuwe KNAW/NWO-instituut op het terrein van data-infrastructuur voor de alfa- en gammawetenschappen, zou hierbij goede intermediaire diensten kunnen bewijzen.

Zouden we er door middel van deze grote coalitie in slagen om voor een reeks van onderwerpen omvangrijke gekoppelde (administratieve en survey-) databestanden te construeren, dan zou dit de Nederlandse data-infrastructuur tot een van de beste ter wereld te maken. Onze wat gemankeerde concurrentiepositie op de internationale markt voor Research & Development kunnen we hiermee in één klap structureel verbeteren. En niets stemt de Nederlander gelukkiger dan Nederland blijvend op de wereldkaart te zetten.

Referenties

- Bekker, S., P. Ester & T. Wilthagen (2005). Jong en oud op de arbeidsmarkt. Generaties, transitie en levensloop. Den Haag: Reed Business Information.
- Blossfeld, H.-P. (1986). Career opportunities in the Federal Republic of Germany: A dynamic approach to the study of life-course, cohort, and period effects. *European Sociological Review* 2: 208-225.
- Hagenaars, J.A.P. & N.P. Cobben (1978). Age, cohort and period: A general model for the analysis of social change. *The Netherlands' Journal of Sociology* 14: 59-91.
- Leeuwe, P. van, D. Scheele, J. Theeuwes & T. Woortman (2005). Arbeidsmarktdynamiek en vergrijzing. (deze publicatie)
- Riley, M.W. (1973). Aging and cohort successions: Interpretations and misinterpretations. *Public Opinion Quarterly* 37: 35-49.
- Vinken, H., P. Ester, H. Dekkers & L. van Dun (2002). Aan ons de toekomst. Toekomstverwachtingen van jongeren in Nederland. Assen: Van Gorcum.





Jan Latten (CBS/UvA), Bas Hamers (RPB) en Han Nicolaas (CBS)

De prijs van migratie. Selectieve verhuisstromen van de vier grote steden.

De leeftijden en de etnische achtergrond van degenen die naar de vier grote steden verhuizen verschillen sterk van die van degenen die de steden verlaten. In dit artikel wordt nagegaan in welke mate de verhuisstromen wat dit betreft selectief zijn en of er ook sociaal-economische selectie plaatsvindt. Dan wordt zichtbaar dat het vertrek van de autochtonen in financieel opzicht niet wordt opgevangen met de komst van allochtonen. In kwantitatief demografisch opzicht is dat wel het geval. Het wordt inzichtelijk dat continuering van de selectieve verhuistrends een druk legt op de economische kracht van de stedelijke bevolking.

Inleiding

De ontwikkeling van de vier grote steden baart zorgen. Verpaupering, onveiligheid en ontbrekende sociale cohesie worden sterk geassocieerd met stedelijke gebieden en hun selectieve bevolkingssamenstelling. De bevolkingssamenstelling van de steden is niet doorsnee te noemen. Het aandeel personen met lage inkomens of met een allochtone achtergrond is in de grote steden aanzienlijk. Bovendien vallen deze kenmerken nogal eens samen. Selectiviteit in verhuisstromen kan dergelijke concentraties in stand houden of versterken.

De totale bevolkingssomvang van Amsterdam, Rotterdam, Den Haag en Utrecht is afgelopen jaren maar zeer beperkt gegroeid. De bevolking verandert desondanks ingrijpend qua samenstelling. Dat heeft onder andere te maken met de immigratie vanuit het buitenland, die zich vooral richt op de grote steden. Anderzijds is er sprake van een binnenlandse trek uit de stad naar randgemeenten. Naast de aantallen die verhuizen, zijn ook de demografische en sociaal-economische achtergrondkenmerken van degenen die verhuizen van belang. De vraag is bijvoorbeeld in hoeverre de leeftijden, etnische achtergronden, arbeidsparticipatie en inkomens van degenen die naar de stad toekomen verschillen met die van degenen die vertrekken?

De selectie in migratiestromen, zowel in de buitenlandse als binnenlandse migratie, draagt er toe bij dat er per saldo sprake is van een uittocht van de autochtone bevolking uit de grote steden (Latten e.a., 2004). Hun plaats wordt grotendeels door allochtone immigranten ingenomen. Sinds een aantal decennia richt een internationale immigratie van niet-westerse bevolkingsgroepen zich vooral op de grote steden. De concentratie van niet-westerse allochtonen in Amsterdam, Den Haag, Rotterdam en Utrecht nam daardoor sterk toe. Dit ondanks een in de laatste jaren op gang komende verhuisstroom van niet-westerse immigranten naar randgemeenten van de grote steden. Instromers hebben daarnaast overwegend een zwakke sociaal-economische positie. Deze instromers zijn vaak niet in staat direct aansluiting te vinden op de bestaande arbeidsmarkt (Sprangers e.a., 2004). Bovendien, als de banen zich in de stad bevinden, maar degenen die een baan krijgen buiten de stad gaan wonen, hoeft groei van stedelijke bedrijvigheid of werkgelegenheid niet samen te vallen met verbetering van de sociaal-economische positie van de stedelijke bevolking. In deze bijdrage worden demografische verhuisstromen naar en uit de vier grote steden economisch getypeerd.

17) In het Sociaal Statistisch Bestand zijn economische variabelen behorend bij de demografische verhuisstromen beschikbaar vanaf 1999. De hier gebruikte data beslaan de periode tot 2003. De verhuisstromen hebben telkens betrekking op een

verhuizing tussen twee peilmomenten. Daarvoor is in het Sociaal Statistisch Bestand de laatste vrijdag in september als peilmoment gekozen. Zowel de uitstroom als de instroom van de vier grote steden betreft telkens een verhuizing die zich tussen die

peilmomenten voordoet. De inkomensgegevens betreffen het jaarincome dat hoort bij het eerste peilmoment. In de periode 1999 tot en met 2003 hebben er in Den Haag en Utrecht grenscorrecties plaatsgevon-

Het Ruimtelijk Planbureau (De Vries, 2005) constateert als manco in de informatievoorziening dat de samenhang tussen inkomensontwikkelingen in de steden en verhuispatronen niet bekend is. Bij gebrek aan data wordt het beeld meestal weergegeven op basis van inkomens van de woonachtige bevolking. Met het beschikbaar komen van het Sociaal Statistisch Bestand van het Centraal Bureau voor de Statistiek is het voor het eerst mogelijk individuele ruimtelijke mobiliteit te koppelen aan inkomensgegevens en andere persoonlijke kenmerken, zoals etnische achtergrond of levensfase¹⁷. Vanuit dat perspectief wordt hier nagegaan hoe personen die naar of uit de grote stad verhuizen, kunnen worden getypeerd naar levensfase, etnische achtergrond en inkomen.

De economische variabele die in dit onderzoek wordt gehanteerd is een individueel fiscaal jaarinkomen. Dit is gebaseerd op het totaal aan inkomenscomponenten waarover loonbelasting wordt ingehouden. Hiertoe horen zowel inkomen uit arbeid, inkomsten uit onderneming als inkomen uit uitkeringen. Het betreft echter niet inkomen uit vermogen. Bij iedere persoon die in de periode september 1999 tot en met september 2003 is verhuisd, is nagegaan wat hij of zij aan jaarinkomen had omstreeks het moment van verhuizing (daarbij is uitgegaan van het jaarinkomen bij het laatste peilmoment dat men nog in de stad woonde). Ook personen zonder zo'n inkomen, zoals studenten, huisvrouwen en kinderen tellen in de berekeningen mee. Zij typeren immers ook de verhuisstromen. Het totale inkomenssaldo van de verhuisstromen bepaalt de winst- of verliesrekening van de verhuismobiliteit.

2 Selectiviteit in verhuisstromen

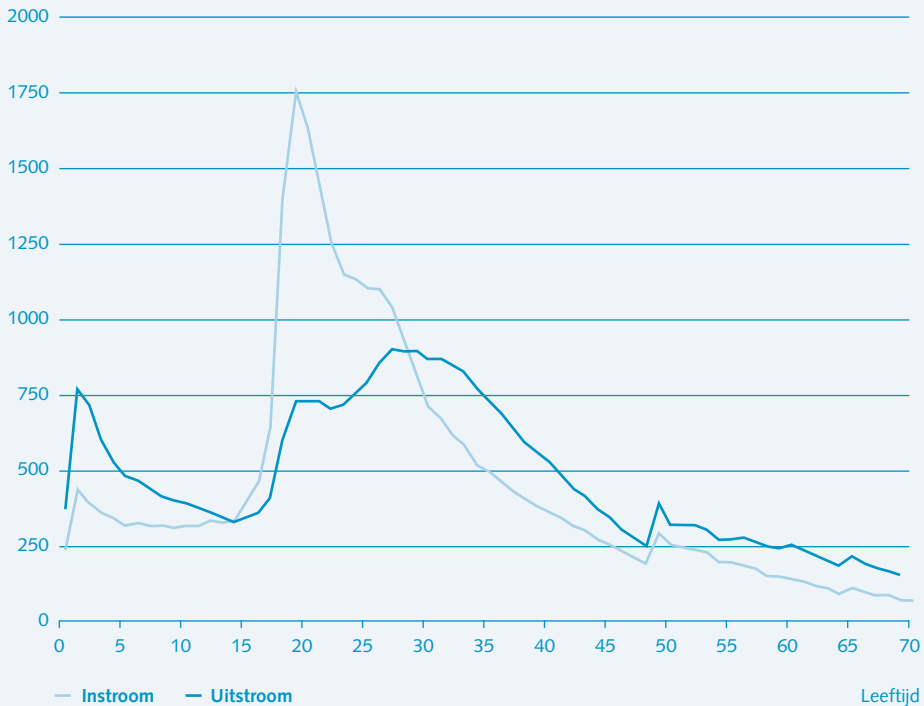
2.1 Selectiviteit naar levensfase

De verhuismobiliteit verschilt per levensfase. Dat blijkt uit figuur 1. Het gaat om de totale in- en uitstroom voor de periode 1995 tot en met 2003 (daarmee is een toevallig effect van een bepaald meetjaar uitgeschakeld). Door, per leeftijd, het aantal personen dat verhuist te relateren aan het aantal inwoners, wordt vertekening door leeftijdsopbouw vermeden. Als referentie geldt het aantal inwoners per leeftijd in 1995. Door de berekeningswijze kunnen in de figuren de promillages boven de 1000 voorkomen. De leeftijdsrange is beperkt tot leeftijd 70 omdat de mobiliteit op hogere leeftijden uiterst laag blijft.

den. Op 1 januari 2001 is Vleuten-De Meern bij de gemeente Utrecht gevoegd en op 1 januari 2002 zijn gedeeltes van Leidschendam, Nootdorp, Pijnacker en Rijswijk door de gemeente Den Haag geannexeerd. Deze correcties worden in het Sociaal Statistisch

Bestand niet als verhuizing geteld, omdat er geen sprake is van een daadwerkelijke verplaatsing.

Figuur 1. In- en uitstroom in de vier grote steden in de periode 1995-2003 ten opzichte van het inwonertal op 1-1-1995, per 1.000 van elk leeftijdsjaar, totale bevolking



Uit de figuur valt direct af te lezen dat in de periode 1995-2003 de hoogste instroom plaatsvond bij twintigers. Een piek is waarneembaar voor 19-jarigen. De hoogste uitstroom ligt rond het dertigste levensjaar en bij kinderen onder de vijf. Bovendien overtreft de uitstroom vanaf 30-jarige leeftijd de instroom. Voor 30-plussers en 15-minners is er in de vier grote steden sprake geweest van een vertreksaldo. Uitsluitend in de tussenliggende leeftijden was er sprake van een positief vestigingsaldo. Het zou dus kunnen dat de dertigplussers die zich eerder als twintigers hebben gevestigd weer vertrekken uit de grote steden. In hoeverre het inderdaad dezelfde personen betreft zal in een volgende analyse aan de orde komen. Hier gebruiken we het leeftijdsprofiel als criterium voor afbakening van leeftijdsgroepen. Verderop zal de in- en uitstroom daarom worden beschreven voor 18-29-jarigen en dertigplussers en kinderen jonger dan 18 jaar afzonderlijk.

2.2 Selectiviteit naar levensfase en etnische achtergrond

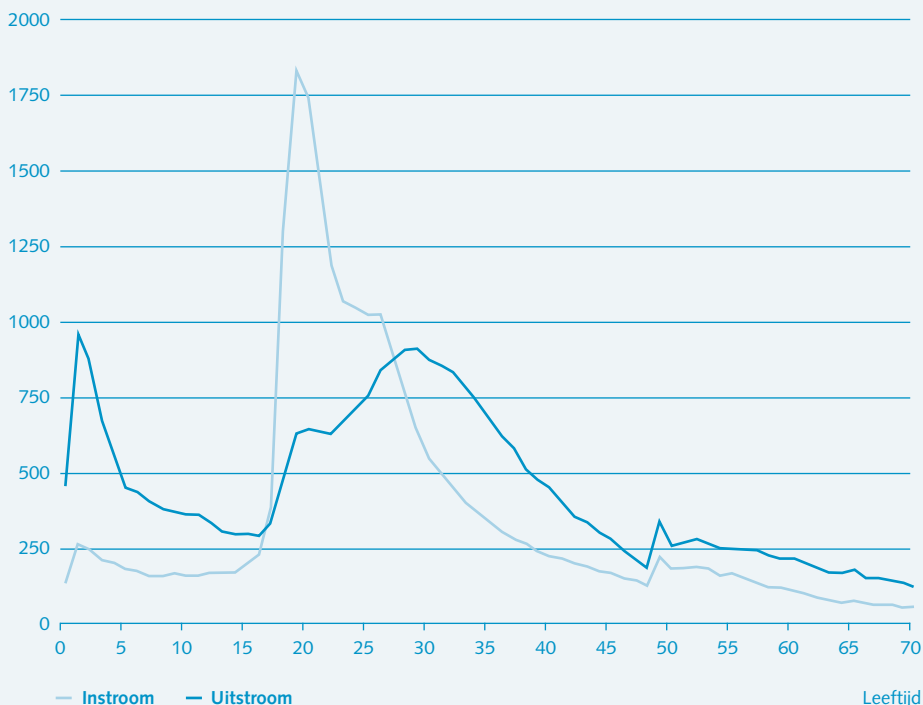
Immigranten die vanuit het buitenland arriveren verkeren in andere situaties dan binnenlandse nieuwkomers. Niet-westerse allochtonen komen bijvoorbeeld vanwege een huwelijk, gezinshereniging, vanwege werk of asiel. Westerse allochtonen zullen nauwelijks vanwege asiel, maar wel voor werk of studie naar Nederlandse steden komen. Ook voor

binnenlandse verhuizers kunnen de motieven weer anders zijn. Uiteraard kan dit consequenties hebben voor het leeftijds patroon in de mobiliteit.

2.2.1 Autochtonen

In figuur 2 is te zien dat onder autochtonen enkel voor de 17-27-jarigen de instroom de uitstroom overtreft. Dit is een scherpere concentratie dan voor 17-27-jarigen zonder etnische detaillering. Uitgaande van duizend 19-jarige autochtonen in 1995 zijn er de daaropvolgende 8 jaren circa 600 19-jarigen vertrokken en er zo'n 1800 bijgekomen. Dat geeft aan dat er in de periode 1995 tot en met 2003 bijna twee maal zo veel 19-jarige autochtonen instroomden dan er op 1 januari 1995 aan 19-jarigen in de vier grote steden woonden. Als dit leeftijds patroon als indicatie wordt gezien voor de aantrekkelijkheid van de grote stad als woonplaats voor autochtonen dan heeft de stad blijkbaar uitsluitend nog onder 17-27-jarige autochtonen een vestigingsoverschot. Het is aan te nemen dat het voor een groot deel studenten en starters op de arbeidsmarkt betreft. In iedere andere levensfase (kinderen, tieners tot 18-jarige leeftijd en dertigplussers) trekken meer autochtonen uit de vier grote steden weg dan er naartoe verhuizen. De uitstroom van dertigers en kinderen betreft uiteraard vaak gezinnen.

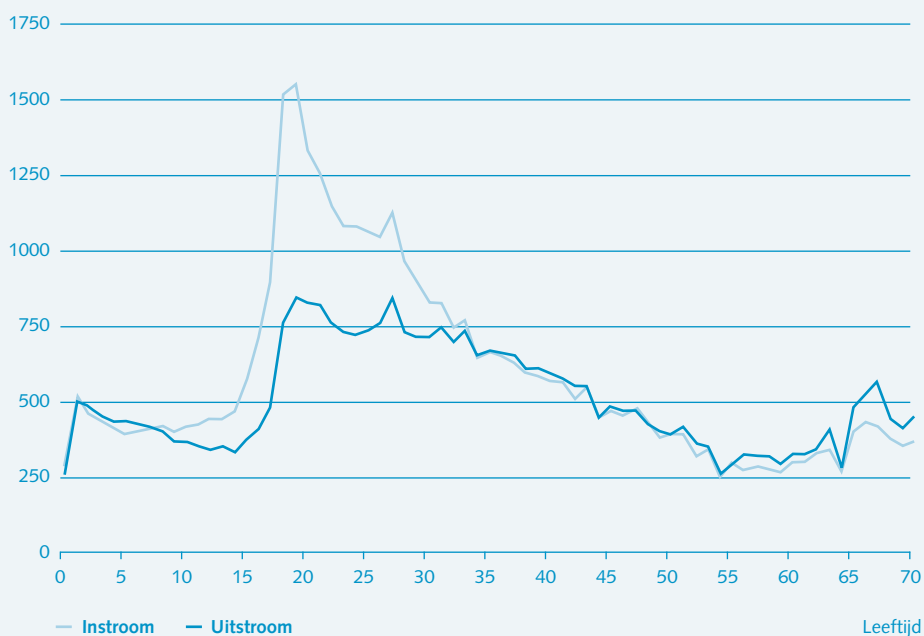
Figuur 2. In- en uitstroom van autochtonen in de vier grote steden in de periode 1995-2003 ten opzichte van het totale aantal autochtonen op 1-1-1995, per 1.000 van elk leeftijdsjaar



2.2.2 Niet-westerse allochtonen

Bij niet-westerse allochtonen is er duidelijk sprake van een ander patroon. Weliswaar is er ook een concentratie van in- en uitstroom onder jongeren: de niet-westerse allochtonen hebben evenals de autochtonen een piek van 19-jarige instromers (zie figuur 3), maar het vestigingssaldo is over een grotere leeftijdsrange gespreid. Het aantal uitstromers onder dertigplussers ligt vergeleken met autochtonen lager. Niet-westerse allochtonen hebben per saldo in alle levensfasen een sterkere neiging om zich in de stad te vestigen dan er uit weg te trekken. De stad heeft blijkbaar voor niet-westerse allochtonen een vestigingsfunctie op alle leeftijden.

Figuur 3. In- en uitstroom van niet-westerse allochtonen in de vier grote steden in de periode 1995-2003 ten opzichte van het totale aantal niet-westerse allochtonen op 1-1-1995, per 1000 van elk leeftijdsjaar

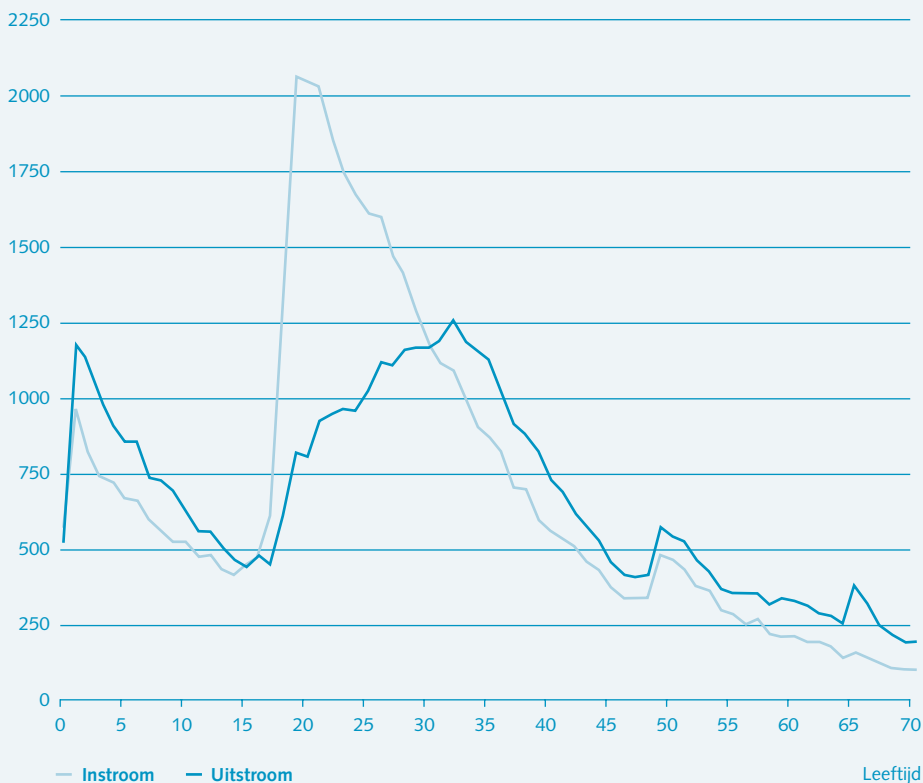


2.2.3 Westerse allochtonen

Tenslotte de westerse allochtonen. Voor westerse allochtonen is eenzelfde beeld te zien als onder de autochtonen. De instroom overtreft de uitstroom op alle leeftijden van 18 tot en met 30 jaar. Op andere leeftijden is de uitstroom groter dan de instroom. Ook bij de westerse allochtonen is sprake van een piek in vestiging op jonge leeftijd. Deze piek ligt echter nog hoger dan bij de autochtonen. Zowel op 19, 20 als 21-jarige leeftijd is sprake van

een instroom boven de 2000 (zie figuur 4). Dat wil dus zeggen dat er in de periode vanaf 1995 tot en met 2003 ruim twee maal zo veel westerse allochtonen in deze levensfase instroomden dan dat er in 1995 in een van de vier grote steden woonden. Voor een deel kan dat te maken hebben met instroom van buitenlandse studenten, een tendens die de afgelopen jaren in de vorm van internationalisering van het onderwijs alleen maar is toegenomen. Er is echter in vergelijking met de autochtonen ook sprake van een relatief grote instroom van eind twintigers. Dit kan mede verklaard worden door het grote aantal buitenlandse werknemers van internationale bedrijven, organisaties en vertegenwoordigingen dat zich meestal tijdelijk in Nederland vestigt. De internationale bedrijvigheid in de grote steden genereert immers een aanzienlijke stroom tijdelijke arbeidsmigranten (Cortie, 2003) met een hoge mobiliteit.

Figuur 4. In- en uitstroom van westerse allochtonen in de vier grote steden in de periode 1995-2003 ten opzichte van het totale aantal westerse allochtonen op 1-1-1995, per 1.000 van elk leeftijdsjaar



Resumerend blijkt aan de hand van leeftijd in combinatie met etnische achtergrond dat de vier grote steden in de afgelopen jaren voor autochtonen uitsluitend voor twintigers een positief vestigingssaldo hebben gekend. Voor autochtone dertigplussers en kinderen heeft de stad als woonplaats aan betekenis verloren: het aantal autochtone inwoners dat de stad verlaat, is in laatstgenoemde levensfasen telkens groter dan het aantal dat zich in de vier grote steden vestigt. Voor allochtonen lijkt de stad een andere rol te vervullen. Althans dat is op te maken uit het feit dat de stad voor allochtonen een vestigingsplaats is voor alle leeftijden.

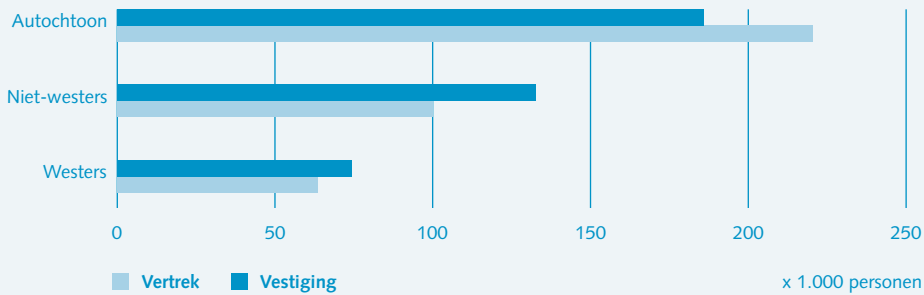
3 De demografische rekening

Los van de leeftijds patronen is uiteraard van belang hoe groot de absolute omvang van de stromen is. Vanaf september 1999 tot en met september 2003 hebben zich in de vier grote steden in totaal ruim 393 duizend personen gevestigd. Daarvan was bijna de helft, ruim 186 duizend, autochtoon (zie figuur 5). Het aantal niet-westerse allochtonen dat zich in deze periode in een van de vier grote steden vestigde, bedroeg 133 duizend. Daarnaast vestigden 74 duizend westerse allochtonen zich in de grote stad.

Het migratiesaldo laat voor de onderscheiden bevolkingsgroepen een verschillend beeld zien. Zowel bij de niet-westerse allochtonen als bij de westerse allochtonen kwamen er meer mensen naar de stad toe dan er weggingen. Niet-westerse allochtonen en westerse allochtonen hadden een vestigingsoverschot van achtereenvolgens 33 duizend en 11 duizend inwoners. Autochtonen daarentegen vertrokken in groten getale uit de vier grote steden, Ondanks de sterke instroom van twintigers leverde dat per saldo een verlies van 35 duizend autochtone inwoners. Wat opvalt, is dat de tegengestelde stromen van de onderscheiden bevolkingsgroepen elkaar sterk compenseren.

Gesteld kan worden dat de uittocht van autochtonen wordt opgevangen door de vestiging van niet westerse en westerse allochtonen. Deze tendens wijst op een 'verkleuring' en 'internationalisering' van Amsterdam, Rotterdam, Den Haag en Utrecht als geheel.

Figuur 5. Vestiging en vertrek van de vier grote steden naar herkomstgroepering, sept. 1999-sept. 2003

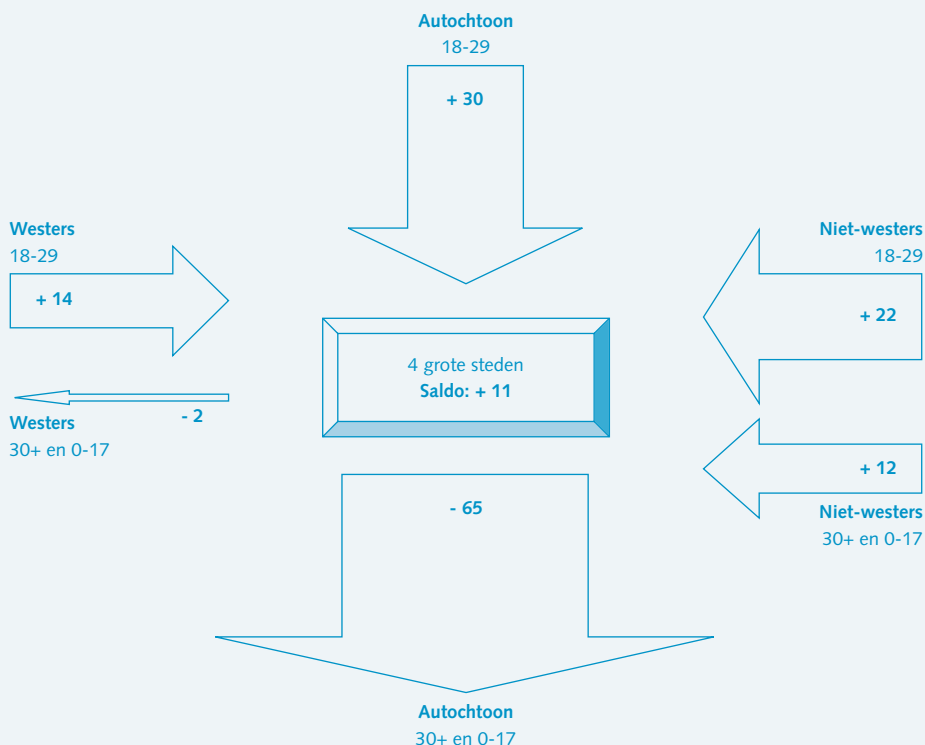


In het vorige hoofdstuk hebben we kunnen lezen en zien dat de instroom in de grote steden vooral bestaat uit twintigers. In figuur 6 is te zien dat de aantrekkingskracht van de grote steden op 18-29-jarigen inderdaad enorm is. Ongeacht etnische achtergrond kwamen er 66 duizend meer twintigers naar de stad toe dan er weggingen. Opvallend is daarbij dat het ook om allochtone twintigers gaat. Er komen per saldo 36 duizend allochtone twintigers de stad binnen versus 30 duizend autochtone jongeren.

Een vertreksaldo doet zich in wezen uitsluitend voor onder autochtone dertigplussers en kinderen jonger dan 18 jaar. Autochtone dertigplussers zoeken, in de levensfase met kinderen, een woning buiten de stad. Maar voor de niet-westerse allochtone gezinnen geldt dit minder. Over het algemeen behoren zij tot de lagere inkomensgroepen. Ze maken gebruik van de woningvoorraad in de vier grote steden die relatief sterk bestaat uit een oudere voorraad sociale huurwoningen. In combinatie met het positieve saldo voor de 18-29-jarigen betekent dit dat de stad voor niet-westers allochtonen op alle leeftijden een vestigingsfunctie heeft. Bij de westerse allochtone dertigplussers en kinderen jonger dan 18 jaar is, net als bij de autochtonen, het migratiesaldo negatief.

Kortom, tegenover een omvangrijk vertrek van autochtonen uit de grote stad, vooral naar randgemeenten, is er een inkomende verhuisstroom van niet-westerse allochtonen, die deze uitstroom voor een groot deel compenseert. Dit beeld van bevolkingswisseling door binnenlandse verhuizingen en internationale migratie wordt versterkt door selectieve ontwikkelingen in natuurlijke aanwas: voor autochtone inwoners is er sprake van een natuurlijke krimp en voor allochtone inwoners van een natuurlijke groei (Bontje en Latten, 2005). Deze tegengestelde trend versterkt nog eens de bevolkingsverandering vanwege verhuizingen. Voor de vier grote steden afzonderlijk wordt dit in de laatste paragraaf weergegeven.

Figuur 6. Migratiesaldo van in- en uitstroom in de vier grote steden naar levensfase en etnische achtergrond, uitgedrukt in aantal personen (x 1.000), sept. 1999-sept. 2003

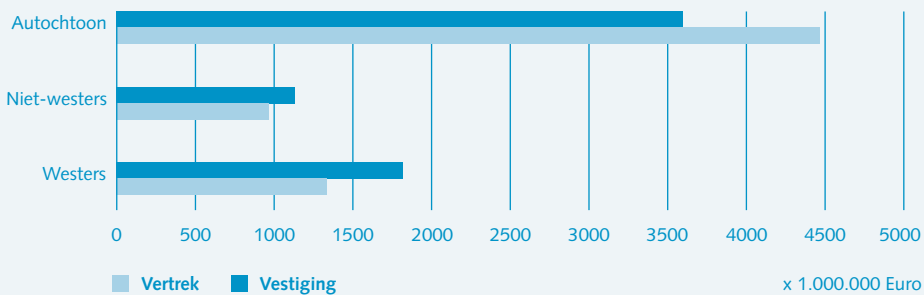


4 De sociaal-economische rekening

4.1 Selectieve migratiestromen

Door de demografische verhuisstromen uit te drukken in stromen van inkomens, kan een indicatie worden verkregen van het economisch effect van selectieve verhuizingen. Net als in de demografische benadering wordt ook nu onderscheid gemaakt naar herkomstgroepering en levensfase. Bij iedere persoon die in de periode september 1999 tot en met september 2003 uit of naar een van de vier grote steden verhuist, wordt gekeken naar zijn of haar fiscaal jaarinkomen rond het jaar van verhuizing. De som van alle inkomens van degenen die naar de stad toekomen en van hen die vertrekken bepaalt de winst- of verliesrekening aan inkomens vanwege verhuizingen voor de grote stad.

Figuur 7. Totaal fiscaal jaarinkomen bij in- en uitstroom in de vier grote steden naar levensfase en etnische achtergrond, uitgedrukt in euro's (x 1.000.000), sept. 1999-sept. 2003



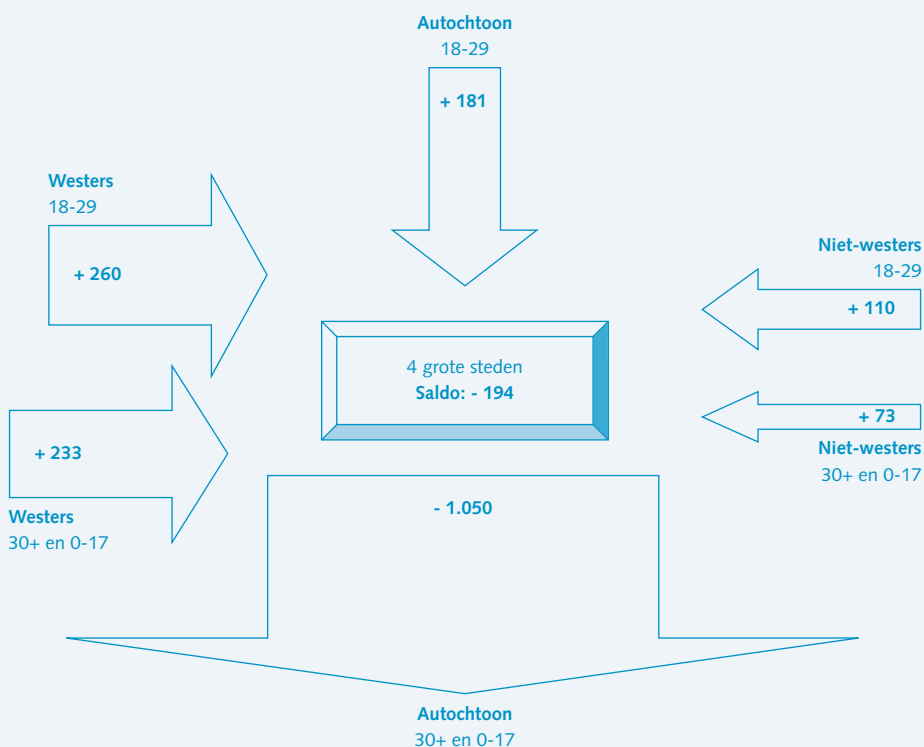
In figuur 7 zien we dat door middel van verhuisstromen van autochtonen 3,6 miljard euro aan jaarinkomens de vier grote steden binnenkomt. Vertrekkende autochtonen nemen daarentegen 4,5 miljard euro aan jaarinkomens mee de stad uit. Dat betekent een verlies vanwege uitstroom van autochtonen van 0,9 miljard euro aan jaarinkomens.

Westerse allochtonen laten een plussaldo zien. Nieuwkomers hebben ruim 1,8 miljard euro aan jaarinkomens, terwijl vertrekkenden een bedrag van 1,3 miljard euro aan jaarinkomens meenemen. Dat geeft een positief inkomenssaldo voor de grote steden van 0,5 miljard euro. Niet-westerse allochtonen brengen het minst mee aan jaarinkomens, namelijk een bedrag van 1,1 miljard euro. Dit terwijl de laatsten in absolute zin veel meer instromers tellen (er verhuizen bijna 60 duizend niet-westerse allochtonen meer naar de grote stad dan westerse allochtonen). Niet-westerse allochtonen brengen dus relatief gezien het minst mee aan jaarinkomens. Per saldo resulteren de verhuisstromen in een licht negatief inkomenssaldo van 0,2 miljard (figuur 8). Hoewel uit de bevindingen geen directe consequenties kunnen worden getrokken voor de inkomensontwikkeling van de totale stadsbevolking, mag duidelijk zijn dat de selectiviteit in de inkomenskenmerken van de verhuisstromen een eigen rol in die ontwikkeling kunnen spelen.

Als nu onderscheid wordt gemaakt naar leeftijdsgroepen, is te zien dat het verlies van bijna 0,2 miljard euro door verhuisstromen geheel voor rekening komt van de autochtone dertigplussers en kinderen jonger dan 18 jaar. Per saldo nemen zij ruim 1 miljard euro aan jaarinkomens mee de stad uit (figuur 8). Opmerkelijk is verder het positieve saldo van de westerse dertigplussers en kinderen jonger dan 18 jaar. Hoewel er in genoemde leeftijdsgroep meer westerse allochtonen uit de vier grote steden wegtrekken dan er naartoe verhuizen, is er voor deze leeftijdsgroep financieel toch een positief saldo van ruim 0,2 miljard euro te zien. Blijkbaar brengen de westerse dertigers en kinderen die naar de stad komen meer aan jaarinkomens mee dan degenen die uit de stad wegtrekken.

Geconcludeerd kan worden dat, als het op jaarinkomens aankomt, de allochtonen de omvangrijke autochtone uittocht uit de stad niet compenseren. Per saldo is er namelijk een verlies voor de grote stad van bijna 0,2 miljard euro aan jaarinkomens. Blijkbaar is het van belang *wie* uit de stad vertrekt of er naar toe verhuist. De selectiviteit in verhuisstromen speelt daarmee een rol in de economische ontwikkeling van de stad. Hoewel dit resultaat strikt genomen geen indicatie geeft voor de inkomensontwikkeling van de totale stedelijke bevolking, maakt het wel aannemelijk dat de selectieve verhuisstromen een rol spelen bij de relatief ongunstige inkomensontwikkeling in de grote steden. Duidelijk wordt dat de etnische samenstelling van verhuisstromen invloed heeft op de economische kenmerken van de stadsbevolking.

Figuur 8. Totaal aan fiscaal jaarinkomen bij in- en uitstroom in de vier grote steden naar levensfase en etnische achtergrond, uitgedrukt in euro's (x 1.000.000), sept. 1999-sept. 2003



4.2 Het gemiddeld fiscaal jaarinkomen

Het verschil aan totale jaarinkomens tussen niet-westers en westers allochtone verhuistromen blijkt aanzienlijk te zijn. Hoewel niet-westerse allochtonen qua aantallen een groter vestigingsoverschot laten zien dan westerse allochtonen, brengen de laatsten als groep aanzienlijk meer aan jaarinkomens met zich mee. Dat wordt nog duidelijker als per herkomstgroepering het migratiesaldo wordt gerelateerd aan het inkomenssaldo. In figuur 9 valt af te lezen dat een extra westers allochtoon die zich in een van de vier grote steden vestigt per hoofd een extra jaarinkomen van 42 duizend euro genereert. Bij niet-westerse allochtonen ligt dit bedrag op 5 duizend euro. Dat kan te maken hebben met het feit dat het bij niet-westers allochtonen vaker om gezinnen met kinderen gaat en minder vaak om tweeverdieners. Per hoofd omgerekend brengt een westerse allochtoon dus een veelvoud aan inkomen van een niet-westerse allochtoon mee. Het maakt duidelijk nogal uit welke groepen in de migratiestromen domineren. De autochtone stromen kennen zowel in demografisch opzicht als in sociaal-economisch opzicht een negatief saldo voor de grote steden. Bij elke extra autochtoon die per saldo uit de grote stad wegtrekt, gaat er een bedrag van 25 duizend euro aan jaarinkomen mee. Geconcludeerd kan worden dat, per hoofd gerekend, het bij elke extra westerse allochtoon om de hoogste inkomensbedragen gaat.

