

De stand van het Sociaal Statistisch Bestand

Bart F.M. Bakker

Naar aanleiding van een studiemiddag van de Nederlandse Vereniging voor Demografie over de mogelijkheden en problemen van het gebruik van het Sociaal Statistisch Bestand, is het tijd om de stand van zaken op te maken. Het bestand is de laatste jaren flink in omvang toegenomen, vooral op het terrein van gezondheid, integratie van allochtonen, het verband tussen ruimtelijke en sociale mobiliteit, onderwijs en sociale dynamiek. Daarvoor zijn nieuwe registers en enquêtes toegevoegd. Helaas zijn het opleidingsniveau en de inkomens van personen en huishoudens nog niet integraal opgenomen. Hoewel bij het opleidingsniveau hiertoe wel een poging is gedaan, leidt het gebruik van de betreffende gegevens niet in alle gevallen tot plausibele resultaten. Het aantal toepassingen in wetenschappelijk onderzoek is eveneens toegenomen, zoals onder meer blijkt uit drie artikelen die in deze aflevering van Bevolkingstrends zijn opgenomen.

1. Registers als bron voor statistisch onderzoek

De Nederlandse Vereniging voor Demografie (NVD) heeft de goede gewoonte om enkele keren per jaar een studiemiddag te organiseren, waar het werk van haar leden wordt gepresenteerd voor een wetenschappelijk forum. Meestal wordt daarvoor een specifiek thema gekozen. Op 24 april 2008 is een studiemiddag georganiseerd om de mogelijkheden en problemen van het gebruik van informatie uit registers te laten zien, in het bijzonder de informatie uit het Sociaal Statistisch Bestand (SSB). Dit bestand is een stelsel van koppelbare registers en enquêtes waaruit door het CBS consistente statistische informatie wordt gepubliceerd. Daarnaast wordt het SSB door sociaalwetenschappelijke onderzoekers gebruikt voor onderzoek op verschillende terreinen. Eerder is op 22 april een symposium georganiseerd rond het thema sociale dynamiek. Ook onderzoek op dit terrein maakt veel gebruik van gegevens uit het SSB.

In december 2005 is voor het laatst de stand van zaken van het SSB opgemaakt in een conferentie met veel deelnemers uit de sociale wetenschappen (CBS, 2006), waaronder demografie en sociale geografie. De belangrijkste algemene conclusies van deze conferentie waren:

- De mogelijkheden voor onderzoek in tal van disciplines (demografie, economie, sociologie, sociale geografie, etc.) zijn groot. Het SSB maakt het bijvoorbeeld mogelijk om onderzoek te doen naar familieverbanden, naar selectieve verhuisstromen en naar crimineel gedrag. Dergelijk onderzoek kan niet of moeizaam met behulp van surveys worden uitgevoerd. Een ander voordeel is dat de geregistreerde kenmerken ook longitudinaal en in hun onderlinge samenhang kunnen worden bestudeerd.

- De in het SSB opgenomen registers bevatten veel relevante informatie voor sociaalwetenschappelijk onderzoek. Toch is het zinvol te blijven speuren naar geschikt registermateriaal dat kan worden toegevoegd. Zo werd de variabele opleidingsniveau, waarvoor veel relevante registers beschikbaar zijn, sterk gemist.
- Daarnaast zijn en blijven surveys noodzakelijk voor het uitvoeren van veel onderzoek in het algemeen en theoretisch gestuurd onderzoek in het bijzonder. Het SSB kan aanzienlijk aan belang winnen door het opnemen van enquêtegegevens van binnen en buiten het CBS. Deze gegevens kunnen dan verrijkt worden met de integrale registerinformatie.
- De toegankelijkheid van het SSB kan op punten nog verder worden verbeterd. Dit geldt niet alleen voor de gegevens, maar ook voor de meta-informatie.

In dit artikel wil ik kort de huidige stand van zaken rond het SSB opmaken. Daarbij ga ik na welke stappen we hebben gezet om de informatie uit te breiden, zowel wat betreft registers als surveys. Verder wordt ingegaan op de verbeteringen die het CBS heeft doorgevoerd in de borging van de kwaliteit van de informatie en de meta-informatie. Ik sluit af met een overzicht van de artikelen die in dit nummer worden gepubliceerd uit het SSB. Deze artikelen zijn gepresenteerd op de studiemiddag van de NVD.

