

HOE DE ARBEIDSMARKT EVALUEREN OM ER LERING UIT TE TREKKEN: BOUWSTENEN VOOR EEN BELEIDSRELEVANT ANALYSEKADER?

Verslag van een literatuurstudie

Walter Van Trier (HIVA)

Sinds het midden van de jaren '70 wordt de Westerse geïndustrialiseerde wereld getroffen door een massale en aanhoudende werkloosheid. Zo aanhoudend en zo massaal dat de bekende econoom en Nobelprijswinnaar Franco Modigliani ooit (en met een ironische verwijzing naar de theorie van de natuurlijke graad van werkloosheid - dé dominerende visie onder mainstream economen) sprak van een 'beschamende graad van werkloosheid'.

Om verdere uitdeining van de werkloosheid te voorkomen en om de negatieve sociale effecten van de huidige situatie te bestrijden, hebben beleidsinstanties op lokaal, regionaal, nationaal en internationaal vlak gedurende de voorbije kwart eeuw talloze initiatieven opgezet en instrumenten ontwikkeld. Het meest in het oog springen zeker de inspanningen die men heeft gedaan op het vlak van wat men het 'actief' arbeidsmarktbeleid noemt.

Studiewerk over de vraag of het arbeidsmarktbeleid heeft gewerkt, waarom het al dan niet heeft gewerkt, op welke domeinen het wel en op welke domeinen het eventueel niet heeft gewerkt - dit alles is in twee opzichten van eminent belang.

In de eerste plaats krijgen de overheden zelf door goede evaluaties van het arbeidsmarktbeleid een beter zicht op de mate waarin de aanzienlijke geldstromen, die met een arbeidsmarktbeleid gemoeid zijn, efficiënt en effectief zijn aangewend. Ze kunnen een richtsnoer vormen voor toekomstig beleid omdat ze aangeven welke maatregelen beter de gestelde objectieven halen.

Maar de beleidsvoerders zijn niet de enigen, die voordeel halen uit studiewerk over de werking van het arbeidsmarktbeleid. Zeker in een tijdperk, waarin de arbeidsmarkt aanzienlijke verschuivingen ondergaat ten gevolge van globalisering, technologische ontwikkelingen en institutionele herschikkingen, vormt het arbeidsmarktbeleid een soort laboratorium dat wetenschappers een uitstekende waarnemingspost geeft om processen van verandering te bestuderen en gegevens te verzamelen die anders moeilijk waarneembaar zouden zijn.

Gezien dit grote belang van de resultaten van het arbeidsmarktbeleid en de groeiende potentiële vraag naar evaluaties, zowel vanuit de wetenschappelijke wereld als vanuit de beleidswereld, wekt het enige verwondering te moeten vaststellen dat de literatuuroverzichten vrij eensluidend concluderen dat er op dit terrein al bij al weinig goede studies ter beschikking staan en dat dit soort oefeningen weinig tastbare en bruikbare resultaten oplevert.

Een overzicht van de Vlaamse studies, die men als 'evaluaties van het arbeidsmarktbeleid' kan catalogeren, bevestigt deze algemene vaststelling en leidt tot de volgende meer specifieke, grote conclusies:

1. Hoewel het arbeidsmarktbeleid in Vlaanderen/België de laatste twee decennia gekenmerkt is door een continue stroom van initiatieven, zowel op gewestelijk als op federaal vlak is er nauwelijks sprake van een aangehouden inspanning om de verschillende instrumenten van het beleid te evalueren. Kennelijk werden oude maatregelen ingetrokken of aangepast, en nieuwe maatregelen veelal ingevoerd zonder dat een 'externe' evaluatie voorhanden was of dat de mogelijkheden ertoe werden gebruikt.
2. Van de studies die zijn gebeurd dringen de resultaten weinig door in de gekende vaktijdschriften, zoals het Tijdschrift voor Economie en Management, de Recherches Economiques de Louvain of het Tijdschrift voor Sociologie. Enkel de Cahiers Economiques de Bruxelles heeft een zekere traditie in het publiceren van artikels, die

het arbeidsmarktbeleid evalueren. Noch bij sociologen noch bij economen kan men derhalve spreken van een 'peer-review' van het bestaande materiaal.