Figuur 9. Gemiddeld fiscaal jaarinkomen per extra persoon in de vier grote steden naar etnische achtergrond, sept. 1999-sept. 2003



5 Verschillen tussen de grote steden onderling

5.1 De demografische rekening

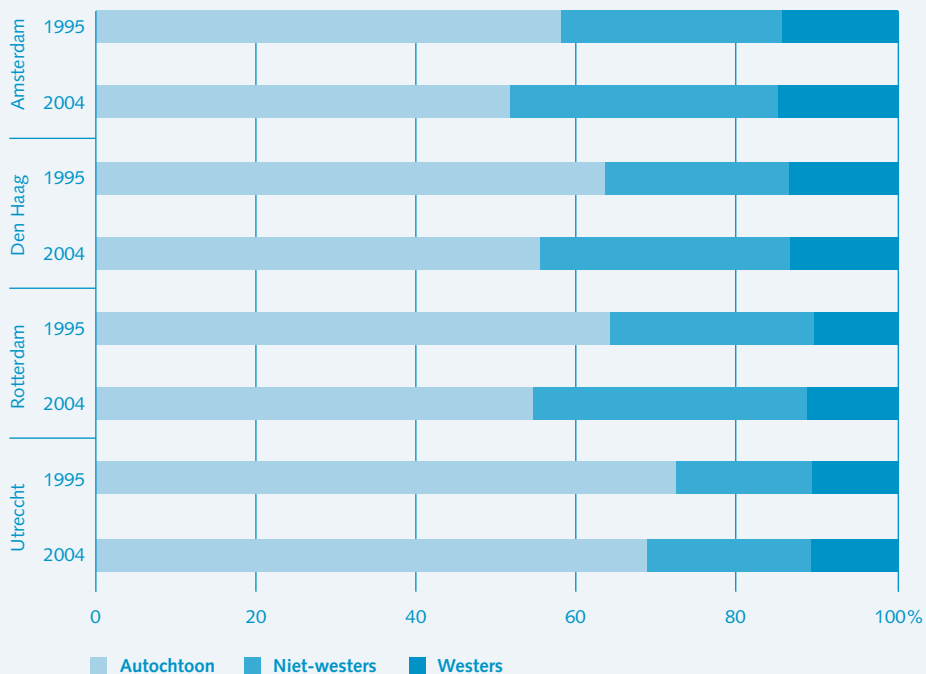
Zowel voor de demografische stromen als voor de sociaal-economische stromen verschillen Amsterdam, Rotterdam, Den Haag en Utrecht in de onderzochte periode van elkaar. Zo heeft alleen Amsterdam een licht negatief migratiesaldo. Het vertrekoverschot van Amsterdam bedroeg in de periode september 1999 tot en met september 2003 ruim 6 duizend personen. Den Haag, Rotterdam en Utrecht hadden een positief migratiesaldo. Het vestigingsoverschot van Utrecht was in genoemde periode het grootst, en bedroeg 9 duizend personen. Den Haag en Rotterdam volgden met een vestigingsoverschot van achtereenvolgens 7 duizend en 2 duizend. Ondanks deze positieve migratiesaldo's vertrokken in alle grote steden, met uitzondering van de stad Utrecht, meer autochtonen dan er bijkwamen. Het positieve vestigingssaldo van autochtonen is voor Utrecht wellicht te verklaren door de relatief grote toestroom van studenten of door omvangrijke nieuwbouwlocaties, zoals de VINEX-locatie Vleuten-De Meern waardoor de uittocht van autochtone dertigers beperkt kon blijven. Ook Den Haag kende een kleiner vertrekoverschot van autochtonen dan de steden Amsterdam en Rotterdam. Dit kan erop duiden dat ook Den Haag autochtonen beter wist 'vast te houden' vanwege haar grote uitbreidingswijken.

Het migratiesaldo van zowel niet-westerse allochtonen als westerse allochtonen was, zoals verwacht, in alle vier de steden positief. Opvallend voor het migratiesaldo van *westerse* allochtonen is het verschil tussen de steden Amsterdam en Den Haag aan de ene kant en Rotterdam en Utrecht aan de andere kant. Het vestigingsoverschot voor eerstgenoemde steden is bijna twee maal zo groot als voor Rotterdam of Utrecht. Wellicht wordt dit veroorzaakt door de aantrekkingskracht van beide steden op internationale arbeidsmigranten. Hierover meer in de volgende paragraaf. De stad Rotterdam is, met een positief migratiesaldo van bijna 14 duizend, duidelijk de meest populaire vestigingsplaats onder *niet-westerse* allochtonen. Rotterdam wordt op de voet gevolgd door Den Haag. Den Haag kent in de periode september 1999 tot en met september 2003 een vestigingsoverschot van 10 duizend niet-westerse allochtonen. Op afstand volgen de steden Amsterdam en Utrecht, met achtereenvolgens een migratiesaldo van bijna 7 duizend en 3,5 duizend niet-westerse allochtonen.

Als gevolg van deze migratiestromen is de bevolkingssamenstelling in de vier grote steden sterk veranderd. Het aandeel niet-westerse allochtonen in de stedelijke bevolking is sterk toegenomen, zoals te zien valt in figuur 10. Dit leidt tot een al eerder geconstateerde 'verkleuring' van de grote steden. Het beeld wordt nog duidelijker als gekeken wordt naar de

verandering in de bevolkingssamenstelling van de grote steden in de afgelopen tien jaar. Voor elke stad is een toename van het aantal niet-westerse allochtonen zichtbaar. In Rotterdam is het aandeel niet-westerse allochtonen op 1 januari 2004 zelfs toegenomen tot 35 procent van het totale aantal inwoners. Op 1 januari 1995 bedroeg het aandeel niet-westerse allochtonen daar nog 26 procent. Rotterdam kent daarmee de grootste relatieve stijging niet-westerse allochtonen van de vier grote steden. Rotterdam wordt hierin gevolgd door Den Haag. De stijging van het aandeel niet-westerse allochtonen was in Amsterdam en met name in Utrecht veel minder sterk, zoals is af te lezen uit de figuur. Toch is ook in de stad Amsterdam op 1 januari 2004 34 procent van alle inwoners van niet-westerse herkomst. In Utrecht ligt dit percentage beduidend lager. Utrecht kent van de vier grote steden ook relatief het meeste aantal autochtonen. In Utrecht is op 1 januari 2004 70 procent van alle inwoners autochtoon. Amsterdam kent met een aandeel van 52 procent relatief het minste aantal autochtonen. Het aandeel westerse allochtonen is in de onderzochte periode aan weinig veranderingen onderhevig. Wel is terug te zien dat het aandeel westerse allochtonen het grootst is in Amsterdam en Den Haag.

Figuur 10. Bevolking naar herkomstgroepering in de vier grote steden op 1-1-1995 en 1-1-2004



5.2 De economische rekening

Ook als het gaat om de sociaal-economische stromen zijn er opvallende verschillen tussen de grote steden. Den Haag is de enige stad, waar in de onderzochte periode het totaal aan verhuisstromen een positief inkomenssaldo liet zien (zie tabel). Het vestigingsoverschot van Den Haag lijkt zich uit te betalen. Den Haag kent namelijk qua bijbehorende inkomens een positief saldo van 232 miljoen euro. Dit verschil ten opzichte van de andere grote steden moet waarschijnlijk mede verklaard worden door het al eerder genoemde lagere vertrekoverschot van autochtonen uit de stad. Uitgesplitst naar etnische achtergrond blijkt namelijk dat in Den Haag het verlies aan jaarinkomens door het vertrekoverschot van autochtonen per saldo slechts een min van 23 miljoen euro bedraagt. In Amsterdam en Rotterdam gaat het om negatieve saldi van achtereenvolgens 442 miljoen euro en 349 miljoen euro.

Tabel 1. Migratiesaldo en saldo fiscaal jaarinkomen van de vier grote steden, sept. 1999-sept. 2003

	Saldo aantal personen	Saldo fiscale jaarinkomens			Totale bevolking
		<i>Autochtoon</i>	<i>Niet-westers allochtoon</i>	<i>Westers allochtoon</i>	
	x 1.000	x € 1.000.000			
Den Haag	+6,9	-23	+88	+167	+232
Utrecht	+9,2	-55	+11	+18	-26
Amsterdam	-6,4	-442	+15	+278	-148
Rotterdam	+1,8	-349	+68	+29	-252

De verhuissaldi van Amsterdam, Rotterdam en Utrecht worden, in tegenstelling tot Den Haag, vergezeld van negatieve inkomensaldi. Voor alle duidelijkheid, hierbij gaat het enkel om de totalen aan jaarinkomens van personen die verhuizen; de zittende bevolking wordt buiten beschouwing gelaten. Rotterdam kent verreweg het grootste verlies aan jaarinkomens. Ondanks een positief demografisch vestigingssaldo, bedraagt het verlies aan jaarinkomens door verhuizingen in de periode september 1999 tot en met september 2003 maar liefst 252 miljoen euro. Dit komt omdat de 349 miljoen euro aan jaarinkomens die autochtonen met zich meenemen, nauwelijks wordt opgevangen door het instromende allochtonen. Allochtonen brengen per saldo tezamen 97 miljoen euro aan jaarinkomens mee naar de stad. Daarvan komt slechts een bescheiden bedrag van 29 miljoen euro voor rekening van westerse allochtonen.

Ook Amsterdam lijdt een verlies aan jaarinkomens door verhuisstromen. In Amsterdam bedraagt het verlies 148 miljoen euro. Ook dit verlies wordt veroorzaakt door de omvangrijke trek uit de stad van de autochtone bevolking. Per saldo levert dit een negatief inko-

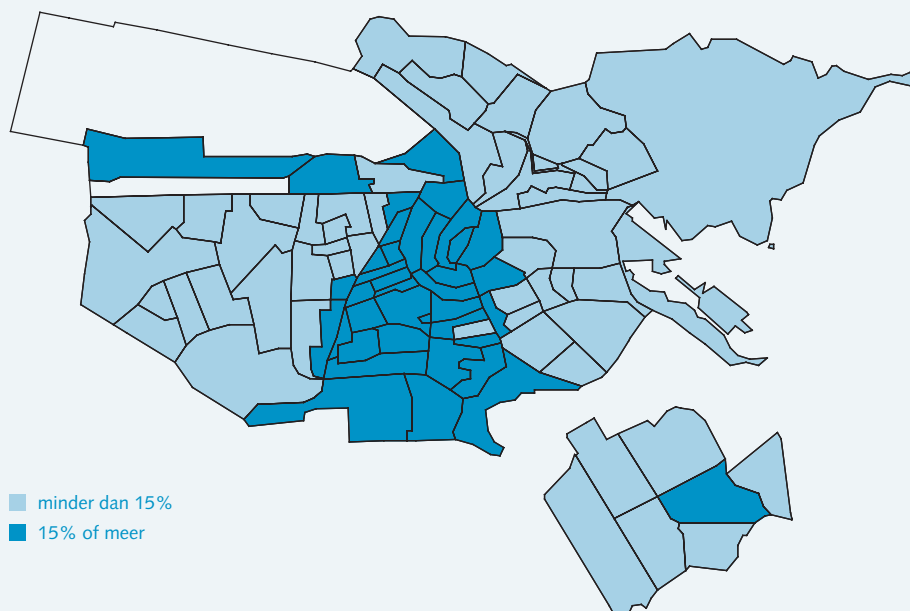
menssaldo van 442 miljoen euro aan jaarinkomens. In Amsterdam wordt dit negatieve saldo echter veel meer dan in Rotterdam gecompenseerd door de inkomens van westers allochtone verhuisstromen. Waar het saldo aan jaarinkomens dat westerse allochtonen met zich meebrachten in de periode 1999 tot en met september 2003 in Rotterdam slechts 29 miljoen euro bedroeg, bedroeg dat in Amsterdam maar liefst 278 miljoen euro. Eenzelfde beeld is te zien bij Den Haag, waar de westerse allochtonen per saldo een bedrag van 167 miljoen euro meebrachten.

Het is niet uit te sluiten dat de verschillen voortkomen uit de aantrekkingskracht op een ander type arbeidsmigrant. Volgens een studie van Burgers en Musterd (2002) raakt Rotterdam, in vergelijking met Amsterdam, achterop met betrekking tot de groei van de zakelijke dienstverlening. Amsterdam beschikt volgens hen over een gevarieerde economische structuur, terwijl Rotterdam zich nog vooral focust op haar haven. Hierdoor is een ander profiel waarneembaar van mensen die naar Rotterdam toetrekken. De zogenaamde professionals trekken naar Amsterdam. Dit is terug te zien in het grote aantal buitenlandse hoofdkantoren dat in Amsterdam gevestigd is. Zo is het internationale bankwezen vooral terug te vinden in Amsterdam, met hoofdkantoren aan de Zuidas. Den Haag is bij uitstek de Nederlandse diplomatenstad. Zij is gastheer aan ten minste zeventig internationale organisaties, meer dan tachtig ambassades en consulaten, tien internationale scholen en tal van hoofdvestigingen van grote multinationals. De internationale gemeenschap is volgens officiële schattingen ten minste 26 duizend mensen groot. (Stokmans, 2005) Het gemeentebestuur hecht veel waarde aan deze internationale inbreng en noemt zich graag de juridische hoofdstad van de wereld.

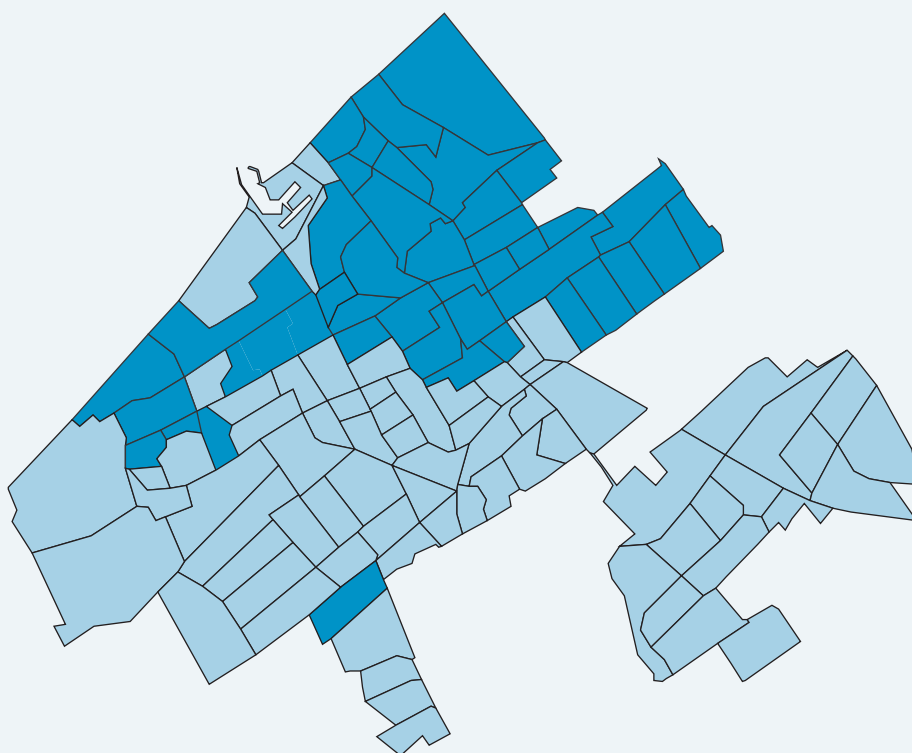
Het verschil in sociaal-economische stromen tussen de steden Amsterdam en Den Haag aan de ene kant en Rotterdam aan de andere kant wordt ook duidelijk als we het demografisch migratiesaldo opnieuw in ogenschouw nemen. Hoewel Rotterdam in genoemde periode een vestigingsoverschot had van 2 duizend personen en Amsterdam een demografisch vertrekoverschot van ruim 6 duizend personen staat Amsterdam er qua inkomenssaldo beter voor. De indruk uit deze analyse is dat het voor een belangrijk deel te maken heeft met het feit dat het migratiesaldo van westerlingen in Amsterdam bijna twee keer zo groot is als in Rotterdam. Ook Den Haag scoort in deze gunstig. Rotterdam lijdt, zelfs met een positief demografisch vestigingsoverschot, het grootste verlies aan jaarinkomens vanwege de verhuispatronen. Het gewicht van de westerse migratie naar de grote steden mag duidelijk zijn, het heeft ook verdergaande consequenties. Een sterke gerichtheid op internationalisering van een stad kan bijvoorbeeld het winkelaanbod beïnvloeden, kan afgeleide werkgelegenheid in de dienstverlening creëren maar blijft ook niet ongemerkt op de woningmarkt.

Dit is onder andere te zien aan de vestigingspatronen van westerse allochtonen in Amsterdam en Den Haag (zie figuren 11 en 12). In Den Haag wonen in bijna tweederde van alle buurten minimaal 10 procent westerse allochtonen. Maar men woont niet 'at random'. Er is sprake van concentratie in duurdere buurten. In Den Haag zijn er concentraties ten noorden van het centrum, in de statige klassieke woonbuurten als Statenkwartier, Duinoord, Archipel en de nieuwere Vogelbuurt dicht tegen het duingebied aan of in Scheveningen-Bad. In Amsterdam wonen westerse allochtonen relatief vaak in de historische centrumbuurten en in het zuiden van de stad, waar de woningwaarde ook het hoogst is. (Raets, 2005)

Figuur 11. Spreiding westerse allochtonen in Amsterdam, 1-1-2004



Figuur 12. Spreiding westerse allochtonen in Den Haag, 1-1-2004



Westerse allochtonen per buurt, 2004

- minder dan 15% westerse allochtonen
- 15% of meer westerse allochtonen

Samenvattend

- De toestroom naar de vier grote steden Amsterdam, Den Haag, Rotterdam en Utrecht bestaat voornamelijk uit personen in de leeftijd van 18 tot en met 29 jaar. Er kan dus wel gesteld worden dat er sprake is van een ‘vertwintiging’ van de grote steden. Per saldo gaat het hierbij zowel om allochtone als autochtone twintigers.
- Autochtone dertigplussers en kinderen jonger dan 18 jaar laten per saldo een vertrek zien.
- De grote steden hebben voor niet-westerse allochtonen op alle leeftijden een vestigingsfunctie. Dit betekent dat er sprake is van een ‘verkleuring’ van de grote steden. Deze ‘verkleuring’ wordt nog versterkt doordat er per saldo meer autochtonen de stad verlaten dan er zich vestigen.
- Behalve Den Haag lijden de grote steden door verhuisstromen per saldo een inkomensverlies.
- Het positieve migratiesaldo van allochtonen compenseert het inkomensverlies vanwege de uitstroom van autochtonen niet, hoewel dit in demografisch opzicht nog wel het geval bleek te zijn.
- Westerse allochtonen die zich in een van de vier grote steden vestigen hebben omgekeerd per hoofd aanzienlijk hogere jaarinkomens dan niet-westers allochtonen toestromers.
- Een voortgaande bevolkingswisseling, waarin autochtonen grotendeels worden opgevolgd door niet-westerse allochtonen, legt een druk op de inkomensontwikkeling van de stedelijke bevolking.
- Omgekeerd: westers allochtone instroom gaat gepaard met hogere inkomens.

- Bontje, M. en J. Latten (2005). Selectieve migratie in grote steden. In: *Rooilijn: tijdschrift voor wetenschap en beleid in de ruimtelijke ordening*, 1, januari 2005, 17-22. Universiteit van Amsterdam, afdeling Geografie en Plaanologie, Amsterdam.
- Burgers, J. en S. Musterd (2002). Understanding urban inequality: A model based on existing theories and an empirical illustration. In: *International Journal of Urban and Regional Research*, 26, 2, Juni 2002, 403-413. Blackwell Publishers.
- Cortie, C. (2003). The metropolitan population, origin and mobility. In: Musterd, S. en W. Salet (Eds.), *Amsterdam Human Capital*, 199-216. Amsterdam University Press, Amsterdam.
- Latten, J., M. Bontje en H. Nicolaas (2004). Bevolkingsdynamiek in de vier grote steden. In: *Bevolkingstrends: statistisch kwartaalblad over de demografie van Nederland*, 2e kwartaal 2004. Centraal Bureau voor de Statistiek, Voorburg/Heerlen.
- Raets, B. (2005). Wonen aan de rand. In: *Webmagazine*, 12 september. Centraal Bureau voor de Statistiek, Voorburg/Heerlen.
- Sprangers, A., A. Zorlu, J. Hartog en H. Nicolaas (2004). Immigranten op de arbeidsmarkt. In: *Sociaal-economische trends: statistisch kwartaalblad over arbeidsmarkt, sociale zekerheid en inkomen*, 2e kwartaal 2004. Centraal Bureau voor de Statistiek, Voorburg/Heerlen.
- Stokmans, D. (2005). "Nederland is duur, moeilijk en ongastvrij." In: *NRC Handelsblad*, 25 september 2005.
- Vries A. de, (2005). Segregatie en menging: Inkomensdynamiek in buurt en stad. In: *Stedenbouw & Ruimtelijke Ordening*, 86, 4, 66-69. Nederlands Instituut voor Ruimtelijke Ordening en Volkshuisvesting, Den Haag.





Jeroen Slot (Gemeente Amsterdam)

In balans?
Grote steden op weg naar
een nieuw evenwicht.*

* Reactie op: 'De prijs van migratie. Selectieve verhuisstromen van de vier grote steden'.

‘Uit balans, Selectieve verhuisstromen naar en uit de grote stad’¹⁸ levert een prachtig verhaal op. En dat komt natuurlijk omdat het SSB een prachtig bestand is, zeker als het om demografische kenmerken gaat. Latten, Hamers en Nicolaas laten zien dat de vier grote steden ieder een ‘liftfunctie’¹⁹ hebben in de Nederlandse samenleving. Precies die liftfunctie is een paar jaar geleden sterk benadrukt in een manifest van de G4, waarin zij het kabinet om aandacht vroegen - en om geld natuurlijk. De verwijzing naar de liftfunctie als onmisbaar mechanisme was onderdeel van een beeld dat steden neerzette als innovatiecentra van de economie. In de jaren tachtig domineerden vooral probleemverhalen bij de zoektocht naar aandacht en geld. Toentertijd karakteriseerde het SCP Amsterdam als ‘hoofdstad der marginalen’ en de plaatselijke GG&GD schatte zijn werkvoorraad losjes door uit te gaan van ‘een gek per trap’.

Uit Balans is prachtig onderzoek omdat het de liftfunctie beter zichtbaar maakt en daarmee een goede basis vormt voor de discussie die in alle grote steden wordt gevoerd over de samenstelling van de bevolking in relatie tot de leefbaarheid en de economie van de stad. Alle steden willen kansrijke bewoners aantrekken, ze willen ook allemaal middengroepen behouden en sommige willen sociaal-economisch zwakke groepen weren.

Het SSB maakt de belofte een databank met ongekeerde analysemogelijkheden te zijn helemaal waar. De keuze om migratie van en naar de stad te beschrijven in termen van in- en uitgaande inkomensstromen is een gouden greep. Nederlanders willen overal de prijs van weten, dus dit zal zeker publiciteit opleveren en de discussie over de bevolkingssamenstelling van steden een nieuwe impuls geven.

Maar er is ook wel iets op te merken over de manier waarop de gegevens gepresenteerd zijn. En aangezien ik als referent ben ingehuurd, zal ik dat ook doen.

Iedereen weet dat de universiteit niet dommer wordt als er afgestudeerden vertrekken en eerstejaars instromen. In die zin hoeft het niet te verbazen dat steden ‘inkomen’ verliezen aan het ommeland. Sterker nog: het is in zekere zin de functie van steden. Dat dit in Den Haag niet of in veel mindere mate gebeurt, is bijzonder. Als we even voorbij gaan aan de mogelijkheid dat de data het definitieve bewijs vormen dat Den Haag een dorp is en ook schuivende gemeentegrenzen geen roet in het eten hebben gegooid, is die uitkomst opmerkelijk. Het zou kunnen betekenen dat de onderste roltrappen in ‘de liftfunctie’ van Den Haag ontbreken.

18) Redactie: Slot reageert in zijn referaat op een eerdere versie van de paper ‘De prijs van migratie, selectieve verhuisstromen van de vier grote steden’ van Latten, Hamers en Nicolaas, toen nog ‘Uit Balans, selectieve verhuisstromen naar en uit de grote stad’ geheten.

19) Redactie: Slot refereert aan de zin ‘Uiteraard kan een grote stad fungeren als een roltrap voor individuele sociale en economische stijging. Men volgt een opleiding, krijgt een eerste baan, een hoger inkomen en vertrekt na stijging op de maatschappelijke ladder weer uit de stad’ uit de eerdere versie

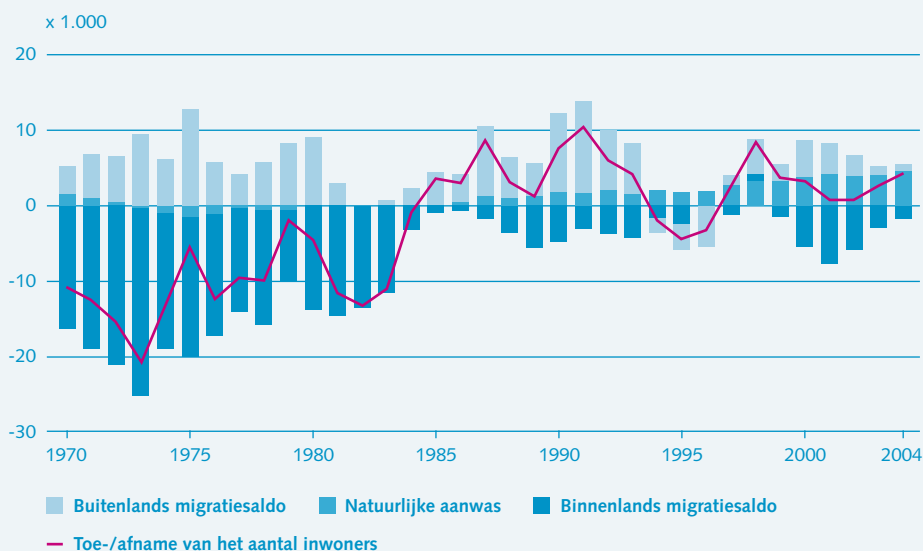
van de paper van Latten, Hamers en Nicolaas (zie ook voetnoot 18).

Om dat goed te kunnen beoordelen is meer kennis nodig van de mate waarin stadsbewoners er tijdens hun verblijf in de stad in slagen om extra inkomen te verwerven door carrière te maken. Zonder die kennis suggereert de titel *Uit Balans* misschien meer dan in werkelijkheid het geval is. De opmerking dat een ‘duurzame’ liftfunctie in- en uitstroom binnen ‘bepaalde grenzen’ veronderstelt, versterkt die suggestie. Dit soort formulering doet een beetje denken aan de gretigheid waarmee media melden wanneer het percentage allochtonen de 50% zal passeren. Zowel op de teller als de noemer die aan dit percentage ten grondslag ligt, is wel wat af te dingen. Alle allochtonen - uit verschillende herkomstgebieden, met verschillende migratiemotieven en uit verschillende generaties - worden op een hoop gegooid, terwijl de gemeentegrens bestuurlijk overeind blijft maar in de sociaal-economische ontwikkeling van stedelijke regio's steeds minder betekenis heeft. Een eerste opmerking is dan ook dat een aanvullende balans waarin centrumstad, buitenwijken en suburbs van stedelijke regio's worden onderscheiden nodig is voor een beter beeld van de ontwikkeling. Ook zou het de moeite waard zijn de Noord- en Zuidvleugel van de Randstad te vergelijken.

Een tweede opmerking is dat in een wat langer tijdsperspectief de betekenis van de selectiviteit in de verhuisstromen beter te begrijpen is. De grote steden zijn niet zozeer uit balans, maar eerder op weg naar een nieuwe balans. Ik zal dat illustreren aan de hand van cijfermateriaal over de stad die ik het beste ken, Amsterdam (figuur 1).

Figuur 1. Binnen- en buitenlands migratiesaldo van Amsterdam, 1970-2004

Bron: O+S Amsterdam, 2005

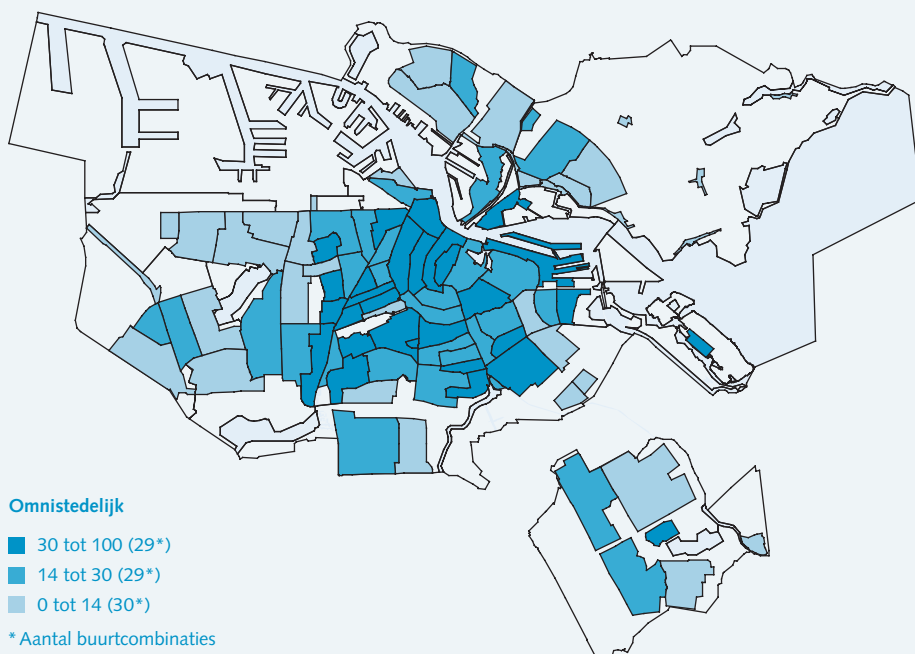


Bron: O+S Amsterdam, 2005

Uit de cijfers blijkt dat de ‘grote verstoring’ in de Amsterdamse bevolkingssamenstelling al lang achter ons ligt: de periode van overloop uit de jaren zeventig, gevolgd door een buitenlands migratieoverschot.

Sinds het midden van de jaren tachtig is de stad weer een aantrekkelijke plek geworden om te wonen. Vooral jonge, kleine huishoudens die in menig opzicht aan het begin van hun carrière staan, waarderen het stedelijk klimaat. Huishoudens met een - in de termen van Van Diepen en Musterd - omnistedelijke oriëntatie (zowel sociaal als economisch) zijn gericht op de stad, en vooral op het centrum van de stad (figuur 2). Deze huishoudens zijn voor hun sociaal vertier vooral aangewezen op derden en missen de vooral intern gerichte oriëntatie van gezinnen met kinderen. Bovendien nemen ze, anders dan bijvoorbeeld de meeste senioren, deel aan het arbeidsproces en zijn aangewezen op de stedelijke arbeidsmarkt.

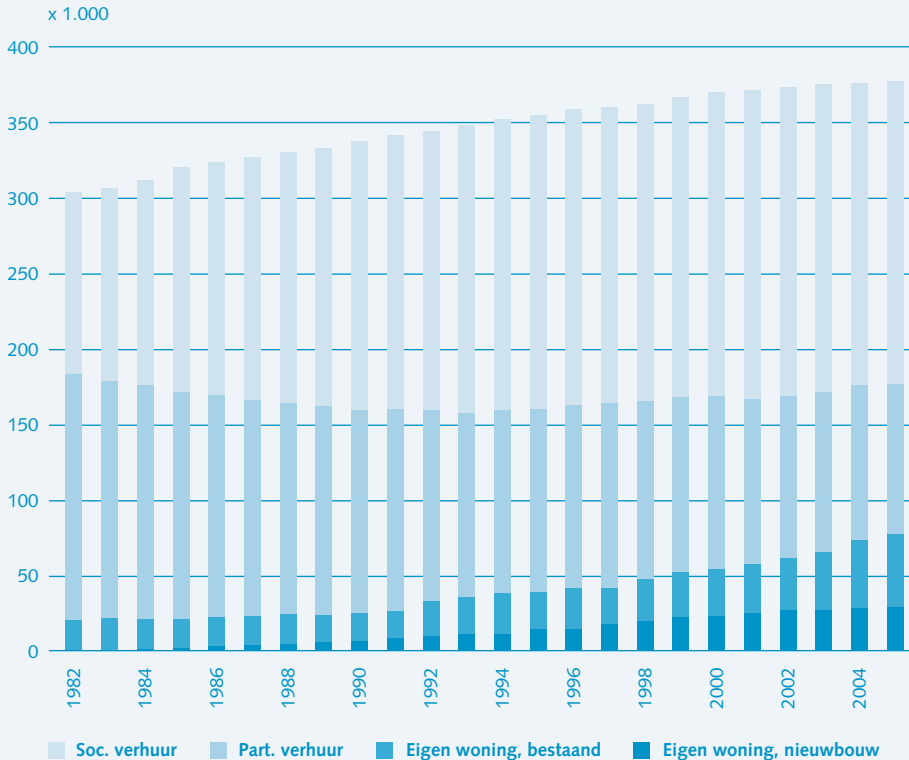
Figuur 2. Percentage huishoudens met een omnistedelijke oriëntatie per buurtcombinatie, 2003



Bron: O+S Amsterdam, 2005

De ontwikkeling van de woningvoorraad helpt hierbij. Niet alleen zijn er in het centrale deel van de stad in de afgelopen decennia veel woningen bijgekomen, ook is het aandeel koopwoningen toegenomen (figuur 3). Dat betekent dat de stad meer ruimte is gaan bieden aan mensen met een sterke stedelijke oriëntatie.

Figuur 3. Ontwikkeling woningvoorraad Amsterdam 1982-2005

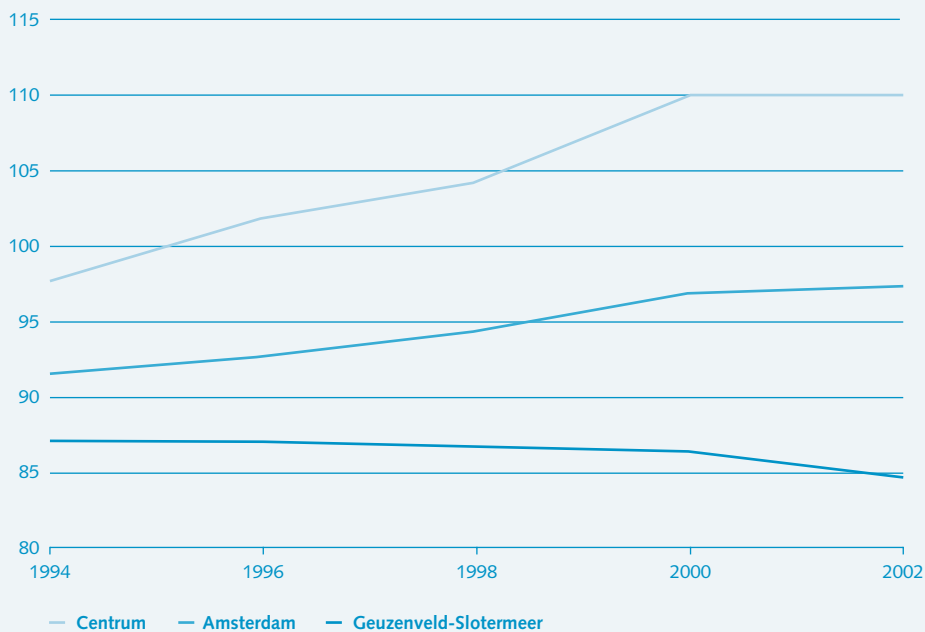


Bron: O+S Amsterdam, 2005

Deze koopwoningen bieden eerder huisvesting voor omnistedelijk en economisch georiënteerde huishoudens dan andere huishoudens. Zo wordt 41% van de Amsterdamse koopwoningen bewoond door omnistedelijk georiënteerde huishoudens en slechts 26% van de huurwoningen; bij economisch stedelijk georiënteerde huishoudens is dat 21 respectievelijk 11%.

Ik schets deze achtergrond, omdat wie kijkt naar de ontwikkeling van het gemiddeld huishoudensinkomen in Amsterdam (figuur 4) vaststelt dat Amsterdam naar het landelijk gemiddelde toekruipt: hoezo uit balans?

Figuur 4. Ontwikkeling gemiddeld huishoudensinkomen Amsterdam, 1994-2002
(gestandaardiseerd huishoudeninkomen [NL=100])

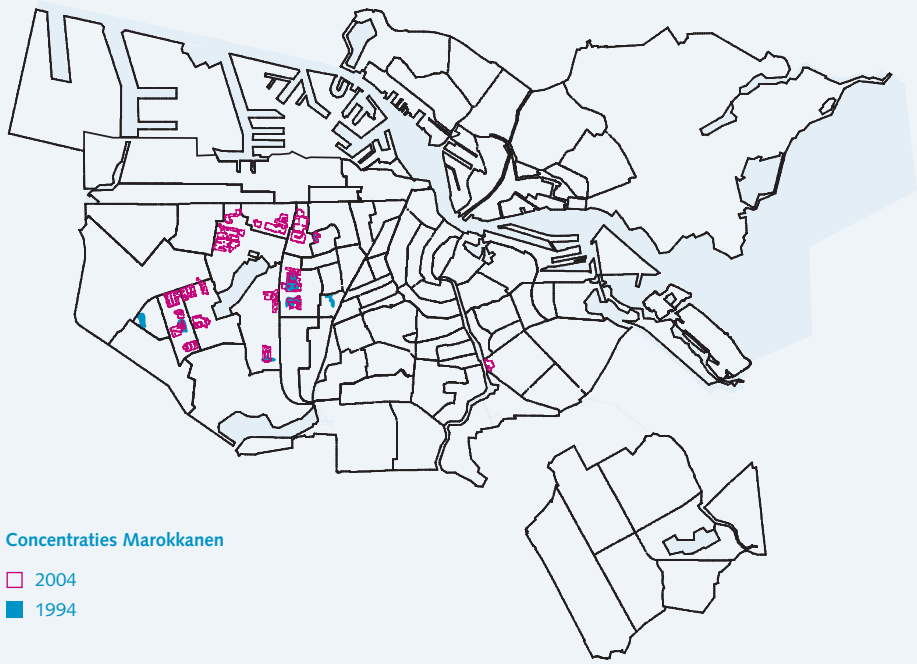


Bron: CBS

De figuur laat overigens ook zien dat de ontwikkeling niet overal in Amsterdam op dezelfde wijze verloopt. De ontwikkeling in het centrum is heel anders dan de ontwikkeling in Geuzenveld-Slotermeer, een van de stadsdelen in Nieuw-West.

