

Vertekening door non-respons

Hoe nauwkeurig zijn de uitkomsten van persoonsenquêtes?

Saskia te Riele

Door het CBS is veel onderzoek gedaan naar vertekening in de uitkomsten van persoonsenquêtes die ontstaat als gevolg van selectieve non-respons. Dit artikel vat de belangrijkste conclusies uit deze studies samen. Allereerst blijkt dat mensen met een minder gunstige maatschappelijke positie in persoonsenquêtes zijn ondervertegenwoordigd. Daarnaast veroorzaakt het onderwerp van de enquête selectiviteit in de non-respons. Voor deze selectiviteit kan veelal niet in voldoende mate gecorrigeerd worden, waardoor uitkomsten vertekend zijn. De uitkomsten zouden nauwkeuriger kunnen worden als de selectiviteit in de non-respons wordt teruggebracht. Het generiek verhogen van de respons biedt hiervoor waarschijnlijk geen afdoende oplossing. Nader onderzoek zou gericht moeten zijn op het vinden van strategieën om de respons onder ondervertegenwoordigde groepen te verhogen. Zolang de non-respons selectief is, is het noodzakelijk aandacht te blijven besteden aan het ontwikkelen van betere weegmodellen om de vertekening te reduceren. Het Sociaal Statistisch Bestand biedt hiervoor mogelijkheden. Onderzocht zal worden of de variabelen die in dit bestand beschikbaar komen ook daadwerkelijk beter voor vertekening corrigeren dan de nu veelal gebruikte demografische kenmerken.

1. Inleiding

De non-respons bij de persoonsenquêtes van het CBS is hoog. De Enquête beroepsbevolking en het Permanent Onderzoek Leefsituatie kennen bijvoorbeeld al jarenlang een non-respons van 40 à 45%. Non-respons kan negatieve gevolgen hebben voor de nauwkeurigheid van de uitkomsten van persoonsenquêtes. Deze uitkomsten kunnen er door vertekend worden. De geschatte waarde van een variabele wijkt dan af van de werkelijke waarde, doordat bepaalde groepen in de steekproef over- of juist ondervertegenwoordigd zijn.

Door het CBS zijn vele studies verricht naar de nauwkeurigheid van uitkomsten van persoonsenquêtes en naar de vertekening die als gevolg van selectiviteit in de respons ontstaat. Dit artikel bevat een overzicht van deze studies. Daaruit blijkt welke groepen systematisch zijn ondervertegenwoordigd in de persoonsenquêtes van het CBS en in welke mate dit leidt tot vertekening van de uitkomsten.

Na een korte uiteenzetting over vertekening door non-respons (paragraaf 2), volgen op basis van een aantal studies de bevindingen over ondervertegenwoordigde groepen in de respons en in de uitkomsten van enquêtes (paragraaf 3). Vervolgens wordt aandacht besteed aan selectiviteit als gevolg van het onderwerp van de enquête (paragraaf 4), gevolgd door aanbevelingen voor het terugdringen van vertekening door responsverhogende maatregelen (paragraaf 5) en door weging (paragraaf 6). Afgesloten wordt met conclusies en aanbevelingen (paragraaf 7).

2. Vertekening door non-respons

Non-respons ontstaat wanneer bij een steekproefpersoon geen geslaagd interview is afgenomen. Allereerst is bij een deel van de getrokken adressen geen contact geweest met de steekproefpersoon. Dit heeft verschillende oorzaken. Vanwege capaciteitstekorten bij het interviewcorps kunnen niet altijd alle steekproefadressen worden uitgezet. Verder worden te enquêteren personen soms niet bezocht, bijvoorbeeld door ziekte van de interviewer, of worden

mensen niet bereikt doordat ze telkens niet thuis zijn. Bij een ander deel van de adressen is er wel contact geweest, maar weigert de persoon in kwestie mee te doen. De steekproefpersoon heeft bijvoorbeeld geen tijd of geen zin. Van de totale uitval bestaat ongeveer tweederde uit weigeringen.

Non-respons kan een negatieve invloed hebben op de nauwkeurigheid van de uitkomsten. Door selectieve uitval kunnen de uitkomsten vertekenen. De geschatte waarde van een variabele wijkt dan af van de werkelijke waarde. De mate van vertekening, hangt af van de hoogte van de non-respons en van de mate waarin de non-responsgroep afwijkt van de responsgroep. Naarmate er meer steekproefpersonen niet responderen en naarmate de non-responderenten meer afwijken van de respondenten, zal de vertekening groter worden (Groves en Couper, 1998).

Non-respons is problematisch wanneer deze selectief is ten aanzien van de variabelen waarop het onderzoek gericht is. Dit laatste is niet eenvoudig te achterhalen. De waarden van de onderzoeksvariabelen zijn immers niet bekend voor de non-responderenten. Om hier toch uitspraken over te kunnen doen, wordt gebruik gemaakt van kennis over andere variabelen waarvoor populatiegegevens beschikbaar zijn of waarvan de waarden bekend zijn voor zowel respondenten als non-responderenten. Deze informatie kan bijvoorbeeld uit registers gehaald worden. Naarmate deze variabelen sterker samenhangen met de onderzoeksvariabelen, kunnen betere conclusies getrokken worden over de selectiviteit in de non-respons ten aanzien van de onderzoeksvariabelen.

Voor vertekening kan in een aantal gevallen goed gecorrigeerd worden. Dit gebeurt door middel van wegen. Daartoe wordt de respons gestratificeerd naar een aantal variabelen, waarvan de populatietotalen bekend zijn uit bijvoorbeeld registers. Dit worden hulpvariabelen genoemd. Vervolgens wordt aan de respondenten binnen elke categorie een gewicht toegekend. Respondenten die vallen in een categorie die is ondervertegenwoordigd in de steekproef krijgen een groter gewicht dan respondenten die in een categorie vallen die juist goed heeft gerespondeerd. De gewichten worden zodanig bepaald dat de respondenten, ingeteld met de gewichten, optellen tot de totalen volgens het register.