3. De bestaande evaluatiestudies behoren nagenoeg allen tot het type 'impactstudie'. Ze meten op één of andere manier het resultaat van een bepaald programma. Hoe het beleid vorm krijgt en op welke manier de verschillende maatregelen werden geïmplementeerd blijft tot hiertoe buiten het evaluatieproces. Dit is spijtig want in de internationale literatuur wijst men er geregeld op dat de effectiviteit van een maatregel dikwijls sterk afhangt van factoren, die de invoering van de maatregel bepalen. Vandaar dat men er sterk voor pleit om elementen van 'procesevaluaties' op te nemen, zelfs als het om impactstudies gaat.
4. Niet alleen zijn de bestaande studies bijna uitsluitend impactstudies, de meting van de effecten van de maatregelen centreert zich hoofdzakelijk op, wat we gemakshalve zouden willen noemen, 'administratieve criteria'. Het gaat meestal om een evaluatie op basis van de effectiviteit (meestal gemeten via zogeheten 'schakelresultaten' of doorstromingskansen) en het bereik (ofwel het aantal en de aard van de deelnemers) van bepaalde programma's. Dit betekent dat men slechts een beperkt zicht krijgt op de vier belangrijke criteria waarop men arbeidsmarktmaatregelen kan evalueren:
 - het werkgelegenheidseffect (leidt de maatregel tot meer werkgelegenheid?),
 - het verdelingseffect (verdeelt de maatregel de lasten van de werkloosheid gelijkmatiger?),
 - productiviteitseffect (te meten via kortere duurtijd van vacatures, verminderde mismatch, effecten op de later verdiende lonen, enz.)
 - efficiencyeffect (worden de vooropgestelde doelstellingen bereikt en met welke opportuniteitskost? En welke kosten-effectiviteit?).Ook krijgt men slechts in uitzonderlijke gevallen een zicht op de het netto-impact van de maatregelen - over deadweight-loss, substitutie en verdringing vindt men weinig gegevens.
5. De meeste evaluatiestudies beperken zich tot een 'partiële analyse' in twee betekenissen van deze term. Ze bekijken de individuele maatregelen onafhankelijk van hun relatie tot het bredere institutionele raamwerk. Als het zo is dat 'coördinatie' en 'pakketten maatregelen' een belangrijke rol te spelen hebben in het beleid dan duidt deze invalshoek op een ernstig tekort. De meeste studies gaan trouwens niet uit van een uitgewerkt (of geëxpliciteerd) theoretisch model van de arbeidsmarkt. Enigszins verassend is trouwens dat er (ook bij economen) weinig gebruik gemaakt wordt van de veelheid aan modellen om beleidsmaatregelen ex-ante theoretisch door te lichten.
6. Er zijn (vooralsnog) geen evaluatiestudies die de effecten van beleidsmaatregelen op langere termijn bekijken. De meeste studies bekijken de resultaten van een maatregel binnen één à twee jaar na het opstarten. Op basis van het bestaande materiaal is het derhalve onmogelijk om na te gaan hoe lang de effecten van bepaalde programma's blijven doorwerken.

Deze vaststellingen zijn ten dele het gevolg van de specifieke toestand van het arbeidseconomisch onderzoek in Vlaanderen, maar voor de meest essentiële aspecten van deze beoordeling gaat het om problemen die zich op een breder internationaal vlak ook voordoen.

Wat daar vooral opvalt en dit wordt opgemerkt door de meeste waarnemers die overzichten maken van de internationale literatuur inzake evaluaties van het arbeidsmarktbeleid is het grote gebrek aan consistente en robuuste resultaten.

Zowel voor beleidsmakers als voor wetenschappers volgen hieruit minstens drie algemene conclusies. Vooreerst, empirische resultaten van deze studies dient men met nog meer zorg te interpreteren dan resultaten van studies in andere domeinen. Bovendien, de belangrijkste vraag lijkt niet te zijn of een (actief) arbeidsmarktbeleid werkt, maar wel: welk (actief) arbeidsmarktbeleid werkt?, voor wie?, onder welke omstandigheden? En dan rest nog de vraag, die vanuit het oogpunt van de overdraagbaarheid van het geëvalueerde beleid wellicht het meeste belang heeft: weet men wel waarom iets werkt? Ten slotte, een belangrijke kwestie die nader te onderzoeken valt, is waarom de resultaten van dit onderzoekswerk niet tot meer bevredigende resultaten leiden.

Het uitgangspunt van deze studie is dat deze zorgelijke situatie noopt tot een grondige en diepgaande reflectie op de wijze waarop, de visie van waaruit en de condities waarin evaluaties van het arbeidsmarktbeleid in het recente verleden werden en nog steeds worden doorgevoerd.

