



Het meten van weten

Sinds een aantal jaren wordt het economenvolk via *ESB* en *Intermediair* onrustig gemaakt met publikaties van de economen-'top 30' en -'top 40'. Gaat het hier om een hogere vorm van zelfbevrediging (en leedvermaak!) of maakt het ons echt iets wijzer over de opbrengst van wetenschappelijk onderzoek? Om hierop een antwoord te kunnen geven, wordt in dit artikel een aantal methoden om de output van economen en economisch onderzoek vast te stellen met elkaar vergeleken en geplaatst in het ruimere kader van de groeiende behoefte aan maatstaven om de produktie en de waarde van het universitaire bedrijf te evalueren.

Een van de terreinen waarop - vooral door de invoering van de z.g. voorwaardelijke financiering 1) - steeds meer wordt gedaan aan outputmeting is het uit de collectieve middelen gefinancierde universitaire onderzoek. Een adviescommissie 2) adviseerde vorig jaar zelfs aan de minister van Onderwijs en Wetenschappen om waar mogelijk output als grondslag voor de financiering van het wetenschappelijk onderzoek te gaan hanteren. Ook in de recente ministeriële beleidsnota *Hoger onderwijs; autonomie en kwaliteit* 3), wordt aangedrongen op de ontwikkeling van 'performance indicators' als noodzakelijk hulpmiddel in de evaluaties van onderzoek en onderwijs van en door instellingen. Dat vereist tenminste enige consensus over de maatstaf voor die 'performance'. We zullen de volgende vier ontwikkelingen op het gebied van outputmeting van economisch onderzoek beoordelen:

- de 'economen-top 40', gebaseerd op de produktie van individuen, gemeten aan de hand van artikelen;
- de 'economen-top 30', eveneens gebaseerd op de produktie van individuen, maar gemeten aan de hand van citaten;
- het rapport van de economische faculteit van de Rijks Universiteit Groningen, waarin de produktiviteit van grote onderzoekclusters wordt gemeten aan de hand van een groot aantal publikatiecategorieën; en
- het rapport van de Verkenningcommissie Economische Wetenschappen, waarin de produktie van onderzoeksgroepen wordt gemeten aan de hand van publikaties (artikelen, boeken, dissertaties).

Om het methodologische probleem zo zuiver mogelijk te stellen laten we de pikante vraag hoe verschillende onderzoekers, onderzoeksgroepen of faculteiten 'scoren' in deze verschillende systemen hier graag rusten. De dis-

cussie over deze vraag kan het best worden samengevat met het woord van de dichter: „Beauty is in the eye of the beholder”.

De top 40

In 1980 werd de eerste economen top 40 gepubliceerd 4), eerst in *ESB*, daarna jaarlijks in *Intermediair*. De top-40 wordt samengesteld door A.D.S. de Schuite, nom de plume van twee vooraanstaande Nederlandse economen. De samenstelling is gebaseerd op het idee dat de produktiviteit van economen blijkt uit het aantal, de omvang en de kwaliteit van hun artikelen. Boeken zijn niet opgenomen omdat de kwaliteit ervan moeilijk objectief valt vast te stellen (althans moeilijker dan bij tijdschriften). Bovendien worden nieuwe ideeën doorgaans eerst in de vorm van een wetenschappelijk artikel gepubliceerd en pas later eventueel in boekvorm uitgebreid herhaald.

Hoe wordt dan de kwaliteit van artikelen gemeten? De Schuite benadert de kwaliteit door de bladen waarin ze verschijnen te rangschikken in een drietal categorieën: 'zeer goede tijdschriften', 'uitstekende tijdschriften' en 'toptijdschriften'. Alleen internationale tijdschriften worden in beschouwing genomen. De reden hiervoor is dat alleen bijdragen aan de internationale literatuur wetenschappelijk gezien betekenis hebben. „Bijdragen in Nederlandse tijdschriften hebben dan ook vaak het karakter van een bijdrage aan een politieke discussie, van oplossing van een praktisch probleem, of van een inleidend overzicht van nieuwe internationale ontwikkelingen op het vakgebied” 5). Uit oogpunt van kwaliteitsbeoordeling is verder van belang dat deze bladen een strikt proces van 'refereering' hanteren. Een aangeboden bijdrage wordt alleen voor publikatie geaccepteerd als de 'referees' (beoordelaars) het goed genoeg vinden. Het ene blad is daarbij selectiever dan het ander. Dat bepaalt de plaats van het blad in één van de drie kwaliteitscategorieën.

