



Oorbaar optellen

Auteur(s):

Cherchye, L.
Moesen, W.
Puyenbroeck, T. van

De auteurs zijn verbonden aan economiedepartementen van de Associatie ku Leuven: Cherchye aan de KU Leuven en FWO-Vlaanderen, Moesen aan de KU Leuven en Van Puyenbroeck aan de Europese Hogeschool Brussel. Ton.vanpuyenbroeck@pi.be

Verschenen in:

ESB, 88e jaargang, nr. 4413, pagina 415, 5 september 2003

Rubriek:**Trefwoord(en):**

europa

De Europese coördinatie van sociaal beleid vraagt om indicatoren waarmee beleidsresultaten kunnen worden vergeleken. Hoe combineer je verschillende indicatoren?

Hoe doe je aan globale resultaatvergelijking als prestaties in uiteenlopende beleidsdimensies relevant zijn? Dit is een pertinente vraag in het licht van de golf van 'benchmarking' en 'peer review' die nu ook de Europese beleidsmakers overspoelt. De voordeel-van-de-twijfel-benadering die we hier bespreken, verzwakt alvast de argumentatie van diegenen die menen dat één samengestelde indicator de eigenheid van elke welvaartsstaat over het hoofd ziet. We hangen ons verhaal op aan een concrete casus, maar wie tussen de regels door leest, ziet allicht snel alternatieve toepassingsmogelijkheden.

De open coördinatie-methode

Op de Europese top van Lissabon van maart 2000 lanceerden de staatshoofden en regeringsleiders een grootse agenda die van de eu "de meest competitieve en dynamische kenniseconomie ter wereld wil maken, bestemd tot duurzame economische groei met meer en betere banen en sociale cohesie". Ondanks de gezwollen en tegelijk wollige formulering, hebben de leiders met deze intentieverklaring het sociale beleid meer dan ooit tot een aandachtspunt voor Europese samenwerking gepromoveerd. Met zulk vaag mandaat in handen tekenen diplomaten en technici gestadig een beleid uit, waarbij elke tussentijdse blauwdruk de zegen krijgt van de verzamelde premiers en presidenten. Zo formuleerden de regeringsleiders nog in 2000 in Nice én een lijst met nauwer afgebakende doelstellingen én een feitelijke werkwijze om de strijd tegen armoede en sociale uitsluiting te voeren.

Die werkwijze staat, in Euro-speak, bekend als de open coördinatiemethode. Wezenlijk gaat het om een proces waarbij eerst gezamenlijk overeengekomen doelstellingen worden opgesteld. De lidstaten stellen daarna, elk voor zich, zogeheten Nationale Actieplannen op. Die zijn onderworpen aan 'peer review', een bespreking onder gelijken die in principe de mogelijkheid biedt nationale praktijken te vergelijken om zo van elkaar te leren. 'Zachte' aanbevelingen van de regeringsleiders en van de Commissie - die een gemeenschappelijk rapport redigeert - en vergelijkbare en gezamenlijk afgesproken indicatoren, laten toe de vorderingen op weg naar de doelstellingen te evalueren¹.

Misschien lijkt die methode nogal fluwelig, zeker in vergelijking met de strakke aanpak die we kennen als het om begrotingsdoelstellingen gaat. De omzichtigheid heeft echter een goede reden: alle lidstaten reageren uiterst defensief wanneer 'eu' en 'sociaal beleid' in één adem worden genoemd. Ze wantrouwen Brusselse inmenging in wat ze traditioneel als hun hoogsteigen beleidsdomein beschouwen. Dát welvaartsstaten onderling grondig kunnen en mogen verschillen en dat de ultieme controle over het sociale beleid niet bij Europa maar bij de lidstaten ligt, is precies een van de premissen waarop deze zachte methode steunt. Kan je, gegeven deze rechtmatige verscheidenheid, qua onderliggende opvattingen, instituties en concrete beleidsinvulling, toch werken aan gezamenlijke vooruitgang? Die vraag vat de filosofie achter de open coördinatiemethode.

Sociale indicatoren als prikkel?

Juist omdat het hier zo op eieren lopen is, is het opmerkelijk dat de lidstaten zich sowieso achter gezamenlijke doelstellingen hebben geschaard. Nog sterker: eind 2001, op de top van Laken, kwam men aanzetten met een lijst van kengetallen die verschillende aspecten van sociale uitsluiting in kaart brengen. Van meet af aan speelde inderdaad de kernidee dat sociale uitsluiting een multidimensioneel verschijnsel is, dat niet herleidbaar is tot bijvoorbeeld dat ene armoedecijfer. Even belangrijk is dat het om resultaatindicatoren gaat, om outputs en niet om inputs, wat opnieuw de zienswijze weerspiegelt dat lidstaten vrij zijn in de wijze waarop ze hun sociaal beleid voeren.