Een overzicht van de literatuur maakt duidelijk dat het evalueren van arbeidsmarktmaatregelen op zich een moeilijke en al bij al nog een jonge onderneming is. De basisartikels zijn immers pas geschreven begin de jaren '90. Dat de meeste inspanningen op dit vlak van erg recente datum zijn, zou dus kunnen verklaren waarom de resultaten van deze onderzoeken vooralsnog erg beperkt blijven.

Maar naar onze mening is dit slechts gedeeltelijk het geval.

De belangrijke oorzaak voor het gebrek aan robuuste en consistente resultaten lijkt ons de eerder de werkwijze die men doorgaans volgt bij evaluaties van het beleid op het terrein van de arbeidsmarkt. Meer in het bijzonder lijkt het zo te zijn dat zich in toenemende mate een discrepantie ontwikkelt tussen de visie die schuil gaat achter de meeste evaluatiestudies en de diagnose die zich meer en meer opdringt als belangrijkste verklaring voor de Europese werkloosheid.

Omdat deze hypothese zo belangrijk is, willen we hier kort wat dieper op ingaan.

Er lijkt zich onder economen een consensus te vormen over het feit dat men in de huidige diagnoses van de werkloosheid een centrale plaats moet geven aan twee soorten problemen, die men respectievelijk kan aanduiden met de termen 'coördinatie-mechanismen' en 'transformationele groei'.

Een diagnose die wijst op het belang van de coördinatie tussen de bestaande instituties, regelgeving of beleidsmaatregelen zegt in feite dat het resultaat van beleidsingrepen onvermijdelijk afhangt van de context, waarin ze worden doorgevoerd en van de synergie tussen de nieuwe maatregelen en andere onderdelen van het bestaande beleidsinstrumentarium.

Een diagnose die een rol toekent aan processen van transformationele groei, houdt rekening met het feit dat groei en ontwikkeling steeds meebrengen dat er zich niet enkel kwantitatieve maar ook kwalitatieve wijzigingen kunnen voordoen in de socio-economische toestand en dat een goede coördinatie tussen de verschillende onderdelen van de regelgeving derhalve niet blijvend gegarandeerd kan worden.

Een belangrijk gevolg van dit alles is, bijvoorbeeld, dat het tijdsplan en de timing, gevolgd bij de invoering van maatregelen, of de manier waarop de nieuwe maatregel zich (al dan niet) inpast in de bestaande structuren van groot belang kan zijn op de effectiviteit en de efficiëntie van een maatregel. Welnu, als dit het geval is dan is het niet onmogelijk dat de evaluatie van op zich staande maatregelen een verkeerd en misleidend beeld geeft van het

effect van deze maatregelen. Anders gezegd: in een context waarin de samenhang tussen maatregelen en hun geschiedenis van belang is om de werking er van te begrijpen is het erg waarschijnlijk dat de resultaten van een bepaalde maatregel ook de effecten van deze samenhang en dit tijdspad weerspiegelen.

Deze redenering maakt het begrijpelijk waarom Günther Schmid en Klaus Schömann uit een uitgebreid overzicht van de beschikbare socio-economische evaluaties van het arbeidsmarktbeleid de conclusie moeten trekken:

- a. dat het onmogelijk is om één eenduidig succesverhaal te vertellen; er bestaan er meerdere afhankelijk van welk criterium men in de evaluatie doet doorwegen;
- b. dat men onmogelijk een bepaald institutioneel model als garantie voor succes kan aanwijzen; het geheim van het succes lijkt te zitten in een institutioneel raamwerk dat de juiste incentieven geeft opdat de verschillende centrale actoren op de arbeidsmarkt en in het werkgelegenheidsbeleid bereid zijn om op een competente manier samen te werken.

Onze hypothese over het belang van 'coördinatie' en 'transformationele groei' maken deze diagnose van Schmit en Schömann inzichtelijk en kan ook verklaren waarom dat de meeste evaluaties van arbeidsmarktbeleid tot inconsistente en weinig robuuste resultaten leiden.

Het effect van individuele maatregelen en zelfs van volledige programma's hangt teveel af van de context, waarin ze ingevoerd worden om eenduidig te zijn. Meer zelfs, ook als men voor een bepaald type maatregelen in een bepaald land robuuste en significante resultaten zou vinden, is het lang niet zeker dat deze maatregelen in een andere context dezelfde resultaten zou opleveren of op een zelfde (positieve) manier geëvalueerd worden.

In zekere zin kan men configuraties van instellingen beschouwen als functionele equivalenten en zelfs als het zo is dat de huidige samenstelling ervan voor een aantal landen erg identiek lijkt, kan het vervangen van één van de elementen van dit geheel meebrengen dat men dit op een verschillende manier moet aanpakken, wil het nieuwe geheel efficiënte coördinatie mogelijk maken.