2. De uitbreidingen van het SSB

Sinds 2005 is het SSB sterk in omvang toegenomen. In december 2005 bestond het SSB uit de SSB-kern waarin de jaargangen 1999–2003 waren opgenomen en zes satellieten, waarvan er enkele nog in een embryonaal stadium verkeerden. Aan de SSB-kern zijn nadien de jaargangen 2004 en 2005 toegevoegd. De toevoeging van de jaargang 2006 laat vanwege leveringsproblemen van gegevens uit de Polisadministratie van het Uitvoeringsinstituut Werknemersverzekeringen (UWV) en de Belastingdienst nog op zich wachten. Voorlopig wordt verwacht dat niet eerder dan in de herfst van 2009 het verslagjaar 2006 kan worden toegevoegd. Als de inhoud en de kwaliteit van de informatie voor de latere verslagjaren niet sterk wijzigt, kan daarna een inhaalslag worden gemaakt.

De SSB-kern is ook uitgebreid met het opleidingsniveau, zij het dat dit nog niet altijd toegepast kan worden. In december 2005 was een eerste prototype van deze variabele beschikbaar. Hoewel deze voor de bevolking van Nederland plausibele resultaten opleverde, was de plausibiliteit voor kleinere subpopulaties niet altijd voldoende. Dit bleek vooral voor te komen bij subpopulaties waarvoor veel informatie afkomstig was van de bestanden van de Centra voor Werk en Inkomen (CWI). Daarom is besloten om de gegevens

van de CWI niet langer voor dit doel te gebruiken. Hiermee is overigens niet gezegd dat de informatie van de CWI niet van goede kwaliteit is: er is voornamelijk een conceptueel verschil.

Het opleidingsniveau wordt volgens de CBS-definitie geoperationaliseerd als het niveau van het hoogst behaalde diploma. Bij de werkzoekenden van de CWI wordt eerder een werk- en denkniveau vastgesteld. Dit verklaart waarschijnlijk een groot deel van de verschillen tussen de CWI-informatie over opleidingen en die in andere bronnen. Verder zijn de regels voor het accepteren van 'oudere' gegevens aangescherpt en is de weging verbeterd. De kwaliteit van de variabele is op dit moment nog niet van dien aard dat deze algemeen beschikbaar kan worden gesteld. Er wordt nog onderzoek gedaan naar de condities waaronder deze variabele mag worden toegepast. Zolang dit onderzoek niet is afgerond, wordt de variabele alleen beschikbaar gesteld binnen het CBS en mag deze pas gebruikt worden na een vooronderzoek met geslaagde uitkomsten.

Een andere verwachting is helaas ook niet waargemaakt: het SSB bevat nog steeds alleen informatie over een aantal inkomensbestanddelen van personen en niet over het totaal besteedbaar inkomen van personen en huishoudens. Deze uitbreiding en de beoogde versnelling daarin zijn nog niet gerealiseerd. Dit heeft voornamelijk technische oorzaken. Momenteel wordt gewerkt aan een structurele verbetering die naar verwachting eind 2008 zal worden geïmplementeerd. Gelijktijdig worden met de dataleveranciers, voornamelijk de Belastingdienst, de technische systemen ontwikkeld om de aanlevering van de basisgegevens stabiel te maken en te versnellen.

De inhoud van de verschillende SSB-satellieten is flink uitgebreid. De verdere ontwikkeling van de SSB-satelliet Integratie heeft zich voornamelijk gericht op publicatie van gegevens over de integratie van allochtone groepen in de Nederlandse samenleving en het faciliteren van hieraan gerelateerd onderzoek. Dit heeft onder andere geleid tot vervolgpublishings van de Integratiekaart in 2006 en 2007 die in samenwerking met het WODC zijn samengesteld (Oudhof et al., 2006; Jennissen en Oudhof, 2007) en de ontwikkeling van de Integratiekaart op lokaal niveau in de vorm van vijf tabellen op *Statline*. In deze publicaties staat de demografische en sociaaleconomische integratie van niet-westerse allochtonen centraal. In een samenvattend artikel is dit onlangs uitgebreid met meer sociaal-culturele aspecten, gemeten in enquêtes. Een verdere uitwerking zal plaatsvinden in de eind 2008 te verschijnen CBS-publicatie Jaarrapport Integratie 2008, waarin een breed spectrum aan onderwerpen aan bod zal komen. De SSB-satelliet Integratie is een belangrijke leverancier van de gegevens voor deze publicatie. Alle CBS-informatie over onderzoek onder allochtonen is bij elkaar gebracht in het Dossier Allochtonen op de CBS-website.

