

“Het meten van de effecten van beleid”

Inleiding¹

Karel Van den Bosch en Jef Breda

In Vlaanderen en elders wordt veel onderzoek gedaan naar de effecten – bedoelde en onbedoelde, gewenste en ongewenste – van beleid op de maatschappelijke werkelijkheid. Dit beleid kan uitgaan van nationale, plaatselijke of supra-nationale overheden, maar ook van bedrijven en non-profit organisaties. Afhankelijk van het terrein, de probleemstelling en de beschikbare gegevens worden in dit onderzoek zeer verschillende methoden gebruikt. De keuze van de methode wordt niet zelden ingegeven door hetgeen gangbaar is in een bepaalde onderzoekslijn, en de vooronderstellingen en beperkingen worden niet altijd expliciet gemaakt. Er is wellicht ruimte voor reflectie en kruisbestuiving. Dit was de gedachte die aan de basis lag van dit themanummer.

Het zou overdreven zijn om te zeggen dat de zeven artikels, verzameld in dit nummer, een staalkaart van het onderzoek naar de effecten van beleid in Vlaanderen zouden vormen. Tenslotte is de samenstelling van het nummer het resultaat van spontane respons op onze oproep voor papers, en zoals uit de bijdrage van Glorieux, Moens en Van Thielen in dit nummer blijkt, zo’n spontane respons kan geen representatief beeld opleveren. Niet alle beleidsgericht onderzoek komt aan bod. De sociaal-wetenschappelijke analyse van beleid reikt immers veel verder dan alleen maar de effecten van beleid. Beleidsvoering omvat immers ook de kadering van waaruit het beleid vorm krijgt, de doelgroep waarop het beleid is gericht, de middelenkeuze en het tempo van de uitvoering. Over elk van deze aspecten wordt er onderzoek verricht en het is duidelijk dat de keuzes die beleidsvoerende instanties terzake maken, ook een impact hebben op de effecten. Een andere beperking is dat bijna alle auteurs afkomstig zijn uit de sociale wetenschappen. Sociaal-wetenschappelijke onderzoekers zijn uiteraard lang niet de enigen die de effecten van beleid bestuderen: economen verrichten dergelijk onderzoek evenzeer, en ook de politieke wetenschappen en de rechtswetenschap hebben oog voor deze onderwerpen.

Desalniettemin zijn de artikels zeer divers van aard. Sommige artikels passen een experimentele (Breda en Schoenmaekers) of quasi-experimentele (De Lathouwer, Bogaerts en Van den Bosch) methode toe, andere gebruiken cross-sectionele gegevens. Beleid uitgaande van zowel nationale, regionale als van lokale overheden wordt behandeld, terwijl De Prins, Van den Audenaerde en Henderickx het beleid van zorginstellingen bestuderen. De terreinen van beleid die aan de orde komen betreffen onder meer zorg (Breda en Schoenmaekers, De Prins en collega's), fiscaliteit en sociale zekerheid (Van Mechelen en Verbist) en arbeidsmarkt (De Lathouwer en collega's).

De meeste artikels presenteren de resultaten van een concreet onderzoek, waarbij min of meer uitgebreid wordt stilgestaan bij de gebruikte methode. Sommige bijdragen zijn echter eerder opgevat als overzicht artikelen over een bepaalde benadering of methode. Zo breekt Marx een lans voor comparatief case onderzoek, terwijl Van Mechelen en Verbist de mogelijkheden van micro-simulatiemodellen voorstellen. Verlet, Devos en Reynaert geven in het eerste deel van hun artikel een overzicht van de voor- en nadelen van subjectieve indicatoren van de kwaliteit van het leven.