Het systeem werkt verder als volgt. Auteurs van artikelen in deze tijdschriften verdienen met hun publikaties punten. Die punten worden opgeteld over een periode van vijf jaar. Per publikatie krijgt een auteur vijf punten plus één punt per bladzijde (in het geval van een grote bladspiegel anderhalf punt per bladzijde). Dit getal wordt vervolgens vermenigvuldigd met twee als het om een 'zeer goed tijdschrift' gaat, met

drie als het een 'uitstekend tijdschrift' betreft, en met vier als het een 'toptijdschrift' is.

Ten slotte vindt nog een correctie plaats voor het aantal auteurs van een artikel. Als een artikel n auteurs heeft

krijgt elke auteur $\frac{2}{n+1}$ de deel van de punten.

De top 30

Sinds 1981 publiceert *ESB* jaarlijks een economen top-30. De eerste jaren werd deze verzorgd door J.J. van Duijn. De recentste top-30 werd opgesteld door T.M. Buitendorp en M.A. Langman. Evenals bij de top-40 gaat het hierbij om een rangschikking van personen. De top 30 is niet gebaseerd op publikaties, maar op citaten. Daar ligt de volgende redenering aan ten grondslag. Om bij te dragen aan de publieke kennis zal een onderzoeker moeten publiceren. Dat is één ding; of andere onderzoekers uit die publikatie waardevolle kennis en inzichten putten en daar in eigen geschriften blijik van geven, is een tweede. Dit laatste bepaalt volgens de redenering van Van

1) Zie C.A. Hazeu, Voorwaardelijke financiering van onderzoek, *ESB*, 16 februari 1983, blz. 163-165.

2) Commissie Financieringsstructuur Onderzoek en Ontwikkeling, *Rekenschap van creativiteit* (eindrapport), 's-Gravenhage, september 1985 en Werkgroep Bedrijfs-economisch Beleid en Beheer van de Commissie Financieringsstructuur Onderzoek en Ontwikkeling, *Ins & Outs. Achtergrondstudie over bedrijfs-economisch beleid en beheer van onderzoeksorganisaties bij input-, throughput- en output financiering*, 's-Gravenhage, juni 1985. Zie ook D.J. Wolfson, *Rekenschap van creativiteit*, *ESB*, 4 september 1985, blz. 867.

3) Ministerie van O en W, *Hoger onderwijs; autonomie en kwaliteit. Een andere bestuurswijze*, 16 oktober 1985, en W.J. Deetman, *Doelmatigheid en hoger onderwijs vanuit politiek perspectief*, in: P.B. Boorsma en J.B.J. Koelman, *Doelmatigheid in het hoger onderwijs*, 's-Gravenhage, 1986, blz. 3-12.

4) *ESB*, 17 december 1980, blz. 1414-1415, *Intermediair*, 21 mei 1982, blz. 11-13, 16 december 1983, blz. 13-17, 14 december 1984, blz. 27-29 en 27 december 1985, blz. 23-25.

5) A.D.S. de Schuite, *Top-40 der Nederlandse economen*, *Intermediair*, 21 mei 1982, blz. 11-13.

6) *ESB*, 18 november 1981, blz. 1143-1145, 15 december 1982, blz. 1338-1340, 14 december 1983, blz. 1158-1159, 12 december 1984, blz. 1174-1175 en 18/25 december 1985, blz. 1305-1307.

Duijn de waarde van een publikatie 7).

Hoe meer citaten dus, hoe beter het onderzoek of de onderzoeker. Een ranglijst moet dus gebaseerd zijn op aantallen citaten. Daarvoor maakt hij gebruik van de *Social Sciences Citation Index* (SSCI), die wordt uitgegeven door het Amerikaanse Institute for Scientific Information. De SSCI omvat alle maatschappijwetenschappen. Wat de economische discipline betreft wordt van 150 tijdschriften (waaronder drie Nederlandse, t.w. *De Economist*, *het Tijdschrift voor Economische en Sociale Geografie* en de *European Economic Review*) alle in artikelen genoemde referenties genoteerd. Verder wordt nog een beperkt aantal boeken nagetrokken, als regel congresbundels.