Naast de eis dat populatietotalen van hulpvariabelen bekend moeten zijn, zijn er nog twee belangrijke criteria waaraan deze variabelen moeten voldoen. In de eerste plaats worden veelal variabelen gebruikt waarnaar men wil publiceren, zoals leeftijd en geslacht. Door te wegen naar deze hulpvariabelen wordt consistentie tussen de publicatietabellen van verschillende statistieken verkregen. Hiervoor worden voornamelijk variabelen uit de bevolkingsadministratie gebruikt. Daarnaast wordt gezocht naar variabelen die zo hoog mogelijk correleren met de onderzoeksvariabelen. Hoe groter de samenhang, hoe meer de selectiviteit ten aanzien van de onderzoeksvariabelen rechtgetrokken kan worden door correctie van de hulpvariabele. Dit betekent ook dat het effect van een weging een indicatie kan zijn voor selectieve non-respons. Worden voor de weging hulpvariabelen gebruikt die hoog correleren met de onderzoeksvariabelen en is het effect ervan groot, dan vormt dit een aanwijzing voor vertekening in de ongewogen uitkomsten.

Met behulp van een weging kan niet in alle gevallen in voldoende mate voor vertekening gecorrigeerd worden. Wanneer de hulpvariabelen slechts een zwakke relatie hebben met de onderzoeksvariabelen, zal nauwelijks voor de vertekening gecorrigeerd worden. Uit het resultaat van de weging kan in dat geval ook geen conclusie verbonden worden over eventuele selectiviteit in de non-respons. Daarnaast wordt door weging de vertekening alleen

gereduceerd wanneer de samenhang tussen onderzoeksvariabelen en de gekozen hulpvariabelen voor respondenten en non-respondenten even hoog is (Van den Brakel, Diesveld, De Jong, Luiten en Neve, 1996).

Aangezien niet in alle gevallen in voldoende mate gecorrigeerd zal kunnen worden voor vertekening, is het van belang de uitkomsten van enquêtes te toetsen op nauwkeurigheid. Het vaststellen van vertekening kan door de uitkomsten van de enquête op macroniveau te vergelijken met relevante informatie uit andere bronnen, zoals registers of uitkomsten van ander onderzoek, waarvan de betrouwbaarheid vaststaat. Wanneer de geschatte populatietotalen of -verdelingen van een onderzoeksvariabele overeenkomen met gegevens uit een register, kan geconcludeerd worden dat er waarschijnlijk geen sprake is van vertekening als gevolg van selectieve non-respons. Wijken de uitkomsten van de enquête sterk af van de andere, vergelijkbare bron, dan zijn de resultaten vertekend. Het is echter vaak niet mogelijk aan te geven waar deze vertekening precies door is ontstaan. Behalve door selectieve non-respons kan deze ook veroorzaakt worden door fouten die ontstaan tijdens het interviewproces of tijdens de verwerking van de gegevens. Daarnaast kan sprake zijn van kaderfouten in geval het steekproefkader geen goede afspiegeling van de doelpopulatie vormt.

Door de uitkomsten van enquêtes ook op microniveau te vergelijken met andere bronnen, kan meer inzicht in de vertekening worden verkregen. Voor dit type onderzoek wordt de brutosteekproef op persoons- of adresniveau gekoppeld aan een registratie. Op deze manier wordt gedetailleerde informatie verkregen over zowel de respondenten als de non-respondenten. Daaruit kan niet alleen afgeleid worden of de non-respons selectief is en hoe groot de vertekening als gevolg daarvan is, maar is het ook mogelijk te achterhalen of de door de respondenten verstrekte antwoorden overeenkomen met de gegevens in de registers. Dit laatste kan een indicatie geven van de mate van vertekening die ontstaat als gevolg van meet- en verwerkingsfouten. Maar ook aan koppelingsonderzoek zijn nadelen verbonden. Registers kunnen fouten bevatten en niet voor elk individu zal een koppeling tot stand kunnen komen. Daarnaast zijn registers niet altijd voorhanden. In dat geval is het ook mogelijk de non-respondenten nog eens te benaderen en hen een verkorte vragenlijst of alleen een kernvraag voor te leggen. Grote verschillen tussen respondenten en herbenederde non-respondenten vormen een aanwijzing voor vertekening in de uitkomsten.

De hier genoemde methoden zijn bij het CBS toegepast om uitkomsten van persoonsenquêtes te toetsen. In de afgelopen tien jaar is een aantal van deze vertekeningsonderzoeken uitgevoerd. In dit artikel zijn deze studies samengebracht en zijn de resultaten ervan naast elkaar gelegd. In de volgende paragrafen worden de belangrijkste bevindingen samengevat.

3. Ondervertegenwoordigde groepen

Uit diverse vertekeningsonderzoeken komt naar voren dat mensen met een minder gunstige maatschappelijke positie zijn ondervertegenwoordigd in de respons en de uitkomsten van persoonsenquêtes van het CBS. De studies die in deze paragraaf besproken zullen worden, geven aan dat lager opgeleiden, mensen met een uitkering, allochtonen, mensen met een lager inkomen en bewoners van huur- en etagewoningen relatief slecht responderen. Daarnaast wordt aangetoond dat met een weging van de steekproeven veelal niet voldoende voor deze ondervertegenwoordiging wordt gecorrigeerd.

Knops en Kragt (1997) laten zien dat het aantal lager opgeleiden in de Enquête Beroepsbevolking wordt onderschat. Zij maken daarvoor gebruik van kennis over selectieve non-respons in de statistiek Geregistreerde Werkloosheid. Voor deze statistiek worden respondenten van de Enquête Beroepsbevolking op persoonsniveau gekoppeld met de bestanden van de arbeidsbureaus. De koppelingssteekproef die op deze manier ontstaat, wordt opgehoogd naar de totale populatie ingeschrevenen bij het arbeidsbureau. In de weging wordt onder andere gebruik gemaakt van de

verdeling van het opleidingsniveau zoals dat bekend is bij het arbeidsbureau. Analyse van de gewichten wijst uit dat ingeschrevenen met een lage opleiding gemiddeld een hoger gewicht krijgen toegekend dan ingeschrevenen met een hoge opleiding. Dit wijst erop dat lager opgeleide ingeschrevenen minder responderen dan hoger opgeleide ingeschrevenen. Onder de aanname dat het responsgedrag van niet bij een arbeidsbureau ingeschreven mensen overeenkomt met dat van ingeschrevenen, wordt het aantal mensen van 15 tot en met 64 jaar met alleen basisonderwijs in deze enquête met ongeveer 110 duizend, ofwel 1 procentpunt, onderschat. Het aantal mensen met MAVO/VBO wordt met ongeveer 90 duizend onderschat. Het aantal mensen met een middelbare of hogere opleiding, daarentegen, wordt met ongeveer 200 duizend overschat.