Ondanks deze diversiteit in methoden en benaderingen zijn er ook enkele gemeenschappelijke kenmerken in de verschillende bijdragen, die misschien ook wijzen op tekortkomingen in het onderzoek naar de effecten van beleid in Vlaanderen. Zo werken alle empirische artikels met enquêtegegevens. Administratieve gegevens worden vooralsnog zeer beperkt aangewend voor beleidsanalyse. De overvloedige neerslag van gegevens bij allerlei beleidsactoren in het brede sociale beleid is erg weinig gesystematiseerd en beschikbaar voor onderzoekers. Hun belang is erg onderschat en kan alleen maar toenemen in de toekomst. Daarnaast lijkt veel beleidseffect-onderzoek vooral aandacht te hebben voor korte termijn-effecten, daar waar er ongetwijfeld dikwijls vooral lange termijn-effecten zijn te verwachten. Onderzoek daarover vereist bijvoorbeeld ook een longitudinaal opzet.

Onderzoekers gebruiken een grote waaier van methoden om de effecten van beleid vast te kunnen stellen. Impliciet in de meting van de effecten van beleid is er echter altijd een vergelijking met een situatie zonder dat beleid. Hieronder in deze inleiding wordt getracht de diverse methoden te ordenen naargelang de manier waarop die vergelijking – impliciet of expliciet – wordt gemaakt. Op deze manier komen een aantal beperkingen en problemen van de diverse methoden aan het licht. Uiteraard zal waar mogelijk verwezen worden naar artikels in dit themanummer. Niet alle benaderingen zijn echter vertegenwoordigd; en ook passen niet alle artikels in dit raamwerk. We hopen toch dat deze manier van voorstellen verhelderend is.

Vaak wordt het sociaal experiment naar voren geschoven als de ideale manier om de effecten van beleid te evalueren. In zulke experimenten worden personen (of andere onderzoeksgevallen) toevalsgewijs aan een onderzoeksgroep,

of een controlegroep toegewezen. In beide groepen vindt zowel vóór als na de ingreep of behandeling (bijvoorbeeld, een cursus solliciteren, of het toekennen van een uitkering) een meting plaats van variabelen waarvoor een effect wordt verwacht (vóór- en nameting). Het effect van de ingreep of behandeling kan dan in principe geschat worden als het verschil tussen de verschillen: het verschil tussen de na- en voormeting bij de onderzoeksgroep, met daarvan afgetrokken het eventuele verschil tussen de na- en voormeting bij de controlegroep. Onder invloed van buitenexperimentele factoren kan de variabele waarbij effecten worden verwacht immers veranderen ook zonder experimentele ingreep.

Een voorbeeld van deze benadering in dit themanummer is het artikel door Breda en Schoenmaekers, die rapporteren over een van de weinige sociaal-wetenschappelijke experimentele onderzoeken in Vlaanderen, waarin het persoonlijk assistentiebudget (PAB) voor personen met een handicap werd geëvalueerd. Ondanks de kleine aantallen, het relatief beperkte budget, en andere problemen (zie hieronder), konden zij de experimentele opzet grotendeels waarmaken, en daarmee aantonen dat zo'n persoonlijk assistentiebudget voor bepaalde groepen van personen met een handicap duidelijk positieve effecten op het welbevinden heeft: meer activiteit, betere zorgverlening.

Ondanks hun grote aantrekkingskracht hebben sociale experimenten grote beperkingen, zoals Heckman e.a. (1999) benadrukken, en zoals ook blijkt uit het onderzoek van Breda en Schoenmaekers. Ten eerste kunnen zij eigenlijk alleen ingezet worden voor beleidsmaatregelen die nog niet in voege zijn (zoals bij de PAB het geval was). Dit is evident zo bij maatregelen die een gehele populatie betreffen, zoals bijvoorbeeld het recht op een uitkering. Om morele en wettelijke redenen kan men moeilijk op willekeurige wijze aan sommigen een recht ontnemen, en anderen dit laten behouden. Dit probleem geldt echter ook bij programma's waaraan normaliter een selectie deelneemt van de populatie die in aanmerking komt. Aan werkervaringsprogramma's zal bijvoorbeeld slechts een deel van de personen deelnemen die in principe in aanmerking komen. Wanneer een onderzoeker uit de gegadigden op willekeurige wijze een experimentele groep zou samenstellen, kunnen de niet-geselecteerden echter in veel gevallen gelijkaardige programma's vinden, zodat de controlegroep niet meer zuiver is.