Bij de samenstelling van zijn ranglijst heeft Van Duijn nog een aantal additionele regels toegepast:

- alleen in Nederland opererende Nederlandse economen komen in aanmerking;
- boekbesprekingen, die de SSCI wel opneemt, doen niet mee;
- verwijzingen naar eigen werk tellen niet mee; verwijzingen van onderzoekers die elkaar binnen een kring veelvuldig citeren zijn daarentegen niet uitgesloten;
- een beperking tot de algemene economie, incl. de econometrie en de wiskundige economie; de bedrijfseconomie is buiten beschouwing gelaten 8).

Een bezwaar van de SSCI, waarvoor Van Duijn niet corrigeert, is dat bij publikaties door meer dan een auteur alleen de eerste geïndexeerd wordt.

Het Groningse rapport

Anticiperend op het rapport van de Verkenningcommissie Economische Wetenschappen is onlangs onder auspiciën van het bestuur van de economische faculteit van de Rijks Universiteit Groningen door Szirmai en Corsel 9) een produktiviteitsstudie verricht. De onderzoekoutput van de verschillende economische faculteiten is geïnventariseerd op basis van publikaties. Bron vormen de wetenschappelijke verslagen van de instellingen. De in beschouwing genomen periode is 1980-1983. De volgende acht publicatiecategoriën en bijbehorende wegingsfactoren zijn onderscheiden:

1. Dissertaties	
- intern	5
- elders	4
- extern	1
2. Boeken	4
3. Geredigeerde bundels	1
4. Herdrukken	1
5. Artikelen	1
6. Rapporten	0,5
7. Annotaties	0,5
8. Vakpublicaties	0,25

Opmerkelijk is dat de artikelen niet

nader geclassificeerd zijn naar de kwaliteit van het tijdschrift. De output is vervolgens toegerekend aan vier categorieën, t.w.:

- Algemene economie;
- Bedrijfseconomie;
- Econometrie en wiskunde; en
- Overig.

In de categorie 'overig' zit o.a. het onderzoek van de niet-economische vakgroepen.

Verkenningcommissie Economische Wetenschappen

Ging het bij de top-30 en de top-40 om initiatieven van geïnvolveerde individuen, de Verkenningcommissie Economische Wetenschappen heeft een andere achtergrond. Ze werd in 1985 ingesteld door de minister van O en W als uitvloeisel van de Taakverdelings- en Concentratie-exercitie in de economische discipline. De commissie had een driedelige taak:

- a. het beschrijven van de aard en de omvang van het door de overheid gefinancierde wetenschappelijk onderzoek op het gebied van de economie, de econometrie en de actuariële wetenschappen dat wordt verricht aan Nederlandse instellingen van wetenschappelijk onderwijs;
- b. het doen van uitspraken over de kwaliteit, de doelmatigheid en de samenhang van het onderzoek op het gebied van de economische wetenschappen; en
- c. het formuleren van aanbevelingen met het oog op de gewenste ontwikkeling, waarbij voor het onderzoek een profilering per instelling wordt uitgewerkt, rekening houdende met eventuele mogelijkheden van verdeling van taakgebieden.

In het kader van de onder a. genoemde taak was er op het moment dat dit artikel geschreven werd nog slechts een concept-deelrapport 10) beschikbaar. Daarin heeft een inventarisatie van onderzoekoutput plaatsgevonden op basis van publikaties. Daarbij is een ruime publikatiemaatstaf gehanteerd: naast tijdschriftartikelen zijn ook boeken en bijdragen aan boeken, zowel nationaal als internationaal meegenomen. In tegenstelling tot de top-30 en de top-40 wordt de output niet per individu gemeten, maar per instelling (d.w.z. per economische faculteit en daarbinnen per vakgroep; een afzonderlijke beschouwing is gewijd aan de economische vakgroepen van niet-economische faculteiten). De output is verder toegerekend naar onderwerp, waarbij op basis van de *Journal of Economic Literature* een tiental hoofdcategorieën wordt aangehouden. De in beschouwing genomen periode is 1979-1983. De bron wordt gevormd

door de wetenschappelijke verslagen van de instellingen.

