



Goed beleid begint met goede toetsing

Auteur(s):

Klaauw, B., van der Berg, G.J., van den

Bas van der Klaauw is werkzaam bij de afdeling Algemene Economie, Vrije Universiteit Amsterdam, en is verbonden aan nwo prioriteitsprogramma Scholar. Gerard J. van den Berg is werkzaam bij de afdeling Algemene Economie, Vrije Universiteit Amsterdam, en is verbonden aan OCFE.

Verschenen in:

ESB, 86e jaargang, nr. 4335, pagina 904, 23 november 2001

Rubriek:**Trefwoord(en):**

arbeidsmarkt

Beleidsambtenaren en politici hebben geen idee van de effectiviteit van het gevoerde activerend arbeidsmarktbeleid. Kennis hierover is van cruciaal belang voor de ontwikkeling van goed beleid. Betere beleidsevaluatie is daarom nodig.

Half oktober is het door een aantal topambtenaren geschreven rapport Aan de slag openbaar gemaakt¹. Het rapport geeft een overzicht van het activerend arbeidsmarktbeleid dat het afgelopen decennium in Nederland is gevoerd en het formuleert richtlijnen voor toekomstig activerend arbeidsmarktbeleid. Enkele conclusies uit het rapport waren eind augustus al uitgelekt. Gesuggereerd werd dat met name de Melkertbanen en loonkostensubsidies voor langdurig werklozen nauwelijks extra mensen aan regulier werk helpen. Deze conclusie werd als bijzonder schadelijk gezien voor de voormalig minister van Sociale Zaken en Werkgelegenheid Ad Melkert. In dit artikel bespreken we het rapport en gaan we in op de vraag of de kritiek op Melkert terecht is. We komen daarbij als vanzelf op de vraag hoe de kwaliteit van beleidsevaluaties verbeterd kan worden.

Aan de slag?

In het rapport *Aan de slag* wordt ingegaan op de vraag hoe effectief de verschillende beleidsinstrumenten de afgelopen jaren zijn geweest. Op basis van het rapport kan worden geconcludeerd dat beleidsmakers daarvan eigenlijk geen idee hebben. Een indrukwekkende lijst beleidsevaluatierapporten wordt besproken, maar een nadere blik op de rapporten leert dat de informatieve waarde van de meeste rapporten vrijwel nihil is.

Grofweg wordt onderscheid gemaakt tussen twee vormen van beleidsevaluatie. Dat zijn ten eerste empirische studies op microniveau en ten tweede analyses op macro-niveau met behulp van MIMIC, een algemeen evenwichtsmodel van het Centraal Planbureau. De kwaliteit van de empirische studies op micro-niveau laat vaak te wensen over. Meestal wordt een groep deelnemers aan een bepaald type beleid vergeleken met een ad hoc geconstrueerde controlegroep van niet-deelnemers. Je hoeft geen briljant onderzoeker te zijn om te onderkennen dat een groep deelnemers niet vergelijkbaar is met zo maar een groep niet-deelnemers. Veel verschillen tussen deelnemers en niet-deelnemers kunnen worden waargenomen, zoals verschillen in opleidingsniveau en werkervaring. In een empirische analyse kan eenvoudig rekening worden gehouden met verschillen in zulke waargenomen kenmerken. Er zijn echter ook kenmerken die niet worden waargenomen, bijvoorbeeld sociale vaardigheid en motivatie om werk te vinden.

In sommige micro-studies wordt een poging gedaan te corrigeren voor niet-waargenomen kenmerken. Uit wetenschappelijke literatuur is echter bekend dat correctie voor niet-waargenomen kenmerken slechts in een beperkt aantal gevallen mogelijk is en dan nog onder erg strikte aannames. De gevonden resultaten kunnen daarom evengoed het gevolg zijn van verschillen in niet-waargenomen kenmerken tussen deelnemers en niet-deelnemers als van blootstelling aan het beleid.

De tweede vorm van beleidsevaluatie geeft evenmin houvast. Het is nogal ambitieus om algemene evenwichtseffecten te onderzoeken als de partiële individuele effecten onbekend zijn. Het rommelen aan de parameters van MIMIC moet eerder worden opgevat als een uit de hand gelopen liefhebberij dan als een serieuze poging activerend arbeidsmarktbeleid te evalueren.

Beleidsvaluatie in het buitenland

Hoe kan worden voorkomen dat over tien jaar een soortgelijk rapport geschreven wordt? De eerste aanzet hiervoor wordt in *Aan de slag* gegeven. In de toekomst zal een protocol ontwikkeld moeten worden dat zorg draagt voor een zorgvuldige evaluatie van beleidsinstrumenten. Zo'n protocol zal in ieder geval gecontroleerde experimenten moeten bevatten waarbij individuen willekeurig toegewezen worden aan verschillende vormen van beleid.

In het buitenland bestaat al ruime ervaring met beleidsevaluatie aan de hand van gecontroleerde experimenten. Een goed voorbeeld is de Maryland Unemployment Insurance Work Force Demonstration (1998). Gedurende een jaar werden 27.000 werklozen willekeurig toegewezen aan een aantal groepen die verschilden in de verplichtingen omtrent wekelijkse baanzoekinspanning en de controle hierop. Eén van de groepen, de controlegroep, werd volledig met rust gelaten. De uitkomsten ten aanzien van de werkloosheidsduur van de verschillende groepen werden met elkaar en met de controlegroep vergeleken, waarna het meest kosteneffectieve beleid eenvoudig te bepalen was. Natuurlijk werd hierbij ook gekeken naar lonen, kwaliteit van gevonden banen en dergelijke.