De SSB-satelliet Veiligheid is minder sterk gegroeid dan we aanvankelijk verwachtten. De inhoud op het terrein van criminaliteit bestaat uit de gegevens uit het Herkeningsdienststelsel (HKS) van de politie. Daarin zijn alle processen verbaal opgenomen waarin sprake is van een misdrijf. Er zijn helaas nog geen slachtoffergegevens in de satelliet

opgenomen en ook gegevens over rechtbankstrafzaken worden nog gemist. Wel wordt op dit moment gewerkt aan het opnemen van de zogenaamde HALT-afdoeningen (gegevens over jeugdige verdachten) en aan het opnemen van gegevens over slachtofferhulp. De verwachting is dat de SSB-satelliet Veiligheid nog in 2008 met deze gegevens wordt uitgebreid als de kwaliteit van deze informatie voldoende is.

In de satelliet Gezondheidsstatistisch Bestand (GSB) zijn gegevens uit externe gezondheidszorgregistraties opgenomen. Het GSB is begonnen met de ontwikkeling van ziekenhuisstatistieken op basis van de Landelijke Medische Registratie (LMR) van ziekenhuisopnamen. Deze statistieken zijn verder ontwikkeld met diverse *Statline* tabellen en publicaties als resultaat, onder andere in de jaarlijkse CBS-publicatie 'Gezondheid en Zorg in Cijfers'. Personen met ziekenhuisopnamen worden in deze statistieken gevolgd in de tijd, onder andere met betrekking tot sterfte. Verder vindt uitsplitsing naar kenmerken als herkomstgroep en inkomen plaats. In 2006 en 2007 is gestart met de ontwikkeling van nieuwe statistieken op basis van registraties van het gebruik van AWBZ-zorg (CAK-bestanden) en huisartsenzorg (LINH). In 2008 worden hiervan de eerste resultaten gepubliceerd. Ook wordt in 2008 gestart met de ontwikkeling van nieuwe perinatale statistieken op basis van gegevens van de Stichting Perinatale Registratie Nederland. In 2009 worden nog andere zorgregistraties aan het GSB toegevoegd: registraties over medicijnverstrekking, gedeclareerde zorgkosten (Vektis) en de geestelijke gezondheidszorg. Ook wordt gekeken naar de statistische mogelijkheden van de nieuwe registratie van Diagnose-behandelingscombinaties (DBC) in de ziekenhuiszorg.

De SSB-satelliet Ruimtelijke en Sociale Mobiliteit heeft een stormachtige ontwikkeling doorgemaakt. Er zijn drie enquêtes gekoppeld: de Woningbehoefteonderzoeken van 1998 en 2002 en het Netherlands Kinship Panel Study (NKPS). In deze drie enquêtes zijn onder meer de verhuiscriteria en de rol van familierelaties bij verhuizingen waargenomen. Door de enquêtes te verrijken met de al aanwezige registerinformatie is een scala aan nieuwe onderzoeksmogelijkheden ontstaan. Binnen deze satelliet worden vooral demografische en sociaalgeografische onderzoeken verricht, waarvoor het CBS enkele samenwerkingsverbanden heeft gesloten. Met de Universiteit van Amsterdam worden de verbanden tussen familierelaties en verhuisgedrag, de problematiek van de vergrijzende regio Parkstad Limburg en het verband tussen de kenmerken van de buurt en de individuele loopbanen onderzocht. Met het Ruimtelijk Planbureau wordt de realisatie van verhuiscriteria en onverwachte verhuizingen onderzocht. Ten slotte wordt met het Sociaal en Cultureel Planbureau het verband tussen verhuisgedrag, de sociaaleconomische positie van de bewoners en de effecten daarvan op de zogenaamde Vogelaarwijken onderzocht. Er zijn publicaties verschenen over onder andere de kloof tussen arme en rijke wijken en buurten en de invloed van verhuizingen daarop, de inkomstendynamiek in de vier grote steden, verhuisstromen in Noord-Nederland en de functie van de stad Groningen voor sociale stijging, de uitsortering naar herkomstgroep door selectieve verhuisstromen en de zogenaamde zwarte vlucht.