Ten tweede zijn experimenten onvermijdelijk beperkt in omvang, tijd en ruimte. (Het onderzoek van Breda en Schoenmaekers duurde, dankzij verlengingen, uiteindelijk 3 jaar, een vrij lange periode voor beleidsgericht onderzoek.) Lange-termijn effecten zijn daarmee moeilijk meetbaar. Daarenboven zullen de deelnemers wellicht al anticiperen op het einde van het experiment, en daarom niet altijd de beoogde gedragseffecten tonen. Anderzijds, wanneer een experiment wordt gezien als de voorbode van werkelijk nieuw beleid, is het mogelijk dat ook de controlegroep zijn gedrag al bijstelt.

Een derde beperking is dat personen wel uitgesloten kunnen worden van een programma of behandeling, maar meestal niet verplicht kunnen worden om mee te doen. Vanwege de vrijwilligheid van deelname is de onderzoeksgroep in veel praktisch onderzoek toch selectief samengesteld, ondanks alle inspanningen van de onderzoekers. Daar het uiteraard veelal de sterker gemotiveerden of beter geschikt zijn die zullen meedoen, kan deze selectieve deelname de resultaten positief vertekenen. Ook het onderzoek naar de PAB's trok een kleine groep van sociale voortrekkers aan, die beter voorbereid en geïnformeerd waren dan de gemiddelde deelnemer zal zijn. Tenslotte zijn experimenten een dure vorm van onderzoek, en kosten zij veel tijd en energie aan onderzoekers én veldwerkers. De wetenschappelijke vereiste van strikt toevalsgewijze toekenning aan onderzoeksgroep of controlegroep kan conflicteren met de professionele houding van programmaverantwoordelijken en veldwerkers, die het liefst zouden zien dat degenen met de grootste behoefte of de sterkste motivatie zouden deelnemen.

Een tweede methode is de 'verschil van de verschillen' benadering (Engels: *'difference-in-difference'*). Hierbij worden de uitkomsten voor personen die deelnemen aan een bepaald programma of een bepaalde behandeling ondergaan, vergeleken met de uitkomsten voor overigens gelijkaardige personen die niet deelnemen of niet ondergaan, en dit zowel vóór als na het programma of de behandeling. De methode kan daarom beschreven worden als een quasi-experimentele methode; het grote verschil met de 'echte' experimentele methode is dat het veelal gaat om bestaande programma's, zodat de onderzoeker geen zeggenschap heeft over de toewijzing van personen of gevallen. Het grote probleem is dan ook – bij gebrek aan een echte controlegroep – het vinden van een goede vergelijkingsgroep. Per definitie kunnen personen in de vergelijkingsgroep niet volledig gelijk zijn aan de onderzoeksgroep – anders zouden zij immers in aanmerking komen voor het programma in kwestie. Soms nemen onderzoekers daarom hun toevlucht tot een minder strenge assumptie: de vergelijkingsgroep is niet echt vergelijkbaar met de onderzoeksgroep, maar alle trends en factoren (buiten het programma in kwestie) die de uitkomsten in de onderzoeksgroep zouden kunnen beïnvloeden, worden geacht de vergelijkingsgroep in even sterke mate te beïnvloeden. Op deze manier gebruiken Francesconi en Van der Klaauw (2004) alleenstaande vrouwen zonder kinderen als vergelijkingsgroep in hun evaluatie van de impact van het *Working Families Tax Credit* op de tewerkstellingsgraad van alleenstaande moeders (Het *WFTC* is een terugbetaalbaar belastingskrediet voor gezinnen met kinderen, bedoeld om herintrede op de arbeidsmarkt te stimuleren.) De tewerkstellingsgraad van de eerste groep ligt uiteraard veel hoger dan die van de eerste, maar bijvoorbeeld veranderingen in de vraag naar arbeid vanwege de economische conjunctuur zouden beide groepen op identieke wijze moeten beïnvloeden (na controle voor onderwijsniveau, regio etc.).