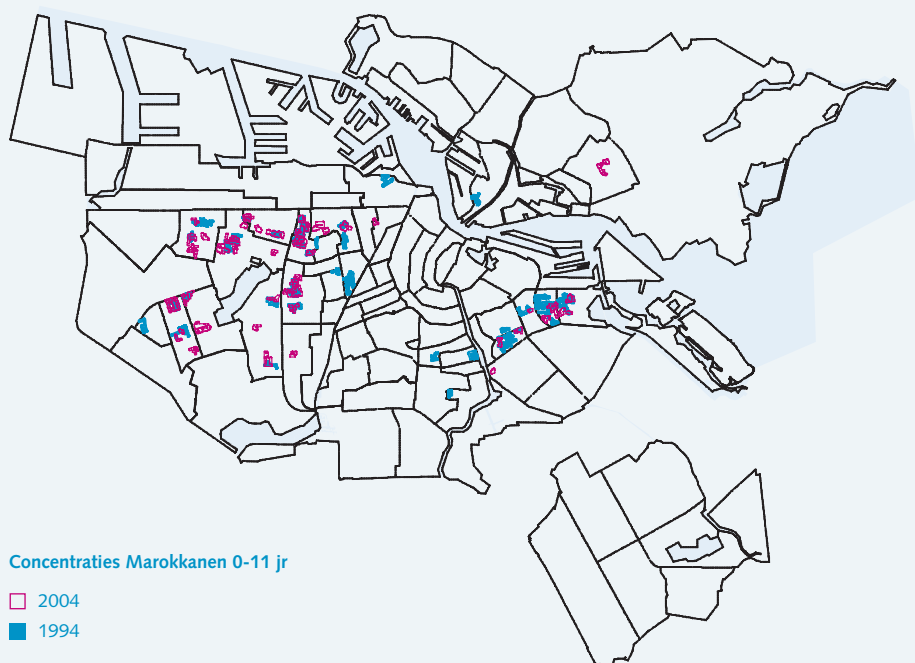
De figuren 5 en 6 laten zien dat in Nieuw-West sprake is van een groeiend aandeel Marokkaanse Amsterdammers.

Figuur 5. Concentraties Marokkanen (>35%), 1994 en 2004



Bron: Stadsmonitor Amsterdam, O+S Amsterdam / Universiteit van Amsterdam

Figuur 6. Concentraties Marokkanen (0-11 jaar), 1999 en 2004

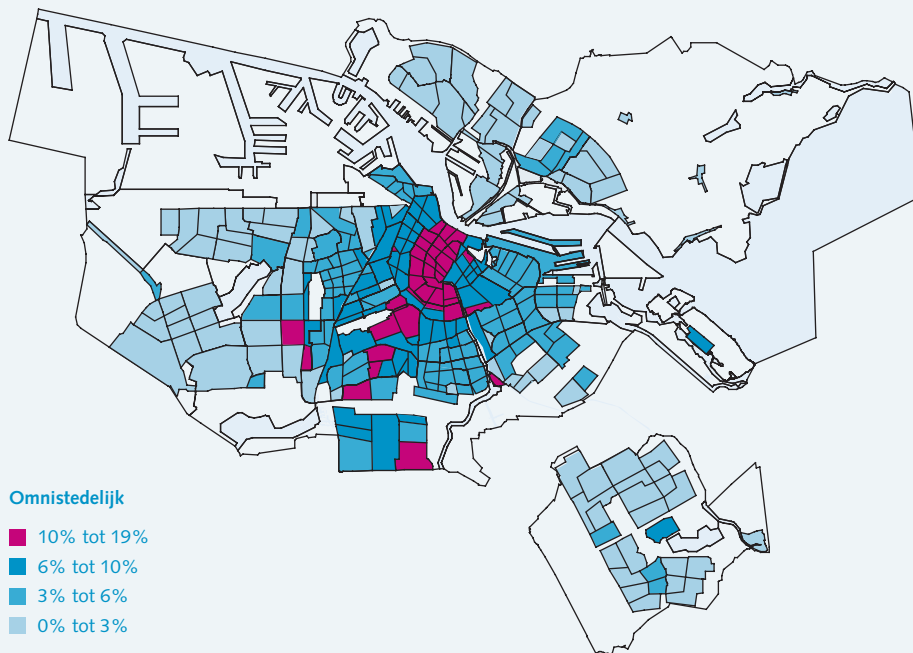


Bron: Stadsmonitor Amsterdam, O+S Amsterdam / Universiteit van Amsterdam

Bij de interpretatie van deze beweging moet ook aan de levensfase van deze huishoudens gedacht worden: juist voor huishoudens met kinderen is de voorraad relatief grote en goedkope woningen in de westelijke tuinsteden aantrekkelijk. Figuur 6 laat zien dat in Amsterdam-west een toename te zien is van gebieden waarin het aandeel Marokkaanse kinderen twee standaardafwijkingen hoger dan gemiddeld is: zogenaamde concentratiegebieden.

De binnenstad daarentegen is bij uitstek aantrekkelijk voor westerse allochtonen, in figuur 7 zichtbaar gemaakt aan de hand van het aandeel westerse allochtonen met twee buitenlandse ouders die twintig jaar of korter in de stad verblijven. Van deze groep mag verondersteld worden dat ze naar Amsterdam gekomen zijn vanwege het economisch en cultureel klimaat.

Figuur 7. Percentage westerse allochtonen per buurtcombinatie, 2004



Bron: O+S Amsterdam, 2004

Samenvattend:

- Het SSB is een prachtig instrument, ook voor lokale beleidsmakers
- Meer aandacht is nodig voor het regionale en binnenstedelijke niveau
- Meer aandacht is nodig voor ‘vermijding’ als begin van veel selectiviteit

De eerste conclusie is dat het SSB prachtige mogelijkheden biedt, juist ook voor lokale beleidsmakers. Dat de grote stad als roltrap functioneert, maken de cijfers wel duidelijk. De vraag in welke mate een stad als Amsterdam in balans verkeert, is wat mij betreft nog niet beantwoord. Daarvoor is niet alleen het regionaal stedelijk perspectief onontbeerlijk, maar ook meer zicht nodig op de vraag hoe het ‘roltrapmechanisme’ voor verschillende bewonersgroepen functioneert. Dat leidt dan weer tot vervolgvragen: wat voor type stad wil je zijn, welke roltrapmechanismen versterken elkaar en welke rol speelt de woningvoorraad bij het accommoderen en sturen van deze ontwikkelingen?

Ook voor het vroegtijdig signaleren van ontwikkelingen biedt het SSB mogelijkheden. Vaak begint selectiviteit met vermijdingsgedrag: bepaalde groepen bewoners beginnen de stad als geheel of bepaalde delen te mijden. Een interessante vraag in dit verband is dan hoe Amsterdam-Zuidoost en Almere over een langere periode bezien in de stedelijke regio functioneren.

Kortom, veel vervolgvragen. Maar ook één zekerheid: het volgende SSB-symposium belooft weer erg interessant te worden. Misschien moet het maar jaarlijks plaatsvinden.

Referenties

Amsterdam in Cijfers 2005, O+S Amsterdam, 2005

Demografische ontwikkelingen: blijvende groei Amsterdamse bevolking, fact sheet O+S Amsterdam en dRO, 2005

Diepen, A. van en S. Musterd, Stedelijke leefstijlen en woonmilieus in Amsterdam, Universiteit van Amsterdam en Dienst Ruimtelijke Ordening Amsterdam (dRO), 2001

Staat van de Stad Amsterdam III, O+S Amsterdam en Dienst Maatschappelijke Ontwikkeling Amsterdam, 2005

... Tijdens
Amsterdam Institute for
Advanced Labour Studies
Universiteit van Amsterdam





Kea Tijdens (UvA/EUR)

SSB-tabellen in StatLine.

Introductie StatLine

StatLine is de elektronische databank van het CBS op Internet, in de vorm van tabellen, grafieken en kaarten. Zoeken in de databank kan op trefwoorden, via themaboom en via de kaart. Mijn ervaringen als web-gebruiker zijn dat StatLine snel werkt, dat het gemakkelijk is, dat de tabellen 'downloadable' zijn.

Eén opmerking: bij mijn eerste kennismaking met StatLine had ik niet direct in de gaten dat voor het gebruik van StatLine Java nodig was. Dat zou misschien nog ergens vermeld kunnen worden.

StatLine tabellen uit SSB

Recent is een dertiental StatLine tabellen gemaakt met gegevens uit het Sociaal Statistisch Bestand (SSB). Het gaat om:

- 1 Bevolking naar inkomensbron
- 2 Werknemers en loon
- 3 Uitkeringsontvangers (kenmerken)
- 4 Personen met werk, uitkering of pensioen
- 5 Werknemers en loon (economische activiteit)
- 6 Loonverdeling werknemers naar geslacht e.a.
- 7 Ondernemers en loon/winst
- 8 Arbeidsstromen blijvers naar economische activiteit
- 9 Inkomensbron vóór instroom/na uitstroom uit arbeid
- 10 In- en uitstroom uitkeringen
- 11 Stromen in de arbeidsdeelname
- 12 Niet-Nederlandse immigranten, inkomstenbron en geboorteland
- 13 Niet-Nederlandse immigranten, inkomstenbron en migratiemotief

Eigen tabellen kiezen

Het is nu in StatLine mogelijk om zelf tabellen te detailleren door uitsplitsingen te vragen naar leeftijdsgroep, geslacht of arbeidsperiode.

Stromen in arbeidsdeelname

Met deze SSB tabellen zijn ook nieuwe begrippen in StatLine geïntroduceerd, waaronder begrippen als begin- en eind-periode en sluitende stromen.

Van tabel 11 van de SSB-StatLinetabellen wil ik een screenshot laten zien hoe dit werkt.

Tabel 11

Arbidsdeelname totale bevolking naar leeftijd en geslacht: stromen				
Geslacht	Totaal mannen en vrouwen	Leeftijd begin periode	Totaal leeftijd	Periode
				27 september 2002 tot 26 september 2003
Onderwerpen	Bevolking, begin periode	Integral gebouwtijd begin periode		
	Totaal bevolking, begin periode	Verlozen, begin periode	Niet verlozen, begin periode	
Arbidsaansluitpunten, einde periode	3000 personen			
Totaal arbeidsaansluitpunten einde periode		15 181,3	7 768,6	8 414,8
Werkloze, einde periode		7 888,2	7 188,4	501,8
Niet werkloze, einde periode		8 249,6	533,1	7 716,4
Emigratie/sterfte tijdens periode		243,6	47,8	195,8

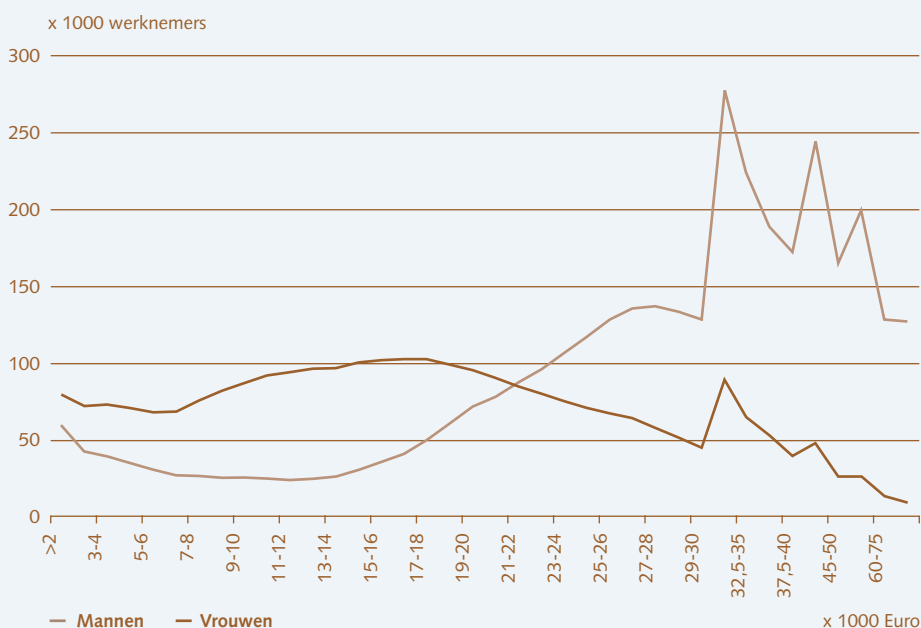
Sexeverschillen onder werknemers met een jaarloon

Uit StatLine heb ik een tabel kunnen samenstellen over verschillen in lonen tussen de sexen. De figuur laat de verdeling van werknemers over jaarloongroepen zien. Een jaarloon is gedefinieerd als het bruto loon voor sociale verzekeringen (BLSV). Zoals verwacht is de verdeling over de groepen zeer ongelijk tussen mannen en vrouwen. De data zijn zeer betrouwbaar, want de gegevens waarop de cijfers uit de gebruikte tabel gebaseerd zijn, waren integraal beschikbaar uit het SSB. Ongeveer 10% van de werknemers had geen bruto jaarloon of van deze werknemers waren deze gegevens niet bekend. Voor deze werknemers is een theoretisch jaarloon geschat aan de hand van hun fiscale loon.

Tabel 6

Werknemers: loonverdeling naar geslacht, leeftijd en herkomstgroepering				
Geslacht	Vrouwen	Leeftijd	Arbidsperiode	
		0 tot 15 jaar		Volledig jaar werknemer
				Onderwerpen
Burgerlijke staat		Herkomstgroepering	Loonklasse	Aantal werknemers, totaal
				x 1 000
Totaal burgerlijke staat	Totaal herkomstgroeperingen	Totaal alle loonklassen	2002, vrijdag 27 september	2,7
		Minder dan 2 duizend euro	2002, vrijdag 27 september	2,5
		2 tot 3 duizend euro	2002, vrijdag 27 september	0,1

Figuur 1. Verdeling werknemers met een jaarloon over jaarloongroepen

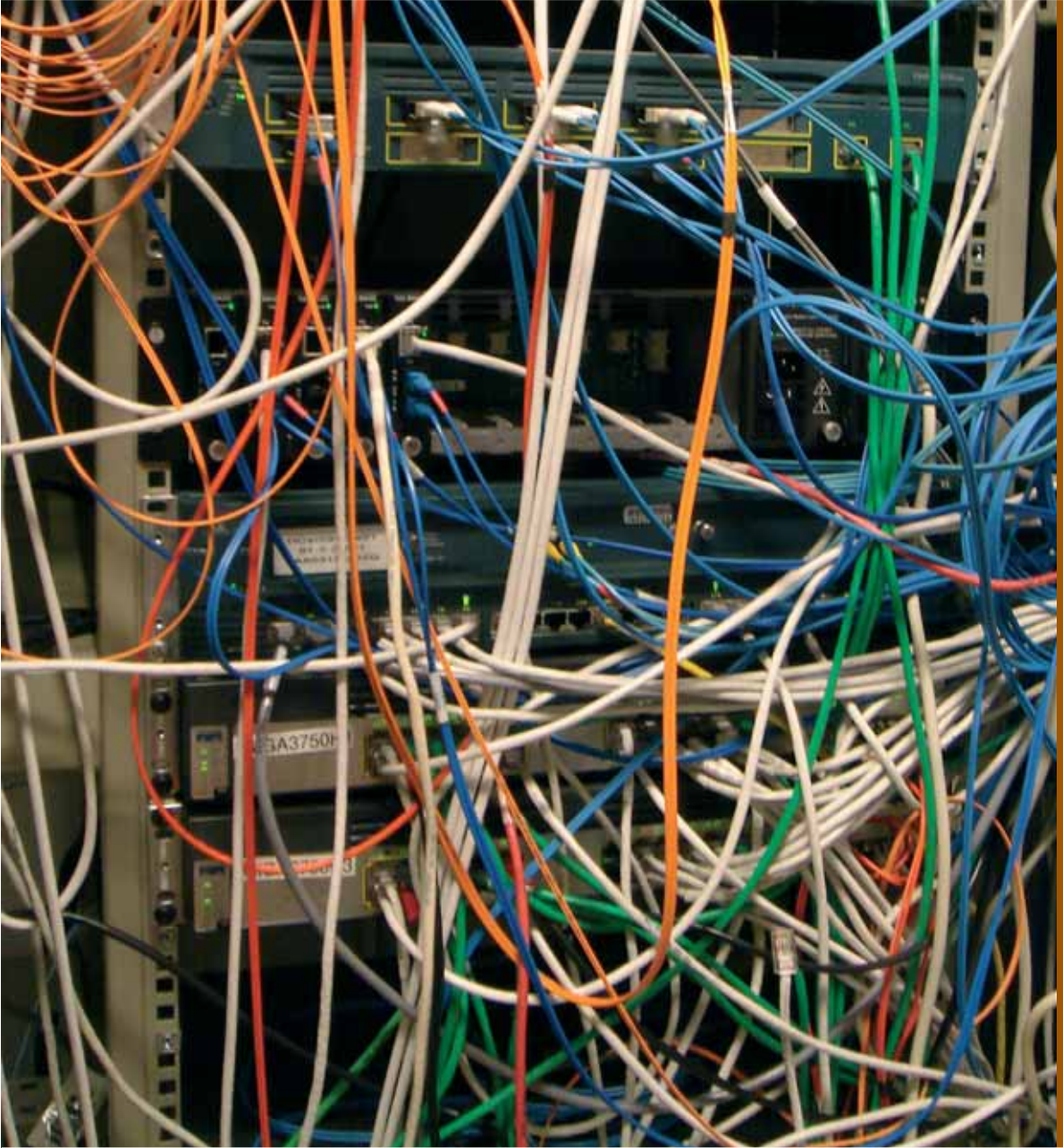


Gebruik voor onderzoek

De data-tabellen van StatLine bieden zeer diverse mogelijkheden voor onderzoek. Alleen al de tabellen bieden vele toepassingen voor beschrijvend onderzoek. Om bijvoorbeeld onderzoek met micro-datacontext data te genereren, zoals werkloosheidscijfers per regio of de verdeling van de gestandaardiseerde gezondheidsindex naar regio. Maar ook voor exploratie van bias in een dataset en voor eventuele weging, bieden de StatLinetabellen de mogelijkheid van een snelle en goedkope check. Ikzelf gebruik de data uit StatLine bijvoorbeeld voor de dataset van de doorlopende Loonwijzer websurvey. Omdat ik verwacht dat er een bias is naar beroep, doe ik op dit moment een onderzoekje daarnaar, waarbij ik een tabel gebruik, bestaande uit beroep op 3-digit niveau * geslacht * jaartal.

Conclusie

De SSB tabellen uit StatLinebruikbare data zijn heel bruikbaar voor diverse onderzoeksdoeleinden. Ze bevatten betrouwbare data, maar ook gedetailleerde data. En niet in de laatste plaats: StatLine zelf is erg gebruikersvriendelijk. Het is prettig om zelf tabellen in een gewenste 'gelaagdheid' te kunnen aanmaken, zonder tussenkomst van anderen. En tenslotte is het heel prettig dat de tabellen naar eigen ontwerp ook downloadable zijn, en nog wel 'for free'!





Koos Arts (CBS)

Reactie op: SSB-tabellen in StatLine.

Inleiding

Mevrouw Tijdens concludeert dat de SSB-tabellen bruikbare data opleveren voor diverse doeleinden. Kenmerken van de data zijn volgens haar betrouwbaarheid en veel detail.

Toch zijn de SSB-StatLine tabellen slechts een *beperkte* representatie van de beschikbare data. Ik wil graag ingaan op de uitgangspunten die wij hebben gekozen bij het ontwerpen van de tabellen. Daarna geef ik aan welke beperkingen dat oplegt aan de te publiceren Stat-Line tabellen uit het SSB. Tot slot komen nieuwe tabellen die volgend jaar verschijnen aan de orde.

Uitgangspunten huidige SSB-StatLine tabellen

Deze uitgangspunten zijn:

- Er is gekozen voor een peilmoment, namelijk de laatste vrijdag van september in een jaar (einde derde kwartaal). De keuze is gebaseerd op het feit dat de periode september/oktober voor de arbeidsmarkt een relatief stabiele periode is. Voor vrijdag is gekozen om week-seizoensinvloeden tussen jaren uit te schakelen. Zo worden met name in het weekend banen beëindigd terwijl op maandag relatief veel banen starten. Dit effect wil je niet in de jaar-op-jaar vergelijking hebben.
- De gepresenteerde gegevens moeten “nieuw” zijn. Daarbij gaat het om:
 - thema-overstijgende uitkomsten (bijvoorbeeld de bevolking naar inkomensbron)
 - stroomgegevens op basis van transities tussen de laatste vrijdag van september in een jaar en de laatste vrijdag van september een jaar later
 - nog niet eerder beschreven populaties (bijvoorbeeld ondernemers of werknemers naar herkomstgroepering).
- De tabellen moeten relatief veel detail bevatten. Dit kan vanwege het integrale karakter van de SSB-gegevens.
- De SSB-tabellen zijn thematisch gerangschikt op StatLine.

Deze uitgangspunten leggen beperkingen op aan de huidige tabellen op StatLine.

- In principe zijn er per jaar 365 (of 366) peilmomenten per verslagjaar mogelijk. Met name kunnen genoemd worden de laatste vrijdagen van andere kwartalen (maart, juli en december). We sluiten niet uit dat we in de toekomst ook andere kwartalen publiceren. Naast een peilmoment is er ook de mogelijkheid om volumecijfers of gemiddelden te publiceren.
- We willen bijvoorbeeld geen “nieuwe” werkgelegenheidsstatistiek maken. Daarvoor zijn recentere bronnen beschikbaar, zoals de Enquête Werkgelegenheid en Lonen (EWL), de Enquête Beroepsbevolking (EBB) en de Arbeidsrekeningen (AR). De toegevoegde waarde van het SSB zit in verdere detaillering van bijvoorbeeld de werkgelegenheid, zoals de verdeling naar herkomstgroepering en transities. Overigens zullen standgegevens over werkgelegenheid uit het SSB afwijken van bijvoorbeeld de EWL en de EBB als gevolg van populatie-, definitie- en peilmomentverschillen. Stroomgegevens zijn nu beperkt tot jaar-op-jaar-transities. Door de keuze van een aantal andere peilmomenten neemt het aantal mogelijkheden om transities te bepalen flink toe. Voor sommige dynamische verschijnselen, zoals de stromen van en naar werkloosheidsuitkeringen is een kwartaal-op-kwartaal-transitie tabel wellicht een goede aanvulling op de jaar-op-jaar tabel.
- Alhoewel er al veel detail zit in de tabellen, zijn er altijd beperkingen. Zo is het aantal selectievariabelen beperkt tot maximaal 5 à 6. Bovendien moeten keuzes gemaakt worden bij het aantal categorieën van variabelen. Zo wordt op dit moment de economische activiteit in 16 bedrijfstakken opgesplitst terwijl er ruim 1100 categorieën beschikbaar zijn. Om u een idee te geven van het detail dat nu wordt gegeven:
 - het aantal werknemers van Marokkaanse herkomst dat tussen september 2002 en 2003 is ingestroomd op de arbeidsmarkt in de landbouw en visserij bedraagt afgerond vier honderd. Driehonderd hiervan hadden eind september 2002 geen inkomensbron.
 - de dertien gepubliceerde SSB-StatLinetabellen bevatten ruim 25 miljoen cellen.
- Door de thematische opzet van StatLine staan de SSB-tabellen niet bij elkaar en zijn lastig te vinden. Er is een lijst met links beschikbaar.

Nieuwe SSB-tabellen

Concreet komen in de eerste helft van 2006 de volgende gegevens uit het SSB op StatLine:

- de bevolking per wijk/buurt naar inkomensbron (in het kader van de Kerncijfers Wijken en Buurten),
- arbeidsmarktpositie van afgestudeerden in het hoger onderwijs,
- verdachten naar herkomstgroepering.

In de tweede helft van 2006 komt er een update van 2004 van de eerder beschreven tabellen.

Er komen nog steeds nieuwe gegevens in het SSB beschikbaar, zoals opleiding, ouder-kindrelaties en gegevens over huishoudens. Ook hiervoor zullen nieuwe tabellen ontworpen worden. Wij nodigen u uit om uw informatiebehoefte kenbaar te maken.

De gegevens uit de geplaatste StatLinetabellen zijn geheel gratis, uiteraard onder bronvermelding. Wanneer u een verdere detaillering wilt van een StatLinetabel, dan is dat tegen vergoeding mogelijk (1 à 2 weken levertijd). Ook andere maatwerkopdrachten behoren tot de mogelijkheden. De levertijden daarvoor zijn onder andere afhankelijk van de omvang en complexiteit van de opdracht en de beschikbare capaciteit.





Bart F.M. Bakker (CBS), Anne Marthe Bouman (Readers Digest)
& Leo van Toor (CBS)

Opleidingsniveau uit registers: nieuwe bronnen, maar nog niet compleet.

In deze bijdrage wordt de ontwikkeling van een variabele opleidingsniveau gepresenteerd op basis van een combinatie van register- en enquête-informatie. Nader onderzoek moet uitwijzen of de resulterende variabele bruikbaar is voor onderzoek naar de rol van het onderwijs in uiteenlopende maatschappelijke processen.

1 Inleiding

Indien maatschappelijke problemen moeten worden opgelost, wordt door beleidsmakers dikwijls verwezen naar het onderwijs. Het onderwijs kan bijvoorbeeld bijdragen aan het verbeteren van de arbeidsmarktposities van bevolkingsgroepen door het verhogen van de kwalificaties, het terugdringen van criminaliteit en ander ongewenst gedrag door het overdragen van normen en waarden, of het leren omgaan met ingrijpende levensgebeurtenissen zoals een echtscheiding van ouders. Hoewel de reacties uit het onderwijsveld niet altijd positief zijn, is het inzetten van onderwijs voor het realiseren van uiteenlopende beleidsdoelstellingen niet vreemd. Het opleidingsniveau is een van de beste voorspellers voor de verwerving van lucratieve posities op de arbeidsmarkt en het hangt samen met de kans op ongewenst gedrag en de kwaliteit van de leefsituatie.

In het Sociaal Statistisch Bestand (SSB) van het CBS wordt een groot aantal registers gekoppeld en gecorrigeerd. Daarnaast worden aan deze gekoppelde registers surveys gekoppeld, zodat de informatie uit de enquêtes wordt verrijkt met de registerinformatie (Al & Bakker, 2000; Bakker, Arts & Van Toor, 2005). Het SSB wordt steeds vaker voor allerlei onderzoeksdoeleinden gebruikt. In het SSB is echter tot nog toe geen opleidingsniveau opgenomen op integrale basis, maar alleen op basis van enquêtes. De voordelen die het SSB biedt zoals de hoge mate van regionaal detail en het inzoomen op kleine maatschappelijke groepen vallen daarmee grotendeels weg. Aangezien een registratie van de door een ieder gevolgde opleidingen in Nederland ontbreekt, wordt er momenteel een opleidingsregister opgebouwd vanuit de wel beschikbare informatie uit registers en enquêtes.

Op de lange termijn is het de bedoeling om voor iedereen in Nederland de opleidingshistorie vast te leggen in één register. De opleidingshistorie geeft aan welke opleidingen achtereenvolgens gevolgd en afgerond zijn en wanneer deze opleidingen zijn gestart en beëindigd. Daarvan kunnen op ieder gewenst peilmoment de volgende variabelen worden afgeleid: de opleiding die de persoon op het peilmoment volgt, de hoogst gevolgde opleiding (zonder diploma) en de hoogst behaalde opleiding (met diploma). Deze opleidingen worden allen getypeerd naar het niveau en de richting volgens de Standaard Onderwijs Indeling 2003 (CBS, 2005).

Gezien de beschikbare gegevens is het nu nog niet mogelijk om voor de gehele populatie een complete opleidingshistorie vast te stellen, omdat daarvoor te weinig gegevens uit administraties en enquêtes beschikbaar zijn. In deze bijdrage wordt daarmee slechts een begin gemaakt door de omvangrijkste informatie waarover het CBS beschikt bij elkaar te

brengen en te koppelen aan de Gemeentelijke Basisadministratie persoonsgegevens (GBA). Daarnaast concentreren we ons op het bepalen van de hoogst behaalde opleiding op één peilmoment (ultimo derde kwartaal 2003), omdat dat de meeste recente datum is waarover uit het SSB is gepubliceerd. De doelstelling van dit project is dan ook om voor zoveel mogelijk personen het hoogst behaalde opleidingsniveau op dat peilmoment te bepalen zodat een variabele ontstaat die na koppeling met andere gegevens bruikbaar is als achtergrondvariabele in sociaal-wetenschappelijk onderzoek.

2 De beschikbare bronnen en het opleidingsregister

2.1 Het opleidingsregister

We maken gebruik van vijf verschillende bronnen die verschillen in omvang, kwaliteit en de periode die beschreven wordt. De bronnen zijn in een opleidingsregister opgenomen. Ieder record uit het register is een combinatie van een persoon en een opleiding die door die persoon is gevolgd met een begindatum, een einddatum en de datum waarop dat gegeven is geregistreerd. Ieder record is voorzien van een koppelsleutel, zodat de informatie altijd weer koppelbaar is met de informatie uit andere records en informatie uit andere bronnen die voorzien zijn van deze koppelsleutel. De opleidingsinformatie wordt in eerste instantie in het register opgenomen zoals deze in de oorspronkelijke bron is opgenomen. Dit maakt het mogelijk om ook in de (verre) toekomst bij verandering van onderwijssystemen of –classificaties nieuwe afleidingen te maken voor bijvoorbeeld de hoogst behaalde opleiding. Hieronder worden de afzonderlijke bronnen besproken.

2.2 CRIHO

De IB-groep beheert het Centraal Register Inschrijvingen Hoger Onderwijs (CRIHO) dat alle inschrijvingen van personen in het hoger onderwijs registreert. Zij leveren jaarlijks een bestand aan het CBS met daarin gegevens over ingeschrevenen bij een hbo- of wo-opleiding. Dit omvat naast gegevens als in- en uitschrijfdatum, geslacht en geboortedatum ook gegevens over opleidingsfase (propedeuse, kandidaats, doctoraal, bachelor, master, etc.), opleidingsvorm (voltijd, deeltijd) en examen resultaten. Naast niveau is ook de richting van het onderwijs te bepalen. Naast de opleiding waarvoor iemand is ingeschreven bevatten de CRIHO-bestanden ook beperkte gegevens over de vooropleiding. Ook hiervan wordt gebruik gemaakt.

Het CBS beschikt over CRIHO-bestanden vanaf 1986. De jaargangen zijn in een samenwerkingsproject tussen het Ministerie van Onderwijs, Cultuur en Wetenschappen (OCW) en het CBS onderling gekoppeld en gecorrigeerd. Iedere combinatie van studiejaar en ingeschreven persoon levert een record op. In totaal gaat het om ongeveer 5,5 miljoen records en 1,7 miljoen personen die hoger onderwijs hebben gevolgd en 2,9 miljoen records en 1,6 miljoen personen die hoger onderwijs hebben gevolgd waarvan een vooropleiding is geregistreerd.

2.3 ERR

Het Examen Resultaten Register (ERR) registreert alle examenresultaten van leerlingen in het voortgezet onderwijs. Het Ministerie van OCW levert sinds het schooljaar 1998-1999 jaarlijks een bestand aan het CBS met daarin gegevens over de leerling, de vakkenpakketten en de soort examen binnen het vbo, de mavo, het vmbo, de havo, en het vwo. De schooljaren van 1998-1999 tot en met 2002-2003 zijn in het opleidingsregister opgenomen. In totaal gaat het om ongeveer 853 duizend records en 793 duizend personen waarvan een opleidingsgegeven uit het ERR bekend is.

2.4 EBB

De Enquête beroepsbevolking (EBB) is een steekproefonderzoek onder personen, van 15 jaar en ouder die in Nederland wonen, met uitzondering van personen in instellingen en tehuizen. Voor de EBB wordt elk jaar een steekproef van iets minder dan 1% uit de bevolking getrokken. Elke maand worden ruim 8 duizend huishoudens benaderd. Aan iedereen worden vragen gesteld over de opleiding die gevolgd wordt en de opleidingen die gevolgd zijn met een maximum van zes. Het gaat om vragen over de naam van de opleiding, de soort onderwijs, de richting, actuele fase, laatste behaalde fase en of het diploma behaald is. De hoogst behaalde opleiding en de actuele opleiding zijn vastgelegd met behulp van de SOI-2003. De EBB's van 1995-2003 zijn in het opleidingsregister opgenomen. In totaal is opleidingsinformatie beschikbaar van 883 duizend personen die in totaal 2,2 miljoen opleidingen hebben gevolgd.

2.5 CWI

De Centra voor Werk en Inkomen (CWI) registreren werkzoekenden. Zij leveren maandelijks een bestand aan het CBS met daarin gegevens over de hoogst behaalde opleiding van de werkzoekenden. Ze onderscheiden hierbij vijf categorieën: geen diploma behaald in het voortgezet onderwijs, lbo/mavo/vmbo, mbo/havo/vwo, hbo, en wo. Het CBS beschikt over CWI-bestanden vanaf 1990. Vanaf 1990 zijn er 4,7 miljoen personen ingeschreven (geweest) bij de CWI die gekoppeld kunnen worden. Daarvan zijn 22,9 miljoen opleidingen geregistreerd.

2.6 WSF

Het bestand Wet op Studiefinanciering (WSF) is afkomstig van de Informatie Beheer Groep (IBG) en omvat gegevens over mensen die studiefinanciering ontvangen. Naast studenten in het hoger onderwijs, zijn ook studenten van 18 jaar en ouder die een Beroepsopleidende leerweg (BOL) van het mbo volgen daarin opgenomen. Er is alleen bekend dat ze een opleiding volgen, maar er is geen informatie over de examenresultaten. Vanaf 1999 t/m 2003 zijn de WSF-gegevens gebruikt. De koppelingsrendementen van de verschillende jaarbestanden is nagenoeg 100%. Een record is een combinatie van een studiejaar en een student. In totaal gaat het om ongeveer 3,3 miljoen records en 1,2 miljoen personen.

De bovengenoemde bronnen zijn vrijwel allemaal goed te koppelen. De koppelingsrendementen zijn ruim boven de 90%. Uitzondering vormen de eerste jaargangen van de CWI die minder goed gekoppeld kunnen worden. Omdat onder meer op adresgegevens gekoppeld wordt die pas sinds 1995 integraal beschikbaar zijn, leidt dit waarschijnlijk tot selectiviteit binnen de categorie werkzoekenden. Personen die voor 1995 op een ander adres woonden dan erna worden immers minder goed gekoppeld. De koppelingsrendementen variëren van 70% voor 1990 tot 85% voor 1994.

In de volgende paragrafen worden achtereenvolgens de bewerkingen die hebben plaatsgevonden op de gegevens in het opleidingsregister, de afleiding van de hoogst behaalde opleiding en de validering daarvan besproken.

3 De micro-integratie van de opleidingsgegevens

3.1 Micro-integratie

De in het opleidingsregister opgenomen gegevens zijn afkomstig uit bronnen die verschillen in de beschreven populaties, die verschillen in de wijze waarop de opleidings(historie) bekend is met name in de mate van detail, en die verschillen in kwaliteit. Om het doel te verwezenlijken moeten de gegevens zodanig worden bewerkt dat voor zoveel mogelijk mensen een geldige hoogst behaalde opleiding op het gekozen peilmoment bekend is. Na weging zou dit geschikt moeten zijn als achtergrondvariabele voor sociaal-wetenschappelijk onderzoek. Dit proces van bewerken van gegevens wordt micro-integratie genoemd. De aangrijpingspunten voor het bewerken van de gegevens vinden we in de tegenspraken tussen de bronnen. Dat kan worden veroorzaakt door definitieverschillen, meetfouten en populatieverschillen.

3.2 Definitieverschillen: harmonisatie

Bij de harmonisatie wordt gecorrigeerd voor verschillen in definities van het gemeten kenmerk, de daarbij gebruikte indelingen, de waarnemingsperiode en detaillering. Allereerst moet daarbij vastgesteld worden welke definities toegepast dienen te worden, met andere woorden wat het ankerpunt is waarop we de gegevens uit de verschillende bronnen zullen integreren.

Voor het bepalen van de hoogst behaalde opleiding zijn per persoon drie ingrediënten nodig: de gevolgde opleiding(en), of daarvan het diploma of getuigschrift behaald is en het moment waarop dat is gebeurd. Het harmonisatieproces van deze drie variabelen wordt hieronder beschreven.

Het niveau van de opleiding

Aangezien de CWI-bestanden op dit moment nog een belangrijke bron zijn voor het vaststellen van het opleidingsniveau van zoveel mogelijk personen en de opleidingsinformatie in deze bron zoals die aan het CBS geleverd wordt globaler is dan de Standaard Onderwijs Indeling-2003 (SOI-2003), moeten we genoeg nemen met een globale indeling voor het niveau. Ook kunnen we de richting in de CWI-gegevens nog niet vaststellen. Bij de indeling naar het niveau van de hoogst behaalde opleiding wordt wel zoveel mogelijk aangesloten bij de terminologie van de SOI-2003:

- Primair onderwijs. In de praktijk wordt verondersteld dat iedereen basisonderwijs gevolgd heeft. Als iemand geen diploma in het voortgezet onderwijs heeft behaald, dan wordt het opleidingsniveau 'primair onderwijs'. In de SOI-2003 vallen hier de niveaus 1 (onderwijs aan kleuters) en 2 (primair onderwijs) onder.
- Secundair onderwijs, eerste fase. Algemeen vormend en beroepsvoorbereidend onderwijs en voorbereidend middelbaar beroepsonderwijs dat in directe aansluiting op het basisonderwijs gevolgd kan worden. Daaronder vallen vmbo, vbo, lbo, mavo, en de eerste drie leerjaren van havo en vwo. Hoewel er enige afwijkingen zijn komt dit grotendeels overeen met niveau 3 in de SOI-2003.
- Secundair onderwijs, tweede fase. Algemeen vormend en beroepsonderwijs dat aansluit op het secundair onderwijs lagere trap. Hieronder vallen het mbo, de havo, het vwo en de propedeuse van het wetenschappelijk onderwijs. In de SOI-2003 is dit vrijwel gelijk aan niveau 4.
- Hoger onderwijs, eerste fase. Hieronder vallen hbo-opleidingen (exclusief hbo-masters-opleidingen) en kandidaats en bachelors van het wetenschappelijk onderwijs. In de SOI-2003 is dit niveau 5.

- Hoger onderwijs, tweede en derde fase. Hieronder vallen de mastersopleidingen in het hbo, doctoraal en mastersopleidingen in het wetenschappelijk onderwijs, postdoctoraal onderwijs en onderwijs dat leidt tot een promotie. In de SOI-2003 zijn dit de niveaus 6 en 7.