Aanwijzingen voor een ondervertegenwoordiging van lager opgeleide werknemers in de Enquête Beroepsbevolking worden gevonden in het Loonstructuuronderzoek (Boerdam, Loeve en Ruijs, 1998; Schulte Nordholt en Ruijs, 2000). Dit onderzoek gaat uit van de Enquête Werkgelegenheid en Lonen, een enquête bij bedrijven naar loon en arbeidsduur van werknemers. Deze enquête is op persoonsniveau gekoppeld aan de Enquête Beroepsbevolking, om zo de gegevens over werknemers te verrijken met opleidingsgegevens. De ontstane koppelingssteekproef is vervolgens opgehoogd naar het totale aantal werknemers. Daartoe is, anders dan in de Enquête Beroepsbevolking, gewogen naar variabelen die sterk met opleidingsniveau samenhangen, zoals dienstverband, loonklasse en economische activiteit. Dit levert een hogere schatting van het percentage laag opgeleide werknemers op dan in de Enquête Beroepsbevolking.

Ook in de Gezondheidsenquête en het Doorlopend Leefsituatieonderzoek van 1995 zijn lager opgeleiden ondervertegenwoordigd. De schatting van het percentage lager opgeleiden is in deze onderzoeken wel hoger dan in de Enquête Beroepsbevolking. Dit is echter het gevolg van verschillen in de manier waarop het opleidingsniveau is gemeten. In de Enquête Beroepsbevolking wordt een veel uitgebreidere inventarisatie gemaakt van de gevolgde opleidingen dan in de Gezondheidsenquête en het Doorlopend Leefsituatieonderzoek. Als gevolg daarvan wordt het opleidingsniveau in de laatste twee enquêtes onderschat. Na correctie hiervoor komt de onderschatting van laagopgeleiden vrijwel even hoog uit als die in de Enquête Beroepsbevolking (Te Riele, 2001). Uit een onderzoek van Van Gils en Thijssen (1989) blijkt dat ook het aantal mensen met een bijstandsuitkering wordt onderschat. De Enquête Beroepsbevolking en het Woningbehoefte Onderzoek tellen ongeveer 20% minder mensen met een bijstandsuitkering dan de uitkeringsadministraties. Een koppeling van de Enquête Beroepsbevolking met bestanden van de Sociale Diensten laat zien dat de vertekening voor een groot deel het gevolg is van het niet of onjuist opgeven van de uitkering (Arts, Imbens en Kragt, 1995). Dit veroorzaakt een onderschatting van ongeveer 15 procent. Daarnaast blijkt dat mensen met een bijstandsuitkering ondervertegenwoordigd zijn. Door deze selectiviteit in de non-respons wordt het aantal mensen met een bijstandsuitkering met ongeveer 5% onderschat.

Een vergelijkbaar beeld komt naar voren uit een onderzoek waarvoor gegevens van mensen met een arbeidsongeschiktheidsuitkering uit de Enquête Beroepsbevolking zijn gekoppeld met gegevens van de uitkeringsadministraties in 1995 (Kartopawiro en Ploeger, 1998). Ook hier is de vertekening voornamelijk een gevolg van het niet of onjuist opgeven van de uitkering. Het gaat hier met name om mensen die van hun werkgever loon doorbetaald krijgen, terwijl de werkgever hun uitkering ontvangt. Dit veroorzaakt een vertekening van 15 tot 19 procent. Daarnaast responderen mensen met een arbeidsongeschiktheidsuitkering relatief slecht in de Enquête Beroepsbevolking. Door deze selectiviteit in de respons ontstaat een onderschatting van 7 tot 11 procent. Dit is een groter aandeel dan bij de bijstandsuitkeringen. Een mogelijke verklaring hiervoor is dat arbeidsongeschiktheid te maken heeft met ziekte en psychische klachten. Deze mensen zijn waarschijnlijk minder goed in staat om aan een enquête deel te nemen.

Aanwijzingen voor een ondervertegenwoordiging van groepen met een minder gunstige maatschappelijke positie worden ook gevon-

den in het Woningbehoefte Onderzoek. Non-responsanalyses laten zien dat de respons onder bewoners van huur- en etagewoningen in dit onderzoek achterblijft (Beckers, Brakenhoff, De Ree, Van Til en Meeuwissen, 2000). Huurders zijn ook ondervertegenwoordigd in het Permanent Onderzoek Leefsituatie, waarvan het Woningbehoefte Onderzoek deel uitmaakt. Dit blijkt uit een onderzoek onder een deel van de respondenten en non-respondenten (Van Baal, 2000). Onder de groep huurders blijken zich verhoudingsgewijs veel lager opgeleiden, alleenstaanden, en mensen met lagere inkomens te bevinden. Daarnaast komen huurwoningen vaker voor in minder goede buurten, en is de verhuishwens onder huurders daardoor groter. Een ondervertegenwoordiging van huurders kan daarom leiden tot een onderschatting van het aantal alleenstaanden, lager opgeleiden, lagere inkomens en potentiële verhuizers. Een weging met alleen demografische kenmerken blijkt niet voldoende te corrigeren voor de ondervertegenwoordiging van huurders en bewoners van flats. Als gevolg daarvan zijn ook de uitkomsten ten aanzien van andere onderzoeksvariabelen van het Woningbehoefte Onderzoek en het Permanent Onderzoek Leefsituatie waarschijnlijk vertekend.

Dit wordt bevestigd wanneer de uitkomsten van het Woningbehoefte Onderzoek met betrekking tot inkomens vergeleken worden met het Inkomenspanelonderzoek, een steekproef van huishoudens waarvan informatie verzameld is bij de administraties van de belastingdienst, huursubsidie en studiefinanciering. Nadat zo goed mogelijk voor definitie- en waarnemingsverschillen gecorrigeerd is, blijkt het Woningbehoefte Onderzoek iets minder huishoudens met een laag inkomen te schatten. Het aandeel huishoudens met huursubsidie wordt eveneens onderschat in het Woningbehoefte Onderzoek. Hoewel dit deels toch te wijten is aan definitieverschillen, vormt ook dit een aanwijzing voor een ondervertegenwoordiging van huurders met een laag inkomen (WBO documentatie inkomen, 1998). De vertekening in de uitkomsten van het Woningbehoefte Onderzoek heeft geleid tot het ontwerp van een nieuwe weging, waarvoor andere dan alleen demografische hulpvariabelen zijn gebruikt. Deze weging wordt in paragraaf 6 verder besproken.