Bij het opstellen van die lijst ging men niet over één nacht ijs. Een select kransje van experts leverde aanbevelingen die vervolgens kritisch werden getoetst². Uiteindelijk werden in Laken tien 'primaire' kengetallen naar voren geschoven.

In principe objectiveren zulke naakte cijfers het discours over wederzijds leren, onderlinge vergelijking en gezamenlijke vooruitgang in de strijd tegen armoede. Voor sommigen is de meerwaarde van dit 'sociale scorebord' dat politici precies hierdoor worden aangezet om metterdaad werk te maken van sociale insluiting. Dwingende richtlijnen, laat staan Europese sancties, ontbreken immers in dit gevoelige

beleidsdomein. Tot nader order is het scorebord - overigens niet toevallig een politiek weinig gesmaakte term - het minst zachte element in de open coördinatie rond armoede-beleid.

Dat resultatenvergelijking (politieke) agenten kan aansporen tot een betere prestatie is voor economen geen nieuws. Zulk mechanisme vooronderstelt echter dat er een zekere dreiging uitgaat van de principaal annex volmachtgever, zijnde de kiezers. Dit vergt dat er genoeg ruchtbaarheid wordt gegeven aan het hele proces, maar net dat ontbreekt volgens sommigen. Na het verschijnen van de eerste actieplannen merkte het European Anti-Poverty Network op dat er omzeggens geen publieke aandacht was voor het gebeuren, zelfs niet in de nationale parlementen, en dat de eerste rapporten veeleer 'statistiekjes voor Brussel' dan concrete beleidsinstrumenten leken.

Samengestelde indicatoren

Hier hanteren we de normalisatiemethode die ook de Verenigde Naties toepassen bij hun Human Development Index. Wanneer een lagere waarde voor een deelindicator x staat voor een betere prestatie - bijvoorbeeld het aandeel van vroegtijdige schoolverlaters -, herschrijven we die als $(x_{\max} - x) / (x_{\max} - x_{\min})$, waarbij x_{\max} en x_{\min} de hoogste respectievelijk laagste geobserveerde waarde zijn over alle landen in de steekproef. Waar 'meer beter is', gebruiken we de transformatie $(x - x_{\min}) / (x_{\max} - x_{\min})$. Telkens bekomt men een getal tussen 0 (de slechtste geobserveerde prestatie) en 1 (de 'best practice' voor deze dimensie). Conform de politieke argumentatie kijken we dus naar de prestatie binnen de steekproef. Noodzakelijk is dat niet: men kan de grenswaarden makkelijk exogeen maken. Door deze transformatie omzeilt men mogelijke problemen met betrekking tot aggregatie van oorspronkelijk verschillende meeteenheden en blijven intensiteitverschillen in de relatieve prestaties behouden. Dat zou niet zo zijn wanneer men bijvoorbeeld werkt met rangordes.

Eén cijfer?

Het scorebord objectiveert wel een en ander, maar leent zich moeilijk tot een snelle globale positionering van de lidstaten. Gewoonlijk doet land a het op bepaalde vlakken beter dan b en op andere dimensies niet. Een vergelijking met land c levert weer een ander beeld op. Een vergelijkingsoefening zou aan doorzichtigheid winnen als er slechts 1 of 2 leidende indicatoren zouden zijn, maar dat hypothekeert dan weer danig de kernidee dat sociale uitsluiting multidimensioneel is.

Een mogelijke oplossing voor de afruil tussen transparantie en de veelheid aan dimensies bestaat erin die dimensies te combineren tot één synthetisch kengetal. Dat is niet zo esoterisch: op andere beleidsdomeinen die in Lissabon aandacht kregen, bijvoorbeeld de kenniseconomie, gebruikt de Europese Commissie nu reeds zulke samengestelde indicatoren.

Samengestelde indicatoren zijn gebaseerd op subindicatoren die doorgaans een verschillende meeteenheid hebben en waarbij een voor de hand liggende wijze van gezamenlijke weging ontbreekt. Met andere woorden: hoe aggregeren we langdurige werkloosheid, levensverwachting, armoede en andere indicatoren? Gewoonlijk zijn er twee stappen. Eerst worden de oorspronkelijke getallen genormaliseerd (zie kader). Dat is de minst controversiële stap. Vervolgens worden de genormaliseerde deelindicatoren gewogen en geaggregeerd.