De 'verschil van de verschillen' benadering kan ook worden gebruikt op een hoger niveau van aggregatie. Om de gevolgen van hervormingen in sociale ze-

kerheid en de bijstand in de USA na te gaan, kunnen de Amerikaanse deelstaten onderling met elkaar worden vergeleken (Blank 2002). Niet alle deelstaten voeren altijd alle maatregelen door, of doen dat niet op hetzelfde moment. Verschillen in uitkomsten tussen staten kunnen dan worden toegeschreven aan de hervorming, vooropgesteld dat de effecten van andere veranderlijke factoren voor alle staten gelijk zijn.

In dit themanummer kan het onderzoek van De Lathouwer, Bogaerts en Van den Bosch beschouwd worden als een poging om de 'verschil van de verschillen' benadering toe te passen. Zij bestuderen de gevolgen van de schorsing van werklozen uit de werkloosheidsverzekering wegens 'abnormaal langdurige werkloosheid' volgens het voormalige Artikel 80 op de mogelijke herintrede op de arbeidsmarkt. Daartoe vergelijken zij een steekproef van werklozen die in een bepaalde periode van drie maanden geschorst werden met niet-geschorste langdurig werklozen, en kijken naar de evolutie van de tewerkstelling in de twee groepen over een periode van 18 maanden. Met behulp van een logit-model controleren zij voor gemeten verschillen tussen de onderzoeksgroep en de vergelijkingsgroep. Schorsing volgens Artikel 80 lijkt een positieve, maar beperkte, invloed te hebben op de kans op herintrede op de arbeidsmarkt. In hun genuanceerde bijdrage bespreken De Lathouwer e.a. de beperkingen die aan de gebruikte methode kleven, en de voorwaarden waaronder deze een valide schatting kan opleveren van het werkelijke effect van schorsing volgens Artikel 80. Een bijkomende complicatie was dat er geen echte voormeting had plaatsgevonden: er werd bij de onderzoeksopzet vanuit gegaan dat het moment van steekproeftrekking een soort natuurlijk nulpunt was, waar geen enkele werkloze werk zou hebben, noch in de onderzoeksgroep, noch in de vergelijkingsgroep. De sociale werkelijkheid was – zoals altijd – gecompliceerder dan onderzoekers op voorhand konden bevroeden, en deze assumptie bleek achteraf niet houdbaar.

Wellicht de meest directe strategie ter meting van de effecten van beleidsingrepen is een vergelijking tussen relevante uitkomsten vóór en na de invoering van een maatregel. Als gegevens beschikbaar zijn voor een langere tijdsperiode kan men eventueel controleren voor langlopende dalende of stijgende trends in deze uitkomsten. Men zoekt dan als het ware naar een 'sprong' in de trend, die samenvalt met de beleidsingreep.

Ondanks de intuïtieve plausibiliteit kan deze methode toch misleidende resultaten opleveren. Andere gebeurtenissen kunnen samenvallen met de beleidsingreep. Zo daalde het aantal mensen in de bijstand zeer fors, en steeg de tewerkstellingsratio onder alleenstaande moeders sterk na de hervormingen van het Amerikaans bijstandstelsel onder president Clinton. (Beide waren centrale doelstellingen van deze hervorming.) Tegelijk was dit echter een periode van zeer sterke groei in de Amerikaanse economie, zodat het niet eenvoudig is

uit te maken in welke mate de genoemde positieve evoluties toe te schrijven zijn aan de hervormingen (Blank en Ellwood 2001).