Tenslotte blijken allochtonen in persoonsenquêtes te zijn ondervertegenwoordigd. Janssen en Schmeets (2001) laten zien dat het aandeel allochtonen in de steekproeven van de Enquête Beroepsbevolking en het Permanent Onderzoek Leefsituatie lager is dan in de totale populatie. Onder deze groep bevinden zich verhoudingsgewijs meer werklozen en inactieven dan onder de autochtone bevolking. Hiervoor kan in de weging echter goed gecorrigeerd worden door de variabele land van herkomst mee te nemen. Het is dan wel noodzakelijk dat de allochtonen die wel responderen representatief zijn voor de hele groep. Er zijn verschillende aanwijzingen dat dit het geval is. Schmeets en Janssen (2001) laten zien dat allochtonen uit buurten waar de gemiddelde WOZ-waarde laag is, niet zijn ondervertegenwoordigd. De verdeling van allochtonen naar postcodegebieden met een bepaalde WOZ-waarde in het Permanent Onderzoek Leefsituatie komt overeen met de verdeling in de populatie. De WOZ-waarde geeft een indicatie van de waarde van de woning van de respondent en is daardoor een goede indicator voor zijn of haar sociaal-economische positie. Daarnaast blijkt uit een vergelijking van de Enquête beroepsbevolking met het Inkomenspanelonderzoek dat het verschil tussen het percentage werkende niet-westerse allochtonen en autochtonen in beide bronnen goed overeenkomt. De mate waarin autochtonen vaker werk hebben dan allochtonen wordt dus voldoende nauwkeurig geschat in de Enquête Beroepsbevolking. Dit betekent dat selectiviteit ten aanzien van het hebben van betaald werk, voor zover daar al sprake van is, niet verschillend is voor niet-westerse allochtonen en autochtonen (Te Riele, 2001).

De geconstateerde ondervertegenwoordiging van groepen met een minder gunstige maatschappelijke positie leidt niet in alle gevallen tot vertekende uitkomsten. Zo wordt het aantal mensen met een baan voldoende nauwkeurig geschat in de Enquête Beroepsbevolking. Dit blijkt uit confrontaties van de Enquête Beroepsbevolking met de Enquête Werkgelegenheid en Lonen, die in het kader van de Arbeidsrekeningen worden uitgevoerd. Voor de Arbeidsrekeningen worden de twee bronnen eerst grondig geana-

lyseerd. Op basis daarvan kan elke bron goed op zijn kwaliteit beoordeeld worden en kan bepaald worden welke aanpassingen nodig zijn om de bronnen beter vergelijkbaar te maken (Leunis en Verhage, 1996; CBS 1998). Na bewerking van beide bronnen blijkt het geschatte aantal werkenden volgens de Enquête Beroepsbevolking goed overeen te komen met de schattingen van de Enquête Werkgelegenheid en Lonen. Een koppeling van de Enquête Beroepsbevolking met de Verzekerdenadministratie geeft een zelfde beeld (Van Toor en Schaafsma-Harteveld, 1997).

Deze resultaten zijn enigszins onverwacht als men bedenkt dat mensen met een uitkering ondervertegenwoordigd zijn in de Enquête Beroepsbevolking. Men zou dan verwachten dat het aantal werkenden wordt overschat. Het is echter ook mogelijk dat mensen met een uitkering en niet-werkenden die geen uitkering ontvangen per saldo goed in de steekproef vertegenwoordigd zijn. De onderschatting van het aantal uitkeringen wordt dan gecompenseerd door een overschatting van niet-werkenden zonder uitkering.

Ook de gesubsidieerde arbeid wordt voldoende nauwkeurig geschat. Het gaat hier om banen voor zwakkere groepen op de arbeidsmarkt. De uitkomsten van de Enquête Beroepsbevolking komen goed overeen met cijfers van het Ministerie van Sociale Zaken en Werkgelegenheid (Van Cruchten en de Vries, 1997, 1998). Deze resultaten geven aan dat zich onder de minder gunstig gesitueerden ook groepen bevinden die wel in voldoende mate responderen.

De hoge non-respons in de Enquête Beroepsbevolking lijkt eveneens niet van invloed op de schatting van de geregistreerde werkloosheid. Reeksen van uitkomsten van de geregistreerde werkloosheid hangen niet significant samen met reeksen van responscijfers (Huisman, 1999). Ook binnen categorieën van kenmerken als stedelijkheidsgraad, regio en inkomensklasse is geen duidelijk verband zichtbaar tussen het verloop van de geregistreerde werkloosheid enerzijds en het verloop van de netto respons anderzijds (Michiels, 1999). Tenslotte blijkt de geregistreerde werkloosheid nauwelijks gevoelig voor de gebruikte weging. Het totale aantal geregistreerde werklozen komt voor 1996 met de gebruikelijke weging op 440 duizend uit. Wanneer de factor opleidingsniveau buiten het weegmodel wordt gehouden, komt de schatting uit op 439 duizend (Van Bochove en Thijssen, 1998). Dit betekent overigens niet dat een weging naar opleidingsniveau ook voor andere groepen dan ingeschrevenen bij het arbeidsbureau geen effect zal hebben.

Tenslotte zijn er ook geen aanwijzingen voor ernstig ondervertegenwoordigde groepen in het Sociaal-economisch Panelonderzoek. Dit is niet vanzelfsprekend. Het panel is in 1984 gestart en zowel de uitval als de bijwerving kunnen selectief zijn. Het is daarom zeer goed mogelijk dat het panel geen goede afspiegeling meer vormt van de huidige populatie. Nadere analyse van het panel laat zien dat het aandeel allochtonen achterblijft. Dit wordt voor een belangrijk deel veroorzaakt doordat sinds het startjaar de steekproef niet gericht is aangevuld met immigranten omdat de mogelijkheden daartoe ontbraken. Daarnaast, en gedeeltelijk als gevolg daarvan, is het aandeel werklozen en inactieven iets lager dan in de Enquête Beroepsbevolking. De ontwikkelingen in het percentage werklozen door de jaren heen komen in beide onderzoeken echter wel weer goed overeen (Mikulic, 1995; Linden, 1998). Ook de schatting van het gemiddeld netto jaarinkomen ligt dicht bij dat van het Inkomenspanelonderzoek, net als de verdeling van huishoudens naar inkomensklasse (Mikulic, 1998; Sleijpen en Dirven, 1999). Uit deze resultaten kan worden geconcludeerd dat er sprake is van enige selectiviteit in het panel. Deze selectiviteit is echter beperkt.