Dat net bij Europese sociale insluiting de idee om indicatoren te aggregeren op weerstand stuit, behoeft nauwelijks betoog. Naast principiële politieke weerzin, die trouwens verklaarbaar is vanuit een principaal-agentperspectief, zijn er pragmatische tegenwerpingen die precies gericht zijn op de moeilijkheid om een universeel aanvaardbaar wegingschema te vinden in een context waar de eigenheid van elke lidstaat van den beginne af zo'n zwaarwichtige rol speelt.

Schaduwrijzen

Gelijke weging van elke component lijkt hier allesbehalve opportuun. Bij uitbreiding geldt dat voor alle vooraf en van buiten af opgelegde gewichten. Daarom is een flexibele en endogene weegprocedure verkieslijk.

Met flexibel en endogeen bedoelen we respectievelijk dat het met een welbepaalde component verbonden gewicht voor elke lidstaat kan verschillen en dat de gewichten zodanig worden gekozen dat de uiteindelijke samengestelde indicator van een lidstaat een zo hoog mogelijke waarde aanneemt. Concreet geformuleerd: wie ten opzichte van andere lidstaten goed presteert op het vlak van pakweg inkomensverdeling, mag aan die component een hoog gewicht toekennen. Goede (respectievelijk inferieure) relatieve prestaties in een welbepaalde dimensie worden beschouwd als evidentie voor meer (minder) lokale beleidsaandacht voor die dimensie en wegen daarom meer (minder) door. Met andere woorden: we gunnen elke lidstaat het voordeel van de twijfel. Juist omdat we de feitelijke ordening van nationale beleidsprioriteiten niet kennen, lijkt deze constructiewijze, waarin we geobserveerde prestaties conceptueel als gebleken beleidsvoorkeuren beschouwen, ons het meest gepast.

Een onmiskenbaar voordeel is dat hierdoor veel zure oprispingen van de lidstaten worden vermeden: elke andere gewichtenreeks zou immers een lagere waarde opleveren voor hun synthetische indicator. Ten tweede kan men op een intuïtieve manier aan allround vergelijking tussen lidstaten doen. Want het is niet uitgesloten dat een andere lidstaat, die in verschillende dimensies beter presteert, zelfs met 'jouw' gunstigste gewichten toch een nog hogere score haalt. We sluiten de cirkel door de maximale totaalscore vast te prikken op honderd procent. Ofwel behaalt men dan zelf dat maximum, ofwel is er een ander land dat ondanks de op eigen maat gesneden weging aantoonbaar beter presteert en dan is de globale score slechts een fractie van de honderd procent waarmee het endogeen geïdentificeerde 'benchmarkland' gaat lopen.

Wat is hier in wezen aan de orde? Het hele probleem draait rond een gebrek aan eensgezindheid over de relatieve waardering van de oorspronkelijke dimensies. Wat ontbreekt zijn uniforme ruilverhoudingen. Zo bekeken komt de geschetste procedure neer op het zoeken naar schaduwrijzen die de beleidsprestaties van een land, in het licht van de prestaties van andere landen, kunnen rationaliseren als een optimaal beleid. Die idee gaat ver terug en wordt vandaag bijvoorbeeld toegepast bij de meting van technische efficiëntie, hetgeen niet toevallig een probleem is waarbij marktprijsinformatie eveneens ontbreekt³.

Voorbeeld

We illustreren deze weegprocedure met een eenvoudige toepassing, waarbij we ons beperken tot de aggregatie van zes deelindicatoren waarvoor in alle landen gegevens beschikbaar zijn in hetzelfde basisjaar, namelijk 1999. Het gaat daarbij om grademeters voor de volgende factoren:

- » het aantal mensen dat is blootgesteld aan armoederisico in dat jaar;
- » het aantal mensen dat daaraan over een langere periode is blootgesteld;
- » de armoedekloof, dat is de mate waarin het mediaaninkomen van armen onder de armoeddrempel ligt;
- » de inkomensongelijkheid;
- » het risico op langdurige werkloosheid;
- » de levensverwachting⁴.