Op micro-niveau kent de vóór-na methode het probleem dat deelname aan een programma het resultaat kan zijn van een tijdelijk probleem of situatie (de zogeheten *'Ashenfelter dip'*, zie Heckman e.a. 1999). Een persoon kan werkloos worden, instappen in een scholingsprogramma of sollicitatiecursus, en vervolgens weer werk vinden. Het laatste is echter niet noodzakelijk het resultaat van dat programma of die cursus. Op geaggregeerd niveau is er het probleem dat de beleidsingreep endogeen kan zijn: programma's ter bestrijding van de werkloosheid worden in gang gezet als de werkloosheid stijgt. Als op het moment dat de programma's op kruissnelheid geraken de conjunctuur ondertussen weer is gekeerd, kan het zeer moeilijk zijn om de effecten van de programma's te onderscheiden van die van de verbeterende economische omgeving.

Het complement van de vóór-na methode is de cross-sectionele benadering. Toegepast op micro-niveau houdt deze benadering in dat uitkomsten voor deelnemers aan een programma of worden vergeleken met personen of gevallen die niet aan dat programma hebben deelgenomen, of, vaker, aan een ánder programma hebben deelgenomen. Op macro-niveau, waar grote eenheden, zoals maatschappijen, met elkaar worden vergeleken, heeft de cross-sectionele methode een grote populariteit binnen de politieke wetenschappen onder de naam van 'de comparatieve methode'. Ook op het meso-niveau van bedrijven, organisaties en verenigingen kan de benadering uiteraard toegepast worden.

De cross-sectionele methode kan als een gekortwiekte versie van de 'verschil-van-de-verschillen' methode worden gezien, waarbij de voormeting is verval- len. Dit heeft uiteraard grote nadelen, en in het licht van wat hierboven al is gezegd zijn de beperkingen van de methode dan ook evident. Desalniettemin wordt zij in praktisch onderzoek veel toegepast, omdat de onderzoeker vaak pas op het toneel verschijnt als het beleid al lang en breed in voege is.

In dit themanummer wordt de cross-sectionele benadering vertegenwoordigd door De Prins, Van den Audenaerde en Hendrickx, en door Verlet, Devos en Reynaert. De Prins en collega's bestuderen de arbeidskwaliteit in een groot aantal Vlaamse zorginstellingen in twee sectoren: rusthuizen en instellingen voor bijzondere jeugdbijstand. Zij delen de instellingen in in drie groepen: 'excellente' instellingen waar de personeelsleden relatief weinig te kampen hebben met stress en burn-out, 'risico'-instellingen, waar de verplegenden en verzorgenden juist blijk gaven van hoge niveaus van stress en burn-out, en de 'doorsnee' instellingen daar tussenin. Het onderzoek toont aan dat 'excellente' instellingen zich van andere onderscheiden door meer mogelijkheden voor beleidsoverleg en - ondersteuning. Zij verschillen echter niet significant van de andere instellingen inzake de toepassing van innovatieve managementtechnieken. De onderlinge arbeidsverhoudingen, en de kwaliteit van de leiding op de

werkvloer (tot uiting komend in zaken als het op tijd bekend maken van dienstroosters) lijken belangrijker om stress en burn-out te vermijden.

In zekere zin draaien De Prins en collega's de logica van de meeste onderzoeken naar de effecten van beleid om: zij meten niet wat de gevolgen zijn van een bepaalde beleidstechniek of -innovatie, maar kijken naar de uitkomsten, en zoeken dan naar factoren die verschillen in deze uitkomsten kunnen verklaren, of, zoals zij het zelf formuleren, "door niet langer uit te gaan van de normatieve theorie, maar van de 'excellente' praktijk kunnen belangrijke lessen worden geleerd." In het geval van dit onderzoek levert deze meer exploratieve benadering inderdaad vruchten op.