Tussen de ondervertegenwoordigde groepen bestaat een beduidende overlap. Mensen met een lage opleiding hebben vaker een bijstands- of arbeidsongeschiktheidsuitkering, een laag inkomen en een huurwoning in vergelijking met hoogopgeleiden. Door deze overlap is de vertekening in de uitkomsten als gevolg van selectieve non-respons beperkter dan de opsomming van afwijkingen in eerste instantie doet vermoeden. Zo behoort de schatting van de verdeling van het opleidingsniveau tot de minst nauwkeurige. De onderschatting van het aantal laagopgeleiden bedraagt echter

slechts 2 procentpunten. Dit acht het CBS acceptabel. Een grote vertekening ontstaat meestal als selectieve non-respons samen gaat met meetfouten. In dat geval wordt de validiteit van de uitkomsten meestal niet toereikend geacht en ziet het CBS af van publicatie. Dit geldt bijvoorbeeld voor de schatting van mensen met een uitkering uit de Enquête Beroepsbevolking. Deze vertekening wordt voornamelijk veroorzaakt door het niet of onjuist opgeven van de uitkering. Hieruit is geconcludeerd dat het niet goed mogelijk is het aantal uitkeringsgerechtigden via een enquête te meten. Daarom zijn de vragen daarover na 1999 uit de vragenlijst geschrapt.

4. Selectiviteit als gevolg van het onderwerp van de enquête

Naast selectiviteit in de non-respons ten aanzien van de maatschappelijke positie van steekproefpersonen, blijkt ook het onderwerp van het onderzoek selectiviteit te introduceren. Een duidelijk bewijs hiervoor wordt gevonden in studies naar aanleiding van de revisie van het Onderzoek Verplaatsingsgedrag. Deze revisie was noodzakelijk omdat de respons onder huishoudens tot onder de 50 procent gedaald was. Met de nieuwe opzet werd inderdaad een veel hogere respons behaald. Hierdoor ontstond echter wel een trendbreuk. Het aantal verplaatsingen in de oude opzet lag ongeveer 30 procent hoger dan in de nieuwe opzet. Dit bleek het gevolg van een groter aandeel respondenten met verplaatsingen in de oude opzet. Deze resultaten geven aan dat mensen zonder verplaatsingen minder geneigd zijn om aan het Onderzoek Verplaatsingsgedrag mee te doen. Deze laatste groep vindt het onderwerp van de enquête waarschijnlijk niet op zichzelf van toepassing (Moritz en van Evert, 1998; Brög en Erl, 1999; Moritz en Brög, 1999).

Ook voor andere enquêtes worden aanwijzingen gevonden voor selectie als gevolg van het onderwerp of de naam van de enquête. Beukenhorst (1999) laat zien dat personen die deelname aan het Woningbehoefte Onderzoek weigeren over het algemeen niet willen verhuizen. Weigeraars die later alsnog zijn overgehaald om aan het onderzoek mee te doen, hebben veel minder vaak een verhuiscens dan de overige respondenten. Daarnaast geven weigeraars aan wie aan de deur nog de vraag gesteld is of men wil verhuizen, in 85 procent van de gevallen aan niet te willen verhuizen binnen nu en 2 jaar. Voor alle respondenten ligt dit percentage op slechts 70 procent.

Voor de Gezondheidsenquête geldt dat mensen met gezondheidsproblemen relatief vaak mee lijken te doen. Uit een vergelijking met gegevens van ziekenhuizen blijkt dat de Gezondheidsenquête het aantal ligdagen in het ziekenhuis overschat (Bakker en Van Rooijen, 2000). Na toepassing van een andere weging, die gericht is op het reduceren van vertekening door non-respons, komt de schatting van het aantal ligdagen dan ook lager uit (Geuzinge, Van Rooijen en Bakker, 2000). Deze resultaten suggereren dat mensen met een hoge medische consumptie zijn oververtegenwoordigd in de Gezondheidsenquête.

De uitkomsten van het Nationaal Kiezersonderzoek geven een aanwijzing voor oververtegenwoordiging van mensen met een grote politieke belangstelling. Het geschatte opkomstpercentage ligt veel hoger dan het werkelijke opkomstpercentage bij verkiezingen. Door vragen over politieke voorkeur en stemgedrag onder te brengen bij het Permanent Onderzoek Leefsituatie, waarin de nadruk minder sterk op politiek ligt, en door vragen over opkomst alleen nog na de verkiezingen te stellen, komen de geschatte opkomstcijfers dicht bij de werkelijke cijfers te liggen (Schmeets en Janssen, 2001).

Tenslotte zijn er aanwijzingen dat maatschappelijke interesse in het algemeen een rol speelt bij de beslissing om aan een enquête deel te nemen. In de Enquête Beroepsbevolking wordt het aantal vakbondsleden overschat. Dit kan het gevolg zijn van een oververtegenwoordiging van vakbondsleden. Deze groep is maatschappelijk meer geëngageerd en is daarom waarschijnlijk meer dan gemiddeld bereid om aan persoonsenquêtes van het CBS mee te doen (Van Cleef, 1995). Dit gaat ook op voor het Permanent Onderzoek Leefsituatie. Ondanks dat de schatting van het

opkomstpercentage bij verkiezingen dicht bij de werkelijke percentages is komen te liggen na onderbrenging van dit type vragen in het Permanent Onderzoek Leefsituatie, blijft deze enigszins te hoog (Schmeets en Janssen, 2001).

5. Voorkomen van vertekening: het effect van responsverhogende maatregelen

Het beeld dat uit dit overzicht van vertekeningsonderzoeken naar voren komt, is dat personen met een minder gunstige maatschappelijke positie zijn ondervertegenwoordigd in de persoonsenquêtes van het CBS. Ook het onderwerp van de enquête veroorzaakt selectiviteit in de non-respons. Voor deze selectiviteit kan niet in alle gevallen in voldoende mate gecorrigeerd worden met een weging. Daardoor zijn uitkomsten van persoonsenquêtes vertekend. Om tot betrouwbaardere uitkomsten te komen, is het noodzakelijk de selectiviteit in de respons zo ver mogelijk terug te brengen. Een veel voorgestelde maatregel om dit te bewerkstelligen, is het verhogen van de respons.