Laat ons het belang van deze redenering met een voorbeeld illustreren.

Stel dat men in twee landen een wijziging wil aanbrengen aan het niveau van het minimumloon. Stel bovendien dat het huidige minimumloon in beide landen hetzelfde niveau heeft en dat de situaties verder in alle andere relevante karakteristieken identiek zijn, zodat men ze enkel kan onderscheiden in de manier waarop ze historisch tot dit punt in de tijd geëvolueerd zijn. Als dit zo is zo kan men aantonen dan garandeert een gelijke wijziging in het minimumloon niet dat de effecten van deze maatregel in de beide landen identiek zijn.

Daartegenover staat dat de dominerende traditie in het evaluatie-onderzoek nauwelijks aandacht heeft voor deze twee aspecten - het tijdspad en de institutionele inbedding van beleidsmaatregelen. Meestal bekijkt en evalueert men maatregelen als op zichzelf staande en wordt de wijze waarop de maatregel vorm krijgt buiten beschouwing gelaten.

Om een beter zicht te krijgen op de redenen waarom deze toestand zich voordoet, brengen wij in een eerste deel van dit rapport een overzicht van de manier waarop conventioneel evaluaties van het arbeidsmarktbeleid worden geproduceerd.

We overlopen achtereenvolgens de belangrijkste onderzoekstradities (macro-econometrische modelbouw, studies gebaseerd op experimenten, micro-econometrische werkwijze) en geven de moeilijkheden aan, waarmee elk van deze tradities te kampen heeft. We geven ook aan dat verscheidenheid van resultaten ook kan voorkomen uit het gebruik van verschillende methodologieën. Experimenteel en

quasi-experimenteel onderzoek kunnen immers leiden tot substantieel verschillende conclusies, zelfs als de achterliggende realiteit dezelfde is.

Daarna bekijken we welk model standaard (Beveridge-curve en Layard-Nickell model) wordt gebruikt als leidraad bij de interpretatie van de resultaten en geven we aan hoe dit model slechts een deel van de economische realiteit als relevant beschouwt voor het verklaren van de werking van de arbeidsmarkt.

Ten slotte staan we stil bij de criteria (economische, sociale of ethische criteria) die men meestal bij de evaluatie van arbeidsmarktmaatregelen hanteert.

Globaal leidt dit tot het volgende beeld.

De term 'beleidsevaluatie' verwijst in het gewone spraakgebruik naar een erg breed veld van onderzoeksactiviteiten, dat zowel beleidsformulering en -uitwerking, beleidsimplementatie en 'impact-assessment' omvat. Deze uiteenlopende disciplines zijn erg verschillend en hoewel dikwijls complementair vrij heterogeen. Toch kan men stellen dat de kern van al deze activiteiten gevormd wordt door 'impact-assessment'. Dat men als de harde kern van beleidsevaluatie kan beschouwen - beleidsuitwerking en implementatie spelen in deze visie slechts een rol in de mate dat ze een bijdrage leveren tot de verklaring van de differentiële impact van verschillende beleidscontexten en beleidsopties.

Traditioneel werden evaluaties van arbeidsmarktbeleid gekenmerkt door een sterke focus op individuele beleidsinstrumenten en programma's. De impact van deze op enkelvoudige doelstellingen gerichte programma's werd op een top-down manier gemeten door de expliciete doelstellingen van deze programma's te vergelijken met meetbare uitkomsten van het programma.

In deze conventionele aanpak werden beleidsuitwerking en implementatie in hoge mate beschouwd als een zwarte doos. Verwachte effecten werden gemeten in monetaire termen - bijvoorbeeld bij middel van hogere lonen of inkomens en vervolgens vergeleken met de kosten gemaakt door het programma.

Binnen dit model zijn de onderliggende regels van besluitvorming erg simpel. Als de verhouding van de kosten en de baten kleiner is dan of als de balans negatief is beoordeeld men het programma als niet waardevol. Als deze verhouding groter was of deze balans positief voor een programma A in vergelijking met programma's B en C dan werd de voorkeur gegeven aan A.

Omwillen van dit zwarte doos aspect verwaarloosden programma-evaluaties veelal zowel de incentieven die de interactie van verscheidene beleidsinterventies teweeg bracht als het cumulatieve effect ervan op de houdingen en het gedrag van de relevante actoren in het betreffende domein.