De opleidingsinformatie uit de onderscheiden bronnen moet worden vertaald naar de bovenstaande definitie. Daarbij wordt voor iedere opleiding die in het opleidingsregister is opgenomen, en dat kunnen voor een persoon meerdere opleidingen zijn, een niveau volgens bovenstaande definitie bepaald. De CWI heeft als enige informatie over de opleiding het niveau zoals boven is beschreven. Dit wordt dus direct overgenomen. Het CRIHO en het ERR hebben hun eigen codelijsten die ook nog eens per jaargang kunnen verschillen. Van deze opleidingsinformatie wordt eerst de SOI-2003 versie 2004-2005 afgeleid, waarbij zowel het niveau als de richting wordt bepaald. Hiervan wordt vervolgens het opleidingsniveau volgens de bovenstaande definitie afgeleid. De opleidingsgegevens uit de EBB zijn voor alle gevolgde opleidingen voorzien van een zogenaamd opleidingsnummer. Hiervan is het niveau en de richting volgens de SOI-2003 afgeleid. Op zijn beurt wordt hiervan het globalere niveau afgeleid. De WSF bevat slechts een driedeling in mbo, hbo, en wo, waarbij sinds de invoering van de bachelor-masterstructuur ook dit gegeven bekend is. Dit wordt direct vertaald in respectievelijk niveau 3, 4 en 5.

Het behalen van diploma of getuigschrift

Iedere opleiding die door personen gevolgd is, is nu in het register voorzien van dezelfde geharmoniseerde niveau-indicatie. Daarmee is nog niet de hoogst behaalde opleiding daarvan bepaald. Daarom is van de in de bronnen beschikbare opleidingsinformatie afgeleid of men van de betreffende opleiding een diploma of getuigschrift heeft behaald. In de CWI ontbreekt dit gegeven. Omdat de CWI echter in beginsel het hoogst behaalde opleiding levert, is steeds aangenomen dat men het diploma behaald heeft. In het CRIHO is het gegeven afgeleid uit de informatie over de opleidingsfase en het examenresultaat. In het ERR is het gegeven afgeleid uit de variabele of men geslaagd is voor het examen waarvoor men op is gegaan. Indien sprake was van het behalen van deelcertificaten, is aangenomen dat men het betreffende opleidingsniveau niet heeft behaald. In de EBB is de informatie afgeleid uit het antwoord op de vraag of men het diploma behaald heeft van de genoemde opleiding. Voor het hoger onderwijs is in de EBB's gevraagd naar het behaalde diploma (bachelor, master, propedeuse, kandidaats, doctoraal, ingenieur, doctoraat, etc). In de WSF ontbreekt het gegeven helemaal. Dit betekent dat we in de WSF alleen kunnen vaststellen of iemand een opleiding van een bepaald niveau heeft gevolgd.

Het moment waarop het diploma of getuigschrift is behaald

Op grond van de combinatie van het niveau van de opleiding en het al dan niet behaald hebben van het diploma of getuigschrift kan nu de hoogste worden bepaald. Om de hoogst behaalde opleiding op een bepaald peilmoment vast te stellen, dient ook te worden beoordeeld op welk moment dit niveau is bereikt. In de CWI is dit gegeven niet opgenomen. De enige informatie die we daarover hebben is dat het diploma behaald is voor de datum waarop iemand in de CWI is geregistreerd. Voor de afleiding wordt daarom de datum waarop het diploma behaald is gelijk gesteld aan de registratiedatum. In het CRIHO is de maand en het jaar waarop het betreffende diploma is behaald geregistreerd. In het ERR is het schooljaar geregistreerd waarin het diploma in het voortgezet onderwijs is behaald. We nemen aan dat dat steeds aan het einde van het schooljaar gebeurt. In de EBB is het jaar vastgelegd waarin met de betreffende opleiding is begonnen en wanneer deze is beëindigd. In de WSF is niet bekend of een diploma behaald is en kan daarvan dus ook niet het moment bepaald worden.

3.3 Minimalisering van meetfouten: de afleiding van hoogst behaalde opleiding

Voor het bepalen van de hoogst behaalde opleiding op een bepaald peilmoment per persoon hebben we nu alle ingrediënten vastgesteld en voor zover mogelijk geharmoniseerd tussen de bronnen. We kunnen nu de afleiding gaan maken voor de hoogst behaalde opleiding voor ultimo derde kwartaal 2003. Hiervoor moeten nog een aantal problemen opgelost worden: inconsistenties tussen de bronnen en het voorkomen van verouderde gegevens.

Inconsistenties tussen de bronnen

We hebben tot nu toe nog niet gekeken of de informatie die over één persoon in het register is opgenomen ook consistent is. Inconsistenties wijzen op fouten van allerlei aard (administratieve fouten, meetfouten in surveys). Deze inconsistenties worden opgelost door naar de kwaliteit van de bronnen te kijken voor de betreffende variabele. Voor registers geldt dat als een variabele voor de registerhouder minder relevant is, deze vaak relatief slecht van kwaliteit is. Verder geldt dat het doel waarmee de registratie wordt bijgehouden van belang is. Als dit doel afwijkt van datgene dat voor de statistiek van belang is, wordt feitelijk iets anders gemeten (Bakker, 2003; Bakker & Arts, 2003).

Voor de bepaling van het opleidingsniveau van personen zijn de registers CRIHO (inclusief de vooropleiding) en ERR het beste van kwaliteit. Ze zijn integraal voor hun doelpopulatie en ze zijn opgezet om de examenresultaten in het voortgezet en hoger onderwijs te registreren. De registerhouder heeft er daarom alle belang bij om deze informatie goed te registreren. De EBB heeft een tamelijk uitgebreide vragenlijst voor de vaststelling van

het opleidingsniveau, waardoor we een behoorlijk vertrouwen daarin hebben. Als echter een keuze gemaakt moet worden tussen CRIHO of ERR en de EBB, gaan de registers voor.

De CWI-gegevens zijn voor het vaststellen van de hoogst behaalde opleiding minder geschikt. Het vermoeden bestaat dat de arbeidsbureaus de opleiding van mensen registreren die op dat moment voor de arbeidsmarkt het meest relevant is. Dit kan in een aantal gevallen een lagere opleiding zijn dan de hoogst behaalde. Verder is uit eerder onderzoek gebleken dat de aanname dat van de geregistreerde opleiding het diploma is behaald niet terecht is (zie tabel 1)

Tabel 1. Hoogst behaalde opleiding volgens de CWI naar hoogst behaalde opleiding volgens de EBB, 1995-1998 (kolompercentages)

	<i>CWI-opleidingsniveau</i>			
	<i>primair onderwijs</i>	<i>secundair onderwijs, eerste fase</i>	<i>secundair onderwijs, tweede fase</i>	<i>hoger onderwijs</i>
	%			
EBB-opleidingsniveau				
primair onderwijs	65	19	2	0
secundair onderwijs, eerste fase	23	49	17	2
secundair onderwijs, tweede fase	10	31	73	18
hoger onderwijs	1	1	7	80
Totaal	100	100	100	100

Bron: Arts (2000)

Als de CWI- en EBB-gegevens van 1995-1998 aan elkaar worden gekoppeld blijkt dat in hetzelfde jaar ongeveer eenderde afwijkt in het opleidingsniveau (Arts, 2000). Het komt veelvuldig voor dat de CWI één niveau lager registreert dan de EBB. Van de personen die volgens de CWI een diploma op het niveau van het VMBO hebben, blijken in de EBB opvallend vaak een diploma op mbo-niveau te hebben. Dit wordt mogelijk veroorzaakt doordat mensen combinaties van werken en leren (vroeger leerlingwezen, tegenwoordig Beroepsbegeleidende leerweg, BBL) niet altijd als een mbo-opleiding zien. Dit kan overigens twee kanten opwerken: in de beantwoording van de EBB of bij de balie van het arbeidsbureau kan dat misgaan. Verder blijkt dat de CWI dikwijls één niveau hoger registreert dan de EBB. Het blijkt dat in 60% van deze gevallen de hoogst behaalde opleiding volgens de CWI correspondeert met de hoogst gevolgde opleiding volgens de EBB. In die gevallen is de opleiding niet afgerond met een diploma. Overigens moet worden opgepast om de resultaten van de CWI per definitie als fout te bestempelen als deze afwijken van de EBB. Ook de EBB bevat fouten, bijvoorbeeld vanwege geheugeneffecten en sociaal-wettelijk antwoordgedrag.

De WSF bevat geen informatie over het behalen van een diploma. Uit de WSF-bestanden kunnen de personen worden gehaald die een mbo-, hbo- of wo-opleiding hebben gevolgd. Als sprake is van inconsistenties tussen bronnen dan is bij de afleiding de volgende volgorde gehanteerd: CRIHO (inclusief de vooropleidingen), ERR, EBB, CWI, WSF.

Ten minste houdbaar tot

Voor de mensen van wie een opleidingsgegeven bekend is, dateren gegevens soms van jaren her. Dit hoeft geen juiste weerspiegeling te zijn van het bereikte opleidingsniveau op peilmoment. Er dient daarom nagegaan te worden voor wie het opleidingsgegeven een geldige meting is van het opleidingsniveau op peilmoment. Voor deze bijdrage kiezen we voor het peilmoment van ultimo derde kwartaal 2003.

In wezen gaat het hierbij om de vraag hoe lang een bepaald opleidingsgegeven geldig blijft. Voor mensen die ooit een wetenschappelijke opleiding hebben afgesloten met een doctoraal diploma geldt dat natuurlijk, omdat je nooit een hoger diploma kunt halen. Kun je echter aannemen dat iemand die vijftig was toen zijn hoogst behaalde opleiding mbo was daarna geen hoger opleidingsniveau meer behaalt? En iemand van dertig? En iemand van vijftentwintig? Dit is een vraag die empirisch kan worden beantwoord. Daarvoor hebben we voor iedere combinatie van opleidingsniveau en leeftijd op registratiedatum bepaald hoe groot de kans is dat iemand een hoger opleidingsniveau behaalt. Zo kunnen we de leeftijdsgrenzen bepalen om een gegeven geldig te verklaren of niet. We stellen de eis dat aan niet meer dan 5% van de mensen een te laag opleidingsniveau mag worden toegekend. Als blijkt dat na een bepaalde leeftijd de kans dat iemand nog een hoger opleidingsniveau haalt kleiner is dan 5%, dan worden alle gevallen waarop dat opleidingsniveau als hoogste is waargenomen voorbij die leeftijdsgrens als geldig bestempeld.

Om de leeftijdsgrenzen empirisch vast te stellen gebruiken we de EBB's van 2001-2003, waarin de complete opleidingshistorie is opgenomen. Deze grenzen zijn bepaald voor de verschillende overgangen bo-vo; vmbo-mbo/havo/vwo; mbo/havo/vwo-hbo/wo en hbo-wo. De resultaten zijn weergegeven in tabel 2. De opleidingsgegevens van de mensen van wie na hun twintigste geregistreerd is dat zij geen diploma's in het voortgezet onderwijs hebben behaald worden geldig verklaard. De opleidingsgegevens van de mensen van wie na hun 32^e is geregistreerd dat zij een hbo-diploma hebben behaald worden geldig verklaard, uiteraard steeds als er geen hogere opleiding van die mensen geregistreerd is.

Niet alle opleidingsgegevens van de mensen die op de registratiedatum jonger zijn dan de in tabel 2 aangegeven leeftijdsgrenzen zijn ongeldig. Van personen die jonger zijn, zijn de gegevens geldig als die zo recent zijn dat het onwaarschijnlijk is dat een hoger diploma is

behaald. Ook deze grens is empirisch vastgesteld: per opleidingsniveau is de grens bepaald waarbij minder dan 5% een hogere opleiding heeft behaald:

Tabel 2. Leeftijdsgrenzen boven welke de kans op overgang naar het naast hogere opleidingsniveau kleiner is dan 5%

Van	Naar	<i>Leeftijdsgrens in jaren</i>
basisonderwijs	voortgezet onderwijs	20
vmbo/lbo/vbo/mavo	mbo/havo/vwo	30
mbo/havo/vwo	hbo/wo	30
hbo	wo	32

Bron: EBB 2001-2003

- Voor de mensen met primair onderwijs: als de gegevens korter dan 2 jaar geleden zijn geregistreerd is de kans kleiner dan 5% dat al een hogere opleiding afgerond is. Dit wordt overigens alleen maar waargenomen via de EBB. Deze gevallen worden alsnog geldig verklaard.
- Vmbo, lbo, vbo, mavo: als de gegevens korter dan 2 jaar geleden geregistreerd zijn, is de kans kleiner dan 5% dat een hogere opleiding behaald is. Is het gegeven recenter dan twee jaar dan is het wel geldig verklaard.
- Mbo/havo/vwo: als de diploma's korter dan 4 jaar geleden zijn behaald dan is de kans kleiner dan 5% dat een hogere opleiding behaald is. Is de registratie van het mbo/havo/vwo-diploma als hoogst behaalde opleiding vier jaar of ouder, dan is dit gegeven niet geldig. Is het gegeven recenter dan is het gegeven wel geldig.

Aangezien we gebruik maken van integrale registers van het hoger onderwijs sinds 1986 en het algemeen voortgezet onderwijs sinds 1999, kunnen we drie aanvullende regels formuleren:

- Voor iedereen met niveau 2 (vmbo, lbo, vbo, mavo) als hoogst behaalde niveau geldt dat als het diploma is behaald in of na 1998, deze alsnog geldig worden verklaard.
- Voor iedereen met niveau 3 (mbo/havo/vwo) als hoogst behaald niveau geldt dat als het diploma in 1986 of later is behaald, deze alsnog geldig worden verklaard. Als deze personen een hogere opleiding behaald zouden hebben dan was dit bekend geweest (uit de CRIHO-gegevens).
- Voor iedereen met niveau 5 (wo) als hoogst gevolgd en behaald niveau geldt dat dit gegeven geldig is. Een hoger opleidingsniveau onderscheiden we immers niet.

Verder dient de hoogst gevolgde opleiding voor de peildatum afgerond te zijn met een diploma. Dit geeft maar een beperkt aantal correcties, omdat alleen van de CWI en CRIHO-register er informatie is van een recentere datum dan 2003. Omdat in de CWI-bestanden geen informatie is opgenomen over de datum waarop het diploma is behaald en we alleen weten dat een diploma voor de registratiedatum is behaald, wordt de jaargang 2004 niet meegenomen in de bepaling van het opleidingsniveau.

In het opleidingsregister is informatie opgenomen over de opleidingen die door personen zijn gevolgd. Deze hoeven niet altijd afgesloten te zijn met een diploma. Voor die gevallen kan het echter wel iets zeggen over het hoogst behaalde opleidingsniveau, omdat voor het volgen van opleidingen op een bepaald niveau ook een vooropleiding nodig is die toegang geeft tot dat niveau. Daarom hebben we aanvullende regels opgesteld:

- Als iemand een opleiding heeft gevolgd op niveau 2 maar daarvan geen diploma heeft behaald, wordt het opleidingsniveau gesteld op 1.
- Als iemand examen heeft gedaan voor de havo of het vwo maar daarvan geen diploma heeft behaald, wordt het niveau bepaald op 2. Ze hebben immers de overgang van klas 3 naar klas 4 meegemaakt van havo/vwo of zijn ingestroomd vanuit de mavo of vmbo.
- Als iemand hoger onderwijs heeft gevolgd en daar geen diploma van heeft behaald is het niveau op 3 gesteld (mbo/havo/vwo). Dit is immers het niveau dat nodig is om toegang te verkrijgen tot hoger onderwijs.

De laatste afleidingsregel is de meest simpele: alle mensen van 16 jaar en jonger krijgen het laagste opleidingsniveau toegekend, tenzij uit de aanwezige opleidingsinformatie blijkt dat ze een hoger niveau hebben.

Uit de aanwezige register- en enquête-informatie is aan 5,9 miljoen mensen een opleidingsniveau toegekend. Dat betekent dat voor ongeveer 430 duizend personen de aanwezige opleidingsinformatie niet geldig is verklaard. In tabel 3 is de bron van de geldig verklaarde opleidingsniveaus aangegeven. In totaal wordt aan 9,2 miljoen mensen (=57%) van de bevolking in Nederland een opleidingsniveau toegekend. Daarvan is 3,3 miljoen geïmputeerd op basis van de leeftijd. Voor de bevolking van 15 jaar en ouder is van 5,9 miljoen mensen een geldige score bekend (=45%).

Tabel 3. Bronnen van geldige opleidingsniveaus

	Opleidingsniveau						% uit bron
	<i>bo</i>	<i>vmbo, lbo, vbo, mavo</i>	<i>mbo, havo, vwo</i>	<i>hbo</i>	<i>wo</i>	<i>totaal</i>	
	x 1000						
CRIHO	0	0	0	526	289	815	9
Vooropl CRIHO	0	3	663	28	4	698	8
ERR	0	431	20	0	0	451	5
EBB	43	214	314	80	35	686	7
CWI	313	956	1302	314	144	3029	33
hoogst gevolgd	39	146	38	0	0	223	2
leeftijd <17 jaar	3337	0	0	0	0	3337	36
Totaal	3732	1750	2337	948	472	9239	100

Bron: SSB-opleidingsregister.

De validiteit van de hoogst behaalde opleiding op microniveau

De validiteit van de hoogst behaalde opleiding wordt eerst op microniveau beoordeeld. Door voor personen de afleiding te bepalen op grond van uitsluitend de registers, deze te koppelen aan de personen van wie een geldige hoogst behaalde opleiding uit de EBB-2003 bekend is, kunnen deze vergeleken worden. In totaal komt het opleidingsniveau uit de registers en de EBB voor 74% van de gevallen overeen. Gegeven het feit dat een groot deel van de opleidingsgegevens afkomstig is van de CWI en hiervan bekend is dat hiervan ongeveer tweederde overeenstemt met de EBB (zie Arts, 2000) is dit een redelijk resultaat. De overeenstemming wordt groter naarmate het opleidingsniveau volgens de registers hoger is.

Tabel 4. Vergelijking van het hoogst behaalde opleidingsniveau uit registers en EBB Opleidingsniveau

	Opleidingsniveau EBB					
	<i>bo</i>	<i>vmbo, lbo, vbo, mavo</i>	<i>mbo, havo, vwo</i>	<i>hbo</i>	<i>wo</i>	<i>totaal</i>
	%					
Opleidingsniveau registers						
<i>bo</i>	57	31	10	1	0	100
<i>vmbo, lbo, vbo, mavo</i>	10	62	27	1	0	100
<i>mbo, havo, vwo</i>	1	14	77	6	1	100
<i>hbo</i>	0	1	14	80	4	100
<i>wo</i>	0	1	7	9	82	100
Totaal	5	23	43	20	8	100

Bron: SSB-opleidingsregister.

In een latere fase voegen we de van de EBB afgeleide opleidingsniveaus weer toe. Daarbij worden de prioriteitsregels gehanteerd zoals die eerder in deze paragraaf zijn beschreven. Als van een persoon een geldig opleidingsniveau bekend is uit het CRIHO of het ERR, dan wordt geen EBB-informatie toegevoegd. Is het gegeven afkomstig uit de CWI en is het gegeven uit de EBB geldig, dan wordt het EBB-gegeven gebruikt. Door deze werkwijze te volgen wordt het hierboven geschetste resultaat nog verbeterd. Voor de beoordeling van de validiteit van het uiteindelijke resultaat wordt op macroniveau een vergelijking uitgevoerd tussen de uitkomsten van het opleidingsregister en de EBB.

3.4 Populatieverschillen: bepalen gewichten

Bij de completering wordt gecorrigeerd voor populatieverschillen. We stellen eerst de doelpopulatie vast, vergelijken deze met de populatie waarvan waarnemingen bekend zijn en we wegen de gegevens zodanig dat een goede afspiegeling wordt verkregen van de doelpopulatie.

De doelpopulatie

In dit geval definiëren we de bevolking in Nederland ultimo derde kwartaal 2003 als de doelpopulatie. We hebben eerder gesteld dat dit onderzoek een variabele moet opleveren die bruikbaar is in sociaal-wetenschappelijk onderzoek en dat kan dus ook allerlei subpopulaties betreffen. Of de resulterende variabele ook daarvoor geschikt zal zijn, behandelen we verderop.

Verschillen met de waargenomen populatie

De verzameling mensen van wie een geldige opleiding is geregistreerd is selectief op een aantal punten. In de eerste plaats zijn de jongere generaties sterk oververtegenwoordigd, omdat die in de recente registers (CRIHO, ERR) voorkomen. In de tweede plaats zijn hoger opgeleiden oververtegenwoordigd, omdat de met succes afgesloten hbo- en woopleidingen sinds 1986 zijn opgenomen. In de derde plaats zijn de werkzoekenden oververtegenwoordigd die ingeschreven hebben gestaan bij de CWI of zijn voorgangers. In de vierde plaats zijn mensen die een mbo-diploma hebben behaald ondervertegenwoordigd, omdat we niet over een mbo-register beschikken. Deze selectiviteit is echter niet zodanig dat we structureel een aantal groepen missen. Dat komt vooral omdat negen jaargangen van de EBB zijn opgenomen. Hierdoor is het mogelijk om een verdeling van het opleidingsniveau in de Nederlandse bevolking te schatten door alle personen met een geldig opleidingsniveau in het opleidingsregister van een gewicht te voorzien.

Tabel 5. Opleidingsniveau uit het opleidingsregister (zonder weging) en de EBB (na weging) 15-64 jaar, 2003

	<i>Opleidingsregister</i>	<i>EBB</i>
	%	
Opleidingsniveau		
bo	11	10
vmbo, lbo, vbo, mavo	28	25
mbo, havo, vwo	38	41
hbo	15	16
wo	8	8
Totaal	100	100

Bron: SSB-opleidingsregister en EBB

Hoewel we weten dat het resultaat enigszins selectief is, komt de verdeling van de bevolking van 15 jaar en ouder tussen het register en de EBB behoorlijk overeen (tabel 5). In het opleidingsregister zijn in vergelijking met de EBB wat minder mensen op niveau 3 opgenomen en wat meer op niveau 2. Dit zal waarschijnlijk te maken hebben met het feit dat we niet over een register van mbo-afgestudeerden beschikken.

Om de selectiviteit van het uiteindelijke resultaat te onderzoeken hebben we een aantal vergelijkingen uitgevoerd met de resultaten uit de EBB, onder meer naar leeftijd, geslacht, herkomstgroepering, arbeidsmarktpositie, en voor de werkenden de economische activiteit van het bedrijf waar men werkt en gemeente. Daarvan laten we niet alle resultaten zien. We beperken ons tot een vergelijking naar leeftijd, omdat we verwachten dat de schattingen voor de ouderen slechter zijn dan voor de jongeren (tabel 6) en naar arbeidsmarktpositie, omdat we verwachten dat uitkeringsgerechtigden vaker bij de CWI ingeschreven zijn en daarom in het register zijn oververtegenwoordigd (tabel 7).

Tabel 6. Opleidingsniveau uit het opleidingsregister (zonder weging) en de EBB (na weging) 15-64 jaar naar leeftijd, 2003

		<i>Hoogst behaalde opleiding</i>					<i>dekkingsgraad register</i>
		<i>bo</i>	<i>vmbo, lbo, vbo, mavo</i>	<i>mbo, havo, vwo</i>	<i>hbo</i>	<i>wo</i>	
		%					
Leeftijd	bron						
15-24 jaar	EBB	15	39	40	6	1	
	register	21	41	33	5	1	86
25-34 jaar	EBB	5	17	46	20	12	
	register	1	14	49	24	13	67
35-44 jaar	EBB	7	21	44	17	10	
	register	6	22	42	19	12	45
45-54 jaar	EBB	11	24	39	18	8	
	register	12	32	32	17	7	40
55-64 jaar	EBB	15	29	35	14	7	
	register	20	36	26	13	5	30
totaal 15-64	EBB	10	25	41	16	8	
Jaar	register	11	28	38	15	7	53

Bron: SSB-opleidingsregister en EBB.

In tabel 6 zijn enkele patronen te ontdekken. Ten eerste vinden we bij de jongste leeftijdsgroep dat er in het register een oververtegenwoordiging is van mensen zonder diploma. Dit zou te maken kunnen hebben met de beslissing om alle mensen die jonger zijn dan 17 jaar en die niet met diploma zijn geregistreerd in het opleidingsregister het laagste niveau toe te kennen. In ieder geval vinden we bij de leeftijdscategorie 25-34 jaar dit fenomeen in het geheel niet terug. Overigens kan ook selectieve non-respons in de EBB dit verklaren, als het zo is dat juist in de jongste leeftijdscategorie de ongediplomeerden minder responderen dan de ouderen. Verder vinden we duidelijke aanwijzingen dat de mbo-gediplomeerden in het register ondervertegenwoordigd zijn, met name in de leeftijdscategorieën vanaf 35 jaar. Dat dit bij de jongeren niet zichtbaar is, wordt waarschijnlijk verklaard door het feit dat dit wordt gecompenseerd door de havo- en vwo-gediplomeerden die geregistreerd zijn via het ERR. Verder zien we dat de dekkingsgraad van het register lager is naarmate de leeftijd hoger is.

In tabel 7 zien we de verwachte uitkomst dat gepensioneerden ondervertegenwoordigd en uitkeringsgerechtigden in het kader van de ww of abw en onderwijsvolgenden oververtegenwoordigd zijn in het register. Verder vinden we relatief weinig zelfstandigen en arbeidsongeschikten in het opleidingsregister. De verdeling naar opleidingsniveau ziet er echter wel redelijk plausibel uit. Werknemers en zelfstandigen zijn relatief goed opgeleid, uitkerings-

gerechtigden zijn in het algemeen minder hoog opgeleid. Een uitzondering wordt gevormd door de mensen met een ww-uitkering. Deze zijn veel beter opgeleid dan bijvoorbeeld mensen met een abw-uitkering, maar recht op een ww-uitkering is er ook alleen na een werkzame periode van enige duur. Deze resultaten laat zien dat de gegevens uit het opleidingsregister gewogen moeten worden om uitspraken te doen over het opleidingsniveau van de Nederlandse bevolking.

Tabel 7. Opleidingsniveau uit het opleidingsregister (zonder weging) naar arbeidsmarktpositie, 0 jaar en ouder, 2003

	<i>bo</i>	<i>vmbo, lbo, vbo, mavo</i>	<i>mbo, havo, vwo</i>	<i>hbo</i>	<i>wo</i>	<i>Dekkingsgraad register</i>
	%					
Arbeidsmarktpositie						
werknemer	3	26	40	21	10	51
zelfstandige	4	21	42	19	13	35
ao-uitkering	18	37	32	10	3	38
ww-uitkering	7	27	43	16	7	96
abw-uitkering	25	39	27	6	4	80
overige uitkering	12	35	34	12	6	60
gepensioneerd	25	35	26	11	4	10
onderwijsvolgend	78	9	11	1	0	97
Inactief	70	12	13	4	2	59

Bron: SSB-opleidingsregister.

De wegingsprocedure

De weging van de gegevens uit het opleidingsregister is niet eenvoudig. Gebruikelijke methodes waarbij de startgewichten (omgekeerd evenredig met de insluitkansen) zodanig worden aangepast (gekalibreerd) dat bekende populatietotalen worden gereproduceerd kan hier niet worden toegepast. We weten immers de insluitkansen niet op basis van een steekproefdesign. Er bestaan verschillende mogelijkheden om de gewichten te bepalen.

Een eerste mogelijkheid is dat de registers worden beschouwd als een integrale beschrijving van specifieke deelpopulaties. Deze krijgen dan gewicht één vanuit de gedachte dat ze alleen zichzelf representeren en geen anderen. De EBB's zijn op te vatten als steekproeven uit de populatie en worden daarom gebruikt voor de bijschatting van de bevolking in Nederland die niet tot de deelpopulaties behoren die in de registers beschreven worden. Voor deze EBB's worden eerst de startgewichten afgeleid van de in de afzonderlijke EBB's voorkomende gewichten die daarna worden gekalibreerd zodat bekende populatietotalen worden gereproduceerd.

Een tweede mogelijkheid is dat naast de EBB ook de gegevens van de CWI beschouwd worden als een steekproef uit de bevolking in Nederland. Weliswaar weten we dat de CWI-gegevens selectief zijn, maar omdat het gaat om mensen die sinds 1990 zijn ingeschreven als werkzoekende bij de arbeidsbureaus, wordt tot op zekere hoogte ook een bredere populatie beschreven. De selectiviteit is daardoor niet zo groot en misschien komt deze selectiviteit wel in de buurt van die van de EBB waar immers sprake is van selectieve non-respons. Nadat startgewichten zijn toegekend aan de EBB- en CWI-records, kan dezelfde procedure worden gevolgd als bij de eerste mogelijkheid.

De keuze voor een bepaalde wegingsmethode wordt binnenkort gemaakt. De gebruiker die de microdata analyseert zal daarvan overigens weinig merken. Steeds worden bestanden geleverd waarin naast de hoogst behaalde opleiding ook een bijbehorende wegingsfactor is opgenomen. De vraag die nog rest is of de gewichten ook gebruikt kunnen worden om subpopulaties te beschrijven. Dit geldt in ieder geval wel als de subpopulatie wordt afgebakend met variabelen die in de bepaling van de definitieve gewichten zijn betrokken (bijvoorbeeld leeftijd, geslacht, herkomstgroepering, arbeidsmarktpositie). Als dat niet het geval is, kan de methode van consistent en herhaald wegen worden toegepast (Gouweleeuw & Hartgers, 2004; Houbiers, 2004; Kroese, Renssen & Trijssenaar, 2000). Daarbij worden de gewichten die eerder zijn bepaald nogmaals gekalibreerd voor bekende populatietotalen van de betreffende subpopulatie.

4 Discussie

In deze bijdrage wordt de ontwikkeling van een variabele opleidingsniveau gepresenteerd op basis van een combinatie van register- en enquête-informatie. De validiteit op micro-niveau is redelijk. Ongeveer driekwart van de hoogst behaalde opleiding die bepaald wordt op basis van CRIHO, ERR, CWI en WSF komt overeen met die uit de EBB. De uiteindelijke verdeling naar opleidingsniveau nadat de EBB's zijn toegevoegd komt redelijk overeen met hetzelfde gegeven uit de EBB. Toch moeten de gegevens over de geldig verklaarde hoogst behaalde opleiding op peilmoment worden gewogen om een goede schatting van het opleidingsniveau in Nederland te krijgen, omdat de registerinformatie selectief is naar onder meer leeftijd en arbeidsmarktpositie. Daarvoor zijn voorstellen gedaan. Nader onderzoek naar de plausibiliteit moet uitwijzen of de variabele na herweging bruikbaar is in onderzoek als achtergrondvariabele, in het bijzonder voor subpopulaties die worden onderzocht.

Het idee dat het eenvoudig is om op basis van een combinatie van register- en enquête-informatie statistiek te produceren kan definitief als een mythe worden bestempeld. Het is wel relatief eenvoudig om tellingen te maken uit de afzonderlijke registers als deze de doelpopulatie beschrijven, omdat deze meestal niet veel tegenspraken bevatten. De combinatie van registers die gedeeltelijk gedateerd zijn, onder- en overdekking van de populatie in de afzonderlijke bronnen, bronnen van verschillende kwaliteit en mate van detail, maakt dat een ingewikkeld productieproces moet worden ontwikkeld zelfs voor het bepalen van één variabele. Harmonisatie van de aanwezige opleidingsinformatie, het al dan niet geldig verklaren van deze informatie voor de personen en een weging van de personen met geldige informatie vereist een flinke investering.

Toch is deze investering naar ons idee de moeite waard. Niet alleen wordt een op dit moment bruikbare variabele geconstrueerd voor allerhande sociaal-wetenschappelijk onderzoek, tevens komt het uiteindelijke doel een flinke stap dichterbij: een opleidingsvariabele naar niveau en richting die integraal is voor de bevolking in Nederland. Dit verderliggende doel komt ieder jaar een stap dichterbij. In totaal is nu van 9,2 miljoen mensen een opleidingsniveau bekend. Het is dus nog lang niet integraal en zal ook de komende decennia niet integraal worden. Ieder jaar zal wel een behoorlijk aantal personen met een opleidingsgegeven kunnen worden toegevoegd. In 2003 hebben we ongeveer 400 duizend personen kunnen toevoegen. We kunnen het doel sneller dichterbij brengen door andere bronnen te benutten dan we in dit project hebben beschreven.

De onderwijsnummerbestanden komen als eerste in aanmerking. Daarin zijn alle leerlingen die bekostigd onderwijs volgen opgenomen. Momenteel zijn die beschikbaar voor het voortgezet onderwijs vanaf schooljaar 2003-2004. Voor het schooljaar 2004-2005 zijn proefbestanden voor het mbo beschikbaar; voor schooljaar 2006-2007 worden de eerste definitieve gegevens verwacht. Voor het primair onderwijs is schooljaar 2007-2008 een proefjaar en worden naar verwachting het daarop volgende schooljaar de eerste definitieve gegevens geleverd. Met de toevoeging van de onderwijsnummerbestanden wordt voor de jongere generaties een belangrijke leemte opgevuld, omdat we op dit moment vooral gegevens over het mbo missen. Daarnaast kunnen ook andere enquêtes dan de EBB worden benut, zoals de Woningbehoefteonderzoeken. Ook dit zijn enquêtes met een grote steekproef en een goede meting van het opleidingsniveau. Daarmee wordt ook het opleidingsniveau van ouderen beter bekend. Op deze wijze werkend wordt het resultaat ieder jaar beter.

Referenties

- Al, Pieter & Bart F.M. Bakker (eds.), 2000, Re-engineering Social Statistics by micro-integration of different sources. Themanummer Netherlands Official Statistics, jrg. 15, nr. summer.
- Arts, C.H., 2000, GIS en opleidingsniveau (interne CBS-nota 4692-00-SAV).
- Bakker, B.F.M., 2003, Hoe nieuw zijn nieuwe ideeën?, In: J. Nobel, S. Algera, M. Biemans & P. van der Laan (red.), Gedacht en gemeten (Voorburg/Heerlen: CBS), blz. 123-132.
- Bakker, B., & K. Arts, 2003, Dynamiek op de arbeidsmarkt; gegevens over stromen uit het Sociaal Statistisch Bestand. In: B.F.M. Bakker & L. Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS) blz. 59-70.
- Bakker, B., K. Arts & L. van Toor, 2005, Arbeidsmarktstatistieken in Nederland. De opzet van het Sociaal Statistisch Bestand, In: Over.Werk, jrg. 15, nr. 1, blz. 50-54.
- CBS, 2005, Standaard Onderwijs Indeling 2003, versie 2004-2005 <http://www.cbs.nl/nl-NL/menu/methoden/classificaties/overzicht/soi/2003/2004-05/default.htm>.
- Gouweleeuw, J. & M. Hartgers, 2004, The method of repeated weighting in the 2001 Census, In: Eric Schulte Nordholt, Marijke Hartgers, Rita Gircour (eds.), The Dutch Virtual Census of 2001, Analysis and Methodology, (Voorburg/Heerlen: Statistics Netherlands), pp. 261-276.
- Houbiers, M., 2004, Towards a Social Statistical Database and unified estimates at Statistics Netherlands, In: Journal of Official Statistics, jrg. 20, nr. 1, blz. 55-75.
- Kroese, A.H., Renssen R.H. & M. Trijssenaar, 2000, 'Weighting or imputation: constructing a consistent set of estimates based on data from different sources', In: P. Al & B.F.M. Bakker (red.), Re-engineering social statistics by micro-integration of different sources. Themanummer Netherlands Official Statistics, jrg. 15, nr. summer blz. 23-31.





Maarten H.J. Wolbers (VU)

Reactie op: Opleidingsniveau uit registers: nieuwe bronnen, maar nog niet compleet.

Toen ik Bart Bakker onlangs aan de telefoon had – het moet zo eind oktober zijn geweest –, kreeg ik van hem de vraag voorgelegd of het akkoord was dat ik de paper, die hij voor zijn presentatie van vandaag (*redactie: op de dag van het symposium*) aan het voorbereiden was en waarvoor ik als referent was gevraagd, niet op 1 november jongstleden zou krijgen toegestuurd, maar twee weken later. “Want”, zo stelde hij “er zat nog een aantal fouten in.” Ik zei: “Dat is goed, als je er maar niet alle fouten uithaalt, want dan blijft er voor mij niet veel meer over om kritiek op te leveren.” Waarop hij vervolgens antwoordde: “Nou, ik begrijp het, we hebben hier duidelijk tegengestelde belangen!” En daarna beëindigden we ons telefoongesprek.

Vanuit mijn rol als referent is die opmerking vanzelfsprekend te begrijpen. Het is nu eenmaal mijn taak vandaag om kritisch te zijn over hetgeen hij – namens het CBS – hier voor het voetlicht heeft gebracht. Vanuit mijn rol als wetenschapper is dat echter geenszins het geval. Sterker nog, ik ben juist van mening dat wij hier gemeenschappelijke belangen hebben! Zowel vanuit het CBS als vanuit de academische wereld bezien, is het uitermate belangrijk betrouwbare populatiecijfers te hebben over het opleidingspeil van onze (beroeps)bevolking. Want zoals in de inleiding van de paper terecht te lezen valt, speelt opleiding een sleutelrol op tal van terreinen. Zo bepaalt opleiding onze maatschappelijke kansen – niet alleen op de arbeidsmarkt, maar ook op de huwelijksmarkt –, hangt opleiding sterk samen met onze gezondheid en riskant gedrag daaromtrent, bepaalt opleiding waarden en normen die wij uitdragen, maar ook overdragen, etcetera. Populatiegegevens over opleiding zijn er op dit moment niet, maar het is de bedoeling van het CBS om op de lange termijn voor iedereen in ons land de onderwijsloopbaan vast te leggen in één groot register. Zo lang als deze registratie ontbreekt, wordt gepoogd een opleidingsregister samen te stellen gebaseerd op de wel beschikbare informatie uit registers en enquêtes. In de onderhavige paper wordt de ontwikkeling hiervan geschetst en enkele voorlopige resultaten gepresenteerd.

De logische vraag die bij mij opkwam na het aandachtig bestuderen van de paper is of het bepalen van het opleidingsniveau van de bevolking via dit opleidingsregister nou is gelukt. Hoewel het nog te vroeg is daar een definitief antwoord op te geven – dat kan natuurlijk niet van iets dat nog volop in ontwikkeling is –, acht ik het wel mogelijk om te beoordelen of men de goede weg is opgegaan. Het voorlopige antwoord op deze vraag is wat mij betreft bevestigend, ondanks de verschillende haken en ogen die zitten aan een dergelijke, ambitieuze onderneming. De ontwikkeling van een variabele opleidingsniveau is naar mijn idee een degelijke exercitie geweest, waarbij de nodige afwegingen zijn gemaakt (bijvoorbeeld ten aanzien van de harmonisatie van de gebruikte databronnen) en men zich bewust is van de verschillende tekortkomingen (bijvoorbeeld met betrekking tot de over-

vertegenwoordiging van jongeren en hoger opgeleiden). Wel vraag ik me een tweetal zaken af die ik hier achtereenvolgens kort aan de orde wil stellen.