Het generiek ophogen van de respons leidt echter niet altijd tot een reductie van de selectiviteit in de respons. Dit wordt aangetoond in een onderzoek van Beukenhorst (1999). In het Woningbehoefte onderzoek van 1993/94 en in het Woningbehoefte Onderzoek 1998 werd een respons van 75 procent behaald door een intensieve benaderingsstrategie toe te passen. Beukenhorst verlaagde deze respons kunstmatig tot ongeveer 60 procent door alleen die respons mee te nemen die binnen drie benaderingspogingen behaald was. Ook respondenten die bij de eerste benadering geweigerd hadden, werden niet meegenomen. De ongewogen uitkomsten bij een respons van 60 procent bleken nauwelijks te verschillen van de uitkomsten bij een respons van 75 procent. Bovendien werden in dit onderzoek aanwijzingen gevonden voor selectiviteit in de 25 procent non-respons (zie paragraaf 4). Vooral mensen die tevreden zijn met hun woonomgeving lijken deelname te weigeren. Deze resultaten suggereren dat met een intensievere benaderingsstrategie weliswaar een hogere respons wordt behaald, maar dat hiermee voornamelijk meer van hetzelfde wordt binnengehaald. De selectiviteit in de non-respons vermindert in beperkte mate. Dit betekent dat de uitkomsten nog steeds vertekend zijn.

Een vergelijkbare conclusie kan getrokken worden uit een analyse van de respons van het Permanent Onderzoek Leefsituatie (Beukenhorst, 2001). In het Permanent Onderzoek Leefsituatie worden non-respondenten telefonisch herbenaderd. In het telefonische interview wordt een deel van de vragenlijst ingevuld. Daarna wordt toestemming gevraagd om het tweede deel bij de geïnterviewde persoon thuis af te nemen. De respons die via de telefonische herbenadering behaald wordt, blijkt echter nauwelijks te verschillen van de respons uit de eerste veldwerkronde. De uitkomsten zijn vrijwel gelijk op kenmerken als het hebben van betaald werk, het bezit van de woning, het soort woning en het rookgedrag. Ook met deze responsverhogende maatregel lijkt dus voornamelijk meer van hetzelfde te worden binnengehaald.

Waar de respons selectief is, lijkt het daarom eerder aan te bevelen om extra inspanningen te richten op groepen die structureel ondervertegenwoordigd zijn. Op deze manier kan de selectiviteit waarschijnlijk effectiever teruggebracht worden. Naar methoden om gericht te werven zal nader onderzoek gedaan moeten worden. De mogelijkheden hiervoor worden groter met de komst van het Sociaal Statistisch Bestand. Dit is een database waarin informatie uit registers, bedrijfsenquêtes en persoonsenquêtes op het niveau van personen gekoppeld is (Van der Laan, 2000). Hiermee komt veel informatie beschikbaar over zowel respondenten als non-respondenten, waardoor slecht responderende groepen beter geïdentificeerd kunnen worden. Deze kennis kan gebruikt worden om gericht te werven onder bepaalde groepen of om het steekproefontwerp aan te passen. Naast het beschikbaar komen van meer gedetailleerde informatie over de Nederlandse bevolking, kan ook laagregionale informatie gebruikt worden. In postcodegebieden met een lage respons kan bijvoorbeeld intensiever geworven worden dan in andere gebieden.

Daarnaast dient de selectiviteit die ontstaat door interesse of betrokkenheid bij het onderwerp van de enquête teruggebracht te worden. Door de introductie van het Permanent Onderzoek Leefsituatie (POLS) in 1997 is daartoe een eerste aanzet gedaan. In dit onderzoek zijn een aantal afzonderlijke enquêtes zoals over gezondheid, rechtsbescherming en veiligheid, geïntegreerd. Respondenten krijgen eerst een algemene vragenlijst voorgelegd, gevolgd door een module met vragen over een specifiek onderwerp. Steekproefpersonen ontvangen vooraf een algemene aankondigingsbrief, waarin het specifieke onderwerp niet vermeld wordt. Hierdoor is een deel van de selectiviteit teruggebracht. Aanwijzingen hiervoor zijn gevonden in het eerder genoemde onderzoek van Schmeets en Janssen (2001) waarin resultaten van het Nationaal Kiezersonderzoek zijn vergeleken met resultaten gebaseerd op overeenkomstige vragen opgenomen als module in het POLS.

Om deze vorm van selectiviteit verder terug te dringen, zouden ook extra inspanningen verricht kunnen worden om juist die personen over te halen die deelname weigeren vanwege het onderwerp van de enquête. Studies naar aanleiding van de revisie van het Onderzoek Verplaatsingsgedrag laten zien dat zo'n strategie de selectiviteit terug kan brengen. Tijdens de telefonische herbenadering wordt duidelijk gemaakt dat ook deelname van mensen zonder verplaatsingen belangrijk is. Deze strategie leidt ertoe dat meer mensen zonder verplaatsingen deelnemen, waardoor de schatting van het aantal verplaatsingen nauwkeuriger wordt.

6. Corrigeren voor vertekening: mogelijkheden en beperkingen

Zolang de persoonsenquêtes van het CBS te maken hebben met een hoge en selectieve non-respons, moet gezocht worden naar methoden om hier zo goed mogelijk voor te corrigeren. Met de komst van het Sociaal Statistische Bestand komt informatie over andere dan alleen demografische kenmerken beschikbaar. Daarmee wordt het mogelijk nieuwe weegmodellen te ontwikkelen en te toetsen.

Om de mogelijkheden van het Sociaal Statistisch Bestand nader te onderzoeken, is een nieuwe weging ontwikkeld voor de Gezondheidsenquête (Geuzinge, Van Rooijen en Bakker, 2000). Voor deze weging is gezocht naar hulpvariabelen die zowel samenhangen met belangrijke onderzoeksvariabelen van de Gezondheidsenquête als met de kans op non-respons. Deze samenhangen zijn empirisch geverifieerd, waarna er een rangorde in de potentiële hulpvariabelen werd aangebracht. Bij de selectie van hulpvariabelen is rekening gehouden met de onderlinge samenhangen tussen de hulpvariabelen. De gevolgde procedure resulteert in een weegmodel met als kern de variabelen bruto jaarsalaris, leeftijd, gezinstype, etniciteit en geslacht. Na weging met het nieuwe model komt de schatting van het aantal ligdagen in de Gezondheidsenquête dichter bij het aantal ligdagen volgens gegevens van ziekenhuizen te liggen (Bakker en Van Rooijen, 2000).