Verlet en collega's bestuderen de mate van tevredenheid met het lokaal beleid in Antwerpen, Brugge en Gent (over de keuze voor subjectieve indicatoren hieronder meer). Dit betekent dat zij in zekere zin de cross-sectionele methode op twee niveaus toepassen: enerzijds op het niveau van de burgers van een stad, anderzijds op het geaggregeerde niveau van verschillende steden. Gegeven hun onderzoekszopzet is het moeilijk om de verschillen in tevredenheid tussen burgers terug te voeren op de effecten van concreet stedelijk beleid, noch is het eenvoudig om na te gaan in hoeverre de verschillende niveaus van tevredenheid in de drie Vlaamse steden verklaard kunnen worden uit verschillend beleid, dan wel door de verschillende stedelijke contexten. Desalniettemin kunnen Verlet en collega's de belangrijke conclusie trekken dat mensen die invloed hebben – of denken te hebben – op het beleid veel tevredener zijn dan anderen. Gevoelens van machteloosheid daarentegen wekken ontevredenheid op.

Bovenstaande benaderingen hebben alle met elkaar gemeen dat zij de uitkomsten van een programma vergelijken met een situatie zonder dat programma, die werkelijk bestaat of bestond – ofwel de situatie van vergelijkbare gevallen die niet betrokken zijn in het programma, ofwel de situatie van dezelfde gevallen voordat zij aan het programma deelnamen. In modelmatige studies van de effecten van beleid is het vergelijkingspunt geen reëel bestaande toestand, maar een hypothetische of gesimuleerde '*counterfactual*'. Onderzoekers gebruiken een model om de effecten van de invoering van een beleidsingreep (of ook van het ontbreken van een bepaald programma) te kunnen inschatten. De bekendste voorbeelden zijn wellicht de macro-economische modellen die het Planbureau gebruikt om de gevolgen van monetair en globaal economisch beleid, zoals een renteverhoging, te berekenen. Economische modellen bestaan ook op micro-niveau. Bijvoorbeeld, Blundell e.a. (2000) gebruiken enquête gegevens, een microsimulatiemodel van belastingen en uitkeringen en een econometrisch model van arbeidsmarktgedrag om de effecten te voorspellen van het '*Working Families Tax Credit*' (een terugbetaalbaar belastingkrediet voor werkende ouders) in het Verenigd Koninkrijk op de arbeidsmarktparticipatie van alleenstaande ouders. De correctheid van zulke voorspellingen hangt na-

tuurlijk in grote mate af van de kwaliteit van de gegevens, en vooral van die van het model en zijn parameters. Meestal worden gedragsmodellen, bijvoorbeeld betreffende arbeidsmarktgedrag, geschat op basis van enquêtegegevens, zodat de parameters onderhevig zijn aan steekproeffouten, en ook aan specificatiefouten. Bovendien, parameters geschat op steekproefgegevens voor de gehele bevolking, of voor grote groepen (bv. alle gehuwde of samenwonende vrouwen) zijn niet noodzakelijk geldig voor de specifieke groepen waar het beleid zich veelal op richt (bv. laaggeschoolde langdurig werkloze vrouwen).

Een bepaald soort model zijn microsimulatiemodellen van belastingen en uitkeringen (zie de bijdrage van Van Mechelen en Verbist in dit themanummer). In deze modellen wordt de regelgeving inzake belastingen en sociale uitkeringen in zo groot mogelijk mate van detail nagebootst in een computerprogramma, en toegepast op inkomensgegevens van een steekproef van huishoudens en/of individuen afkomstig van een enquête of uit administratieve bron. Men kan bestaande regels vervangen door alternatieve regels, en aldus de effecten simuleren van die wijziging op de inkomensverdeling. Gedragsreacties zijn meestal niet ingebouwd in zulke modellen, zodat het gesimuleerde effect alleen een schatting biedt van de eerste-orde effecten, zonder rekening te houden met tweede-orde effecten via veranderend gedrag of via macro-economische terugkoppelingseffecten. Van de andere kant kunnen zij die eerste-orde effecten in zeer grote mate van detail tonen: welk soort huishoudens gaat er in welke mate op vooruit, wie gaat er hoeveel op achteruit? Voor beleidsmakers is dit al zeer informatief, vooral bij maatregelen waarvoor verwacht kan worden dat de tweede-orde effecten beperkt zullen zijn. Met behulp van microsimulatie kunnen Van Mechelen en Verbist vaststellen dat de belastinghervorming van 2001 tot een zeer geringe stijging van de inkomensongelijkheid zou leiden. Het is zeer twijfelachtig of zulk een vaststelling ook gedaan zou kunnen worden op basis van een tijdreeks van de gemeten inkomensongelijkheid, gezien de jaar-op-jaar schommelingen in deze grootheid.