Deze tekortkoming verklaart mogelijk waarom de meeste programma-georiënteerde evaluatiestudies relatief ontgoochelende resultaten opleverden met betrekking tot de echte beleidsimpact op arbeidsmarktgedrag en resultaten.

Door de banden en de interacties tussen specifieke beleidsmaatregelen zowel als de inbedding in het institutionele regime van het specifieke doelgebied te verwaarlozen, faalden de conventionele aanpak van evaluaties erin om begrippen te ontwikkelen inzake complementariteit, wederzijdse neutralisering en functionele equivalentie of substitueerbaarheid van verschillende beleidsmaatregelen en programma's.

Omdat deze conventionele werkwijze problemen als 'coördinatie' of 'transformationele groei' nauwelijks onderkennen en er in hun methodologie bijgevolg geen rekening mee houden, hebben we in een derde deel van dit rapport getracht een eerste synthese te brengen van een recente stroming in het evaluatieonderzoek, die uitdrukkelijk probeert om deze uitdaging aan te nemen.

Onze leidraad hierbij is het werk van de Günther Schmid en Karl Schömann, twee van de editors van het recent gepubliceerde *International Handbook of Labour Market Policy and Evaluation*.

Dit duidelijke gebrek aan een institutionele invalshoek in de evaluatiepraktijk is in wezen erg verwonderlijk.

Als het de pragmatische bedoeling is van evaluaties van het arbeidsmarktbeleid om te leren welke maatregel of ingreep beter werkt dan een andere dan moeten evaluaties op het niveau van instituties ook vanuit beleidsoogpunt erg relevant zijn, zeker als theoretische inzichten ons meer en meer wijzen op het belang van 'institutional interlock' en 'hysteresis'.

De nood om precies deze banden, interacties en cumulatieve effecten van arbeidsmarktmaatregelen te conceptualiseren is des te meer nodig en urgent in een tijdperk van geïntensiveerde transnationale beleidscoördinatie. In toenemende mate betonen beleidsmakers interesse voor comparatief evaluatieonderzoek. Zij verwachten hieruit te kunnen leren welke beleidsmaatregelen hen toelaten met het meeste succes bepaalde doelstellingen te bereiken.

Hoe komt het, kan men zich terecht afvragen, dat de meeste evaluatiestudies aan deze boodschap weinig waarde lijken te hechten? Of beter: hoe komt het dat evaluatieonderzoekers en wellicht vooral hun financiers voor het overgrote deel vasthouden aan onderzoeksstrategieën, die de effecten van het beleid bekijken alsof het gaat om de gevolgen van aparte en geïsoleerde maatregelen?

Net alsof men zich bij het werpen van een grote steen in de kikkerpoel enkel zorgen moet maken over hoe groot het risico is dat hij de kikker op de bodem raakt en niet over de eventueel verwoestende gevolgen die de erdoor verwekte deining heeft op het geheel van de flora en fauna, waarin kikkers verblijven en die hun overlevingskansen terdege beïnvloeden.

Een eerste verklaring voor deze veronachtzaming van instituties en wat ermee samenhangt is wellicht dat economen en andere sociale wetenschappers in hun modellen lijken te negeren dat de keuze van instituties in belangrijke mate een investeringsaspect heeft en dat bepaalde institutionele patronen omwille van hun samenstelling meer maatschappelijk productief zijn dan andere. Met andere woorden: eens de keuze gemaakt en een bepaalde maatregel doorgevoerd is, valt deze keuze immers meestal erg moeilijk terug te draaien; en als men bij deze keuze geen rekening houdt met de effecten van het geheel van de regelgeving dan zou men op globaal vlak best verstorend kunnen werken, zelfs als men op elementair vlak betere resultaten verkrijgt.

Wij zijn er ons van bewust dat we met dit rapport slechts een beperkte en initiële bijdrage kunnen leveren tot een kentering. Maar zoals voor vele andere kwesties, geldt naar onze mening ook hier de stelling dat oplossingen gemakkelijk(er) worden als de vragen correct gesteld worden - en bijdragen tot het correcter formuleren van de vragen is voorlopig alles wat we hier beogen. Daarom was in dit rapport ook niet in de eerste plaats aan de orde welke resultaten de beschikbare evaluatiestudies opleveren. Wat de beschikbare

evaluaties van het arbeidsmarktbeleid ons feitelijk leren is in dit opzicht bijzaak. Visie-ontwikkeling, zo lijkt ons, is wat men vandaag vooral ontbeert op het terrein van het arbeidsmarktbeleid - ook en vooral in de manier waarop men het beleid evalueert.