In de eerste plaats betreft het een vraag over de gebruikte databronnen. Er zijn drie voor de hand liggende bronnen niet gebruikt, die naar mijn mening wel degelijk aanvullende informatie kunnen opleveren.

Allereerst gaat het om de onderwijscohorten. Het betreft hier de onderzoeken VOCL89 en VOCL93 (maar wellicht ook de voorgangers SMVO en SLVO) die door het CBS zijn gehouden. Hoewel deze cohortenonderzoeken niet meteen voor honderdduizenden, laat staan voor miljoenen, extra individuen adequate opleidingsinformatie zou opleveren, is het op zijn minst merkwaardig dat *het* instrument bij uitstek om onderwijsloopbanen in Nederland vast te stellen, niet is opgenomen in het ontwikkelde opleidingsregister.

Daarnaast is geen gebruik gemaakt van het lesgeldbestand zoals dat bij de Informatie Beheer Groep (IBG) aanwezig is. Dit registerbestand bevat weliswaar geen informatie over diplomabezit, maar goed, het bestand Wet op Studiefinanciering (WSF) wordt wel gebruikt, hoewel dat ook geen informatie verschaft over het behalen van diploma's. Ofschoon leerlingen van het MBO vanaf 18 jaar ook in het WSF-bestand zijn opgenomen, is het met name voor de kortere opleidingen in het MBO de overweging waard om het lesgeldbestand te betrekken bij het opleidingsregister. Moeilijker is het om gegevens over de beroepsbegeleidende leerweg van het MBO – afgekort als BBL – boven tafel te krijgen. Deze leerlingen waren nooit lesgeldplichtig. Voorzover ik weet bestaat er geen integrale registratie van deelnemers en/of gediplomeerden van de BBL, maar wellicht kunnen de afzonderlijke Kenniscentra Beroepsonderwijs Bedrijfsleven (KBB) of hun overkoepelende vereniging Colo toch enigszins uitkomst bieden. Het gaat hier immers wel om een zevende van de jaarlijkse uitstroom van schoolverlaters (ROA, 2004). In ieder geval is dit probleem in de nabije toekomst opgelost met de invoering van het onderwijsnummer. Dit geldt overigens voor het gehele MBO.

Ten slotte vraag ik me af waarom de Volkstelling 1960 en Volkstelling 1971 niet zijn gebruikt. Hoewel de opleidingsgegevens van deze personen niet meer actueel hoeven te zijn, kan met soortgelijke beslisregels als die in de paper zijn toegepast bij de gebruikte registers worden ingeschat in hoeverre individuen hun kwalificaties hebben verhoogd op latere leeftijd. Koppeling met gegevens van de Gemeentelijke Basisadministratie (GBA) maakt het mogelijk te onderzoeken of de leden van de bevolking uit 1960, respectievelijk 1971 in 2003 (het jaar waarop het registerbestand is gebaseerd) nog in Nederland woonden en in leven waren (*redactie: helaas beschikken de individuele gegevens uit de Volkstellingen van 1960 en 1971 over te weinig identificerende kenmerken, zodat ze niet koppelbaar zijn met de GBA-gegevens uit 2003*). Indien nodig, kan voor dit laatste worden gecorrigeerd voor opleidings specifieke sterftekanssen. Voordeel van het gebruiken van

gegevens uit de volkstellingen is dat hiermee een groot deel van de ondervertegenwoordiging van ouderen in het registerbestand kan worden opgeheven (en tegelijkertijd ook die met betrekking tot lager opgeleiden, omdat het met name de ouderen betreft die laag opgeleid zijn).

In de tweede plaats gaat het om het gebruiken van de Enquête Beroepsbevolking (EBB) als vergelijking bij het valideren van de variabele opleidingsniveau uit het register en als bron voor het bepalen van weegfactoren bij de noodzakelijke weging van deze variabele. Mij lijkt de EBB daartoe niet geschikt, omdat vaststaat dat het gerapporteerde opleidingsniveau in deze persoonsenquête systematisch is vertekend als gevolg van selectieve non-respons (Visscher, 1997). Lager opgeleiden zijn over het algemeen minder geneigd om mee te doen aan enquêtes dan hoger opgeleiden, waardoor eerstgenoemde groep is ondervertegenwoordigd in de EBB en daardoor het waargenomen opleidingsniveau van de (beroeps)bevolking is overschat. Is het ontbreken van betrouwbare informatie over de opleidingsverdeling in de EBB destijds niet juist (of op zijn minst gezegd: onder meer) de aanleiding geweest een opleidingsregister te gaan ontwikkelen? En is het dan niet vreemd de cijfers uit dit register te vergelijken met (en daarna eventueel zelfs te wegen op basis van) gegevens uit de EBB waarvan we toch weten dat deze niet juist zijn? Hoewel ik niet een-twee-drie weet welk alternatief voorhanden is – dat bestaat volgens mij niet, want daar is het nou net allemaal om begonnen –, is het op zijn minst gezegd eigenaardig dat over deze problematiek geen enkele aandacht is besteed in de paper.

Samenvattend: ik deel de conclusie getrokken in de paper dat de gedane investering om een opleidingsregister te ontwikkelen de moeite waard is geweest, ondanks de zojuist genoemde kritiekpunten. Hoewel het doel van een integrale registratie van het opleidingspeil van de (beroeps)bevolking nog ver weg is (en wellicht ook nooit wordt bereikt), vormt het huidige opleidingsregister een verrijking van het Sociaal Statistisch Bestand (SSB) en vergroot het de gebruiksmogelijkheden ervan voor onderzoek. Opleiding is nou eenmaal zo'n belangrijk kenmerk van individuen dat het in bijna elk sociaal-wetenschappelijk onderzoek als variabele wordt geanalyseerd.

Referenties

ROA (2004). Schoolverlaters tussen onderwijs en arbeidsmarkt 2003. Maastricht: Researchcentrum voor Onderwijs en Arbeidsmarkt.

Visscher, G. (1997). De blinde vlek van het CBS: systematische vertekening in het opleidingsniveau. De non-response in de Enquête Beroepsbevolking. *Sociologische Gids*, 44(3), 155-179.





Kees Prins (CBS)

Speuren naar ouders. Het ouder-kind-bestand, onmisbaar voor statistieken over intergenerationeel gedrag.

Het ouder-kind-bestand stelt het CBS in staat om gegevens over ouders en kinderen bijeen te brengen om daarmee statistische gegevens over intergenerationeel gedrag samen te stellen. De kracht van het ouder-kind-bestand is gelegen in de omvang. Sinds de laatste volkstelling in 1971 heeft het CBS niet meer de beschikking gehad over een bestand dat grootschalig intergenerationeel onderzoek mogelijk maakt.

1 Doel van het ouder-kind-bestand

Doel van het ouder-kind-bestand is behulpzaam te zijn bij het in beeld brengen van statistische verbanden tussen generaties. Met behulp van dit bestand heeft het CBS namelijk de mogelijkheid gegevens van ouders en kinderen uit verschillende bronnen / registraties aan elkaar te knopen. Het ouder-kind-bestand is een sleutelbestand tussen ouders en kinderen. Per persoon bevat het bestand het eigen zogenaamde A-nummer (zie het Begrippenkader, elders in deze paragraaf) en de A-nummers van de ouders. Ten behoeve van toepassing in het SSB worden deze nummers omgezet naar het RIN-persoon (zie Begrippenkader).

Kennis omtrent het A-nummer en daarmee het SoFi-nummer / Burgerservicenummer van een ouder stelt het CBS in staat door koppeling met andere registratieve gegevens over die ouder een schat aan informatie te verkrijgen.

Het ouder-kind-bestand is helaas niet compleet. Alle personen die in de GBA zijn ingeschreven, komen weliswaar als kind in het ouder-kind-bestand voor, maar niet voor al deze personen zijn de A-nummers van de ouders gevonden. Zo ontbreken in het bestand de A-nummers van ouders die vóór 1 oktober 1994 (dit is de datum waarop de GBA in werking is getreden) zijn overleden en van ouders die nooit in Nederland hebben gewoond. Dit betekent overigens niet dat het CBS helemaal niets weet over de ouders van wie die nummers in het ouder-kind-bestand ontbreken. In de GBA zijn immers van vrijwel iedere inwoner bijvoorbeeld de geboortedatum en het geboorteland van de ouders bekend.

De eerste versie van het ouder-kind-bestand heeft betrekking op de toestand op 1 januari 2005. Het CBS is voornemens het ouder-kind-bestand jaarlijks te actualiseren. Overigens blijven eenmaal aangebrachte relaties bestaan. Dus ook als een ouder of een kind overlijdt of emigreert, blijft in het ouder-kind-bestand de koppeling tussen de desbetreffende A-nummers gehandhaafd.

De term 'kind' wordt gebruikt in de betekenis van zoon of dochter, ongeacht de leeftijd. Er is dus geen beperking aangebracht naar leeftijd. Het ouder-kind-bestand bevat dan ook, zij het uiteraard in beperkte mate, relaties tussen bijvoorbeeld een man van 60 jaar en diens hoogbejaarde ouders.

Begrippen

A-nummer	Uniek nummer dat in de GBA aan een persoon is toegekend.
Allochtoon	Persoon van wie één of beide ouders buiten Nederland zijn geboren.
Autochtoon	Persoon van wie de beide ouders in Nederland zijn geboren.
GBA	Gemeentelijke BasisAdministratie persoonsgegevens. Voorheen: het bevolkingsregister.
Herkomstgroepering	De herkomstgroepering van autochtonen is Nederland. Die van allochtonen is gedefinieerd als het land waar men geboren is, tenzij dat Nederland is. In dat geval is de herkomstgroepering gelijk aan het land waar de moeder is geboren. Is dat land ook Nederland, dan is de herkomstgroepering gedefinieerd als het geboorteland van de vader.
Kind	In het ouder-kind-bestand wordt onder kind verstaan: 'zoon of dochter', ongeacht de leeftijd.
Persoonslijst	Het geheel van gegevens dat over een persoon in de GBA is opgeslagen. Voor iedereen die op 1 oktober 1994 of enig moment daarna tot de bevolking van Nederland behoorde, is er in de GBA een persoonslijst.
PL	Afkorting van Persoonslijst.
RIN-persoon	Uniek nummer dat binnen het CBS aan een persoon is toegekend.

Met behulp van het ouder-kind-bestand kunnen vragen worden beantwoord als:

- Krijgen kinderen van ouders die op jonge leeftijd kinderen kregen, hun kinderen ook op jonge leeftijd?
- Krijgen kinderen uit grote gezinnen zelf ook een groot gezin?
- Hebben kinderen van gescheiden ouders zelf ook een verhoogde kans op echtscheiding?
- Hoe is de relatie tussen de (school-)loopbaan van het kind en de sociaal-economische status van diens ouders?
- Hoe groot is de woonafstand tussen bejaarde ouders en volwassen kinderen? De uitkomst hiervan is van belang voor bijvoorbeeld behoefte aan thuiszorg.
- Wat is de mate van crimineel of anderszins ontspoord gedrag tussen generaties?
- Is er verband tussen doodsoorzaken binnen families, en zo ja, welk verband?
- Bestaan er overeenkomsten en verschillen tussen broers en zussen qua opleiding, inkomen, enz.?

Overigens gaan de toepassingen onder de laatste twee vragen een stapje verder dan de relatie tussen ouders en kinderen, ofwel familie in de eerste graad. Broers en zussen worden immers gekenmerkt door het feit dat ze dezelfde ouder(s) hebben, neven en nichten door dat ze dezelfde grootouders hebben, enz. In paragraaf 4 wordt hierop nader ingegaan.

Aan de hand van RIN-persoon kunnen van zowel de ouders als de kinderen de gegevens aan de diverse op registraties gebaseerde statistische bestanden worden ontleend. Zoals gezegd is het ouder-kind-bestand niet meer dan een sleutelbestand. Het bevat geen andere variabelen dan de A-nummers van het kind en diens ouders.

De kracht van het ouder-kind-bestand is gelegen in de omvang. Sinds de laatste volkstelling in 1971 heeft het CBS niet meer de beschikking gehad over een bestand dat grootschalig intergenerationeel onderzoek mogelijk maakt. Dit maakt het bestand ook aantrekkelijk voor toepassing in SSB-verband.

In een eerste toepassing van het ouder-kind-bestand is onderzocht in hoeverre echtscheidingsgedrag van ouders wordt overgedragen op de kinderen. Zie voor een verslag hiervan Steenhof en Prins (2005).

2 Relaties tussen personen in de GBA

De Gemeentelijke Basisadministratie Persoonsgegevens (GBA) is de geautomatiseerde vorm van wat vroeger werd aangeduid als het bevolkingsregister. De GBA is een gedecentraliseerd register dat, zoals de naam al aangeeft, wordt beheerd door de gemeenten. In de basisadministratie van een gemeente zijn van in principe iedere inwoner van die gemeente gegevens vastgelegd over onder meer de naam, de geboorteplaats, de geboortedatum, het woonadres, de burgerlijke staat, enz. Het geheel aan gegevens dat over een persoon in de GBA is vastgelegd, wordt de Persoonslijst (afgekort: PL) genoemd.

Niet iedereen die zich in Nederland bevindt, wordt in de GBA ingeschreven. Zo hoeven mensen die hier bijvoorbeeld een paar dagen of weken vakantie houden, zich niet in te schrijven. Evenmin hoeven mensen die staan ingeschreven, zich niet af te melden als ze een paar weken naar het buitenland gaan.

Buitenlandse diplomaten en NAVO-militairen zijn vrijgesteld van inschrijving. Indien zij dit wensen, mogen zij zich echter wel laten inschrijven. Het blijkt namelijk in de praktijk dat mensen die niet zijn ingeschreven, soms hinder ondervinden bij het aanvragen van vergunningen, rijbewijzen e.d.

Voor asielzoekers geldt weer een andere regeling. Kort gezegd komt het hierop neer dat zij gedurende de eerste zes maanden van hun verblijf in Nederland buiten de GBA blijven. Pas na die zes maanden wordt tot inschrijving overgegaan, tenzij binnen die periode een beslissing ten aanzien van het verblijf wordt genomen. Als de asielzoeker binnen zes

maanden een verblijfsvergunning krijgt, volgt in dat geval direct daarop inschrijving in de GBA. Als de aanvraag tot toegang binnen zes maanden wordt geweigerd, dan volgt in het geheel geen inschrijving.

In de GBA worden al relaties tussen PL-en gelegd. Behalve over een persoon zelf bevat de PL namelijk gegevens over diens ouder(s), de (ex-)partner(s) en de kinderen. Deze mensen staan bekend onder de verzamelterm 'gerelateerden'. Van een gerelateerde worden onder meer de naam, de geboortedatum, het geslacht en het geboorteland opgenomen. Ook wordt van de gerelateerde het A-nummer vermeld, mits voldaan is aan de volgende voorwaarden:

- de gerelateerde is zelf ingeschreven²⁰ in de GBA;
- de persoon en de gerelateerde stonden op enig moment na 1 oktober 1994 actueel ingeschreven in dezelfde gemeente.

De eerste voorwaarde is evident. Immers, om relaties te kunnen leggen met een gerelateerde, moet van die gerelateerde zelf een PL aanwezig zijn. De tweede voorwaarde heeft er mee te maken dat gemeenten alleen verbanden hoeven leggen tussen PL-en die in hun eigen basisadministratie aanwezig zijn.

Een en ander leidt ertoe dat niet van iedere persoon in de GBA de A-nummers van de ouders bekend zijn. Dat kan zijn veroorzaakt doordat de ouder al overleden was voordat in de GBA relaties tussen ouder- en kind-PL-en zijn gelegd. Soms hebben ouders geen A-nummer, bijvoorbeeld als ze nooit in Nederland hebben gewoond, zoals voor veel immigranten het geval is. In andere gevallen hebben de ouders wel A-nummers, maar zijn ze voor de woongemeente van het kind als het ware onbereikbaar, zoals in het geval dat iemand vóór 1 oktober 1994 het ouderlijk huis heeft verlaten en in een andere gemeente is gaan wonen. Als door verhuizing of door een gemeentelijke herindeling ouders en kind (weer) in dezelfde gemeente terechtkomen, of als de ouders van de immigrant zich bij hun kind in Nederland vestigen, worden de A-nummers op dat moment alsnog toegevoegd.

20) Voor de GBA geldt: eenmaal ingeschreven blijft ingeschreven. Mensen die tot de bevolking behoren, worden aangeduid als actueel ingeschreven. Overleden en geëmigreerde personen zijn niet-actueel ingeschreven.

3 Hoe het ouder-kind-bestand is vervaardigd

In het ouder-kind-bestand zijn op basis van de informatie uit de GBA voor zoveel mogelijk inwoners van Nederland de A-nummers van de persoon zelf en van diens ouders gekoppeld. Uiteraard zijn alle relaties die binnen de GBA zijn gelegd, in stand gehouden. De allereerste stap was dus die mensen van wie in de GBA van beide ouders de A-nummers bekend zijn, in het bestand op te nemen. De eerste uitvoering heeft plaatsgevonden op het bestand van 2002. Deze eerste versie van het ouder-kind-bestand is vervolgens aangevuld met mutaties over 2002-2004. Van de bevolking per 1 januari 2005²¹ zijn de volgende aantallen uit de GBA bekend:

In de GBA van beide ouders het A-nummer bekend:	8,002 miljoen
In de GBA van één ouder het A-nummer bekend:	2,342 miljoen
In de GBA van geen enkele ouder het A-nummer bekend:	5,962 miljoen

Voor de personen uit de twee laatstgenoemde groepen heeft het CBS ten behoeve van het ouder-kind-bestand vervolgens geprobeerd van de beide ouders de A-nummers te vinden. Van beide situaties wordt een voorbeeld gegeven, waarmee tegelijkertijd duidelijk wordt langs welke weg het zoekproces heeft plaatsgevonden.

Voorbeeld 1

De ouders van een vrouw uit Den Haag wonen in Tilburg. De vrouw heeft het ouderlijk huis al in 1985 verlaten, om in Den Haag op zichzelf te gaan wonen. De ouders zijn altijd in Tilburg gebleven.

De gemeenten Tilburg en Den Haag hebben nooit de A-nummers van deze drie mensen uitgewisseld. Weliswaar heeft Tilburg ervoor gezorgd dat de A-nummers van de ouders wederzijds op hun PL-en zijn geplaatst, maar tot een link met hun dochter is het niet gekomen. Met behulp van hun geboortedatums is dat in het kader van het ouder-kind-bestand wél gebeurd, en wel als volgt.

Op de PL van de vrouw staan de geboortedatums van haar beide ouders. Deze zijn bij de aanvang van de GBA in 1994 door Den Haag overgenomen van de toenmalige persoonskaart. Uiteraard staan de geboortedatums van de ouders ook op hun eigen PL in Tilburg. Ten behoeve van het ouder-kind-bestand zijn deze drie personen vervolgens aan de hand van hun respectieve geboortedatums bij elkaar gezocht, en is de link gelegd.

In het hier beschreven voorbeeld wonen de beide ouders bij elkaar, en komen hun beider A-nummers op beide PL-en voor. Er zijn echter ook situaties waarin de ouders niet meer bij elkaar wonen en de A-nummers van de ex-partner niet op hun PL staan. Hoe in dat geval te werk is gegaan, wordt weer toegelicht aan de hand van een voorbeeld.

21) Naast deze aantallen telt het ouder-kind-bestand zoals gezegd ook personen die in op enig moment in 2002-2004 wél, maar op 1 januari 2005 niet meer tot de bevolking van Nederland behoorden.

Voorbeeld 2

Een echtpaar is in 1992 gescheiden, de beide kinderen zijn bij de moeder in Amsterdam blijven wonen. De vader is ten tijde van de echtscheiding verhuisd naar Purmerend. Vanaf dat moment hebben ze niet meer in dezelfde gemeente gewoond.

De gemeente Amsterdam heeft het A-nummer van de moeder op de PL-en van de kinderen geplaatst. Het A-nummer van de vader ontbreekt echter, want diens A-nummer is niet bij Amsterdam bekend. Zijn A-nummer ontbreekt ook op de PL van zijn ex-vrouw, zoals haar A-nummer op zijn PL ontbreekt. Wél staat op hun PL de huwelijksdatum en de geboortedatum van hun ex-partner. Deze zijn bij de aanvang van de GBA in 1994 overgenomen van de respectieve persoonskaarten.

Gegeven die situatie is in het kader van het ouder-kind-bestand eerst een relatie gelegd tussen de twee ex-partners. Dit is gebeurd met behulp van hun beider geboortedatum en de huwelijksdatum. Nadat langs deze weg het gewezen echtpaar is gevormd, kon vervolgens in het kader van het ouder-kind-bestand het A-nummer van de vader aan de kinderen worden toegevoegd.

Beide voorgaande voorbeelden maken duidelijk dat het van belang was, eerst echtparen te formeren, waarna de kinderen aan de echtparen zijn toegewezen. Het werd namelijk niet wenselijk geacht kinderen aan slechts één ouder toe te wijzen. Immers, sinds 1946 zijn er dagelijks gemiddeld tussen 450 en 650 kinderen geboren. De kans dat bij het koppelen met slechts één ouder een mismatch zou ontstaan, is daardoor te groot.

Het ouder-kind-bestand bevat naar de situatie van 1 januari 2005 de volgende aantallen:

- 9,834 miljoen personen van wie de A-nummers van beide ouders bekend zijn.
- 3,194 miljoen personen van wie het A-nummer van één van de twee ouders bekend is.
- 3,227 miljoen personen van wie de A-nummers van beide ouders onbekend zijn.

Vergelijking met de aantallen die al uit de GBA bekend waren, leert dat het aantal mensen van wie A-nummers van beide ouders bekend is, is gestegen van 8,0 naar 9,8 miljoen. De zoekactie naar ouders heeft dus voor bijna twee miljoen personen succes gehad. Het aantal mensen zonder bekende ouders is met bijna de helft gedaald, namelijk van een kleine zes naar iets meer dan drie miljoen.

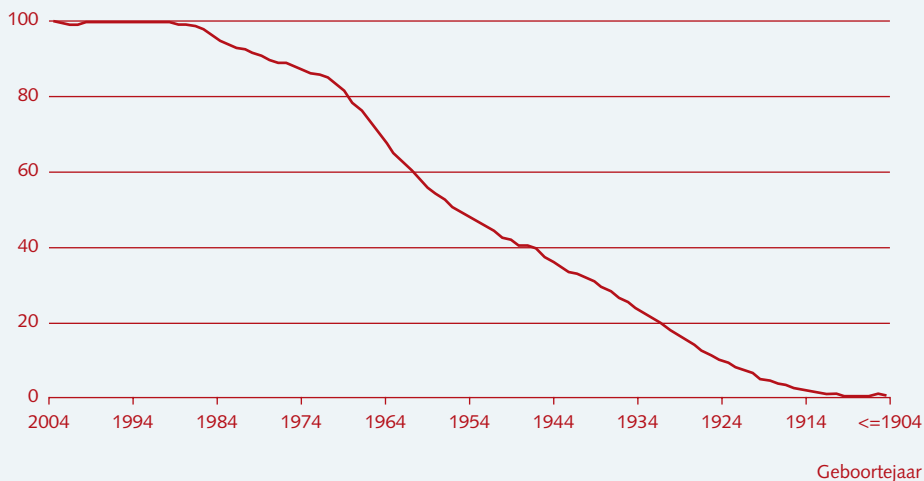
4 Kwaliteit van het ouder-kind-bestand

De kwaliteit van het ouder-kind-bestand is nagegaan aan de hand van de volgende indicatoren:

- leeftijdsverdeling
- buitenechtelijke vruchtbaarheid
- leeftijdsverschil met de moeder
- vrouwen naar kindertal
- personen naar aantal gesloten huwelijken
- leeftijdsverschil tussen huwelijkspartners

Deze indicatoren wijzen uit dat de kwaliteit van het bestand hoog is (Prins, 2004). Ter illustratie wordt hier per leeftijd het percentage personen gegeven van wie van één of van beide ouders het A-nummer is gevonden.

Figuur 1. Percentage personen met een of twee bekende ouders, naar geboortjaar, 1 januari 2005



Voor personen die jonger zijn dan 20 jaar, zijn in vrijwel alle gevallen de A-nummers van de ouder(s) bekend. Vanaf die leeftijd wordt het percentage personen met gevonden ouders steeds kleiner. Bij leeftijd 35 krijgt dit cijfer zelfs een forse tik naar beneden. Van ruim 80% van de 35-jarigen zijn een of twee ouders bekend, terwijl dit voor de 40-jarigen nog maar 67% en voor de 50-jarigen minder dan 50% is.

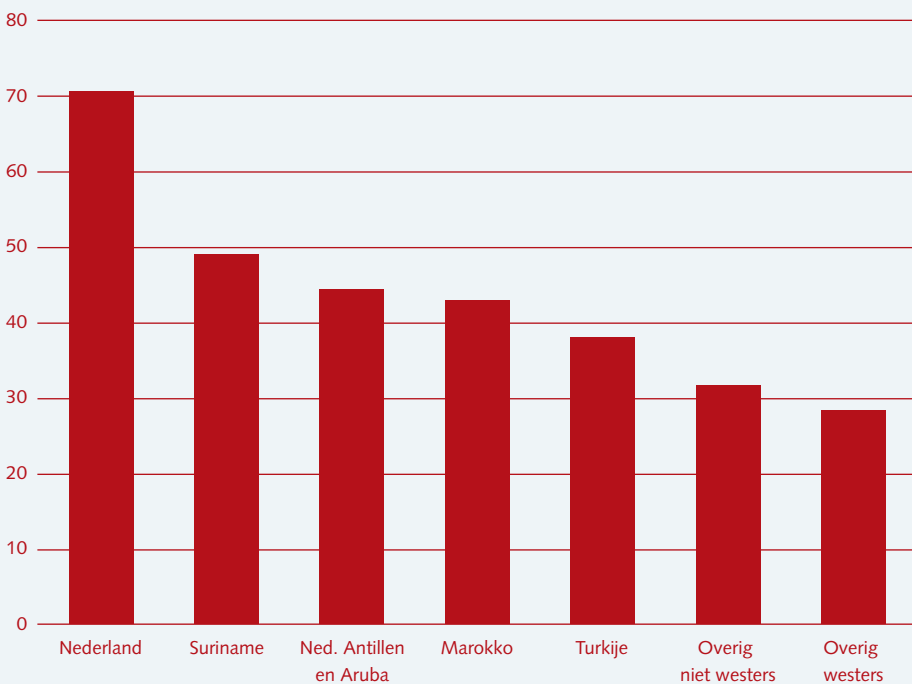
De afnemende percentages naar leeftijd zijn conform de verwachting. Immers, hoe ouder een persoon, hoe kleiner de kans dat ouders nog in leven zijn.

Ook naar herkomstgroepering is de massa in het ouder-kind-bestand selectief. Van 70% van de mensen van wie de beide ouders in Nederland zijn geboren, zijn de A-nummers van die ouders gevonden. Voor de andere hier onderscheiden groepen varieert dit van 49% (Suriname) tot 28% (overig niet-westerse allochtonen).

Met betrekking tot de herkomstgroepering is de selectiviteit eveneens verklaarbaar. De ouders van mensen met herkomstgroepering Nederland zijn immers beiden in Nederland geboren. De kans dat zij in de GBA zijn ingeschreven, is dan ook groter dan die voor allochtonen, van wie minimaal één ouder buiten Nederland is geboren. Dit geldt vooral voor mensen die zelf buiten Nederland zijn geboren.

Uit het voorgaande is duidelijk geworden dat het ouder-kind-bestand niet volledig en in bepaalde opzichten selectief is. Van een aantal mensen is het A-nummer van slechts één ouder bekend, van een aantal anderen ontbreken de A-nummers van beide ouders. De vraag die voorligt, is in hoeverre de bruikbaarheid van het ouder-kind-bestand hierdoor wordt gehinderd.

Figuur 2. Percentage personen met twee bekende ouders naar herkomstgroepering, 1 januari 2005



Tot de groepen die zijn ondervertegenwoordigd in het ouder-kind-bestand, behoren in ieder geval:

- Ouderen. Met het oplopen van de leeftijd van het kind neemt de kans dat diens ouders in leven zijn, af. Zoals gezegd is voor mensen die vóór de inwerkingtreding van de GBA zijn overleden, in de GBA geen PL aanwezig.
- Mensen die buiten Nederland zijn geboren. In veel gevallen hebben hun ouders nooit in Nederland gewoond, waardoor van die ouders geen PL aanwezig is.

Naast deze duidelijk aanwijsbare groepen zijn er subgroepen die in meer of mindere mate kunnen zijn ondervertegenwoordigd. Kinderen die op relatief jonge leeftijd het ouderlijk huis verlaten en buiten de woongemeente van de ouders zijn gaan wonen, zoals studenten, hebben een grotere kans dat in de GBA de A-nummers van de ouders niet op de PL zijn opgenomen. Waarschijnlijk is, als gevolg van de gevolgde procedure, deze ondervertegenwoordiging in het ouder-kind-bestand ongedaan gemaakt, maar helemaal zeker is dat niet. Mocht de ondervertegenwoordiging niet geheel zijn opgeheven, dan lijkt dit vooral de hoger opgeleide bevolking te raken.

De onvolledigheid en selectiviteit van het ouder-kind-bestand leidt ertoe dat resultaten niet zonder meer naar de gehele bevolking kunnen worden gegeneraliseerd. Het lijkt nuttig om zeker ten aanzien van leeftijd en herkomstgroepering een vorm van stratificatie achteraf toe te passen. Overigens zal uit ervaring moeten blijken of hiermee alle systematische afwijkingen zullen verdwijnen.

De kwaliteit van het bestand moet tot slot ook worden beoordeeld op de juistheid van de relaties tussen kind en ouders die wél tot stand zijn gebracht. Er moet immers rekening worden gehouden met de aanwezigheid in het bestand van relaties tussen A-nummers die niet overeenstemmen met de werkelijkheid. Een voorbeeld ter verduidelijking. Stel, de ouders van persoon P zijn geboren op resp. 2 december 1915 en 4 september 1917 en minstens één van hen is vóór de invoering van de GBA overleden. Als er een tweede echtpaar zou zijn waarvan de geboortedata precies gelijk zouden zijn aan die van het eerste echtpaar en die op 1 oktober 1994 beiden nog in leven waren, is persoon P door de vorenbeschreven procedure aan dat echtpaar gekoppeld. Dat kon op basis van de beschikbare gegevens niet worden voorkómen. Immers, het CBS beschikt niet over namen van mensen. Met behulp van de huwelijksdatum van de ouders zou het ontstaan van mismatches zoals hier genoemd wellicht kunnen worden verhinderd, ware het niet dat de huwelijksdatum van de ouders ontbreekt op de PL van de kinderen.

Het aantal aldus ontstane mismatches is echter klein. De belangrijkste aanwijzing hiervoor is dat het aantal gevonden paren met én dezelfde geboortedatum man én dezelfde geboortedatum vrouw én dezelfde huwelijksdatum uiterst klein is.

5 Relaties tussen broers, zussen en andere familieleden

Zoals aangegeven in paragraaf 1 is het ook mogelijk relaties tussen broers en zussen te leggen. Hun A-nummers kunnen langs twee wegen worden gekoppeld. De eerste weg is twee ouder-kindrelaties met elkaar te verbinden; broers en zussen hebben immers dezelfde ouders. Het voordeel van deze route is dat ook halfbroers en halfzussen kunnen worden gevonden.

De andere manier is broers en zussen te definiëren als personen van wie de geboortedatum van de moeders en van de vaders gelijk zijn. Het voordeel van die werkwijze is dat kennis betreffende de A-nummers van de ouders niet nodig is. Daardoor kunnen ook broers en zussen worden gevonden van wie de ouders al zijn overleden. Overigens kunnen op die manier geen halfbroers en halfzussen worden gevonden.

De voornoemde werkwijzen bieden de mogelijkheid om ook relaties tussen andere familieleden op het spoor te komen, zoals die tussen grootouders en kleinkinderen, neven en nichten, enz. Naarmate de 'afstand' tussen familieleden groter is, moet rekening worden gehouden met een geringere mate van volledigheid. Om bijvoorbeeld twee neven (die als kenmerk hebben dat ze dezelfde grootouders hebben) te vinden, is het nodig dat eerst de ouders van de neven worden gevonden. Waar het bijvoorbeeld nog wel mogelijk is twee broers van rond 80 jaar oud te vinden, zal dat voor neven van die leeftijd slechts zelden lukken.

6 Relaties tussen personen in de gemoderniseerde GBA

Over enkele jaren zal het ministerie van BZK volgens planning een moderne versie van de GBA opleveren. Zie voor de stand van zaken de website www.gba.nl van het programma Modernisering GBA. In de nieuwe versie van de GBA wordt de gegevensverstrekking geconcentreerd in GBA-verstrekkingen, waar replica's van alle PL-en worden ondergebracht. De gemeenten blijven in die constructie overigens de beheerders van hun eigen basisadministratie. Die rol wordt dus niet door GBA-verstrekkingen overgenomen. Wel biedt het op één plaats samenbrengen van alle replica's in principe de mogelijkheid om rela-

ties tussen PL-en te leggen die niet in dezelfde gemeente aanwezig zijn. Op die manier zou het aantal relaties binnen de GBA kunnen worden uitgebreid, vergelijkbaar met die in het ouder-kind-bestand. Binnen het programma Modernisering GBA heeft deze toepassing voorlopig echter niet de hoogste prioriteit.

De belangen van gemeenten en het CBS lopen in deze vrijwel geheel parallel. Toch is er een belangwekkend punt van verschil. Voor statistische doeleinden is het namelijk geen probleem als af en toe een onterechte match tussen ouder en kind ontstaat, en elders een match ten onrechte niet tot stand wordt gebracht. Voor organisaties die op individueel niveau met mensen te maken hebben, zoals gemeenten en uitkeringsinstanties, mogen dergelijke fouten uiteraard niet optreden. Zolang in de GBA de relatielegging tussen PL-en niet is verruimd, is het ouder-kind-bestand van grote waarde voor de statistiekproductie, onder meer voor het Sociaal Statistisch Bestand.

In de moderne GBA zal de relatielegging tussen PL-en die zich niet in dezelfde gemeenten bevinden, vooral langs automatische weg moeten gebeuren, met zo weinig mogelijk menselijke tussenkomst. Het is immers lastig als medewerkers van afdelingen burgerzaken uit verschillende gemeenten onderling moeten beslissen of twee mensen wel of niet een ouder-kind-relatie hebben. De automatische relatielegging zal daarom in eerste aanleg waarschijnlijk een conservatief karakter hebben. Alleen die relaties mogen worden gelegd waarvan onomstotelijk vaststaat dat de twee personen ook werkelijk ouder en kind zijn. Technieken als 'intelligent zoeken' zullen hier niet of terughoudend moeten worden toegepast. Een en ander heeft tot gevolg dat voor statistische toepassingen het ouder-kind-bestand voor langere tijd van grote waarde zal zijn en blijven.

Referenties

Prins, Kees, 2004, Speuren naar ouders. Eindrapport. Centraal Bureau voor de Statistiek, Divisie Sociale en Ruimtelijke Statistieken, Sector Ontwikkeling en Ondersteuning, 14 september 2004.

Steenhof, Liesbeth, en Kees Prins 2005, Echtscheiding van ouders en kinderen. Bevolkingstrends, jaargang 53, aflevering 4, blz. 47-52.





Clara H. Mulder (UvA)

Het ouder-kind-bestand van het SSB.*

* Reactie op: 'Speuren naar ouders. Het ouder-kind-bestand, onmisbaar voor statistieken over intergenerationeel gedrag'.

Inleiding

Mijn reactie op ‘Speuren naar ouders’ van Kees Prins spitst zich toe op drie punten. Ten eerste op wat er kan met het ouder-kindbestand: volgens mij kan er nog aanzienlijk meer mee dan Kees Prins in zijn presentatie naar voren heeft laten komen. Ten tweede op de selectiviteit van de beschikbaarheid van informatie. En ten derde op de praktische mogelijkheden voor onderzoekers om het bestand te analyseren.

1 Er kan méér

Kees Prins heeft een aantal voorbeelden genoemd van analyses die met het ouder-kindbestand kunnen worden uitgevoerd. Volgens mij kan niet genoeg worden benadrukt hoe rijk de mogelijkheden zijn van dit bestand, uiteraard in combinatie met de rest van het Sociaal Statistisch Bestand. Ik noem een aantal aanvullende voorbeelden van onderwerpen en mogelijkheden. Voor een deel zijn die ontleend aan de concrete plannen die ik met mijn collega's van de Universiteit van Amsterdam heb om analyses uit te voeren met behulp van het ouder-kindbestand, voor een deel aan onderzoeksinteresses die ik al langer heb en waarvoor de analyses voorheen alleen met surveygegevens konden worden uitgevoerd.

- De rol van ouders, kinderen en broers en zussen in verhuisgedrag: verhuist men eerder in de richting van familieleden of eerder ervandaan?
- De rol van familieleden in etnische of inkomenssegregatie: in hoeverre hangt ruimtelijke uitsortering samen met de locatie van familieleden?
- Intergenerationele overdracht van woonkwaliteit: in hoeverre lijken ouders en kinderen, broers en zussen op elkaar in woonkwaliteitsaspecten als de waarde van de woning en de eigendomssituatie (huur of koop)?
- Koppeling van gegevens uit het ouder-kindbestand met de Netherlands Kinship Panel Study, een surveybestand met gegevens over familierelaties. De mogelijkheden zijn velerlei: nonresponsanalyse, verrijking NKPS met SSB-gegevens, toevoeging van NKPS-gegevens aan SSB.
- Uit huis gaan van jongeren: wat is de invloed van ouderlijke hulpbronnen op het tijdstip van uit huis gaan? Bijvoorbeeld, wat is de rol van het inkomen en het opleidingsniveau van de ouders? Zijn er binnen gezinnen patronen te ontdekken in de tijdstippen van uit huis gaan van broers en zussen en in de locaties waar de broers en zussen naar toe vertrekken?

Dit zijn nog maar een paar voorbeelden van onderwerpen en mogelijkheden, maar met een beetje creativiteit zijn er vast nog veel en veel meer te bedenken. En het mooie is, dat de mogelijkheden alleen maar toenemen door de tijd heen, mits het bestand ook in de toekomst steeds wordt bijgewerkt en uitgebreid. Het aandeel mensen aan wie gegevens van ouders kunnen worden gekoppeld zal steeds verder toenemen, op den duur komen daar grootouders bij, enzovoort. In een paar van mijn voorbeelden is het noodzakelijk gebruik te maken van het longitudinale karakter van het Sociaal Statistisch Bestand. Ook dat zal steeds beter kunnen worden uitgebuit, als de waarnemingsperiode langer wordt.