Van Mechelen en Verbist bespreken in hun bijdrage naast microsimulatiemodellen ook de zogeheten standaardsimulatiemodellen, waarin de sociale en fiscale regelgeving wordt nagebootst voor een aantal hypothetische type-gezinnen. Om deze reden tonen standaardsimulatiemodellen niet zozeer de effecten van beleid, als wel de implicaties van beleid. De sociale en fiscale wetgeving is immers bijzonder ingewikkeld, en het is niet altijd duidelijk wat een bruto-verhoging van een uitkering een gezin netto zal opleveren, rekening houdende met fiscaliteit, maar ook met andere inkomensgetoetste uitkeringen. Van Mechelen en Verbist kunnen aldus met deze techniek de evolutie van de werkelijke belastingdruk in België tussen 1989 en 2003 in kaart brengen voor een aantal type gezinnen.

In de tot nu toe besproken benaderingen is 'meer' altijd 'beter': hoe groter de steekproef of hoe meer gevallen de data-base omvat, hoe betrouwbaarder de

effecten van beleid kunnen worden geschat. In zijn bijdrage breekt Marx een lans voor studies waarin een beperkt aantal gevallen grondig worden vergeleken, en getracht wordt deze gevallen zo omvattend mogelijk te verklaren. Via een analyse van verschillen en gelijkenissen kunnen noodzakelijke voorwaarden voor een bepaalde uitkomst worden opgespoord. De grondige studie van een beperkt aantal gevallen laat toe om causale processen en mechanismen te beschrijven. Aldus kan men 'waarom' vragen beantwoorden: niet 'wat is het effect van dit beleid?', maar 'waarom leidt dit beleid tot dit effect?' Deze benadering blijkt onder meer vruchtbaar als men vraagt waarom sommige landen een sterke economische ontwikkeling kennen, terwijl andere in armoede blijven steken. In dit soort onderzoek kan één nieuwe afwijkende case de conclusies omgooien, en een juiste selectie van de cases is dan ook van groot belang.

De meeste onderzoeken proberen min of meer specifieke effecten van beleid te vinden. De uiteindelijke doelstelling van alle beleid bestaat wellicht uit het bevorderen van de kwaliteit van het leven. De kwaliteit van het leven is echter een veelomvattend begrip, en verschillende mensen zullen verschillende accenten leggen. Verlet, Devos en Reynaert stellen dat de persoonlijke ervaring centraal moet staan in de meting van de kwaliteit van het leven, zodat de eigen beoordeling door het individu van zijn of haar leven essentieel is. Dit betekent dat de meting van de kwaliteit van het leven niet zonder subjectieve indicatoren kan, ter aanvulling van (niet: in plaats van) meer objectieve indicatoren. Deze piste werd uitgewerkt in een empirische studie over de tevredenheid over het lokaal beleid te Gent, Brugge en Antwerpen.

Het blijkt dat het algemeen subjectief welzijn voornamelijk samenhangt met de tevredenheid over meer private facetten van het leven, en ook met kenmerken van de persoonlijkheid. Dit geeft het lokaal beleid weinig aanknopingspunten om het algemeen subjectief welzijn te verhogen, tenzij via de verbetering van het gevoel van verbondenheid met het maatschappelijk gebeuren. De tevredenheid met het stadsbeleid zelf hangt voor een deel af van de beoordeling van afzonderlijke aspecten van beleid, maar wordt voor een groot deel ook bepaald door exogene factoren. Daarnaast oordelen vele mensen over aspecten van beleid, zonder deze werkelijk te kennen. Dit alles betekent dat subjectieve indicatoren moeilijk gebruikt kunnen worden bij de evaluatie van beleid, en zeker geen goed meetinstrument zijn voor de kwaliteit van de stedelijke dienstverlening. Dit neemt niet weg dat subjectieve attitudes op zichzelf een belangrijk gegeven zijn voor stedelijke beleidsmakers.