Het effect van deze weging is ook onderzocht voor de variabele opleidingsniveau. De geschatte verdeling van het opleidingsniveau blijkt echter nauwelijks te veranderen. Dit resultaat is onverwacht. In de weging wordt bijvoorbeeld het bruto jaarinkomen meegenomen, een variabele waarvan vaststaat dat deze sterk correleert met opleidingsniveau. Het effect van dit type variabelen is ook al in eerder onderzoek aangetoond. In het Loonstructuuronderzoek werd het geschatte aantal lager opgeleide werknemers hoger na toepassing van een weging met onder andere de variabelen dienstverband en loonklasse. Het uitblijven van een effect op de verdeling van het opleidingsniveau in de Gezondheidsenquête geeft aan dat ook een weging met andere dan alleen demografische kenmerken niet voor elke vertekening kan corrigeren. Overigens moet opgemerkt worden dat het weegmodel toch nog voornamelijk demografische kenmerken bevat. De extra hulpvariabelen die met het Sociaal Statistisch Bestand beschikbaar komen,

hebben dus niet altijd een groter corrigerend vermogen dan de demografische kenmerken die nu veelal gebruikt worden. Dit betekent dat kennis over meer hulpvariabelen niet per definitie leidt tot betere weegmodellen en een betere correctie voor vertekening door non-respons.

Ook voor het Woningbehoefte Onderzoek is een weegmodel ontwikkeld waarin andere dan alleen demografische kenmerken zijn gebruikt. Voor dit onderzoek is gekozen voor een weging op laagregionaal niveau. Daartoe zijn zes-cijferige postcodegebieden geclusterd naar de grootte van de woonplaats, verhouding huur- en koopwoningen, verhouding etagewoningen en eengezinswoningen, WOZ-waarde en omgevingsadressendichtheid, elk ingedeeld in 5 categorieën. Er is vervolgens zodanig opgehoogd dat de verhouding van de populatie en de respons in een cluster van postcodegebieden gelijk is (Beckers et al, 2000). Na toepassing van dit weegmodel komen de schattingen van het aantal koopwoningen, etagewoningen en éénpersoonshuishoudens dichter bij de verwachte waarden te liggen dan met de oude weging. Het is echter moeilijk te bepalen of ook de schattingen van andere onderzoeksvariabele als de wens om te verhuizen nauwkeuriger worden.

Tenslotte biedt ook het Permanent Onderzoek Leefsituatie mogelijkheden om beter voor selectiviteit te corrigeren. Door de integratie van voorheen afzonderlijke enquêtes, is de steekproefomvang van dit onderzoek beduidend groter in vergelijking met elk van de voorheen afzonderlijke enquêtes. Deze grotere steekproefomvang biedt mogelijkheden een meer gedetailleerde stratificatie te gebruiken bij de weging en op deze wijze beter te corrigeren voor mogelijke vertekening. Een bijkomend aspect is dat in het Permanent Onderzoek Leefsituatie de vragenlijst bestaat uit een algemeen deel en een module met een specifiek onderwerp. In het algemeen deel zijn kernvariabelen op de verschillende terreinen van de leefsituatie opgenomen. Voor deze kernvariabelen komen gewogen resultaten beschikbaar. Deze resultaten worden vervolgens weer gebruikt bij de weging van de afzonderlijke modules. Hierdoor wordt niet alleen de variantie gereduceerd en consistentie tussen de modules verkregen maar ook beter gecorrigeerd voor vertekening als gevolg van selectieve respons in vergelijking met de vroegere situatie van afzonderlijke enquêtes.

7. Conclusies en aanbevelingen

Uit het grote aantal vertekeningsonderzoeken dat de afgelopen jaren bij de persoonsenquêtes is uitgevoerd, komt naar voren dat personen met een relatief ongunstige maatschappelijke positie zijn ondervertegenwoordigd. Dit beeld gaat voor vrijwel alle onderzochte persoonsenquêtes op, en leidt ook na weging tot vertekende uitkomsten. Daarnaast blijkt dat ook het onderwerp van de enquête selectiviteit in de non-respons introduceert. Personen die niet geïnteresseerd zijn in het onderwerp, zijn minder bereid om mee te werken. Verder onderzoek naar het terugdringen van de selectiviteit in de non-respons is noodzakelijk om te komen tot betrouwbaardere uitkomsten.

Zolang de respons van persoonsenquêtes laag blijft, is er een grote kans op vertekening in de uitkomsten. Het blijft daarom nodig naar manieren te zoeken om hiervoor te corrigeren. Nu met de komst van het Sociaal Statistisch Bestand informatie over een groter aantal variabelen voor alle personen beschikbaar komt, kan onderzocht worden of met behulp daarvan effectievere weegmodellen ontwikkeld kunnen worden. De eerste resultaten op dit punt zijn overwegend positief. Wel blijkt dat ook met nieuwe variabelen niet per definitie voor elke vertekening gecorrigeerd kan worden. Deze variabelen corrigeren ook niet in alle gevallen beter voor vertekening dan de nu veel gebruikte demografische kenmerken. Experimenten met een weging op laagregionaal niveau laten eveneens positieve resultaten zien. De weging van het Woningbehoefte Onderzoek op het niveau van geclusterde postcodegebieden levert meer nauwkeurige uitkomsten op. Wellicht dat een dergelijke aanpak ook voor andere enquêtes zinvol is.