Ook methodisch is het ouder-kindbestand uitermate interessant. Door de koppeling van ouders aan kinderen is het mogelijk sibling-modellen te schatten. Met deze modellen kan uiteen worden gerafeld in hoeverre broers en zussen méér op elkaar lijken dan anderen met dezelfde sociaal-economische en geografische achtergrond. Iets dergelijks zou ook mogelijk moeten zijn met neven en nichten. Misschien is het ook mogelijk een onderscheid te maken tussen eigen en geadopteerde kinderen. In principe zouden zelfs tweelingstudies mogelijk moeten zijn, al is de waarde daarvan groter als ook bekend is of het eeneiige of twee-eiige tweelingen betreft. Naar ik aanneem wordt dat niet geregistreerd in de Gemeentelijke Basisadministratie.

Doordat de exacte woonlocatie van de ouders en kinderen bekend is, zijn er ook legio mogelijkheden voor het gebruik van Geografische InformatieSystemen (GIS) en voor het koppelen van informatie over ruimtelijke eenheden aan informatie over individuen: bijvoorbeeld de beschikbaarheid van winkels, parkeerplaatsen, infrastructuur, kinderopvangmogelijkheden, banen, groen, schone lucht, en gegevens over de lokale woningvoorraad en bevolkingssamenstelling.

2 De selectiviteit van de informatie

Het ouder-kindbestand is geen survey. In surveys is er altijd non-respons. De oorzaken daarvan zijn lang niet altijd duidelijk, maar kunnen zeer wel te maken hebben met het onderwerp van de survey of met de onafhankelijke variabelen in de analyse: de non-respons is selectief. Aan de voor veel statistische procedures vereiste aseletheid van de non-respons wordt dan niet voldaan.

In het ouder-kindbestand is de situatie heel wat prettiger dan in surveys. Van non-respons is geen sprake. Toch is er wel selectiviteit in de beschikbaarheid van informatie. Niet alle A-nummers kunnen worden gekoppeld. Om een A-nummer aan dat van een ouder te koppelen moet de ouder in leven zijn of na 1994 zijn overleden, en in Nederland wonen of gewoond hebben. Het is dus bekend waar de selectiviteit van de informatie vandaan komt; dat is een voordeel ten opzichte van surveys. Tegelijk is duidelijk dat zowel het op een bepaalde leeftijd nog in leven zijn als het al dan niet in Nederland wonen uitermate selectief zijn. Weliswaar zijn de resultaten van de kwaliteitsanalyses die Kees Prins heeft genoemd hoopvol, maar garanties dat andere analyses niet te lijden hebben onder de selectiviteit naar leeftijd en geboorteland zijn er niet.

Tot op bepaalde hoogte valt voor deze selectiviteit misschien te corrigeren in analyses, maar helemaal zal dat nooit lukken. Het is daarom van groot belang dat gebruikers van het ouder-kindbestand goed op de hoogte zijn van de valkuilen op dit terrein. Een goede documentatie van de bevindingen van onderzoekers is noodzakelijk. Hier ligt een belangrijke taak voor de onderzoekers zelf, maar ook voor het CBS.

3 Analyse mogelijkheden

Dat er in principe heel veel kan met het ouder-kindbestand is duidelijk. Maar veel staat of valt met de toegankelijkheid van de data en de praktische mogelijkheden voor onderzoekers om ook echt met de data aan de slag te gaan. Die mogelijkheden zijn er gelukkig, in de vorm van *on site* op het CBS met de data werken en voor sommige instellingen in de vorm van *remote execution*. Als deze mogelijkheden tegen redelijke kosten beschikbaar zijn komen we al een heel eind. Maar ik zie één mogelijk probleem: de programmatuur. Voor het soort analyses dat ik eerder noemde, zoals sibling-analyses, statistische analyses op meer ruimtelijke schaalniveaus en GIS-analyses, is gespecialiseerde programmatuur nodig die de mogelijkheden van standaard statistische pakketten overschrijdt. Het gebruik van zulke programmatuur is niet een luxe extra; om werk gepubliceerd te krijgen in internationale tijdschriften kan een wetenschapper er niet omheen. Het CBS zal niet voor al deze programmatuur licenties hebben. Als de data bij het CBS moeten blijven en de softwarelicenties elders berusten kan er een probleem ontstaan. Om het ouder-kindbestand werkelijk optimaal te kunnen benutten zal dit probleem moeten worden opgelost.

(Redactie: Voor het on-site werken (en over enige tijd remote werken) biedt het CBS de mogelijkheid om met een aantal standaardpakketten analyses uit te voeren. Het gaat hierbij om SPSS en Stata, terwijl op aanvraag ook het gebruik van SAS en Oxpro mogelijk is. Deze pakketten bieden vele mogelijkheden tot analyse en dekken daarom het overgrote deel van de wensen van onderzoekers. In bijzondere gevallen kan het voorkomen dat de functionaliteit van deze pakketten onvoldoende is voor het uitvoeren van de benodigde analyses. In dat geval is het mogelijk aanvullende software op de on-site server te installeren, waarbij uiteraard de licenties met de onderzoeker correct moeten worden geregeld. De mogelijkheid om bij de faciliteit van remote access software vanuit de werkplek van de onderzoeker te gebruiken is niet aanwezig, omdat de (bewuste) inrichting van de remote access faciliteit het nodig maakt dat software op CBS server draait).





José Gouweleeuw (CBS)

Hoe toegankelijk is het Sociaal Statistisch Bestand (SSB)?

Onderzoekers van buiten het CBS hebben verschillende mogelijkheden om data uit het Sociaal Statistisch Bestand (SSB) te analyseren. In dit artikel worden alle mogelijkheden met de bijbehorende voor- en nadelen beschreven. Hierbij wordt met name ingegaan op recente nieuwe ontwikkelingen.

Inleiding

Het Sociaal Statistisch Bestand (SSB) is een rijke bron van data die menig onderzoeker graag zou willen analyseren. In de discussie van de conferentie 'De Virtuele Volkstelling en het Sociaal Statistisch Bestand' (CBS en SISWO, 2003) is een aantal gebruikerswensen ten aanzien van de data uit het SSB geïnterpreteerd. Deze wensen hebben zowel betrekking op de inhoud van de data (het aantal variabelen en de mate van detail) als de manier van toegang ertoe.

Het SSB heeft sinds die tijd niet stil gestaan. Niet alleen zijn er nieuwe jaargangen bijgekomen, ook het aantal variabelen in het SSB is uitgebreid. Denk hierbij bijvoorbeeld aan gegevens over huishoudens en aan het in ontwikkeling zijnde opleidingsregister. Er zijn de afgelopen jaren diverse nieuwe tabellen met SSB-gegevens gepubliceerd. De tabellen uit de Virtuele Volkstelling 2001 zijn hier een voorbeeld van. Daarnaast staan er op StatLine 13 uitgebreide tabellen met SSB-gegevens over de verslagjaren 1999 – 2003. Niet alleen binnen het CBS wordt er op basis van het SSB gepubliceerd, ook het aantal externe onderzoekers dat op één of andere manier gebruik maakt van data uit het SSB is toegenomen. Er wordt gewerkt aan nieuwe mogelijkheden voor toegang tot microdata, zoals bijvoorbeeld blijkt uit de pilot remote access. Een andere belangrijke ontwikkeling is dat in 2004 de CBS-wet (Staatsblad, 2004) van kracht is geworden. Als gevolg hiervan wordt de toegang van het CBS tot registraties verbeterd, en is er een wettelijk kader gekomen rond de toegangsmogelijkheden tot microdata voor de externe gebruiker.

In dit artikel wordt beschreven wat de huidige stand van zaken is ten aanzien van de toegang van onderzoekers tot het SSB. Dit varieert van toegang tot het volledige SSB on-site op het CBS tot de beschikking over een publicatiebestand op de werkplek van de gebruiker, zoals ook beschreven in Nobel (2003). In de rest van dit artikel zullen de verschillende mogelijkheden worden geschetst, waarbij telkens zal worden ingegaan op zowel de voordelen voor de gebruiker als de beperkingen en randvoorwaarden.

Merk in dit verband op dat op grond van de CBS-wet (Staatsblad, 2004) niet alle vormen van toegang voor iedere onderzoeker beschikbaar zijn. Er is een voortdurende afweging van de bescherming van persoonlijke individuele gegevens enerzijds en het tegemoetkomen aan de roep om microdata van wetenschappers anderzijds. Dit speelt overigens niet alleen in Nederland, maar is ook in internationaal verband aan de orde, zoals blijkt uit Trewin (2004).

Een onderzoeker kan op verschillende manieren toegang krijgen tot de data in het SSB. Achtereenvolgens kan dit door:

- Gebruik te maken van StatLine. Op StatLine wordt een aantal uitgebreide tabellen met gegevens uit het SSB gepubliceerd. Deze gegevens zijn door iedereen te raadplegen.
- Het aanvragen van maatwerk-tabellen. De tabellen worden door medewerkers van het CBS samengesteld en tegen onthulling beveiligd. De beveiligde resultaten worden vervolgens naar de onderzoeker gestuurd.
- On-site werken op het CBS. De onderzoeker voert de analyses uit op het CBS, en kan beveiligde resultaten van de analyses meenemen.
- Samenwerkingsverbanden. De analyses worden door de onderzoeker in samenwerking met het CBS uitgevoerd. Dit vindt plaats op het CBS. De onderzoeker kan de beveiligde resultaten meenemen naar zijn eigen werkplek.
- Remote execution. De analyseprogramma's worden door de onderzoeker op zijn eigen werkplek gemaakt en naar het CBS gestuurd. Hier worden de programma's uitgevoerd, de uitvoer beveiligd en naar de onderzoeker gestuurd.
- Remote access. De analyses worden door de onderzoeker op zijn eigen werkplek uitgevoerd. De databestanden blijven hierbij op het CBS. De resultaten van de analyses worden op het CBS beveiligd en naar de onderzoeker gestuurd.
- Het gebruik van publicatiebestanden of microbestanden die onder contract worden geleverd. De gebruiker heeft op zijn eigen werkplek de beschikking over (beveiligde) microbestanden en kan hier analyses op uitvoeren. Omdat de data al beveiligd zijn, worden er aan de uitvoer geen eisen gesteld.

De verschillen tussen bovengenoemde methoden bestaan uit de hoeveelheid detail van de data die de gebruiker tot zijn beschikking heeft, en door wie en waar de analyses worden uitgevoerd. In de rest van dit artikel zullen bovengenoemde mogelijkheden nader worden beschreven. Hierbij wordt ook ingegaan op de voorwaarden waaraan moet worden voldaan om een bepaalde vorm van toegang te verkrijgen.

Maatwerk-tabellen

Op basis van het SSB worden op dit moment 13 uitgebreide tabellen via StatLine ontsloten. Deze tabellen zijn beschikbaar voor de verslagjaren 1999 – 2003, en kunnen via het internet door iedereen worden geraadpleegd. Wanneer een onderzoeker een tabel wenst die niet uit deze 13 tabellen kan worden afgeleid, kan er een aanvraag voor maatwerk-tabellen worden ingediend. Deze maatwerk-tabellen worden door medewerkers van het CBS samengesteld. Meestal komt een onderzoeker met een probleemstelling naar het CBS. Er vindt overleg plaats, waarbij wordt gekeken hoe één en ander kan worden geoperationaaliseerd met behulp van de informatie die in het SSB is opgenomen. Als er tot overeenstemming wordt gekomen, worden de kosten begroot en worden de tabellen samengesteld. De tabellen die worden geleverd zijn altijd gecontroleerd op onthullingsrisico.

Iedere onderzoeker komt in aanmerking voor het aanvragen van maatwerk. Er zijn natuurlijk wel praktische beperkingen: zo zijn aan maatwerk altijd kosten verbonden (afhankelijk van het aantal tabellen en de hoeveelheid werk die eraan verbonden is), en moet het maken van de tabellen kunnen worden ingepland op het CBS. Het laatste jaar is ongeveer 20 keer van deze manier van toegang tot SSB-data gebruik gemaakt.

Een voorbeeld van maatwerk-tabellen is het werk dat voor de Wetenschappelijke Raad voor het Regeringsbeleid (WRR) is gedaan. Hierbij zijn tabellen gemaakt over de creatie en vernietiging van banen bij bedrijven en mobiliteit van de betrokken werknemers.

On-site werken

On-site werken houdt in dat de onderzoeker op het CBS komt om zijn analyses uit te voeren. De onderzoeker kan hierbij toegang krijgen tot de volledige SSB-bestanden. Deze bestanden zijn altijd geanonimiseerd, wat betekent dat alle directe identificatoren zoals naam- en adresgegevens, sofinummer en geboortedag uit het bestand zijn verwijderd. De overige variabelen in het bestand worden zonder verdere beveiliging integraal beschikbaar gesteld voor analyses. De onderzoeker heeft hiermee dezelfde data tot zijn beschikking als de CBS-medewerkers die het SSB analyseren. Ook de CBS-medewerkers werken altijd op geanonimiseerde bestanden. Het belang van deze bescherming van persoonlijke gegevens is beschreven in het Volkskrant-artikel van Haan (2005).

Naast de SSB-data heeft de gebruiker documentatie tot zijn beschikking waarin het te gebruiken bestand met alle variabelen wordt beschreven. Er wordt naar gestreefd om per bestand slechts één versie beschikbaar te stellen, om versiebeheerproblemen te voorkomen. Het hebben van slechts één versie per bestand zorgt ervoor dat onderzoek reproduceerbaar is, zoals is aangekaart in Dronkers (2003).

Zoals hierboven beschreven, ziet de gebruiker die on-site komt werken alle gegevens (behalve de direct identificerende). Hij mag hier echter niet zomaar alles uit publiceren. Uit de uitkomsten van het onderzoek mag het niet mogelijk zijn gegevens over afzonderlijke personen, huishoudens, bedrijven of instellingen te onthullen. Om de gebruiker hierop voor te bereiden, wordt gevraagd kennis te nemen van een korte elektronische cursus over geheimhouding. Bovendien worden alle uitkomsten die de onderzoeker mee zou willen nemen naar zijn eigen werkplek eerst gecontroleerd. Pas als de uitkomsten na controle veilig worden bevonden, kan de gebruiker de resultaten meenemen. Er is binnen het CBS een nieuw handboek statistische beveiliging met onder andere regels waaraan de uitvoer van analyses zou moeten voldoen. (CBS, 2006).

On-site werken is niet voor iedereen weggelegd. In artikel 41 van de CBS-wet (Staatsblad, 2004) is geregeld welke organisaties toegang tot de microdata kunnen krijgen. Dit omvat in ieder geval universiteiten, organisaties voor wetenschappelijk onderzoek en planbureaus. Daarnaast kunnen organisaties toestemming aan de Centrale Commissie voor de Statistiek (CCS) vragen. Als het om een niet-commerciële onderzoeksinstelling gaat, is deze toestemming in de praktijk bijna altijd te regelen. Wanneer een onderzoeker on-site mag werken, moet de instelling waartoe deze onderzoeker behoort een contract tekenen waarin is opgenomen dat er geen individuele data naar buiten mag worden gebracht. Daarnaast moeten de onderzoekers zelf een geheimhoudingsverklaring tekenen, waarvoor de leiding van de betreffende onderzoeksinstelling bovendien garant moet staan.

Tot de herfst van 2005 was het on-site werken op verschillende plaatsen binnen het CBS mogelijk. Tegenwoordig is het aanvragen van on-site werken gecentraliseerd bij het Centrum voor Beleidsstatistiek. Een voorbeeld van recente resultaten van on-site gebruik van het SSB, is het onderzoek naar het aantal en de aard van de zelfstandig ondernemers van allochtone herkomst in de detailhandel, dat is uitgevoerd door het Instituut voor Toegepaste Sociale Wetenschappen (ITS) van de Radboud Universiteit (Van den Tillaart en Doesborgh, 2004).

Samenwerkingsverbanden

Bij samenwerkingsverbanden worden de analyses op het SSB door medewerkers van het CBS en externe onderzoekers samen uitgevoerd. Deze vinden plaats op het CBS. Een samenwerkingsverband ligt feitelijk tussen on-site werken en maatwerk-tabellen in. Om in aanmerking te komen voor een samenwerkingsverband, moet een onderzoek binding hebben met het werkprogramma van het CBS. Het moet hier expliciet gaan om het uitvoeren van een project dat opgenomen is (wordt) in het CBS werkprogramma, en dat leidt tot een gezamenlijke publicatie. (Als er sprake is van een project dat leidt tot alleen een publicatie van de onderzoeksinstelling, dient gebruik te worden gemaakt van de on-site constructie.) Samenwerkingsverbanden zijn niet (altijd) kosteloos. Een voordeel van een samenwerkingsverband boven on-site werken is dat de onderzoeker niet alles alleen hoeft te doen. Daarnaast kan de expertise van de CBS-medewerkers ten aanzien van (SSB-)data rechtstreeks worden benut. Toegang tot de data voor de onderzoeker komt overeen met on-site werken. Ook hier geldt dat de uitvoer eerst moet worden gecontroleerd op onthullingsrisico, voordat deze mee naar de eigen werkplek kan worden genomen.

Een voorbeeld van een samenwerkingsproject waarbij het SSB wordt gebruikt is het project over de verdachten van criminaliteit. Hierbij is samengewerkt met het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC) van het Ministerie van Justitie. Het SSB is hierbij gebruikt om achtergronden van verdachten van criminaliteit in kaart te brengen. De resultaten van dit onderzoek zijn beschreven in Blom et al. (2005).

Remote execution

Bij remote execution bevinden de gebruiker van de data en de data zelf zich op twee verschillende locaties. De onderzoeker maakt op zijn eigen werkplek een analyseprogramma (bijvoorbeeld in SPSS) en stuurt dit naar het CBS. Op het CBS wordt het programma uitgevoerd op de betreffende bestanden. De uitkomsten van de analyse worden gecontroleerd op onthullingsrisico. Wanneer er niets wordt onthuld, kan de uitvoer naar de gebruiker worden toegestuurd.

Om met remote execution te kunnen werken moet de onderzoeker goed op de hoogte zijn van de inhoud en de structuur van de gegevensbestanden. Hiervoor is in ieder geval een

goede documentatie nodig. Verder moeten er testbestanden worden aangemaakt waarop de gebruiker zijn analyseprogramma kan testen, voordat deze naar het CBS wordt opgestuurd. Een dergelijk testbestand is een (beveiligde) steekproef uit het te analyseren bestand met exact dezelfde structuur.

Wanneer de onderzoeker op basis van het testbestand een analyseprogramma heeft gemaakt, wordt dit naar het CBS opgestuurd. Na een controle op een juiste technische specificatie van het programma, wordt dit op een proefbestand gedraaid. Wanneer dit probleemloos verloopt, wordt het programma op het volledige bestand gedraaid. De gegenereerde uitvoer wordt gecontroleerd op onthulling. Als de uitvoer niet veilig is, wordt contact opgenomen met de onderzoeker en aangegeven wat het probleem is. Het programma kan dan worden aangepast en opnieuw worden toegestuurd.

Het gebruik van remote execution werkt niet interactief. De onderzoeker stuurt een analyseprogramma op en moet vervolgens wachten op de resultaten, voordat hij kan beginnen met vervolganalyses. Er zijn afspraken over de tijd binnen welke de remote execution opdrachten worden uitgevoerd. De onderzoekers die gebruik mogen maken van remote execution zijn dezelfde als degenen die on-site mogen werken.

Een voorbeeld van gebruik van remote execution binnen het CBS zijn analyses op de data uit de Volkstelling van 1971. De eerste resultaten van het onderzoek moeten op het moment van schrijven van dit artikel nog worden gepubliceerd.

Remote access

Remote access betekent evenals bij remote execution dat de gebruiker op zijn eigen werkplek blijft en de data op het CBS. De gebruiker kan vanaf zijn eigen werkplek analyses op de data uitvoeren. Op basis van de uitkomsten van de analyses, die zichtbaar zijn op het scherm, kunnen er aanpassingen worden gedaan en nieuwe analyses worden uitgevoerd. Remote access heeft dus, in tegenstelling tot remote execution, wel een interactief karakter. Het grote voordeel van remote access voor de gebruiker is dat deze niet meer naar het CBS hoeft af te reizen. Ook is hij niet gebonden aan de openingstijden van het CBS.

Deze manier van toegang tot het SSB is voorlopig voor de meeste gebruikers nog toekomstmuziek. Op dit moment wordt er op het CBS een pilot-project voor remote access uitge-

voerd: CBSOnSite@Home. Deze pilot zal ongeveer een half jaar duren, en is voor één gebruiker: het Network for Pensions, Aging and Retirements (Netspar) van de Universiteit van Tilburg (Netspar, 2005).

De beveiliging bij remote access is gelijkwaardig met on-site werken. Er wordt gewerkt met een Citrix-server en een intensieve authenticatie, waarbij gebruik gemaakt wordt van biometrische gegevens (in dit geval een vingerafdruk). Met behulp van Citrix is het mogelijk om een scheiding te maken tussen het werkgebied op afstand (het CBS) en de lokale omgeving (de werkomgeving van de gebruiker). Het is daardoor niet mogelijk om data van het CBS te downloaden naar de lokale omgeving. De gebruiker kan wel op afstand op de bestanden werken, maar de data wordt niet uitgewisseld. De uitvoer van analyses in SPSS is wel zichtbaar op het scherm, maar kan niet worden afgedrukt.

Het gebruik van biometrische gegevens betekent dat er bij de remote access computers een vingerafdruklezer staat die de vingerafdruk van de gebruiker moet identificeren. Deze lezer wordt bij de gebruiker geïnstalleerd in een aparte ruimte waar speciale PC's staan die alleen voor het remote access werk worden gebruikt. Deze PC's moeten natuurlijk ook goed beveiligd zijn, zodat niet iedereen zomaar erbij kan. Denk bijvoorbeeld aan het automatisch instellen van schermbeveiliging na een bepaalde tijd, of het verbreken van de verbinding als een set-up eenmaal draait.

Een gebruiker wil zijn resultaten natuurlijk uiteindelijk wel in een of andere vorm hebben. Voor het meenemen van resultaten bij remote access geldt (evenals bij on-site werken of remote execution) dat de uitvoer eerst moet worden gecontroleerd op onthullingsrisico en veilig worden bevonden. In de praktijk betekent dit dat de gebruiker zijn uitvoer in een uitbak moet plaatsen, waar de uitvoer door CBS-medewerkers op onthulling wordt gecontroleerd. Wanneer de uitvoer veilig is, wordt deze naar het email-adres van de gebruiker gestuurd.

Wanneer de pilot met remote access succesvol is, zal deze worden uitgebreid naar andere gebruikers en andere databestanden. Hierbij kunnen bijvoorbeeld de huidige WSA-bestanden en data uit het SSB via remote access beschikbaar worden gesteld.

Microbestanden die onder contract worden geleverd

Microbestanden die onder contract worden geleverd zijn steekproefbestanden die worden geleverd via het Wetenschappelijk Statistisch Agentschap (WSA), dat is opgegaan in DANS (Data Archiving and Networked Services). De gebruiker kan deze bestanden op zijn eigen werkplek analyseren. Ook voor deze vorm van toegang tot data is in de wet vastgelegd wie er gebruik van mogen maken. Het betreft hier dezelfde groep onderzoekers als degenen die gemachtigd zijn om on-site te werken.

De microbestanden waarop wordt gewerkt zijn zelf al beveiligd tegen onthulling, zodat de uitvoer hier niet opnieuw hoeft te worden gecontroleerd. Voor het beveiligen van dergelijke microdata onder contract bestaan regels, zoals beschreven in het handboek statistische beveiliging (CBS, 2006). De regels houden in dat de bestanden geen directe identificatoren (bijvoorbeeld naam- en adresgegevens) bevat. Ook mogen in de records in het bestand geen waardencombinaties voorkomen die zeldzaam zijn in de populatie. Dit wordt geoperationaliseerd door te eisen dat bepaalde (driedimensionale) waardencombinaties van variabelen ten minste 100 keer in de populatie moeten voorkomen. Verder wordt het maximale detailleringniveau voor beroep, bedrijf en opleiding bepaald door de meeste gedetailleerde aanduiding van woon-, werk- of opleidingsregio, en moet iedere categorie van een regioaanduiding in het bestand tenminste 10.000 inwoners in de populatie bevatten.

Een voorbeeld van een microbestand onder contract dat aan het SSB gerelateerd is, is de 1%-steekproef uit de Virtuele Volkstelling 2001. Voor een beschrijving van de aan de virtuele volkstelling ten grondslag liggende data, wordt verwezen naar Advokaat et al. (2004). Deze 1%-steekproef is aan DANS geleverd. Het bestand is een personensteekproef, waarin ook huishoudenskenmerken zijn opgenomen. In Gierveld (2003) is de wens voor een dergelijk bestand beschreven. Er zijn gebruikers die liever een bestand met huishoudens als eenheid willen (Dronkers, 2003). De afzonderlijke huishoudens zijn echter in de 1%-steekproef niet als zodanig te herkennen.

De steekproef bevat de variabelen geslacht, leeftijd in 5-jaarsklassen, nationaliteit, geboorteland, woonland in het vorig jaar (de laatste drie variabelen zijn in een beperkt aantal categorieën beschikbaar), positie in het huishouden, huishoudgrootte, burgerlijke staat, economische status (geeft aan of iemand werknemer, zelfstandige, werkloos, gepensionerd, huisman/vrouw, scholier of iets anders is), beroepsniveau, onderwijsniveau, bedrijfs-

tak en woonprovincie. Het bestand is vervolgens beveiligd volgens de eerder beschreven regels. In de praktijk betekent dit dat de woonprovincie in een aantal records op onbekend is gezet. De overige variabelen zijn hierdoor onaangetast gebleven.

Publicatiebestanden

Publicatiebestanden zijn bestanden die voor iedereen tegen een lage vergoeding en nageoog zonder eisen ten aanzien van de wijze van gebruik ter beschikking staan. De gebruikers hoeven niet aan bepaalde voorwaarden te voldoen, zoals het geval is bij microbestanden die onder contract worden geleverd of bij on-site werken. Publicatiebestanden bevatten echter wel minder detail en worden strenger beveiligd tegen spontane herkenning dan bij microdata onder contract het geval is.

Een voorbeeld van dergelijke publicatiebestanden zijn de bestanden die zijn geleverd in het kader van de International Public Use Microdata Series (IPUMS, zie www.ipums.org). Dit betreft een 1%-steekproef uit de volkstellingen van 1960, 1971 en 2001. De 1%-steekproef uit 2001 is hier strenger beveiligd dan het bestand dat als microdata onder contract geleverd mag worden. In de praktijk betekent dit bijvoorbeeld dat de woonprovincie niet in het bestand is opgenomen. Het IPUMS-bestand uit 2001 bevat de internationale indelingen van onderwijsniveau en beroep, zodat de gegevens internationaal vergelijkbaar zijn.

De drie IPUMS-bestanden zijn zodanig beveiligd dat iedereen erbij mag. Dit betekent onder andere dat alle categorieën van alle variabelen in het bestand voldoende gevuld moeten zijn (dit moet overeen komen met 200.000 personen in de populatie). Hiervoor kan altijd worden gezorgd, door eventueel categorieën samen te voegen. Verder moeten alle bivariate combinaties tenminste 1.000 keer in de populatie voorkomen. Wanneer een combinatie hier niet aan voldoet, moet één van de twee variabelen (voor de betreffende combinaties) op onbekend worden gezet. Daarnaast is er nog een aantal beveiligingsvoorwaarden, zoals bijvoorbeeld dat het bestand tenminste 1 jaar oud moet zijn. Het gevolg hiervan is dat de mate van detail beperkt is. De bestanden zijn echter wel vrij en eenvoudig toegankelijk.

Er zijn voor externe onderzoekers verschillende mogelijkheden om gebruik te maken van de data in het SSB. Hierbij moeten zij echter wel gemachtigd zijn door de CBS-wet om toegang tot de data te krijgen (dit kan ook na specifieke toestemming van de CCS). Zolang het niet om een commerciële onderzoeksinstelling gaat, is deze machtiging in de praktijk eenvoudig te regelen. De toegangsmogelijkheden tot het SSB verschillen in de locatie van de data (bij de gebruiker of op het CBS), de locatie van de gebruiker (op zijn eigen werkplek of op het CBS) en de mate van detail van het bestand (van een volledig SSB tot een steekproef met een beperkt aantal ingedikte variabelen).

Bij de gebruikers bestaat de behoefte aan een 1%-steekproef uit de virtuele volkstelling. Deze steekproef is gemaakt en is beschikbaar via DANS. Daarnaast is door gebruikers ook de wens geuit om een 1%-steekproef uit het SSB te trekken. De vraag is dan onmiddellijk welke variabelen (en met welke mate van detail) in dit bestand moeten worden opgenomen. Het uiteindelijke bestand zal overigens wel aan de beveiligingsregels voor microdata onder contract moeten voldoen. Dit betekent dat een sterke detaillering in één variabele leidt tot een indikking in een andere. Om toch zoveel mogelijk gebruikers van dienst te zijn, kunnen er daarom meerdere steekproeven uit het SSB worden getrokken. Deze steekproeven verschillen in variabelen en categorieën. Het is de bedoeling dat de gebruikers inspraak krijgen in de inhoud van deze steekproefbestanden. Een en ander zal worden georganiseerd via DANS.

In de discussie van de conferentie 'De virtuele Volkstelling en het Sociaal Statistisch Bestand' (CBS en SISWO, 2003) is beschreven dat de gebruikers graag willen dat het SSB door uitgebreide tabellen via StatLine wordt ontsloten. Dat is op dit moment gerealiseerd voor de verslagjaren 1999 – 2003. Wanneer een gebruiker andere tabellen wenst kan dit bijvoorbeeld door maatwerk of on-site werken worden gerealiseerd.

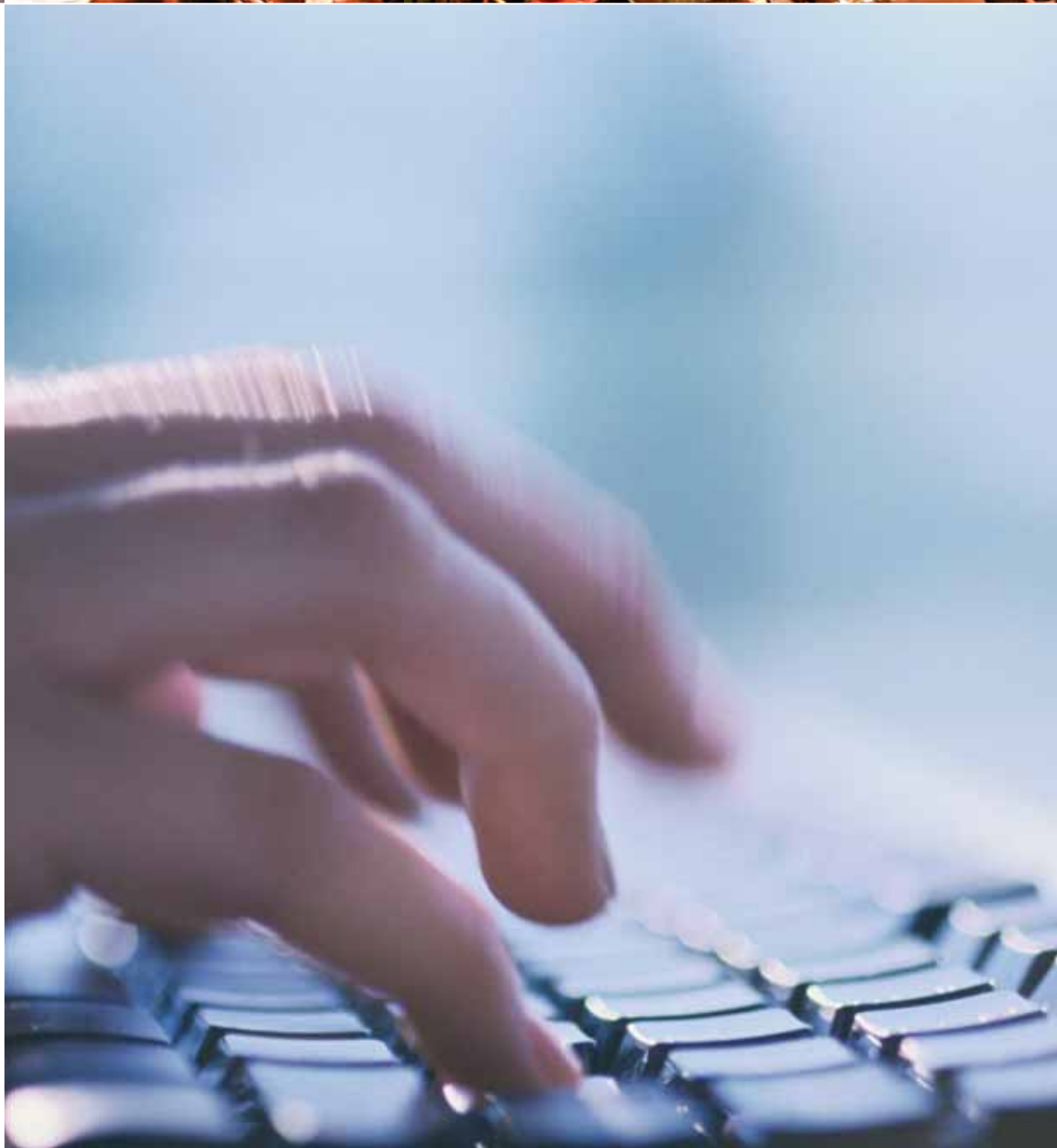
Daarnaast was een van de wensen van gebruikers om bestanden te hebben die internationaal vergelijkbaar zijn. In het kader van het IPUMS-project zijn drie steekproefbestanden uit de Volkstellingen van 1960, 1971 en 2001 gemaakt en deze zijn vrij toegankelijk. De variabelen volgen zoveel mogelijk internationale indelingen en zijn daardoor ook internationaal vergelijkbaar. Via www.ipums.org kan een gebruiker bovendien vergelijkbare bestanden van andere landen aanvragen.

Remote access draait op dit moment alleen nog als pilot-project. Wanneer deze pilot succesvol is, zal de mogelijkheid tot remote access worden uitgebreid tot andere gebruikers en andere microdata, zoals de huidige WSA-bestanden en de data uit het SSB.

Tenslotte is er een wens uitgesproken voor een duidelijk versiebeheer van de bestanden. Van de microdata die onder contract worden geleverd, en de publicatiebestanden is maar één versie beschikbaar, en is versiebeheer geen probleem. On-site werken (en remote execution) wordt binnen het CBS centraal via het Centrum voor Beleidsstatistiek gecoördineerd. Hier wordt voor ieder te analyseren bestand met één versie gewerkt. Het versiebeheer is hierdoor verbeterd.

Referenties

- W. Advokaat, J. van Cruchten, J. Gouweleeuw, C. Harmsen, M. Hartgers, M. Houbiers, F. Linder, H. Oroh en E. Schulte Nordholt. (2004) VT2001 – documentatie van het proces. CBS-rapport, BPA-nr. 476-04SOO/Intern.
- M. Blom, J. Oudhof, R.V. Bijl en B.F.M. Bakker (redactie) (2005) Verdacht van criminaliteit; Allochtonen en autochtonen nader bekeken. Cahier 2005 – 2, Wetenschappelijk Onderzoek- en Documentatiecentrum.
- CBS en SISWO (2003) Discussie. In: De virtuele volkstelling en het Sociaal Statistisch Bestand, B. Bakker en L. Putman (eds), blz. 123 – 126.
- CBS (2006) Handboek Statistische beveiliging, augustus 2006, BPA-nr. 21-06-TMO.
- J. Dronkers (2003) Het belang van een éénprocentsteekproef uit de huishoudens van de virtuele volkstelling. In: De virtuele volkstelling en het Sociaal Statistisch Bestand, B. Bakker en L. Putman (eds), blz. 25 – 30.
- J. Gierveld (2003) De virtuele volkstelling als databron voor sociaal-wetenschappelijk onderzoek; optimaliseringsvragen; Een reactie op de papers van Jacques van Maarseveen en Jaap Dronkers. In: De virtuele volkstelling en het Sociaal Statistisch Bestand, B. Bakker en L. Putman (eds), blz. 31.
- F. Haan (2005) Gezworenen in een statistisch luilekkerland. De Volkskrant, maandag 4 juli 2005, blz. 6.
- Netspar (2005) Newsflash, no. 3, augustus 2005. Zie www.netspar.nl
- J. Nobel (2003) De deur van het demografisch laboratorium. Perspectieven voor externe benutting van de virtuele volkstelling 2001. In: De virtuele volkstelling en het Sociaal Statistisch Bestand, B. Bakker en L. Putman (eds), blz.101 – 120.
- Staatsblad (2004) De Wet op het Centraal Bureau voor de Statistiek. Staatsblad van het Koninkrijk der Nederlanden, nr 695. Zie www.cbs.nl/NR/rdonlyres/F10515CB-91C6-426C-9BD9-177F743F72C6/0/cbswet1512-2004.pdf.
- H. van den Tillaart en J. Doesborgh (2004) Demografie etnisch ondernemerschap in de detailhandel. Zie www.hbd.nl/index.cfm/25,290,81,pdf/Rapport.pdf
- D. Trewin (2004) Task force on confidentiality and microdata – discussion paper. Conference of European Statisticians, Paris, 8 – 10 June, 2004.





Peter Doorn (DANS)

De DANS om het SSB.*

* Reactie op: 'Hoe toegankelijk is het Sociaal Statistisch Bestand (SSB)?'

Inleiding: wat is DANS?

Data Archiving and Networked Services (DANS) is een nieuwe landelijke organisatie die zorgt voor de opslag en blijvende toegankelijkheid van onderzoeksbestanden in de alfa- en gammawetenschappen. DANS is in de zomer van 2005 voortgekomen uit een gezamenlijk initiatief van de Koninklijke Nederlandse Akademie van Wetenschappen (KNAW) en de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). Ofschoon DANS als organisatie nog jong is, bouwt zij voort op de ervaring van het sociaal-wetenschappelijke Steinmetzarchief en het Nederlands Historisch Data Archief (NHDA), die in DANS zijn opgegaan. Ook de taken van het Wetenschappelijk Statistisch Agentschap, destijds opgericht om te bemiddelen bij de toegang tot data van organisaties als het CBS, zijn door DANS overgenomen.

De voorgangers van DANS hadden dus al een traditie in samenwerking met het CBS bij het bieden van toegang tot data. Op het moment van schrijven zijn 78 datasets van het CBS in de DANS-catalogus opgenomen en via DANS verkrijgbaar. Bij het Steinmetzarchief waren al in de jaren zeventig de originele ponskaarten van de Volkstelling 1960 gearchieveerd²². Het NHDA werkt sinds 1996 samen met het CBS aan de digitalisering van de historische volkstellingen vanaf 1795.

DANS heeft de vorm van een netwerk met een centrum dat verantwoordelijk is voor de organisatie van de data-infrastructuur. Dat centrum bestaat uit een team van ca. vijftien mensen die werken op het DANS-kantoor in Den Haag of bij één van de onderzoekcentra in het land. DANS is ook partner in internationale data-organisaties, zodat gegevensbestanden uit het buitenland gemakkelijk te verkrijgen zijn voor Nederlandse onderzoekers.