Een vooraf en van buitenaf opgelegde, gelijke weging van die componenten geeft de resultaten in de eerste kolom. De tweede kolom bevat de resultaten van een eerste voordeel-van-de-twijfel-benadering. Hierbij leggen we geen enkele beperking op aan de gewichtkeuze. Verschillende landen halen in de tweede kolom honderd procent; anderen ondanks de maximale generositeit niet. België's score impliceert bijvoorbeeld concreet dat de Finse globale beleidsprestatie zelfs naar Belgische maat gewogen ruim vijftien procent beter is. De (niet afgebeelde) 'impliciete' Belgische beleidsgewichten hebben betrekking op drie dimensies: de armoedekloof (meer dan vijftig procent), het armoederisico (ongeveer vijftientig procent) en de levensverwachting (bijna twintig procent). Het voorbeeld suggereert meteen een nadeel van de volledig vrije keuze van gewichten: een land kan letterlijk alle gewicht leggen bij slechts enkele beleidsdimensies. Wie dit strijdig vindt met de geest van het insluitingsbeleid heeft wellicht een punt. De beschreven aggregatieprocedure laat echter toe om bijkomende voorwaarden op te leggen, zodat de basisgedachte van flexibele en endogene weging gekoppeld kan worden aan een (expert-) consensus omtrent het minimale, of zelfs het relatieve belang van elke dimensie, zonder daarom specifieke gewichtswaarden vast te leggen. Uiteraard impliceert dat minder generositeit en bijgevolg lagere scores. Kolom drie, het resultaat van een oefening waarbij elk land aan elke beleidsdimensie minstens 'een tiende van zijn volle aandacht' moet schenken, illustreert één van vele mogelijke variaties op dat idee.

En verder...

De beschreven methode is ook toepasbaar op tijdreeksen of panelgegevens, waardoor de beleidsdynamiek op het vlak van sociale insluiting met een oogopslag kan worden gevolgd. Verder kan men, indien gewenst, sociale beleidsinputs uitdrukkelijk mee in de analyse te betrekken, en zo bijvoorbeeld nagaan wie de beste verhouding kan voorleggen tussen aangewende middelen en een samengestelde outputindicator. Natuurlijk zijn al die oefeningen gebaat bij sensitiviteitsanalyse met betrekking tot gewichtsbepalingen, opgenomen beleidsdimensies, gegevenswaarden, enzovoort, waarvoor geëigende technieken bestaan⁵. Hoe robuuster het eindresultaat, hoe beter natuurlijk.

Dat resultaat moet niet overgewaardeerd worden: het vervangt een grondige analyse van het beleid niet. Het wil daar wel een conceptueel eenvoudige samenvatting van geven die, rekening houdend met de individualiteit van de geëvalueerde lidstaten, precies een aanzet kan geven tot verhoogde transparantie en zo misschien tot een meer verbeterde strijd tegen armoede en sociale uitsluiting.

Laurens Cherchye, Willem Moesen en Tom Van Puyenbroeck

[tabel 1](#)

weegprocedure	exogeena	endogeen (onbeperkt)	endogeenb
Zweden	100	100	100
Finland	80,1	100	96,9
Denemarken	79,0	100	91,8
Luxemburg	76,4	100	97,8
Oostenrijk	76,3	100	92,1
Nederland	81,0	99,6	92,5
Duitsland	70,2	87,5	77,1
België	59,8	84,2	76,0
Frankrijk	59,7	94,1	75,8
VK	44,3	86,7	65,5
Ierland	40,1	69,4	56,5
Italië	33,2	87,5	54,8
Spanje	31,2	87,5	53,8
Portugal	20,5	85,5	50,6
Griekenland	16,4	75,0	39,7

a. gelijk: elke dimensie eenzesde

b. minimaal tien procent voor elke dimensie

1 Voor een uitvoeriger bespreking van de methode verwijzen we naar F. Vandenbroucke, Samen werken aan duurzame rechtvaardigheid, *ESB*, 12 oktober 2001, blz. 776-778.

2 A. Atkinson e.a., *Social indicators: the EU and social inclusion*, Oxford University Press, 2002.

3 Zie bijvoorbeeld de enorme literatuur omtrent 'data envelopment analysis'. Een vroege toepassing hiervan is te vinden in: W. Melyn en W. Moesen, *Towards a synthetic indicator for macroeconomic performance*, CES Public Economics research paper, Leuven, 1991.

4 Meer informatie over deze en de andere indicatoren staat in Eurostat, Poverty and social exclusion in the EU after Laeken, part I and II, reeks *Statistics in Focus*, Brussel, 2003.