Het Maryland-experiment is niet uniek. De afgelopen jaren zijn in het buitenland veel gecontroleerde experimenten uitgevoerd om de effectiviteit van beleidsinstrumenten te toetsen. De meeste experimenten vonden plaats in Amerika, onder andere bij de evaluatie van intensieve hulp bij het zoeken naar werk, sollicitatietraining, werkervaringsprojecten, striktere controle van baanzoekinspanningen, bijscholingsprojecten en het verstrekken van bonussen voor werkherleving.

Gedifferentieerd beleid nodig

Wat kan worden geleerd van het buitenlandse beleidsevaluatieonderzoek? Onlangs zijn enkele goede overzichten gepubliceerd van evaluatiestudies in verschillende landen. De uitkomsten ten aanzien van het scheppen van gesubsidieerde banen in de niet-private sectoren zijn onthutsend. Deze maatregelen zijn nutteloos om de doorstroom van werklozen naar regulier werk te stimuleren. Reïntegratiekansen zijn wat groter bij beperkte loonkostensubsidies in de private sector. De gesubsidieerde banen creëren echter nauwelijks extra werk, ze verdringen vooral reguliere banen.

Bij om-, her- en bijscholingsprogramma's voor werklozen is het belangrijk onderscheid te maken tussen mensen met verschillende kenmerken. Het meest profiteren mensen met sluimerende kwaliteiten. Voor herintredende vrouwen zijn scholingsprogramma's succesvol. Ook immigranten met een redelijk opleidingsniveau in het land van herkomst zijn gebaat bij deelname aan trainingsprogramma's.

Wat het beeld somber maakt, is dat de werklozen met de laagste kwaliteiten het minste profiteren. Trainingsprogramma's voor werkloze jongeren zijn nutteloos. Dit laatste is waarschijnlijk het meest consistente resultaat uit de buitenlandse studies.

Eenvoudige doch intensieve hulp bij het vinden van een baan lijkt nog steeds de nuttigste vorm van beleid te zijn. Hierbij moet de nadruk wel op 'intensief' liggen. Een half uurtje per maand is onvoldoende. Opnieuw geldt dat herintredende vrouwen het meest profiteren van dit soort hulp.

Ook strengere controle van werklozen, met een geloofwaardige dreiging dat misbruik tot sancties leidt, heeft haar nut bewezen. Het effect van controle en straf is het grootst bij de minst kansrijken. Dit plaatst de ambtenaar van de sociale dienst voor een dilemma. Moet hij die zielige langdurig werkloze gaan straffen, of zal hij zich bezighouden met het controleren van pas afgestudeerde academici? Het rechtvaardigheidsgevoel zal misschien leiden tot het laatste, maar vanuit beleidsoogpunt is het eerste te prefereren.

De Nederlandse situatie

Buitenlandse resultaten kunnen niet zomaar worden geëxtrapoleerd naar de Nederlandse situatie. Elk land heeft eigen unieke instituties. Beleidsmakers kunnen zich bij nieuw beleid wél laten inspireren door succesvol buitenlands beleid. Om na te gaan of dit beleid ook effectief is voor de Nederlandse situatie is een degelijke beleidsevaluatie noodzakelijk. Gecontroleerde experimenten zijn degelijk en goedkoop. Ze kosten vaak maar een fractie van de kosten van het daadwerkelijke beleidsprogramma. In Nederland is nog weinig ervaring met experimenten. In opdracht van het Landelijk instituut sociale verzekeringen is tweemaal een experiment uitgevoerd naar de effectiviteit van activering en controle. Het eerste experiment betrof ongeveer 1600 werklozen, het tweede zo'n 400. Dit is veel minder dan de 27.000 in het experiment in Maryland, terwijl Nederland drie keer zoveel inwoners heeft.

Bij de introductie van de 'sluitende aanpak van de werkloosheid' enkele jaren geleden, bestond bij het ministerie de intentie tot grondige evaluatie met behulp van gecontroleerde experimenten. Dit is een stille dood gestorven. De vraag is of de beleidsmakers in Den Haag nu wél klaar zijn voor de cultuuromslag waarbij beleidsevaluatie net zo belangrijk wordt als beleidsontwikkeling. Een uitgelezen mogelijkheid voor een beleidsevaluatie met behulp van een gecontroleerd experiment is de recente introductie van bonussen voor langdurig werklozen die werk vinden. Er is nooit gesproken over evaluatie van dit beleidsinstrument. Daarom is onze inschatting dat er volgend jaar een ad hoc rapport zal verschijnen dat in vage termen de effectiviteit van dit beleidsinstrument zal beschrijven.

Conclusie

In onze ogen is de belangrijkste conclusie van het rapport Aan de slag dat de beleidsmakers in Den Haag weinig idee hebben van de effectiviteit van het activerend arbeidsmarktbeleid dat de afgelopen jaren is gevoerd. De belangrijkste kritiek op Ad Melkert zou dus niet moeten zijn dat hij jarenlang ineffectief beleid heeft gevoerd, maar dat hij de effectiviteit van zijn beleid niet grondig heeft getoetst. Deze kritiek is echter evenzeer van toepassing op voorgangers en opvolgers bij het ministerie van Sociale Zaken en Werkgelegenheid. Het positieve van het rapport is dat de relevantie van beleidsevaluatie wordt benadrukt en dat wordt erkend dat daar de afgelopen jaren te weinig aandacht aan is besteed. Het voorgestelde protocol voor systematische grondige evaluatie biedt een klein sprankje hoop voor de toekomst

1 Ministerie van SZW, *Aan de slag*, Den Haag, oktober 2001.