Om in de blijvende toegankelijkheid tot databestanden in de alfa- en gammawetenschappen te voorzien concentreert de organisatie zich op een viertal speerpunten: het beheren van bestaande data-archieven, het stimuleren van nieuwearchieven en netwerkdiensten om die doorzoekbaar te maken, retrospectieve archiveringsprojecten en ontwikkeling van nieuwe manieren om datasets voor de verre toekomst te behouden.

De archieftaak is de eerste speerpunt van DANS: het voor de lange termijn beheren en toegankelijk houden van databestanden in de alfa- en gammawetenschappen. Gedeeltelijk werd deze taak voorheen door het Steinmetzarchief, het Nederlands Historisch Data Archief en het *Wetenschappelijk Statistisch Agentschap* uitgevoerd.

22) Het bestand heeft wel enkele beperkingen. Het is destijds heringelezen vanaf de ponskaarten. Van de ca. 11 miljoen kaarten (één per Nederlander) zijn er zo'n 300.000 verloren gegaan. Voor een

uitgebreide beschrijving van dit bestand zie:
<http://www.volkstelling.nl/nl/documentatie/1960/>

De tweede speerpunt in het DANS-arsenaal zijn de *networked services*. Hiermee *stimuleert* DANS de verbreding van de data-infrastructuur. Deze taak krijgt vorm in TOP's, of Thematische Ontwikkel Programma's. Deze programma's brengen innovatieve netwerken tot stand voor het opslaan en hergebruiken van gegevens voor een bepaalde discipline of onderzoeksgroep. Onderzoekers hebben daarin het voortouw; DANS helpt door het inbrengen van ervaring en deskundigheid, en door te zorgen dat het nieuwe netwerk voldoet aan bepaalde kwaliteitscriteria, zodat de opgeslagen data gemakkelijk voor secundaire analyse beschikbaar zijn.

Kleinschaliger en meer dienstverlenend van aard is de derde speerpunt: Academisch Digitaal Archief (ADA). Samen met een opdrachtgevende faculteit of onderzoeksgroep, wordt in de vorm van een project een aantal bestaande databestanden *geordend* en toegankelijk gemaakt. Er zijn veel mogelijkheden voor dergelijke retro-archivering. Instellingen hebben niet zelden nog ergens oudere databestanden die van wetenschappelijke waarde zijn en blijven. DANS heeft een eigen 'ADA-methode' om ze te inventariseren, selecteren, reconstrueren en eventueel over te zetten naar toekomstbestendige media of software.

Tenslotte vormen Research & Development een speerpunt. De gebruikers willen eenvoudig via internet hun gegevens kunnen zoeken en ophalen, en daarbij gaat het om steeds grotere hoeveelheden. Nieuwe manieren van archiveren zijn nodig – vooral manieren die zo duurzaam en flexibel zijn dat ze toekomstige ontsluitingsmethoden ondersteunen. Het *ontwikkelen* gebeurt op projectbasis, zoveel mogelijk binnen de lopende activiteiten.

Rol van DANS bij toegang tot SSB

José Gouweleeuw maakt in haar bijdrage elders in deze bundel onderscheid naar wel zes vormen van toegang tot het SSB voor onderzoekers: tabellen, on-site werken bij CBS, samenwerkingsverbanden, *remote execution*, *remote access* en beveiligde bestanden. Ik loop ze successievelijk langs en bespreek daarbij de mogelijke rol van DANS.

Tabellen

Hierbij wordt onderscheid gemaakt tussen standaardtabellen en maatwerktabellen. Voor maatwerktabellen kunnen onderzoekers zich rechtstreeks tot het CBS wenden en is geen tussenkomst van DANS nodig. Van de standaardtabellen is een deel via CBS-StatLine toegankelijk, maar er zijn ook veertig uitvoerige tabellen gemaakt voor de virtuele volkstel-

ling 2001. Ook deze zijn als Excel-bestanden via de website van het CBS toegankelijk. Op verzoek van het CBS zijn deze tabellen, met uitvoerige documentatie over de methoden waarop ze uit het SSB zijn samengesteld, bij DANS gearhiveerd.²³ Hierbij zijn de bestanden geconverteerd naar XML, een formaat dat 'zelfbeschrijvend' is en onafhankelijk is van specifieke software zoals Microsoft Excel. Deze bestanden bevinden zich uit overwegingen van veiligheid en digitale duurzaamheid in het DANS-depot en worden niet beschikbaar gesteld aan gebruikers zolang het CBS dat zelf kan blijven doen.

On-site werken bij CBS

Om on-site op bestanden met microgegevens te werken kunnen onderzoekers rechtstreeks bij het CBS terecht en is geen tussenkomst van DANS vereist. Een deel van de statistisch beveiligde microbestanden die onder het datacontract tussen het CBS en het voormalige WSA vallen (zie verderop in deze bijdrage) zijn in de originele versie in de dataruimte van het CBS on-site beschikbaar.

Het CBS is druk bezig de toegangswegen voor on-site werk eenduidiger te formuleren. Tot voor kort waren er verschillende ingangen en voorwaarden voor on-site werk voor verschillende bestanden en gebruikersgroepen. Ook werkt het CBS aan een web-catalogus van beschikbare en gedocumenteerde bestanden. Hiermee zal duidelijk worden welke microbestanden on-site beschikbaar zijn (of kunnen worden gemaakt). Als ook de documentatie van de bestanden (die op papier op de on-site kamer van het CBS beschikbaar is) on-line toegankelijk wordt, kan de rol van DANS beperkt worden tot het beter bekend maken van de on-site mogelijkheden bij het CBS bij onderzoekers.

DANS zou ook graag de garantie willen hebben dat nu beschikbare bestanden in de toekomst toegankelijk blijven door ze goed te archiveren. In internationaal verband draagt DANS bij aan de invoering van een keurmerk voor 'Trusted Digital Repositories', waaraan ook het digitale CBS-archief zou moeten voldoen.

Samenwerkingsverbanden

Het SSB is toegankelijk voor onderzoekers die een samenwerkingsverband met het CBS aangaan. DANS stimuleert het ontstaan van samenwerkingsverbanden tussen het academisch onderzoek en de grote dataleveranciers, zoals het CBS. DANS heeft immers de taak om een samenhangend landelijk aanbod aan data voor sociaal-wetenschappelijk onderzoek tot stand te brengen, dat aansluit bij de vraag naar data. Om een betere afstemming van vraag en aanbod te realiseren is een landelijke data-agenda voor zowel de organisaties van 'academisch' onderzoek (universiteiten, NWO- en KNAW-instituten) als de dataverzamelende organisaties van de overheid (CBS, planbureaus, departementen) het

23) *The Dutch Virtual Census of 2001: Analysis and Methodology*, Eric Schulte Nordholt, Marijke Hartgers en Rita Gircoir (eds.), CBS: Voorburg/Heerlen 2004.

aangewezen instrument. Zo'n landelijke data-agenda kan een kader vormen om de samenwerking tussen het academisch onderzoek en het CBS te stimuleren. Dat dergelijke samenwerkingsverbanden verankerd moeten zijn in het werkprogramma van het CBS en derhalve instemming van de CCS vereisen maakt duidelijk dat er hier sprake is van gedeelde verantwoordelijkheden. Nederland kan in dit verband een voorbeeld nemen aan Engeland, waar een eerste opzet van een 'National Strategy for Data Resources for the Social Sciences' gereed is.²⁴ Deze 'Nationale Data Strategie' heeft tot doel om een samenhangend kader te scheppen voor de ontwikkeling en het onderhoud van een robuuste data-infrastructuur, die garandeert dat relevante en actuele data beschikbaar zijn voor de beantwoording van de belangrijkste maatschappelijke en economische vraagstukken.

Remote execution en remote access

Bij het ontwikkelen van *remote execution* op het SSB heeft DANS (eigenlijk de voorganger ervan: het NIWI) al een rol gespeeld. Dit is gebeurd in de vorm van een test, waarvoor de digitale bestanden van de Volkstelling 1971 zijn gebruikt. Onderzoekers van de universiteiten van Utrecht en Nijmegen hebben de werking van de remote execution in de praktijk beproefd. Naar onze indruk, die mede gebaseerd is op de praktijktest, is remote execution voor onderzoekers (en voor het CBS) onslachtiger dan remote access.

De statistische analyse wordt lokaal (off-line) voorbereid op een beveiligde steekproef van het echte bestand. Een voorbereide 'job' wordt vervolgens opgestuurd naar het CBS en daar uitgevoerd op het totale bestand. De werkwijze doet sterk denken aan de 'batch-verwerking' uit de tijd van de vroegere mainframe-computers. Hierdoor gaan de voordelen die interactieve verwerking met moderne statistische software biedt, verloren. Door de lange wachttijd tussen het aanbieden van een 'job' en het terugkrijgen van de output (die eerst door het CBS moet worden gescreend op onthullingsrisico's; het CBS gaat ervan uit dat de wachttijd standaard een dag zal zijn) is het nagaan van het effect van subtiele varianten in een gehanteerde statistische techniek een langdurige aangelegenheid. Ook wordt de keuze van een toe te passen analysetechniek vaak gebaseerd op de uitkomsten van eerdere bewerkingen, en die zijn dus ook pas na vertraging te beoordelen. De remote access-methode kent deze nadelen niet, maar deze mogelijkheid vergt wel een sophisticated technische oplossing in de vorm van een veilige ruimte met apparatuur en software die ondubbelzinnige authenticatie van de gebruiker mogelijk maakt.

DANS voert in 2006 een vergelijkbaar experiment als het CBS uit, maar dan met het Ministerie van VROM op het WoonOnderzoek Nederland (WoON; de opvolger van het Woningbehoeftenonderzoek WBO). DANS hoopt samen met het CBS om op alle sociaal-wetenschappelijke onderzoeksinstituten 'safe rooms' voor remote access te realiseren.

24) zie: <http://www2.warwick.ac.uk/fac/soc/nds/>

Daarbij werkt DANS aan een eenduidige toegangsregeling voor alle privacygevoelige data, ongeacht of deze nu van het CBS, het Ministerie van VROM of een andere organisatie afkomstig zijn.

DANS streeft in het algemeen naar een zo open mogelijke toegang, mede gebaseerd op de *open access* licenties van Creative Commons.²⁵ Voor toegang tot bestanden met individuele gegevens zijn aanvullende voorwaarden vereist. Uitgangspunten daarbij zijn de ‘Gedragscode voor gebruik van persoonsgegevens in wetenschappelijk onderzoek’ en de ‘Notitie Wetenschappelijke integriteit’.²⁶

Beveiligde bestanden

DANS biedt op dit moment toegang tot een aanzienlijk aantal beveiligde CBS-bestanden. In het kader van het Integrated Public Use Microdata Series-International project, IPUMS-International, heeft het CBS public-use microdata bestanden van de volkstellingen 1960, 1971 en 2001 samengesteld. Het gaat hierbij om beveiligde 1% steekproeven uit de microdata bestanden van de genoemde volkstellingen. Een beperkt aantal geselecteerde variabelen is in de public-use bestanden opgenomen. In eerste instantie zijn deze bestanden gecreëerd voor en geleverd aan het IPUMS-International project van het Minnesota Population Center (MPC) van de University of Minnesota.²⁷

In het verleden hebben het CBS en het WSA (nu onderdeel van DANS) een contract afgesloten voor de levering van beveiligde microbestanden van een aantal survey-onderzoeken aan de wetenschap.²⁸ DANS en het CBS zijn in overleg over dit contract. DANS streeft ernaar de toegang tot deze bestanden gratis te maken voor wetenschappelijk onderzoek. Dit is in lijn met de OESO-verklaring voor toegang tot data, die uit publieke middelen zijn gefinancierd, en met de zogenaamde ‘Verklaring van Berlijn’ over open toegang tot wetenschappelijke kennis.²⁹ Ook vraagt DANS zich af of er in de nabije toekomst, wanneer remote access beschikbaar komt, nog wel vraag zal zijn naar microbestanden. Liever zou DANS een deel van de middelen die met het contract gemoeid zijn besteden aan de verbetering van de toegang tot CBS-data, inclusief het historisch data-archief.

Wat betreft wensen tot specifieke extracties of samenstellingen uit SSB-bestanden leeft bij het CBS de vraag aan welke ‘satellieten’ onderzoekers de voorkeur geven. Het gaat dan om vragen als: welke steekproefomvang, welk regionaal detail en welke variabelen? Het CBS heeft DANS verzocht (bijvoorbeeld via gebruikersbijeentkomsten) te helpen prioriteiten te stellen.

25) “Some rights reserved”. Zie: <http://creativecommons.org/> en <http://sciencecommons.org/>

26) College Bescherming Persoonsgegevens, Gedragscode voor gebruik van persoonsgegevens

in wetenschappelijk onderzoek van de VSNU. Staatscourant 2 januari 2006, nr. 1, pag. 17. Notitie Wetenschappelijke Integriteit - Over normen van wetenschappelijk onderzoek en een Landelijk Orgaan voor Wetenschappelijke Integriteit (LOWI).

KNAW, NWO en Vereniging van Universiteiten (VSNU), 2001.

27) Zie: <http://www.ipums.umn.edu/international/>

Het CBS biedt op zes manieren toegang tot het SSB voor onderzoek. DANS vormt voor de toegang tot CBS-data voor onderzoekers een toegangspoort. De rol van DANS daarbij kan als volgt worden samengevat:

Tabel 1.

Vorm van toegang tot SSB	Rol van DANS
Tabellen	<ul style="list-style-type: none"> • Standaardtabellen: archivering bij DANS • Maatwerktabellen: rechtstreeks bij CBS (archivering bij DANS?)
On-site werken bij CBS	Rechtstreeks bij CBS
Samenwerkingsverbanden	DANS werkt aan landelijke data-agenda voor samenwerking tussen onderzoek en grote data-aanbieders
Remote execution	DANS ondersteunt (test VT71)
Remote access	DANS helpt implementeren
Beveiligde bestanden	<ul style="list-style-type: none"> • 1% steekproef VT60, 71, 2001: Beschikbaar via DANS • Microbestanden op maat ('satellieten'): Welke steekproefomvang, variabelen en regionaal detail? WSA-contract?

Daarnaast biedt DANS nog toegang tot andere databestanden, zowel van het CBS als van andere data-aanbieders, maar ook uit de data-archieven in eigen beheer. Het is niet ondenkbaar dat de vele toegangswegen op onderzoekers overkomen als een moeilijk doordringbaar woud. DANS streeft er samen met het CBS naar om de toegang tot alle CBS-bestanden onder één overeenkomst te brengen met een transparante toegangsregeling en tegen minimale kosten voor het onderzoek. Dit alles onder het motto: *public access to public data!*

28) Voor een overzicht zie: http://www.dans.knaw.nl/nl/data/tarieven_2006/

29) "Ministerial Declaration on Access to Research Data from Public Funding", Science, Technology and

Innovation for the 21st Century. Meeting of the OECD Committee for Scientific and Technological Policy at Ministerial Level, 29-30 January 2004 - Final Communique zie: www.oecd.org. De "Berlin Declaration on Open Access to Knowledge in the

Sciences and Humanities" is getekend door KNAW, NWO, SURF, universiteiten. Zie: <http://www.zim.mpg.de/openaccess-berlin/>





Paneldiscussie

Aan het discussiepanel ter afsluiting van het SSB-seminar namen deel: Bert Schijf (UvA), Bart Bakker (CBS), Kea Tijdens (UvA/EUR), Clara Mulder (UvA), Jan Latten (CBS/UvA) en Harry van den Tillaart (ITS). Huib van de Stadt (CBS) was voorzitter.

Een viertal vragen is aan de orde gekomen:

- in welke richting zou het SSB zich moeten ontwikkelen?
- wat zijn de wensen voor output uit het huidige SSB?
- hoe kunnen de gebruiksmogelijkheden van het SSB worden vergroot?
- vragen naar aanleiding van de presentaties.

De eerste vraag

aan panel en zaal betrof de richting waarin het SSB zich zou moeten ontwikkelen in de toekomst. Aan welke elementen en welke onderwerpen moet prioriteit gegeven worden?

Bert Schijf zou behoefte hebben aan kleine datasets voor specifieke toepassingen over kleine onderwerpsgebieden. Hij noemde het voorbeeld van een antropologisch onderzoek naar een Turkse bezorgpizzeria in Amsterdam, waarin behoefte was aan referentiegegevens over kleine Turkse ondernemers die een pizzeria drijven (hoewel in dit specifieke geval ook de Gouden gids wellicht dienst had kunnen doen). Hoe goed is het mogelijk om informatie over kleinere populaties uit het SSB te halen en wat is dan de kwaliteit daarvan?

Bart Bakker, en in aanvulling daarop Koos Arts, maakten duidelijk dat informatie over ondernemers, de aard van hun economische activiteit en hun locatie zeer gedetailleerd beschikbaar zijn in het SSB. Voor het gewenste niveau van detail in dit voorbeeld zou zelfs de zeer gedetailleerde typering volgens de zogeheten Standaard Bedrijfsindeling (SBI) waarschijnlijk niet toereikend geweest zijn. Huib van de Stadt maakte in aanvulling hierop ook nog gewag van de ontwikkeling van het Economisch Statistisch Bestand, waarin economische gegevens over bedrijven uit registers en enquêtes bij elkaar gebracht zullen gaan worden, en wat wellicht ook gedeeltelijk in deze informatiebehoefte zou kunnen gaan voorzien.

Kea Tijdens vroeg aandacht voor de variabele beroep als een moeilijke variabele die niet voorkomt in de registraties. Wel komen in de registraties van bedrijven aanduidingen voor van functies. Een probleem is echter dat er veel meer functies bij bedrijven zijn dan overeenkomstige beroepen in de standaard beroepenclassificatie (SBC) van het CBS. Zij zou

grote behoefte hebben aan een conversietabel, zodat functies in personeelsadministraties van bedrijven vertaald kunnen worden naar beroepen.

Volgens Bart Bakker is het al heel moeilijk om de functies zoals gespecificeerd in CAO's te vertalen naar de beroepen in de SBC. Gezien het grote aantal CAO's zou dit een erg grote inspanning zijn. Een andere mogelijkheid om meer massa te creëren voor de variabele beroep is om gegevens uit vele surveys (zoals een aantal jaargangen van de Enquête Beroepsbevolking van het CBS) samen te brengen, na te gaan of die op de peildatum nog geldig zijn, en daaraan gegevens over vrije beroepsbeoefenaren uit ondernemingsregisters toe te voegen, volgens een soortgelijke methode als nu bij het opleidingsregister wordt toegepast. Hiermee zou het mogelijk moeten zijn om voor circa 10% van de beroepsbevolking een aanduiding van het beroep te krijgen. Een subpopulatie van die omvang zou veel gewenste analyses over beroep mogelijk maken.

Harry van den Tillaart ondersteunt de suggestie van Kea Tijdens. De aanduiding van de sector lijkt onvoldoende informatief, zoals ook de analyses van Jules Theeuwes (over zeer grote arbeidsmobiliteit tussen sectoren) suggereren.

Clara Mulder vraagt zich af in hoeverre internationale vergelijkbaarheid een aandachtspunt is in het SSB. Er zijn diverse andere landen, zoals de Scandinavische landen, waar het werken met registers ook al ver gevorderd is. Voor onderlinge vergelijkbaarheid is afstemming van de gebruikte indelingen en classificaties hoogst gewenst.

Huib wijst op het samenstellen van gegevens uit de Virtuele Volkstellingsgegevens van het SSB ten behoeve van het IPUMS als een voorbeeld van investeringen in het beschikbaarstellen van internationaal vergelijkbare gegevens via het SSB.

Jan Latten benadrukt het belang van samenwerking in het kader van het gebruik van SSB-gegevens. Het is niet alleen efficiënt, maar ook goed voor grotere inventiviteit en creativiteit in het onderzoek. Aansluitend op kwalificaties als 'SSB-data zijn plat' bepleit hij dat ook in het kader van het SSB aandacht voor enquête-onderzoek moet blijven. Huib van de Stadt zegt daarop dat het belang van opnemen van survey-data in het SSB ook onderkend wordt.

Dorien Manting ondersteunt de oproep voor meer samenwerking. Onderzoekers willen altijd meer, en in dit geval vooral ook een betere toegankelijkheid van het SSB. Samenwerking van het CBS met onderzoekers buiten het CBS bevordert een goede ontwikkeling van het SSB en zijn gebruiksmogelijkheden.

Bart Bakker vult dit aan door te verduidelijken dat aanvullende survey data van onderzoekers van groot belang zijn. Samenwerking (die moet aansluiten op het eigen werkprogramma van het CBS) is niet alleen uitwisseling van kennis, maar is ook nuttig in de vorm van uitwisseling van data.

Clara Mulder wijst op het risico van een elite van onderzoekers die (al dan niet toevallig) passen bij het CBS werkprogramma en daardoor gratis kunnen samenwerken met het CBS en de SSB-data kunnen gebruiken, terwijl anderen voor het gebruik daarvan grote bedragen moeten gaan betalen. Hoe is dat te voorkomen?

Bart Bakker antwoordt dat niet het CBS zijn eigen werkprogramma vaststelt maar de CCS en dat ook derden daarop invloed kunnen uitoefenen. Het risico wordt ook aanzienlijk minder wanneer *remote access* beschikbaar komt en daarmee de meeste data aanzienlijk gemakkelijker beschikbaar kunnen komen. Huib van de Stadt voegt daaraan toe dat het CBS niet zal discrimineren. In de praktijk levert de eis dat het onderzoek 'niet moet mistaan in het werkprogramma' weinig problemen op; er is nog geen enkel verzoek om samenwerking op die grond geweigerd. Een groter risico voor het inwilligen van verzoeken is de beperkte beschikbare capaciteit.

Naar aanleiding van een vraag van Kea Tijdens over het gebruik van SSB-data voor haar eigen Internet-survey antwoordt Huib van de Stadt dat het SSB ook door onderzoekers gebruikt kan worden die hun eigen survey data meenemen en deze vervolgens laten koppelen aan het SSB. De gekoppelde en aldus verrijkte data kunnen dan on-site bij het CBS worden geanalyseerd. Dat is een mogelijkheid waarvan bijv. medische onderzoekers vrij vaak gebruik maken.

Mars Cramer vraagt zich hardop af waarom commerciële partijen van samenwerking worden uitgesloten. Zij beschikken vaak over zeer interessante databestanden.

Volgens Huib van de Stadt is het formele probleem dat in de CBS-wet een omschrijving is gegeven van specifieke gebruikersgroepen, die wel universiteiten en overheid omvat, maar de meeste commerciële gebruiksdoelstellingen uitsluit. Bart Bakker voegt daaraan toe dat er ook wel commerciële samenwerkingspartners zijn, maar dat het daarbij gaat om specifieke doelstellingen. Er zitten haken en ogen aan, maar de mogelijkheden zullen zeker verder worden onderzocht.

Jeroen Smits is van mening dat er diverse gebruiksmogelijkheden zijn van het SSB, waarbij de genoemde plathed van de data en het beperkte aantal variabelen geen probleem zijn.

Voor de toetsing van veel hypothesen is de *context* van verschijnselen van belang. Daarvoor is het nodig een breed scala aan laagregionale gegevens te koppelen aan de microdata van het SSB. Dat kan heel goed via aggregatie van SSB data naar regionale niveaus of ook naar bedrijfsniveau.

Huib van de Stadt beaamt het belang van deze suggestie en verduidelijkt dat het CBS daarmee ook al bezig is. Het samenstellen en beschikbaar maken van een laagregionaal databestand met tal van gegevens is een speerpunt in het CBS-werkprogramma van 2006. Bart Bakker voegt daaraan toe dat naast de mogelijkheden op bedrijfsniveau in 2006 via het gebruik van het onderwijsnummer ook mogelijkheden op schoolniveau gaan ontstaan.

De tweede vraag

die voorgelegd wordt aan panel en zaal betreft gewenste uitbreiding van de output van het huidige SSB. De vraag wordt eerst aan het panel voorgelegd.

Bert Schijf heeft grote behoefte aan meer mogelijkheden om data weer te geven in de vorm van kaartjes en plattegronden. Een tweede invalshoek (ingegeven door zijn eigen onderzoeksproject naar elites en adel samen met Jaap Dronkers) is de behoefte aan gegevens over specifieke beroepsgroepen, zoals bijvoorbeeld het opleidingsniveau van de president-directeuren van de duizend grootste ondernemingen. Terzijde merkt hij op dat meer marketing voor het SSB nodig is, omdat hijzelf merkt dat vrijwel niemand in zijn eigen afdeling op de hoogte is van bestaan en mogelijkheden van het SSB.

Kea Tijdens sluit aan bij de opmerking van Jeroen Smits dat er behoefte bestaat aan context variabelen, en wel op een dusdanige manier dat de toepasbaarheid bij eigen surveys verbeterd wordt. Onderzoekers zouden met deze handreiking nog meer gediend kunnen zijn, wanneer in een vroeg stadium contacten en overleg tussen CBS en onderzoekers tot stand komen: dan kunnen onderzoekers al in de ontwerpfase hun onderzoek inrichten op het gebruik van specifieke contextgegevens.

Clara Mulder pleit voor bredere beschikbaarheid van het SSB als steekproefkader voor onderzoek van derden. Dan is ook een betere aanpak van non-respons problemen in onderzoek mogelijk. Heel veel output moet niet alleen van het CBS komen maar juist ook van onderzoekers die van het SSB als instrument gebruik maken. Daarom is een aanmoediging om door te gaan met de ontwikkeling van remote access op zijn plaats.

Jan Latten vraagt vanuit zijn eigen onderzoeksthema's ook meer aandacht voor de regionale invalshoek. Hij suggereert dat de levensloopbenadering niet alleen op het niveau van

personen, maar ook op meso-niveau, gemeenten en ook van bedrijven een vruchtbare benadering zou kunnen zijn. Ook voor regionale eenheden kunnen structurele ontwikkelingen als levensloopcyclus geïnterpreteerd worden.

Leo Engberts vertelt dat juridische beperkingen bestaan in het gebruik van het SSB als steekproefkader: dit kan alleen wanneer sprake is van samenwerking met het CBS, en het CBS dus in het onderzoek in kwestie participeert. In aanvulling hierop suggereert Bart Bakker dat het aanvragen van een eigen GBA-autorisatie door onderzoekers een alternatief kan zijn omdat deze aanvragen relatief snel en ook vaak worden gehonoreerd.

De derde vraag

voor panel en zaal gaat erover hoe de bruikbaarheid van het SSB vergroot zou kunnen worden voor onderzoekers buiten het CBS. Allereerst wordt het panel gevraagd of het hierover in aanvulling op de voorgaande discussies op dit punt nog iets wil toevoegen.

Bert Schijf acht het bezwaarlijk dat het vaak zo lang duurt om afspraken te maken over het gebruik van gegevens. Het zou makkelijker moeten worden om snel toegang te krijgen tot gegevensbestanden die men zou willen raadplegen. Daarnaast moet ook de PR van het CBS verbeteren, om iets te doen aan de eerder vermelde onbekendheid van SSB en StatLine.

Kea Tijdens benadrukt dat ze de toegankelijkheid en rijkdom van StatLine (vooral de grotere mogelijkheden om eigen tabellen samen te stellen) heel groot vindt en dat is een van de eerste dingen die ze eerstejaars studenten onder de aandacht brengt. Als je net iets anders wilt dan op StatLine staat, is dat in haar ervaring vaak mogelijk, en volstaat vaak een e-mailtje naar het CBS.

Clara Mulder vraagt of iets gezegd kan worden over eventuele problemen bij het gebruik van specialistische software bij remote access. Jan Jonker vertelt dat soortgelijke kwesties bij het on-site werken zich ook wel eens voordoen: tot op heden bleek het steeds mogelijk om een oplossing te vinden door de software op het CBS te installeren en is ook een oplossing te vinden voor eventuele licenties. Hij verwacht daarom ook dat die zaken ook bij remote access in principe geregeld kunnen worden.

Harry van den Tillaart is blij met de verruimde mogelijkheden, en pleit voor het beter benutten van de bestaande on-site mogelijkheden. Huib van de Stadt meldt dat de bestaande ruimte van 12 plaatsen voorlopig ruimschoots aan de behoefte lijkt te voldoen, want er is nog geen ruimtegebrek.

Panel en zaal hebben aan het eind van de discussie geen behoefte meer om verder in te gaan op de presentaties in de loop van de dag.

Na deze discussie wordt de dag afgesloten door Wim van Nunspeet, directeur van de divisie Sociale en Ruimtelijke Statistiek van het CBS. Hij dankt alle organisatoren, sprekers en toehoorders voor hun collectieve participatie in wat hij -als nieuwbakken directeur- ook heeft beleefd als een diepgaande introductie in het SSB.

Slotconclusies

Na alle lezingen, referaten en discussies, lijkt het nuttig om in kort bestek vast te leggen wat de voornaamste standpunten waren, maar ook welke behoeften zijn uitgesproken, als oogst van dit symposium.

We concluderen dat het CBS met de huidige stand van zaken van het SSB, de beloften die tijdens het voorgaande symposium twee jaar geleden werden gedaan, voor het overgrote deel is nagekomen. Er is er een breed gedeelde waardering voor het nut en de bruikbaarheid van het SSB.

Men was unaniem van mening dat het aantal koppelbare databestanden zeker nog kon en moest worden uitgebreid. Maar ook dat de kwaliteit van de huidige informatie, zoals dat van onderwijs, kwalitatief en kwantitatief nog kon worden verbeterd. Dat het CBS op grond van zijn capaciteit prioriteiten moet stellen, was evident.

Een punt van zorg, dat uit verschillende hoeken van de vertegenwoordigde instituties bij dit symposium werd geuit, betrof het gevaar van 'verschraling' als het CBS zich (teveel) zou beperken tot gegevens uit registraties. Bepaalde maatschappelijke aspecten, als meningen en belevingen, kunnen alleen door survey's worden onderzocht. Bovendien zijn er nog andere gegevens noodzakelijk die niet (tijdig) in een registratie zijn opgenomen en dus ook alleen via enquêtes kunnen worden gemeten. Het SSB krijgt een nog groter nut als meer gegevens van surveys daarin worden opgenomen. Een dringend verzoek dan ook vanuit de wetenschappelijke wereld aan het CBS om ook de surveys op het terrein van personen en huishoudens op de agenda te houden. Zo hield Peter Ester tijdens zijn referaat een pleidooi voor een grote coalitie van het CBS met kennisinstituten om grootschalige surveystudies op te zetten. Het CBS en deze instituten brengen hun databestanden in een kennispartnerschap, waarbij het CBS zorgt voor micro-koppeling en voor verrijking vanuit registers. Het instituut DANS kan daarbij een intermediaire rol spelen.

Curriculum vitae - dagvoorzitter, sprekers, referenten en deelnemers paneldiscussie

Symposium Sociale Samenhang in Beeld, het SSB nu en straks

Arts, Koos (Centraal Bureau voor de Statistiek [CBS])

- Senior onderzoeker.
- Projectleider SSB: vanaf het eerste moment betrokken bij de opzet en de verdere ontwikkeling van het Sociaal Statistisch Bestand (SSB).

Bakker, Bart (Centraal Bureau voor de Statistiek [CBS])

- Manager van de taakgroep Sociaal-economisch Totaalbeeld, waar het Sociaal Statistisch Bestand wordt geproduceerd en verder wordt uitgebouwd.
- Een van de grondleggers van het SSB.
 - Specialisatie: o.m. school- en beroepsloopbanen, sociale ongelijkheid, criminaliteit, en methodologie.

Bunt, Henk v.d. (Erasmus Universiteit Rotterdam EUR)

- Hoogleraar Criminologie.
- Voorheen Directeur WODC.
- Heeft als lid van de Begeleidingscommissie een stimulerende rol gespeeld in het stand brengen van de CBS-publicatie “Criminaliteit en Rechtshandhaving”.

Dagevos, Jaco (Sociaal en Cultureel Planbureau SCP)

- Senior onderzoeker bij het SCP.
- Gespecialiseerd in onderzoek naar de structurele en sociaal-culturele integratie van etnische minderheden. Publiceerde in dit verband onder meer over de arbeidsmarktpositie van allochtonen en over de vrijetijdscontacten tussen allochtonen en autochtonen.
- Voorheen werkzaam bij het Instituut voor Sociologisch-Economisch Onderzoek (ISEO).
- Heeft samen met het CBS gewerkt aan het “Jaarrapport integratie”.

Doorn, Peter (Data Archiving and Networked Services DANS)

- Directeur.
- Heeft nauw samengewerkt met het CBS op het terrein van de Volkstellingen 1960 en 1971.

Engbersen, Godfried (Erasmus Universiteit Rotterdam EUR)

- Hoogleraar Algemene Sociologie aan de Faculteit Sociale Wetenschappen Erasmus Universiteit Rotterdam.
- Lid van het bestuur van de Sociaal-Wetenschappelijke Raad van de Koninklijke Nederlandse Academie van Wetenschappen (KNAW).
- Lid van de Adviescommissie voor Vreemdelingenzaken (ACVZ).
- Nederlandse correspondent voor de Reporting Systemen Migration (soPEMI) van de (OECD).

Ester, Peter (Organisatie voor Strategisch Arbeidsonderzoek OSA / Universiteit van Tilburg UvT)

- Hoogleraar Sociologie.
- Directeur OSA: van daaruit nauw betrokken bij strategisch onderzoek Sociale Dynamiek en Arbeidsmarkt.
- Voorzitter van de CBS-Adviescommissie Maatschappij en Economie.
- Lid CBS Raad van Advies.

Gouweleeuw, José (Centraal Bureau voor de Statistiek [CBS])

- Gepromoveerd in de Wiskunde (“On ranges of vector measures and optimal stopping”).
- Statistisch onderzoeker bij het CBS, betrokken bij het Sociaal Statistisch Bestand.
- Gespecialiseerd in toepassingen van steekproeftheorie (trekken van steekproeven en weging).

Latten, Jan (Centraal Bureau voor de Statistiek CBS / Universiteit van Amsterdam UvA)

- Senior Sociaal-demografisch onderzoeker (CBS).
- CBS-woordvoerder sociale statistieken.
- Bijzonder Hoogleraar Demografie aan de UvA, in het bijzonder demografische en ruimtelijke aspecten van relatie- en gezinsvorming.

Leeuw, Frans (Wetenschappelijk Onderzoek en Documentatiecentrum WODC Ministerie van Justitie / Universiteit Maastricht UM)

- Directeur WODC.
- Tot mei 2006 hoogleraar Evaluatie-onderzoek Universiteit Utrecht. Vanaf mei 2006 hoogleraar Recht Openbaar bestuur en Sociaal-Wetenschappelijk onderzoek Universiteit Maastricht.
- Lid van de CBS-Adviescommissie Maatschappij en Economie.
- Voorzitter Dutch Evaluation Society.

- Werkt nauw samen met CBS in het kader van de projecten Integratiekaart en ‘Verdachten naar herkomstgroepering’.

Mulder, Clara H. (Universiteit van Amsterdam UvA)

- Clara H. Mulder (1962) studeerde sociale geografie aan de Universiteit van Amsterdam en promoveerde daar in 1993 op het proefschrift ‘Migration Dynamics: A Life Course Approach’.
- Van 1993 tot 2001 was zij Universitair Docent en later Universitair Hoofddocent aan de Universiteit Utrecht.
- In 2001 werd zij hoogleraar Demografie en Ruimte aan de Universiteit van Amsterdam.
- Zij leidt het onderzoeksprogramma ‘Residential Choice in a Family Context’, waarvoor zij een VICI-beurs voor innovatief onderzoek heeft gewonnen van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). In het kader van dit onderzoeksprogramma heeft zij een samenwerkingsverband samenwerking met het CBS op het terrein van Ruimtelijke en Sociale Mobiliteit (RSM), die o.a. betrekking heeft op de relatie tussen ouders en kinderen.

Nunspeet, Wim van (Centraal Bureau voor de Statistiek [CBS])

- Directeur Sociale en Ruimtelijke Statistieken.
- Voorheen Plv. Directeur Bedrijfseconomische Statistieken.

Prins, Kees (Centraal Bureau voor de Statistiek [CBS])

- Statistisch onderzoeker,
- Specialisatie: demografie en aspecten van de bevolkingsadministratie.

Schijf, Bert (Universiteit van Amsterdam UvA)

- Universitair Docent bij de Afdeling Sociologie/Antropologie van de Universiteit van Amsterdam.
- Publiceerde recent zowel over Nederlandse adel (met Jaap Dronkers), waarbij geavanceerde technieken werden gebruikt, als over een Turkse Bezorgpizzeria (samen met Ilya de Jong), hetgeen op antropologisch veldwerk was gebaseerd.
- Als redactielid van Mens en Maatschappij schreef hij een marginaal getiteld: “Leiden grote datasets tot saaie artikelen?”

Sinderen, Jarig van (NMA / Erasmus Universiteit Rotterdam EUR)

- Voormalig Plv. Directeur-Generaal CBS met portefeuille statistisch beleid.
- Hoogleraar Economische Politiek (EUR).
- Voorheen werkzaam bij het Directoraat Algemene Economische Politiek (AEP) van het Ministerie van Economische Zaken, en destijds vanuit zijn leidinggevende functie nauw betrokken bij CBS-aangelegenheden als o.m. de laatste CBS-wet, de GBA-wet en de Volkstelling.

Slot, Jeroen (Gemeente Amsterdam)

- Adjunct directeur Dienst Onderzoek en Statistiek Gemeente Amsterdam en Hoofd Onderzoek en Beleidsinformatie.
- Deze afdeling heeft van oudsher een werkrelatie met het CBS, met name over demografische ontwikkelingen.

Stadt, Huib van de (Centraal Bureau voor de Statistiek [CBS])

- Tot en met 30 november 2005: hoofd van de sector Statistische Analyse Personen Voorburg (SAV), waar o.m. het SSB en de Demografische en Ruimtelijke Statistieken zijn ondergebracht.
- Vanaf 1 december 2005: hoofd van de sector Publicaties en Communicatie.

Theeuwes, Jules (Universiteit van Amsterdam UvA, SEO)

- Directeur Stichting Economisch Onderzoek (SEO)
- Hoogleraar Toegepast Economisch Onderzoek Universiteit van Amsterdam, bijzondere aandacht voor arbeidsmarkteconomie en vergrijzing.
- Lid van de WRR, in die hoedanigheid het CBS verzocht om werkgevers- en werknemersgegevens te koppelen.
- Algemeen Directeur van ENCORE (International Network for Competition and Regulation).
- Voorheen voorzitter van de CBS-Adviescommissie Arbeidsstatistieken.
- Lid van de CBS-Adviescommissie Sociale en Ruimtelijke Statistieken.

Tijdens, Kea (Universiteit van Amsterdam UvA / Erasmus Universiteit Rotterdam EUR)

- Onderzoekscoördinator Amsterdams Instituut voor Onderwijs en Arbeidsmarkt (AIAS / UvA).
- Hoogleraar ‘Vrouwen en Werkgelegenheid’ aan de EUR.
- Voorheen lid van de CBS-Adviescommissie Leefsituatiesurvey (LSS).