Sinds 2005 heeft de uitgebreide SSB-satelliet Sociale Dynamiek een centrale rol ingenomen binnen het CBS-speerpunt Sociale Dynamiek en Uitbreiding Arbeidsrekeningen (SDUA). In dit speerpunt worden statistieken ontwikkeld over uiteenlopende onderwerpen als arbeid, zorg en kinderopvang, sociaaleconomische integratie van immigranten en dynamiek op de arbeidsmarkt. De nadruk ligt hier voornamelijk op longitudinale analyses, waarbij statistische eenheden in de tijd worden gevolgd. Om dergelijke analyses beter te kunnen uitvoeren, is een belangrijk deel van de informatie uit de SSB-kern op maandbasis beschikbaar gemaakt. Voorheen werd deze informatie doorgaans op één peilmoment beschikbaar gesteld. Zo is informatie over banen, uitkeringen en de status van personen op de arbeidsmarkt in het SSB tegenwoordig op maandbasis vastgelegd, waar deze voorheen op jaarbasis werd vastgelegd. De informatie die wordt ontsloten via de SSB-satelliet Sociale Dynamiek is tot op heden onder meer gebruikt voor publicaties over baandynamiek, duren van uitkeringen en sociaaleconomische integratie van immigranten. Een selectie van deze resultaten is recent gepresenteerd op het symposium 'Dynamiek in de statistiek'.

De SSB-satelliet Onderwijs levert de informatie voor drie belangrijke aandachtsgebieden, namelijk stromen en cohorten binnen het onderwijs en de overgang van onderwijs naar arbeidsmarkt. Stromen binnen het onderwijs hebben betrekking op de overgang van het ene schooljaar naar het volgende schooljaar. Deze stromen worden bestudeerd voor leerlingen in het voortgezet onderwijs en het middelbaar beroepsonderwijs naar alle onderwijssoorten. Het voortijdig schoolverlaten is daarbinnen een belangrijk onderwerp. Binnen het voortgezet onderwijs worden cohorten brugklasleerlingen gevolgd om cijfers over de studievoortgang samen te stellen. Deze cijfers worden onderscheiden naar leerling- en huishoudenskenmerken. Later dit jaar wordt ook gestart met het eerste mbo-cohort. De arbeidsmarktpositie van de voortijdig schoolverlaters wordt bestudeerd na afloop van het schooljaar waarin de school is verlaten. Afgestudeerden in het hoger onderwijs kunnen al wat langer worden gevolgd op de arbeidsmarkt. Het oudste cohort wordt gevolgd tot en met 5 jaar na het afstuderen. Dit jaar wordt tevens gestart met de overgang van het mbo naar de arbeidsmarkt. Deze statistieken worden regulier gepubliceerd op *Statline*, in webmagazineartikelen en in jaarboeken.

In de Europese verordening over de Volks- en Woningtellingen die naar verwachting binnenkort wordt aangenomen, wordt in 2011 weer een inspanning van het CBS op dit terrein verwacht. De zogenaamde Virtuele Volkstelling van 2001 wordt dan voor 2011 herhaald, uiteraard aangepast aan de nieuwe mogelijkheden en de eisen die door het Europese statistische bureau Eurostat worden gesteld. Overigens wordt de SSB-werkwijze waarbij integrale registerinformatie en informatie uit steekproeven worden gecombineerd, voor het eerst ook in een Volkstellingsverordening vastgelegd als valide methode. De verwachting is dat het aantal landen dat deze methode gaat toepassen aanzienlijk wordt uitgebreid ten opzichte van de ronde in 2001. Onder meer Duitsland, Oostenrijk en Zwitserland zullen de volkstelling op deze wijze uitvoeren.