De mening van de bevolking staat ook centraal in wat Glorieux, Moens en Van Thielen het Spontaan Respons Onderzoek noemen. Dit omvat allerlei manieren van gegevensverzameling, zoals internetbevragingen, referenda en televoting, die gemeen hebben dat ze de meningen verzamelen van mensen die zich zelf aanmelden. Is dit een goedkope en snelle manier om de mening van de bevolking te peilen? Glorieux en collega's vergelijken het Spontaan Respons

Onderzoek dat in het kader van de campagne Kleurrijk Vlaanderen via onder meer kranten en een webstek gehouden werd rond het thema mobiliteit, met een onderzoek bij een a-selecte steekproef. Het blijkt dat de resultaten in het geheel niet overeenkomen, ook niet na herweging van het Spontaan Respons Onderzoek. Deze vorm trekt een specifiek staal aan van geïnteresseerde burgers die hun stem willen laten horen, en die niet representatief zijn voor de gehele bevolking. Die geïnteresseerde burgers kunnen wel argumenten en materiaal aandragen om een debat te stofferen en te versterken. Maar voor een diepgaande analyse van argumenten zijn dan weer andere onderzoeksvormen, bij een beperkt en geselecteerd panel van personen, veel beter geschikt. Spontaan Respons Onderzoek lijkt daarmee veeleer zelf onderdeel van een public-relations beleid, dan een manier om beleid te beoordelen.

NOOT

1. Vele mensen hebben aan dit themanummer meegewerkt. Onze dank gaat allereerst uit naar de auteurs voor hun bijdragen, voor de bereidwilligheid om hun bijdragen te herzien, en voor hun geduld om te wachten op publicatie. Ook willen wij graag de anonieme beoordelaars bedanken voor hun vaak uitgebreide en grondige commentaar. Mieke Augustyns verdient onze dank voor de grondigheid waarmee zij alle artikels heeft nagelezen op spel- en typfouten, afwijkingen van de vormvereisten, en ontbrekende of onjuiste referenties. Mieke De Becker en Josiane Van Mele hebben in een laatste fase de artikels eenvormig gemaakt in hun layout. Tenslotte bedanken wij Kurt De Wit voor laatste vormelijke aanpassingen, en voor de efficiënte communicatie met Acco en met de redactie van het Tijdschrift voor Sociologie.

BIBLIOGRAFIE

- Blank, R. (2002), Evaluating Welfare Reform in the United States. *Journal of Economic Literature*, 40: 1105-1166.
- Blank, R. and Ellwood, D. (2001), *The Clinton Legacy for American's Poor*. NBER Working Paper No 6998. Cambridge (Massachusetts): National Bureau of Economic Research.
- Blundell, R., Duncan, A., McCrae, J. and Meghir, C. (2000), The labour market impact of the working families' tax credit. *Fiscal Studies* 21(1): 75-103.
- Francesconi, M. and W. Van der Klaauw (2004), *The Consequences of 'In-work' Benefit Reform in Britain: New Evidence from Panel Data*. ISER Working Paper 2004-13, Colchester: ISER, University of Essex.
- Heckman, J., Lalonde, R. and Smith, J. (1999), The Economics and Econometrics of Active Labor Market Programs. Pp. 1865-2097 in *Handbook of Labor Economics* vol. 3, eds. Ashenfelter, A. and Card, D.. Amsterdam/Lausanne/New York/Oxford/Shannon/Tokyo: Elsevier.