Tillaart, Harry van den (Instituut voor Toegepaste Sociale Wetenschappen ITS Nijmegen)

- Onderzoeker ITS.
- Werkt nauw samen met het CBS: heeft al diverse malen on-site aan microbestanden van het SSB gewerkt, o.m. project etnisch ondernemerschap.

Wolbers, Maarten H.J. (Vrije Universiteit Amsterdam VU)

- Universitair docent Faculteit Sociale Wetenschappen, afd. Methoden en Technieken aan de VU.
- Voorheen: Senior onderzoeker bij het Researchcentrum voor Onderwijs en Arbeidsmarkt (ROA) van de Universiteit Maastricht.
- Werkt met het CBS samen aan het Strategisch Programma 'Sociale dynamiek en Arbeidsmarkt', met name op het terrein van 'schoolverlaters'.

Afkortingenlijst SSB Symposium 2005

<i>Afkorting</i>	<i>Betekenis/omschrijving</i>
ABW	Algemene Bijstandswet
AO	Arbeidsongeschiktheidsuitkering
BBP	Bruto Binnenlands Product
BZK	Binnenlandse Zaken en Koninkrijksrelaties
CAO	Collectieve Arbeidsovereenkomst
CAPI	Computer Assisted Personal Interview
CASI	Computer Assisted Self Interview
CBS	Centraal Bureau voor de Statistiek
CCS	Centrale Commissie voor de Statistiek
CFI	Centrum Financiële Instellingen
CRIHO	Centraal Register Inschrijvingen Hoger Onderwijs
CRV	Centraal Register Vreemdelingen
CWI	Centrum voor Werk en Inkomen
DANS	Data Archiving and Networked Services
DJI	Dienst Justitiële Inrichtingen
EBB	Enquête Beroepsbevolking
ERR	Examen Resultaten Register
ESA	European Space Agency
ESB	Economisch Statistisch Bestand
EUR	Erasmus Universiteit Rotterdam
EWL	Enquête Werkgelegenheid en Lonen
FiBase	Inkomensbestand van de belastingdienst
FIOD	Fiscale Inlichtingen- en Opsporingsdienst
GBA	Gemeenschappelijke Basisadministratie persoonsgegevens
GBR	Geografisch Basisregister
GGZ	Geestelijke Gezondheidszorg
GION	Gronings Instituut voor Onderzoek van Onderwijs, Opvoeding en Ontwikkeling
GIS	Geografisch Informatie Systeem
GW	Gevangeniswezen
HALT	Het Alternatief (voor een straf via justitie)
HAVO	Hoger Algemeen Voortgezet Onderwijs

HBO	Hoger Beroepsonderwijs
HKS	Herkenningsdienst Systeem
IPUMS	International Public Use Microdata Series
ISEO	Instituut voor Sociologisch-Economisch Onderzoek
ITS	Instituut voor Toegepaste Sociale Wetenschappen
JJI	Justitiële Jeugdinstellingen
KLPD	Korps Landelijke Politiediensten
KNAW	Koninklijke Nederlandse Academie voor Wetenschappen
LADIS	Landelijke Alcohol en Drugs Informatie Systeem
LAS	Leefsituatie Allochtone Stedelingen
LBO	Lager Beroepsonderwijs
MAVO	Middelbaar Algemeen Voortgezet Onderwijs
MBO	Middelbaar Beroepsonderwijs
MCB	Monitor Criminaliteit Bedrijfsleven
NIBUD	Nationaal Instituut voor Budgetvoorlichting
NIVEL	Nederlands Instituut voor onderzoek van Eerstelijns gezondheidszorg
NSCR	Nederlands Studiecentrum Criminaliteit en Rechtshandhaving
NWO	Nederlandse organisatie voor Wetenschappelijk Onderzoek
OBJD	Onderzoeks- en Beleidsdatabase Justitiële Documentatie
OGJA	Onderzoek Gezinsvorming Jonge Allochtonen
OM	Openbaar Ministerie
OMDATA	OM Data
OSA	Organisatie voor Strategisch Arbeidsmarktonderzoek
PC	Personal Computer
PL	Persoonslijst
PolMon	Politiemonitor
POLS	Permanent Onderzoek Leefsituatie
PRIMA	Cohortonderzoek Primair Onderwijs
RIO	Regionaal Inkomensonderzoek
RIVM	Rijksinstituut voor de Volksgezondheid en Milieuhygiëne
SBI	Standaard Bedrijfsindeling
SCO	Stichting Centrum voor Onderwijsonderzoek
SCP	Sociaal Cultureel Planbureau
SISWO	Stichting Interuniversitair Instituut voor Sociaal Wetenschappelijk Onderwijs
SOFI	Sociaal Fiscaal
SOI	Standaard Onderwijs Indeling
SPVA	Sociale Positie en Voorzieningengebruik

SSB	Sociaal Statistisch Bestand
SVG	Stichtelijke Verslavingsreclassering GGZ Nederland
TULP	Ten Uitvoerlegging Programma
VBO	Vorbereidend Beroepsonderwijs
VMBO	Vorbereidend Middelbaar Beroepsonderwijs
VO	Voortgezet Onderwijs
VOCL	Voortgezet Onderwijs Cohort Leerlingen
VROM	Volkshuisvesting, Ruimtelijke Ordening en Milieubeheer
VT	Volkstelling
VU	Vrije Universiteit
VWO	Vorbereidend Wetenschappelijk Onderwijs
VWS	Volksgesondheid, Welzijn en Sport
VZA	Verzekerdenadministratie
WALVIS	Wet Administratieve Lastenverlichting en Vereenvoudiging In Sociale verzekeringswetten
WBO	Woningbehoefteonderzoek
WO	Wetenschappelijk Onderwijs
WODC	Wetenschappelijk Onderzoek- en Documentatiecentrum
WoOn	Woononderzoek Nederland
WOZ	Wet Waardering Onroerende Zaken
WRG	Woningregister
WRR	Wetenschappelijke Raad voor het Regeringsbeleid
WSA	Wetenschappelijk Statistisch Agentschap
WSF	Wet op Studiefinanciering
WVC	Welzijn, Volksgezondheid en Cultuur
WW	Werkloosheidswet

Lijst van publicaties over het Sociaal Statistisch Bestand en de Virtuele volkstelling

1996

- Bochove, C.A. van, 1996, Living without a Census: The Dutch View, In: Proceedings of the Expert Group Meeting on Innovative Techniques for Population Censuses and Large-Scale Demographic Surveys, The Hague, Netherlands, 22-26 April 1996, Netherlands Interdisciplinary Demographic Institute and United Nations Population Fund, blz. 119-121. The Hague: NIDI; New York, NY: UNFPA.
- Bochove, C.A. van & P.C.J. Everaers, 1996, Micro-macro and Micro-micro Linkage in Social Statistics, In: The Future of European Social Statistics: Use of Administrative Registers and Dissemination Strategies; Proceedings of the Mondorf Seminar, Third session, Mondorf-les-Bains, Luxembourg, 25 and 26 January 1996, Eurostat, blz. 205-212. Ed. Bernard Grais. Statistical Document. Theme 0: Miscellaneous, Series D: Studies and research. Luxembourg: Office for Official Publications of the European Communities.

1997

- Hen, P. de, 1997, Virtuele volkstelling, In: Elsevier, 29 november 1997, blz. 66-67.
- Kooiman, P., 1997, Een statistisch micro-bestand van de Nederlandse bevolking: Enkele methodologische aandachtspunten. Centraal Bureau voor de Statistiek, Divisie Research en Ontwikkeling, Research Paper Nr. 9752. Voorburg: Centraal Bureau voor de Statistiek, Divisie Research en Ontwikkeling. November 1997.
- Laan, P. van der, 1997, Census based on Integration of Administrative Data and Survey Data: The Dutch Experience, In: Census Belgica 2001: Tweedaagse studiedag over het gebruik en de toekomst van de Volkstelling in België, Brussel, 12 en 19 november 1996, Steunpunt Werkgelegenheid, Arbeid en Vorming en Steunpunt Demografie. Verslagboek deel 2, De toekomst van de Volkstelling in België, blz. 94-107. Leuven: Steunpunt Werkgelegenheid, Arbeid en Vorming.

1999

- Arts, C.H. & F.J. van Lith, Een nieuwe weg voor statistieken over personen: het Sociaal Statistisch Bestand (SSB), In: Sociaal-Economische Maandstatistiek, jrg. 1999, nr. 9, blz. 22-29.
- Vliegen, J.M. & P. van der Laan, 1999a, Der „Zensus“ in den Niederlanden: Eine Integration von Register- und Stichprobendaten, In: Volkszählung 2001: Von der traditionellen Volkszählung zum Registerzensus, hrsg. H. Grohmann, H. Sahner und R. Wiegert, blz. 15-23. Sonderhefte zum Allgemeinen Statistischen Archiv, Heft 33. Göttingen: Vandenhoeck & Ruprecht.
- Vliegen, J.M. & P. van der Laan, 1999b, Methodische und zeitliche Aspekte der Umstellung der amtlichen Statistik auf Register am Beispiel der Niederlande, In: Allgemeines Statistisches Archiv, Bd. 83 (Oktober-Dezember 1999): blz. 434-446.
- Vliegen, J.M. en P. van der Laan, 1999c, Methodische und zeitliche Aspekte der Umstellung der amtlichen Statistik auf Register am Beispiel der Niederlande, In: Statistik und Informationsmanagement, Jg. 58 (10): blz. 256-260.

2000

- Al, P. & J.W. Altena, 2000, Data security, privacy and the SSB, In Netherlands Official Statistics, jrg. 15, nr. summer, blz. 47-50.
- Al, Pieter en Bart F.M. Bakker (eds.), 2000, Re-engineering Social Statistics by micro-integration of different sources. Themanummer Netherlands Official Statistics, jrg. 15, nr. summer.
- Al, Pieter en Bart F.M. Bakker, 2000, Re-engineering Social Statistics by micro-integration of different sources. An introduction. Themanummer Netherlands Official Statistics, jrg. 15, nr. summer, blz. 4-6.
- Arts, Koos, Bart F.M. Bakker, en Erik van Lith, 2000, Linking administrative registers and household surveys, In: Pieter Al en Bart F.M. Bakker (eds.), Re-engineering Social Statistics by micro-integration of different sources. Themanummer Netherlands Official Statistics, jrg. 15, nr. summer, blz. 16-22.

- Bakker, Bart F.M., en Johan van Rooijen, 2000, One figure for the supply and demand of services, In: Pieter Al en Bart F.M. Bakker (eds.), Re-engineering Social Statistics by micro-integration of different sources. Themanummer Netherlands Official Statistics, jrg. 15, nr. summer blz. 40-46.
- Geuzinge, Linda, Johan van Rooijen en Bart F.M. Bakker, 2000, The use of administrative registers to reduce non-response bias in household surveys, In: Pieter Al en Bart F.M. Bakker (eds.), Re-engineering Social Statistics by micro-integration of different sources. Themanummer Netherlands Official Statistics, jrg. 15, nr. summer blz. 32-39.
- Laan, P. van der, 2000, Integrating Administrative Registers and Household Surveys, In: Netherlands Official Statistics, Vol. 15 (Summer 2000): Special Issue, Integrating Administrative Registers and Household Surveys, P.G. Al en B.F.M. Bakker, red., blz. 7-15.
- Trijssenaar, M. en A.H. Kroese. 2000. Estimation of Census Tables from the SSB95: A Weighting Approach. Statistics Netherlands, Division for Research and Development, Research Paper No. 0025. Voorburg: Statistics Netherlands, Division for Research and Development. July 2000.

2001

- Laan, P. van der, 2001, The 2001 Census in the Netherlands: Integration of Registers and Surveys, In: INSÉE-EUROSTAT Seminar on Censuses After 2001, Paris, France, 20 and 21 November 2000, Institut National de la Statistique et des Études Économiques, blz. 39-52. La collection «Regards à l'étranger» N^o 9A. Paris: INSÉE. December 2001.

2002

- Arts, C.H. & E.M.J. Hoogteijling, Het Sociaal Statistisch Bestand 1998 en 1999, In: Sociaal-Economische Maandstatistiek, jrg. 2002, nr. 12, blz. 13-21.
- Arts, C.H., L. van Toor & S. de Vries, Stromen op de arbeidsmarkt, april-oktober 1999, In: Sociaal-Economische Maandstatistiek, jrg. 2002, nr. 12, blz. 66-71.
- Bakker, Bart F.M., 2002. Statistics Netherlands' Approach to Social Statistics: The Social Statistical Dataset, In: OECD Statistics Newsletter, vol. 2002, nr. 11, blz. 4-6.

- Boerdam, A., Hoeveel allochtonen ontvangen een uitkering?, In: Sociaal-Economische Maandstatistiek, jrg. 2002, nr. 12, blz. 35-41.
- Corpeleijn, A.W.F. en M.J. Heerschop, Inkomende en uitgaande pendel, 1998, In: Sociaal-Economische Maandstatistiek, jrg. 2002, nr. 12, blz. 53-58.
- Cruchten, J.M.J. van, M.I. Hartgers en E. Schulte Nordholt, Loon naar opleidings- en beroepsniveau, In: Sociaal-Economische Maandstatistiek, jrg. 2002, nr. 12, blz. 72-81.
- Das, M., L. van Toor & C.H. Arts, Samenloop van banen en uitkeringen, In: Sociaal-Economische Maandstatistiek, jrg. 2002, nr. 12, blz. 22-28.
- Florquin, F., Banen van studenten in het hoger onderwijs, In: Sociaal-Economische Maandstatistiek, jrg. 2002, nr. 12, blz. 59-65.
- Linder, F.S., Bronnen van inkomen in de regio, In: Sociaal-Economische Maandstatistiek, jrg. 2002, nr. 12, blz. 42-52.
- Melser, C. & L. van Toor, Hoeveel allochtonen zijn werknemer?, In: Sociaal-Economische Maandstatistiek, jrg. 2002, nr. 12, blz. 29-34.
- Weidum, J. en F.S. Linder, Arbeidspositie en opleidingsniveau van personen met een uitkering, 1998, In: Sociaal-Economische Maandstatistiek, jrg. 2002, nr. 12, blz. 82-92.

2003

- Bakker, B.F.M., 2003, Hoe nieuw zijn nieuwe ideeën?, In: J. Nobel, S. Algera, M. Biemans & P. van der Laan (red.), Gedacht en gemeten (Voorburg/Heerlen: CBS), blz. 123-132.
- Bakker, B., & K. Arts, 2003, Dynamiek op de arbeidsmarkt; gegevens over stromen uit het Sociaal Statistisch Bestand. In: B.F.M. Bakker & L. Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS) blz. 59-70.
- Bakker, B.F.M. & L. Putman (red.), 2003, De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS).
- Bakker, B.F.M. & L. Putman, 2003, Inleiding, In: B.F.M. Bakker & Lisa Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS), blz. 7-9.

- Beer, P. de, 2003, CBS wees zuinig op de enquêtes! Een reactie op het paper van B. Bakker en K. Arts, In: B.F.M. Bakker & Lisa Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS), blz. 71-72.
- Corpeleijn, A.W.F., Tijdelijke en langdurige banen, 2000, In: Sociaal-economische maandstatistiek, jrg. 2003, nr. 9, blz. 12-18.
- Dronkers, J., 2003, Het belang van een éénprocentsteekproef uit de huishoudens van de virtuele volkstelling, In: B.F.M. Bakker & Lisa Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS), blz. 25-30.
- Everaers, P., 2003, Het Sociaal Statistisch Bestand (SSB): de statistiek van de toekomst, In: B.F.M. Bakker & Lisa Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS), blz. 47-56.
- Gierveld, J., 2003, De 'virtuele volkstelling' als databron voor sociaal-wetenschappelijk onderzoek; optimaliseringsvragen. Een reactie op de papers van J. van Maarseveen en J. Dronkers, In: B.F.M. Bakker & Lisa Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS), blz. 31.
- Hartog, J., 2003, LSO is een mooi bestand, maar...Een reactie op het paper van E. Schulte Nordholt, In: B.F.M. Bakker & Lisa Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS), blz. 81-82.
- Hoog, C. de, 2003, De 'virtuele Volkstelling' is geen kopie van een traditionele Volkstelling. Een reactie op het paper van E. Schulte Nordholt, In: B.F.M. Bakker & Lisa Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS), blz. 43-46.
- Latten, J., 2003, Ruimtelijke spreiding van mensen en hun kenmerken Zicht op autochtonen en allochtonen, In: B.F.M. Bakker & Lisa Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS), blz. 83-97.
- Maarseveen, J. van, 2003, Volkstellingen. De betekenis voor historisch sociaal-wetenschappelijk onderzoek, In: B.F.M. Bakker & Lisa Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS), blz. 11-23.

- Nobel, J., 2003, De deur van het demografisch laboratorium. Perspectieven voor externe benutting van de virtuele volkstelling 2001, In: B.F.M. Bakker & Lisa Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS), blz. 101-120.
- Schulte Nordholt, E., 2003, De 'virtuele volkstelling' 2001, In: B.F.M. Bakker & Lisa Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS), blz. 33-42.
- Schulte Nordholt, E., 2003, Loonstructuuronderzoek – verrijkte data door integratie, In: B.F.M. Bakker & Lisa Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS), blz. 73-80.
- Tesser, P., 2003, Nog geen zicht op allochtonen. Een reactie op het paper van J. Laten, In: B.F.M. Bakker & Lisa Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS), blz. 98-99.
- Tijdens, K., 2003, Hoe ver gaat de deur van het CBS open? Een reactie op het paper van J. Nobel, In: B.F.M. Bakker & Lisa Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS), blz. 121 -122.
- Toor, Leo van & Rik van der Vliet, Wet SAMEN: meeste grote bedrijven hebben weinig allochtone werknemers, In: Webmagazine, 3 november 2003.
- Velden, R. van der, 2003, SSB: wensdroom of nachtmerrie. Een reactie op het paper van P. Everaers, In: B.F.M. Bakker & Lisa Putman (red.), De virtuele volkstelling en SSB (Amsterdam: SISWO/CBS), blz. 57-58.

2004

- Advokaat, Wijnand & Frank Linder, 2004, Active ageing: the senior labour force from work to retirement, In: Eric Schulte Nordholt, Marijke Hartgers, Rita Gircour (eds.), The Dutch Virtual Census of 2001, Analysis and Methodology, Statistics Netherlands, pp. 73-95.
- Bakker, B.F.M., & L. Putman, 2004, De virtuele Volkstelling een feit, In: *Historia & Informatica*, Jrg. 11, nr. 1, blz. 1-2.
- Bakker, B.F.M., 2004, Nederland heeft weer een Volkstelling! In: *Facta*, jrg. 12, nr. 2, blz. 16-19.

- Chamberlain, Jessica & Eric Schulte Nordholt, 2004, The results of the 2001 Census in the Netherlands, the United Kingdom and some other European countries, In: Eric Schulte Nordholt, Marijke Hartgers, Rita Gircour (eds.), *The Dutch Virtual Census of 2001, Analysis and Methodology*, Statistics Netherlands, pp. 225-241.
- Corpeleijn, André & Marcel Kerkhofs, Arbeidsmarkttransities van oudere werknemers 2000-2001, In: *Sociaal Economische Trends*, jrg. 2004, nr. 4, blz. 19-28.
- Corpeleijn, André, Meeste werknemers rond 60 jaar met pensioen, Webmagazine-artikel, In: *Sociaal Economische Trends*, jrg. 2004, nr. 2, blz. 56.
- Das, Marjolijn, Hoge uitstroom uit arbeidsongeschiktheid bij allochtonen, In: *Webmagazine*, 23 augustus 2004.
- Gouweleeuw, José & Carel Harmsen, 2004, Foreigners at work, In: Eric Schulte Nordholt, Marijke Hartgers, Rita Gircour (eds.), *The Dutch Virtual Census of 2001, Analysis and Methodology*, Statistics Netherlands, pp. 97-117.
- Gouweleeuw, José & Carel Harmsen, Allochtonen aan het werk, In: *Bevolkingstrends*, jrg. 2004, nr. 4, blz. 75-84.
- Gouweleeuw, José en Carel Harmsen. Tweede generatie succesvoller dan eerste generatie. In: *Webmagazine*, 27 december 2004.
- Gouweleeuw, José & Marijke Hartgers, 2004, The method of repeated weighting in the 2001 Census, In: Eric Schulte Nordholt, Marijke Hartgers, Rita Gircour (eds.), *The Dutch Virtual Census of 2001, Analysis and Methodology*, Statistics Netherlands, pp. 261-276.
- Hartgers, Marijke, De virtuele Volkstelling 2001: gezin en werk, In: *Bevolkingstrends*, jrg. 2004, nr. 4, blz. 68-74.
- Hartgers, Marijke, 2004, Families at work, In: Eric Schulte Nordholt, Marijke Hartgers, Rita Gircour (eds.), *The Dutch Virtual Census of 2001, Analysis and Methodology*, Statistics Netherlands, pp. 23-39.
- Hartgers, Marijke & José Gouweleeuw, 2004, Working people: what they do, who they are and where they live, In: Eric Schulte Nordholt, Marijke Hartgers, Rita Gircour (eds.), *The Dutch Virtual Census of 2001, Analysis and Methodology*, Statistics Netherlands, pp. 41-58.

- Hartgers, Marijke & Eric Schulte Nordholt, De virtuele Volkstelling 2001, In: Sociaal Economische Trends, jrg. 2004, nr. 4, blz. 29-36.
- Houbiers, M. 2004, Towards a Social Statistical Database and Unified Estimates at Statistics Netherlands, In: Journal of Official Statistics, Vol. 20 (1): blz. 55-75.
- Linder, Frank, 2004, The use of administrative registers and sample surveys in the Dutch Census of 2001, In: Eric Schulte Nordholt, Marijke Hartgers, Rita Gircour (eds.), The Dutch Virtual Census of 2001, Analysis and Methodology, Statistics Netherlands, pp. 243-260.
- Linder, Frank, 2004, The Dutch Virtual Census; a new approach by combining administrative registers and sample surveys, In: Austrian Journal of Statistics, vol. 33, no. 1+2, Special Issue on Data Mining and Record Matching, Proceedings of the DIECOFIS Workshop in Vienna, November 13-14, 2003, pp. 69-88.
- Maarseveen, Jacques van, 2004, The Dutch Virtual Census of 2001 compared to previous censuses, In: Eric Schulte Nordholt, Marijke Hartgers, Rita Gircour (eds.), The Dutch Virtual Census of 2001, Analysis and Methodology, Statistics Netherlands, pp. 119-141.
- Rijn, A.S. van, A. Zorlu, R.V. Bijl & B.F.M. Bakker, 2004, De ontwikkeling van een Integratiekaart WODC-Cahier 2004-9 (Den Haag/Voorburg: WODC/CBS).
- Rijn, A.S. van, A. Zorlu, R.V. Bijl & B.F.M. Bakker, 2004, The development of an integration monitor of first and second generation immigrants in the Netherlands WODC-Cahier 2004-9a (Den Haag/Voorburg: WODC/CBS).
- Schulte Nordholt, Eric, 2004, Introduction to the Dutch Virtual Census of 2001, In: Eric Schulte Nordholt, Marijke Hartgers, Rita Gircour (eds.), The Dutch Virtual Census of 2001, Analysis and Methodology, Statistics Netherlands, pp. 9-22.
- Schulte Nordholt, Eric, 2004, The Dutch Virtual Census of 2001: Interview with Eric Schulte Nordholt, In: ISI Newsletter, Volume 28, no. 3, blz. 14-15.
- Trijssenaar, Marjolijn & Marijke Hartgers, 2004, Young people and their economic activities, In: Eric Schulte Nordholt, Marijke Hartgers, Rita Gircour (eds.), The Dutch Virtual Census of 2001, Analysis and Methodology, Statistics Netherlands, pp. 59-72.

- Vliegen, Mathieu, Werken in het stadsgewest: herkomst en bestemming van forensen, In: Sociaal Economische Trends, jrg. 2004, nr. 4, blz. 37-46.
- Vliegen, Mathieu & Henk Oroh, 2004, The Netherlands: commuter country, In: Eric Schulte Nordholt, Marijke Hartgers, Rita Gircour (eds.), The Dutch Virtual Census of 2001, Analysis and Methodology, Statistics Netherlands, pp. 203-223.

2005

- Bakker, Bart, Registers voor statistiek: het Sociaal Statistisch Bestand, In: SDV-bulletin, 22 februari 2005, blz. 4-5.
- Bakker, B.F.M., A. Walberg & M. Blom, Jeugdige verdachten, In: M. Blom, M., J. Oudhof, R.V. Bijl & B.F.M. Bakker (red.), Verdacht van criminaliteit. Allochtonen en autochtonen nader bekeken. WODC-cahier 2005-2 (Den Haag/Voorburg: WODC/CBS), blz. 45-60.
- Bakker, Bart, Koos Arts en Leo van Toor, 2005, Arbeidsmarktstatistieken in Nederland. De opzet van het Sociaal Statistisch Bestand, In: Over Werk, jrg. 15, nr. 1, blz. 50-54.
- Bakker, Bart & Paulien Giesbertz, 2005, Heeft Cupido een maat(je)? De integratie van allochtonen op de huwelijksmarkt, In: Bevolkingstrends, jrg. 53, nr. 2, blz. 65-74.
- Bijl R.V., A. Zorlu, A.S. van Rijn, R.P.W. Jennissen, M. Blom, 2005, Integratiekaart 2005; de maatschappelijke integratie van migranten in de tijd gevolgd: trend- en cohortanalyses. WODC-Cahier 2005-16 (Den Haag/Voorburg: WODC/CBS).
- Blom, M., J. Oudhof, R.V. Bijl & B.F.M. Bakker (red.), 2005, Verdacht van criminaliteit. Allochtonen en autochtonen nader bekeken. WODC-cahier 2005-2 (Den Haag/Voorburg: WODC/CBS).
- CBS, 2005, De Nederlandse Economie 2004, Nieuwe broninformatie voor de Nationale rekeningen, blz. 31-33; Arbeidsongeschiktheid en werk, blz. 119-120; Aanvullende pensioenen van 65-plussers in Nederland, blz. 169-174; Arbeidsparticipatie, herintreding en uittreding van 50-plussers, blz. 175-182; Arbeidsmarktdynamiek in Nederland, blz. 192-198.

- Corpeleijn, André, Uittreding rond zestigste levensjaar verminderd, In: Economisch Statistische Berichten, 9 september 2005, blz. 402.
- Corpeleijn, André, Weinig vijftigers met baan in grote steden, In: Webmagazine, 20 juni 2005.
- Corpeleijn, André, Stoppen met werk rond 60^e jaar verminderd, In: Webmagazine, 12 september 2005.
- Corpeleijn, André en Koos Arts, Aantal allochtone ondernemers neemt toe, In: Webmagazine, 24 oktober 2005.
- Cramer, Mars, De volkstelling van 2001, Review van Eric Schulte Nordholt, Marijke Hartgers, Rita Gircour (red.), The Dutch Virtual Census of 2001, In: De Gids april 2005.
- Cramer, Mars, Volkstelling: nieuwe en interessante bestanden bij het CBS, In: Stator, december 2005/4.
- Garssen, M.J., H. Nicolaas en A.H. Sprangers, 2005, Demografie van de allochtonen in Nederland, In: Bevolkingstrends jrg. 2005, nr. 3, blz. 96-117.
- Garssen, M.J., H. Nicolaas en A.H. Sprangers, 2005, Demografie van allochtonen in Nederland. In: Handboek Interculturele zorg (Elsevier Bedrijfsinformatie), blz. I1.3-1 – I1.3-74.
- Garssen, Joop, en Aslan Zorlu, 2005, Demografie, In: SCP/WODC/CBS, Jaarrapport Integratie 2005 (Den Haag, SCP), blz. 14-27.
- Latten, Jan, Bas Hamers en Han Nicolaas, 2005, Uit balans. Selectieve verhuisstromen naar en uit de grote stad, In: Bevolkingstrends, jrg. 53, nr. 4, nog te verschijnen.
- Linder, Frank, Dynamiek in de WAO, WAZ en Wajong: een longitudinale analyse van personen met een arbeidsongeschiktheidsuitkering, 1999-2002, In: Sociaal Economische Trends, jrg. 2005, nr. 1, blz. 62-77.

- Mulder, Clara, 2005, The Dutch Virtual Census of 2001, Analysis and Methodology, Eric Schulte Nordholt, Marijke Hartgers, Rita Gircour (red.), boekbespreking, In: Mens en Maatschappij, jaargang 80, nr. 1, blz. 94-95.
- Maarseveen, Jacques van, Twee eeuwen Volkstellingen: de Virtuele Volkstelling 2001, vergeleken met haar voorgangers, In: Sociaal Economische Trends, jrg. 2005, nr. 1, blz. 33-38.
- Nicolaas, Han en Arno Sprangers, Vluchtelingen moeilijker aan het werk, In: Webmagazine, 22 augustus 2005.
- Nicolaas, Han en Arno Sprangers, Meeste gezinsvormende migranten hebben werk, In: Economisch Statistische Berichten, 18 november 2005, blz. 521.
- Oudhof, J., R.V. Bijl & B.F.M. Bakker, 2005, Inleiding, In: M. Blom, M., J. Oudhof, R.V. Bijl & B.F.M. Bakker (red.), Verdacht van criminaliteit. Allochtonen en autochtonen nader bekeken. WODC-cahier 2005-2 (Den Haag/Voorburg: WODC/CBS), blz. 11-26.
- Rijn, Annette van, Frits Huls en Aslan Zorlu, 2005, Jongeren en criminaliteit, In: SCP/WODC/CBS, Jaarrapport Integratie 2005 (Den Haag, SCP), blz. 146-165.
- Schulte Nordholt, Eric, 2005, The Dutch virtual Census 2001: A new approach by combining different sources, In: Statistical Journal of the United Nations ECE (verschijnt binnenkort).
- Sprangers, A.H., H. Nicolaas, A. Zorlu en J. Hartog, 2005, Vluchtelingen en gezinsmigranten vinden moeilijker werk, In: Sociaal-economische trends, jrg. 2005, nr. 4, blz. 29-37.
- Toor, Leo van, en Koos Arts, Een op de elf werknemers heeft auto van de zaak, In: Webmagazine, 12 december 2005.
- Zorlu, Aslan, en Tanja Traag, 2005, Opleidingsniveau en taalvaardigheid, In: SCP/WODC/CBS, Jaarrapport Integratie 2005 (Den Haag, SCP), blz. 44-56.

2006

- Arts, Koos, en André Corpeleijn, Veel minder mensen uit de bijstand aan het werk, In: Webmagazine, 2 januari 2006.
- Bart F.M. Bakker, José M. Gouweleeuw en Leo van Toor, Educational Attainment from different Registers and Surveys (Paper Fourth Seminar on Strategies for Social and Spatial Statistics, Oslo, 27-28 februari 2006).
- Bijl, R.V., M. Blom, J. Oudhof en B.F.M. Bakker, 2006, Criminaliteit, etniciteit en demografische ontwikkeling, in: Demografische Ontwikkelingen, Justitiële Verkenningen, jrg.32, nr.3, pp. 55-74.
- Das, Marjolijn, Tweede-generatie moeders werken vaker, In: Webmagazine, 19 juni 2006.
- Das, Marjolijn, Allochtone vrouwen: arbeidsdeelname en verandering in de gezinssituatie, In: Sociaal Economische Trends, jrg. 2006 nr. 3 (nog te verschijnen).
- Harmsen, Carel, en Koos Arts, Sociaal-economische positie van eerste generatie Antillianen en Arubanen in Nederland, In: Bevolkingstrends, jrg. 2006, nr. 2, blz. 45-47.
- Heijden, Peter van der, Eugene Zwane & Dave Hessen, 2006, Schatting van aantal in Nederland verblijvende Antillianen die niet ingeschreven zijn in de GBA, Een 'capture-recapture'-analyse in opdracht van het Ministerie van Justitie, IOPS-Utrecht, Universiteit van Utrecht.
- Jehoel-Gijsbers, Gerda, Frank Linder en Jurjen Iedema, Werkhervatting door arbeidsongeschikten, In: Trendrapportage ziekteverzuim, arbeidsongeschiktheid en reïntegratie 2006 ('s-Gravenhage: SCP) (verschijnt in 2007).
- Latten, Jan, Han Nicolaas en Bas Hamers, De prijs van migratie, In: Economisch Statistische Berichten, 24 februari 2006.
- Latten, Jan, Han Nicolaas en Bas Hamers, 2006, De prijs van migratie, In: Bevolkingstrends, jrg. 54, nr. 1, 37-44.
- Latten, Jan, Bas Hamers en Han Nicolaas, Selectieve verhuisstromen zetten grote steden onder druk, In: Christen Democratische Verkenningen, jrg. 27, nr. lente 2006, blz. 43-54.

- Latten, Jan, Henk de Feijter, Bas Hamers en Han Nicolaas, Uit Balans. In: WPRB-rapport herfst 2006 (nog te verschijnen).
- Merens, Ans, Saskia Keuzenkamp en Marjolijn Das, 2006, Combinatie van arbeid en zorg, In: Saskia Keuzenkamp en Ans Merens (red.), Sociale atlas van vrouwen uit etnische minderheden (Den Haag: SCP), blz. 91-127.
- Nicolaas, H. en A.H. Sprangers, 2006, Internationale migratie: Nederland in een Europese context. In: Allochtonen in Nederland in Europees perspectief. Boekaflevering van Mens en Maatschappij, Amsterdam University Press.
- Oudhof, Ko, 2006, Naast herkomstgroepering ook nationaliteit? In: Bevolkingstrends, jrg. 54, nr. 2, 63-71.
- Snel, Erik, Frank Linder en Radjesh Manna, Allochtonen en werkhervatting vanuit de WAO, In: Nieuwe ongelijkheden op de transitionele arbeidsmarkt, Justus Veenman (red.), Aksant Amsterdam 2006, blz. 89-125 (nog te verschijnen).

Deelnemers SSB-symposium, Naturalis, 1 dec. 2005

Patty Adelaar (SCP)
H.G. van Andel (Gem. Den Haag)
Freek Arnoldus (CBS)
Koos Arts (CBS)
Bart Bakker (CBS)
Diny Bakker (Nivel)
Frank Bastiaans (CBS)
Joop de Beer (NIDI)
Ernest Berkhout (SEO)
Harry Bierings (CBS)
Rob Bijl (WODC)
Marco Bik (Gem. R'dam)
Jeroen Boelhouwer (SCP)
D.J. Boom (RISBO)
Nicole Bosch (CPB)
Lourens Broersma (RUG)
Henk van de Bunt (VU / EUR)
André Corpeleijn (CBS)
Mars Cramer (UvA-Em.)
Jaco Dagevos (SCP)
Marcel Das (UvT)
Marjolijn Das (CBS)
J.H.C. Deijl (Schiedam)
Henk Jan Dirven (CBS)
Marjolijn Distelbrink (NGR)
Peter Doorn (DANS)
Dick van Eijk (NRC)
Marcel Einerhand (Min. SZW)
Godfried Engbersen (EUR)
Leo Engberts (CBS)
Jean Engelen (Ads.Gr.)
Gerard van Essen (Prismant)
Peter Ester (UvT)
Rob Euwals (CPB)
Hidde Fekkes (Min. SZW)
Tamar Fischer (NIDI)
Hilda Folkerts (CBS)
Linda Franken (SCOOP-Zld)
Ruben van Gaalen (NIDI)
Jenny Gierveld (NIDI)
Merové Gijsberts (SCP)
José Gouweleeuw (CBS)
Dirk-Jan Griffioen (RIVM)
Monique de Haan (UvA)
Bas Hamers (UvA)
Marijke Hartgers (CBS)
Richard Heijink (RIVM)
Onno van Hilten (CBS)
Frans Hoeve (CBS)
Erna Hooghiemstra (NGR)
Els Hoogteijling (CBS)
Frits Huls (CBS)
Hanneke Imbens (CBS)
Evert Janissen (CBS)
Gerda Jehoel-Gijsbers (SCP)
Christine Jol (CBS)
Andries de Jong (RPB)
Jan Jonker (CBS)
Peter Kee (CBS)
Saskia Keuzenkamp (SCP)
Peter Kooiman (CPB)
Koos Koopal (Gem. A'dam)
Hassel Kroes (EUR)
Paul van der Laan (CBS)
Jan Latten (CBS)
Frans Leeuw (WODC)
Paul van Leeuwe (WRR)
Frank Linder (CBS)
Henk van Lith (Gem. R'dam)
Albert Luten (NIBUD)

Dorien Manting (RPB)
Claartje Mulder (UvA)
C. Nieuwenhuijse (KUN)
Henry Nieuwenhuizen (Min. SZW)
Suzanne Nijman (Veiligheid)
Joris Nobel (CBS)
Wim van Nunspeet (CBS)
Adri den Ouden (CPB)
Ko Oudhof (CBS)
Karel Pagrag (NWO)
Annelies te Peele (EIM)
Kees Prins (CBS)
Carla Rothuizen (Gem. Den Haag)
Wiemer Salverda (AIAS, UvA)
Hans Sanders (CBS)
Wim Schaasberg (CBS)
Dirk Scheele (WRR)
Eric Schulte Nordholt (CBS)
Jarig van Sinderen (CBS)
Jeroen Slot (Gem. A'dam)
Jeroen Smits (RUN)
Peter Smulders (TNO)
Jan-Maarten van Sonsbeek (SZW)
Jan Spit (SCP)
Arno Sprangers (CBS)
Huib van de Stadt (CBS)
David van der Steen (KNAW)
Anja Steenbekkers (SCP)
Jules Theeuwes (WRR)
Maarten Tielens (KU Leuven)
Kea Tijdens (UvA / EUR)
Harry van den Tillaart (ITS)
Leo van Toor (CBS)
Veerle Vandersmissen (KU Hasselt)
Gosse van der Veen (CBS)
Caroline Vermandere (KU Leuven)
Mathieu Vliegen (CBS)
Rik van der Vliet (CBS)

Klaas de Vos (UvT)
Jeroen van der Waal (EUR)
Theo van der Weegen (KUN)
Wilmie Weltens (CBS)
Kees de Wit (Min. SZW)
Marion Wittenberg (DANS)
Dick van der Wouw (SCOOP Zld)
Aslan Zorlu (UvA)

Dankwoord referees

De redactie wil hierbij de referees bedanken die veel tijd en moeite hebben gestoken in het deskundig becommentariëren van de papers en referaten die in deze bundel zijn opgenomen:

Thijl Ankersmit, Joop Garsen, Marijke Hartgers, Frans Hoeve, Frits Huls, Peter Kee, Huib van de Stadt en Otto Swertz.



Colofon

Eindredactie:
Frank Bastiaans
Leo Engberts
Frank Linder

Vormgeving:
WAT ontwerpers | Kika Lensvelt

Inlichtingen:
Tel: 0900 0227 (€ 0,50 per minuut)
Fax: (070) 337 59 94
E-mail: infoservice@cbs.nl
Internet: www.cbs.nl

© Centraal Bureau voor de Statistiek, 2006

Prijs: € 66,00 (excl. administratie- en verzendkosten)
ISBN: 978-90-357-1942-2