3. Borging van de kwaliteit van het SSB

3.1 De kwaliteit van registerinformatie

Onder veel beleidsmedewerkers en onderzoekers bestaat nog steeds het idee dat onderzoek op basis van registers een kwestie is van koppelen en tellen. Medewerkers van het CBS en universitaire onderzoekers die uitgebreid gebruik hebben gemaakt van registerinformatie, hebben inmiddels ervaren dat de kwaliteit van registers sterk verschilt. Bovendien kan de ene variabele in een register van uitstekende kwaliteit zijn, terwijl de andere variabele in hetzelfde register van slechte kwaliteit is. Daarom moet tijd worden geïnvesteerd in onderzoek naar die kwaliteit en, bij geconstateerde kwaliteitsproblemen, in het ontwikkelen van procedures om de kwaliteit te verhogen. Overigens mag geen misverstand ontstaan over de kwaliteit van de informatie in registers: deze is vaak uitstekend, en in veel gevallen beter dan de kwaliteit van de informatie die is verkregen via enquêtes (Wallgren en Wallgren, 2007).

Kwaliteitsproblemen van registers (en ook enquêtes) komen veelal aan het licht wanneer informatie uit verschillende bronnen wordt gekoppeld. Vaak blijkt dan dat sprake is van inconsistenties tussen deze informatie. Er kunnen inconsistenties bestaan bij de afbakening en de kenmerken van de eenheden. In de afbakening van de eenheden wordt de dekking van de gewenste populatie beoordeeld. Als sprake is van onder- of overdekking spreken we van *representatiefouten*. Bij de kenmerken van die eenheden kan van inconsistentie sprake zijn vanwege definitieverschillen en meetfouten (inclusief administratieve vervuiling). Zonder nader onderzoek kan op voorhand niet worden gezegd of de inconsistenties worden veroorzaakt door fouten die in de ene, in de andere of in beide bronnen voorkomen. Gegevens uit registers en enquêtes die voor CBS-statistieken worden gebruikt, kunnen allerlei representatiefouten, definitieverschillen en meetfouten bevatten. Consistentie is een kwaliteitsaspect dat nog los gezien kan worden van de onder- of overdekking, of de meetfouten. Een statistische uitkomst wint aan kwaliteit wanneer deze consistent is (gemaakt) met andere statistische uitkomsten, zodat een samenhangende beschrijving van onderwerpen kan worden gegeven.

In een aantal gevallen kan de kwaliteit van de registerinformatie effectief worden verbeterd door micro-integratie toe te passen. Micro-integratie is een methode die op het CBS is ontwikkeld en sinds 2005 in steeds ruimere mate wordt toegepast (Arts en Hoogteijling, 2002; Al en Thijssen, 2003; Bakker, 2003).

3.2 Micro-integratie: completeren, harmoniseren en meetfouten corrigeren

Representatiefouten worden veroorzaakt doordat de populatie waarover uitspraken worden gedaan niet volledig en selectief wordt beschreven. Door middel van het onderdeel completeren worden deze representatiefouten opgespoord en gecorrigeerd. Completeren bestaat eruit dat verschillende bronnen worden gecombineerd om een volledige dekking van de populatie te verkrijgen en daarnaast dat ele-

menten die niet tot de populatie behoren worden verwijderd. Een mooi voorbeeld van completeren wordt gevonden bij de samenstelling van het zogenaamde banenbestand, waarin drie bronnen (de fiscale gegevensbank van aangiftes van bedrijven voor de loonbelastingafdrachten, de verzekerdenadministratie van de UWV en de (bedrijfs)Enquête Werkgelegenheid en Lonen) worden gecombineerd. Al deze bronnen bevatten unieke banen, waardoor alleen door combinatie van de drie bronnen een complete beschrijving kan worden gegeven (Bakker en Arts, 2003).

Voor de verbetering van de consistentie vanwege definitieverschillen wordt harmonisatie toegepast. Voor de verbetering van de consistentie vanwege meetfouten wordt er voor deze meetfouten gecorrigeerd. Bij harmonisatie worden de gewenste statistische begrippen zo goed mogelijk geoperationaliseerd en vergelijkbaar gemaakt tussen de bronnen. Eigenlijk wordt hier niet alleen gezocht naar definitieverschillen tussen bronnen, maar ook naar de verschillen tussen de gewenste en feitelijk gemeten variabelen. Registers worden niet bijgehouden voor de productie van statistische informatie, maar voor de uitvoering van wetten en regelingen en om andere administratieve redenen. De administratieve begrippen kunnen aanzienlijk afwijken van de begrippen zoals die voor de statistiek nodig zijn. Bij de harmonisatiestap wordt de vertaling van de administratieve naar de statistische begrippen gemaakt.