Referenties

- Arts, C.H., Imbens, J.C.M. en Kragt, C. (1995): Vergelijkingsonderzoek werkloosheidsuitkeringen. *Sociaal-economische Maandstatistiek*, nr. 6, p.p. 21–39.
- Baal, M. van (2000) Respons Analysis Survey van POLS: een tus-senrapportage. Intern CBS-rapport, BPA H2994-00-GWM, CBS, Sector Waarnemingsmethodologie, Heerlen.
- Bakker, B.F.M. en J. van Rooijen (2000) One figure for the supply and demand of services. *Netherlands Official Statistics*, vol 15, spring, p.p. 40–46.
- Beckers, Brakenhoff, de Ree, van Til en Meeuwissen. (2001) Herweging WBO 1999. Intern CBS rapport, Sector Statistische Analyse Personen, Heerlen.
- Beukenhorst (1999) Vertekening door non-respons in het WBO. Intern CBS-rapport, Sector Waarnemingsmethodologie, Heerlen.
- Beukenhorst (2001) Rendement van het mixed-mode design van POLS en vertekening. Intern CBS-rapport, Sector Methoden en Ontwikkeling, Heerlen.
- Bochove, C.A. van en J. Thijssen (1998) Het meten van meetfouten. *Economisch-Statistische Berichten*, 23-1-1998, p.p. 60–62.
- Boerdam, A.A., Loeve, J.A. en G.P.C.M. Ruijs (1998) Loon naar opleidingsniveau en beroep: het loonstructuuronderzoek 1995. *Sociaal-economische maandstatistiek*, nr. 3, p.p. 31–57.
- Brakel, J. van den, Diesveld, P., Jong, W. de, Luiten, A. en W. Neve (1996) Het gebruik van registraties bij non-respons. Intern CBS-rapport, Sector Statistische Methoden, Heerlen.
- Brög, W. en E. Erl (1999) Systematic errors in mobility surveys. *23rd ATRF Conference*, Perth, 29 September–1 October 1999.
- CBS (1998) *Arbeidsrekeningen 1996–1997*, p.p. 13–15. CBS: Voorburg/Heerlen.
- CBS (1998) *Woningbehoefte Onderzoek: Documentatie inkomen*. CBS: Heerlen.
- Cleef, B.P.M. van (1995) Kenmerken van vakbondsleden: uitkomsten van de Enquête Beroepsbevolking. *Sociaal-economische Maandstatistiek*, nr. 9, p.p. 19–24.
- Cruchten, J.M.J. van en S. de Vries (1997) JWG'ers en banenpoolers in de Enquête Beroepsbevolking 1996. *Sociaal-economische Maandstatistiek*, nr. 12, p.p. 19–24.
- Cruchten, J.M.J. van en S. de Vries (1998) Gesubsidieerde arbeid in de Enquête Beroepsbevolking. *Sociaal-economische Maandstatistiek*, no. 8, p.p. 21–26.
- Frenken, F. (1997) Bloeddonoren in Nederland, 1996. *Maandbericht Gezondheidsstatistiek*, nr. 11, p.p. 5–12.
- Geuzinge, L. Van Rooijen, J. en Bakker, B. (2000) The use of administrative registers to reduce non-response bias in household surveys. *Netherlands Official Statistics*, vol 15, spring, p.p. 32–39.
- Gils, I. van en J.G.J. Thijssen (1989) Werkloosheidsuitkeringen, ingeschreven personen bij een arbeidsbureau en geregistreerde werklozen. *Supplement bij de Sociaal-economische Maandstatistiek*, nr. 4, p.p. 4–10.
- Groves, R.M. en Couper, M.P. (1998) *Nonresponse in household interview surveys*. Chapter 1, p.p. 1–24. New York: John Wiley and Sons.
- Huisman, M. (1999) Relationships between registered unemployment and response rates in the Labour Force Survey. *Netherlands Official Statistics*, vol 14, autumn, p.p. 13–17.
- Janssen, J. en Schmeets, H. (2001) Onderzoek naar non-respons met informatie uit het SSB. Intern CBS-rapport, Sector Ontwikkeling en Ondersteuning, Divisie Sociale en Ruimtelijke Statistieken, Heerlen.
- Kartopawiro, J.D. en S.A. Ploeger (1998) Vergelijkingsonderzoek arbeidsongeschiktheidsuitkeringen, 1995. *Sociaal-economische maandstatistiek*, nr.1, p.p. 25–42.
- Knops, I. en Kragt, C. (1997) Telt het CBS te weinig laag opgeleiden? *Sociaal-economische Maandstatistiek*, nr. 10, p.p. 18–21.
- Laan, P. van der (2000) Integrating administrative registers and household surveys. *Netherlands Official Statistics*, vol 15, spring, p.p. 7–15.
- Leunis, W.P. en C.G. Verhage (1996) *Arbeidsrekeningen, kern van het statistisch systeem over arbeid*. Dissertatie. CBS: Voorburg.
- Linden (1998) Plausibiliteitscontrole SED98, lange termijn overgangen op de arbeidsmarkt. Intern CBS-rapport, Sector Sociaal-economische Persoonsenquêtes, Heerlen.
- Michiels (1999) Selectieve respons in de Enquête Beroepsbevolking: Effect op niveau en trend van de geregistreerde werkloosheid. Intern CBS-rapport, Sector Sociaal-economische Persoonsenquêtes, Heerlen.
- Mikulic, B. (1995) Comparison of variables from European Community Household Panel Survey and Labour Force Survey: The case of the Netherlands. CBS-rapport, BPA nr.: H02877-95-SIP. Sector Integratie en publicatie, Heerlen: CBS.
- Moritz, G. en W. Brög (1999) Redesign of the Dutch Travel Survey: Response improvement. *Netherlands Official Statistics*, vol. 14, autumn, p.p. 7–12.
- Moritz, G. en H. van Evert (1998) Nieuw OVG. Bijdrage aan het Colloquium Vervoerplanologisch Speurwerk. Amsterdam, november 1998.
- Riele, S.M.M. te (2001) Vertekening door non-respons. Intern CBS-rapport, Sector Ontwikkeling en Ondersteuning, Divisie Sociale en Ruimtelijke Statistieken, Heerlen.
- Schmeets, H. en Janssen, J. (2001) Using national registrations to correct for selective non-response. Political preference and ethnic groups. Paper presented at the symposium on Achieving Data Quality in a Statistical Agency: A Methodological Perspective, Ottawa, 16–19 oktober.
- Schulte-Nordholt, E. en G. Ruijs (2000) Loon naar opleidingsniveau en beroep: het Loonstructuuronderzoek 1997. *Sociaal-Economische maandstatistiek*, nr. 4, p.p. 19–49.
- Toor, L. van en B. Schaafsma-Harteveld (1997) Werknemersgegevens uit de verzekerdienstand: een eerste analyse. *Sociaal-economische Maandstatistiek*, nr. 11, p.p. 24–27.