Een goed voorbeeld van harmonisatie vinden we terug in het Opleidingsarchief. Daarin zijn de belangrijkste bronnen gecombineerd waarin informatie is opgenomen over opleidingen die gevolgd zijn door mensen die ooit tot de bevolking van Nederland hebben behoord. In die registers en enquêtes is de informatie over de opleidingen steeds verschillend weergegeven. Deze informatie is geharmoniseerd door aan iedere opleiding of kleine groep vrijwel identieke opleidingen een zogenaamd opleidingsnummer toe te kennen. Een groot deel van de opleidingen die sinds 1978 in Nederland bestaan hebben en de op dat moment bekende oudere opleidingen, zijn voorzien van een uniek opleidingsnummer. Van deze opleidingsnummers kunnen op hun beurt allerlei classificaties worden afgeleid (Bakker et al., 2006; CBS, 2008).

Bij inconsistenties die veroorzaakt worden door meetfouten worden beslisregels geformuleerd om deze inconsistenties op te lossen. Daarbij gaat het om inconsistenties als:

- variabelen uit twee bronnen hetzelfde verschijnsel beschrijven, maar een verschillende uitkomst hebben;
- er een logische relatie is tussen variabelen die bij toepassing een onjuiste uitkomst oplevert;
- op één peilmoment een eenheid twee of meer eigenschappen heeft die erg onwaarschijnlijk zijn, zoals het hebben van twee voltijdbanen;
- standen en stromen niet op elkaar aansluiten: de stand op peilmoment t plus de mutaties tussen t en peilmoment $t+1$ moet de stand op $t+1$ opleveren.

Alleen als het oplossen van inconsistenties van belang is voor de te publiceren gegevens, wordt naar een manier gezocht om deze op te lossen. Eerst wordt beoordeeld wat voor ieder van de variabelen de beste bron is. Bij de bepaling van de primaire bron dient rekening te worden gehouden met de administratieve praktijk. Vooral als een variabele

voor een berichtgever niet erg belangrijk is (bijvoorbeeld begin- en einddatum van een baan in fiscale gegevens) is de kwaliteit dikwijls twijfelachtig.

Completeren, harmoniseren en corrigeren van meetfouten worden in het verwerkingsproces van de gekoppelde bronnen niet in deze volgorde uitgevoerd. In het productieproces worden de verschillende stappen vele malen herhaald, totdat consistente einduitkomsten bereikt zijn voor de informatie die wordt gepubliceerd.

3.3 Controle van de kwaliteit van de gegevens

Voor gebruikers van de SSB-gegevens wellicht minder interessant, maar wel vermeldenswaard is dat de controle van de kwaliteit van SSB-gegevens is verbeterd. Om dit te kunnen bereiken was een procesverandering nodig. Het proces is gestandaardiseerd en daarmee transparant gemaakt. Tevens is de documentatie van het proces verbeterd. De zogenaamde *componenten* spelen bij deze procesverandering een belangrijke rol. Componenten zijn gegevensbestanden waarin slechts één statistische eenheid (bijvoorbeeld personen) en de daarbij behorende koppelsleutel is opgenomen, met gegevens over een onderwerp die betrekking hebben op een peilmoment of –periode. Componenten hebben afgezien van de koppelsleutel onderling geen overlapping in variabelen. Deze componenten worden uitvoerig gecontroleerd door deskundigen, waarbij de regel wordt gehanteerd dat iemand anders dan de persoon die de component heeft gemaakt de inhoud ervan controleert. Naast de controle en eventuele correcties van de inhoud van de componenten, is ook de documentatie uitgebreid en verbeterd. Door de componenten onderling te koppelen worden de analysebestanden gebouwd waaruit publicaties (*Statlinetabellen*, artikelen, rapporten) worden samengesteld. Inmiddels zijn 1400 componenten in het SSB opgenomen. Dit aantal groeit gestaag.

4. SSB-onderzoek in dit nummer

Het doel van dit alles is uiteraard het mogelijk maken van goed statistisch onderzoek. In deze aflevering van *Bevolkingstrends* zijn drie artikelen opgenomen waarin het SSB wordt gebruikt.

Bert Bunschoten presenteert in zijn artikel *Hoe ver is het naar...* nieuwe mogelijkheden voor de productie van afstandsstatistieken, omdat tegenwoordig van alle adressen de exacte ligging bekend is. Bij wijze van voorbeeld beschrijft hij de afstand tot de dichtstbijzijnde basisschool en het aantal scholen dat dicht bij vervuulende wegen ligt. Door de gegevens per buurt te middelen kan een schat aan nieuwe buurtkenmerken worden gepresenteerd. Deze informatie wordt naar verwachting dit jaar in het SSB opgenomen.

Gregory Besjes en Ruben van Gaalen laten in hun artikel *Jong geleerd, fout gedaan?* het verband zien tussen de geregistreerde criminaliteit van ouders en hun kinderen. Voorgaand onderzoek toont aan dat de zogenaamde *age-*

crime curve een consistent gegeven is: crimineel gedrag komt vooral voor in de leeftijdsgroep 18–22 jaar. Verder komt criminaliteit vaker voor in de lagere dan in de hogere sociaaleconomische klassen. Ook blijkt dat de volwassenen die nog voorkomen in de verdachtenregistraties nadat kinderen werden geboren, een kleine, selectieve groep vormen. De auteurs laten zien dat er een samenhang bestaat tussen de criminaliteit van ouders en kinderen die niet door de huidige huishoudenssituatie of sociaaleconomische positie van ouders en kinderen verklaard kan worden.

Mila van Huis en Elma van Agtmaal-Wobma beschrijven *de relatie tussen vruchtbaarheid en opleidingsniveau van de vrouw*. Daarbij gebruiken ze het opleidingsniveau zoals dat in de SSB-kern is opgenomen. Hoogopgeleide vrouwen zijn gemiddeld ouder als ze voor het eerst moeder worden dan laagopgeleide vrouwen. Dit uitstel van kinderen kan leiden tot afstel: het aandeel kinderloze vrouwen is dan ook groter onder hoog opgeleiden. De auteurs laten zien dat er weliswaar een gedeeltelijke 'inhaalslag' plaatsvindt op latere leeftijd, maar dat de totale vruchtbaarheid van hoogopgeleide vrouwen achterblijft bij die van laagopgeleide vrouwen.

Literatuur

Al, P. en J. Thijssen, 2003, Integratie in de sociale statistieken. In: CBS, *Gedacht en gemeten*, blz. 111–122. CBS, Voorburg/Heerlen.

Arts, C.H. en E.M.J. Hoogteijling, 2002, Het Sociaal Statistisch Bestand 1998 en 1999. In: *Sociaal-economisch Maandstatistiek* 19(12), blz. 13–21.

Bakker, B., 2003, Hoe nieuw zijn nieuwe ideeën? In: CBS, *Gedacht en gemeten*, blz. 123–132. CBS, Voorburg/Heerlen.

Bakker, B.F.M. en C.H. Arts, 2003, Dynamiek op de arbeidsmarkt; gegevens over stromen uit het Sociaal Statistisch Bestand. In: Bakker, B.F.M. en L. Putman (red.), *De virtuele Volkstelling en SSB*, blz. 59–70. SISWO/CBS, Amsterdam.

Bakker, B.F.M., A.M. Bouman en L. van Toor, 2006, Opleidingsniveau uit registers: nieuwe bronnen, maar nog niet compleet. In: CBS, *Sociale Samenhang in Beeld*, het SSB nu en straks, blz. 141–162. CBS, Voorburg/Heerlen.

CBS, 2006, *Sociale samenhang in beeld*, het SSB nu en straks, blz. 212–217. CBS, Voorburg/Heerlen.

CBS, 2008, <http://www.cbs.nl/nl/-/menu/methoden/classificaties/overzicht/soi/2006/default.htm> en dan in het bijzonder paragraaf 5.1 van de methodologische verantwoording.

Jennissen, R.P.W. en J. Oudhof (red.), *Ontwikkelingen in de maatschappelijke participatie van allochtonen. Een theoretische verdieping en een thematische verbreding van de Integratiekaart 2006*. CBS/WODC/Boom Juridische uitgevers, Meppel.

Oudhof, J., R.P.W. Jennissen en M. Blom (red.), *Integratiekaart 2006*. WODC/CBS, Den Haag.

Wallgren, A. en B. Wallgren, 2007, *Register-based statistics: Administrative data for statistical purposes*. Wiley, New York.