De commissie heeft niet alleen aantallen publikaties en aantallen gepubliceerde bladzijden geregistreerd, maar heeft ook gedifferentieerd naar de wetenschappelijke kwaliteit van de geregistreerde publikaties. Wat de tijdschriftartikelen betreft zijn de tijdschriften onderscheiden in een vijftal kwaliteitscategoriën, A t/m E. De hoogte van inschaling van een tijdschrift wordt daarbij bepaald door de mate waarin het geciteerd wordt 11). Ook de Nederlandse tijdschriften zijn ingeschaald, voornamelijk in de twee laagste categorieën. De volgende gewichten zijn gehanteerd:

Categorie	Gewicht
A (voortreffelijk)	4
B (zeer goed)	3
C (goed)	2
D (redelijk)	1
E (van gering wetenschappelijk belang)	1/2

Verder is van belang dat tijdschriften in randgebieden van de economie (wiskunde, 'operations research' e.d.) in categorie D zijn opgenomen. Onderzoek in economische faculteiten in andere disciplines (b.v. sociologie of recht) blijft buiten beschouwing.

In tegenstelling tot de tijdschriftartikelen zijn bij boeken en artikelen in boeken slechts aantallen geregistreerd. (Artikelen in) boeken in de Engelse taal hebben een gewicht van 1,5 gekregen in verhouding tot bijdragen in de Nederlandse (of Duitse of Franse) taal. Verder heeft de commissie de volgende verhouding gehanteerd: vijf artikelen van 20 blz. per artikel in een categorie D-tijdschrift = vijf boekbijdragen in de Nederlandse taal = één boek in de Nederlandse taal. Tenslotte zijn aantallen verdedigde dissertaties afzonderlijk geregistreerd.

Publikaties of citaten?

De belangstelling voor 'performance indicators' in het universitaire wetenschappelijke onderzoek is de afgelopen jaren sterk gegroeid, niet in het minst door de invoering van de voor-

7) Vgl. ook E. Garfield, *Citation indexing - its theory and application in science, technology, and humanities*, New York, 1979.

8) In de laatste top-30 zijn bedrijfseconomen wel meegeteld. Dit was echter nauwelijks van invloed op de ranglijst.

9) A. Szirmai en R.R. Corsel, *Productiviteit van economisch onderzoek aan universitaire instellingen*, Groningen, januari 1986.

10) Verkenningcommissie Economische Wetenschappen, *Concept-rapport van een deel van de door de verkenningcommissie economische wetenschappen verrichte inventarisatie betreffende de periode 1979-1983*, januari 1986.

11) De gehanteerde citatenindexen zijn berekend door S.J. Liebowitz, en J.P. Palmer, *Assessing the relative impacts of economic journals*, *Journal of Economic Literature*, jg.22, nr.1, maart 1984

waardelijke financiering in 1982. Het lijkt daarom nuttig de boven beschreven ontwikkelingen in de economische discipline te beoordelen tegen deze bestuurlijke achtergrond. Thans is de tussentijdse beoordeling (na twee jaar) van de voorwaardelijke financieringsprogramma's aan de orde. Daarbij gaat het dus om het vaststellen van recente output. Vanaf volgend jaar zullen eindbeoordelingen geveld moeten worden. Daarbij is van belang dat de minister na een overgangperiode de mogelijkheid van 'reallocatieve gevolgen' in het vooruitzicht heeft gesteld.

In een recente afspraak tussen minister en de instellingen is besloten dat de externe beoordeling zal uitmonden in een drietal categorieën: + +, + en -. Het reallocatiemechanisme zal er per saldo op neerkomen dat instellingen die relatief veel onderzoek als excellent (+ +) krijgen beoordeeld, er middelen bij kunnen krijgen ten laste van instellingen die minder excellent 'performen'.

Omdat 'excellentie' zich niet laat dwingen lijkt het daarom het meest raadzaam als instellingen en faculteiten hun interne beoordelingsprocessen primair richten op het traceren van hetgeen evident onder de maat gebleven is (de 'bottom-10'). In dat beoordelingsproces kan de prestatie-indicator van (gewogen) aantallen publikaties een bruikbaar hulpmiddel zijn. Aan het hanteren van een effectindicator 12) zoals citatenanalyse kleven daarentegen zowel meetproblemen als meer principiële bezwaren.

De belangrijkste meetproblemen zijn:

- de opgenomen bladen vertegenwoordigen de gevestigde orde; nieuwe tijdschriften, tijdschriften in multidisciplinaire gebieden, 'radical' tijdschriften (die een nieuw of onderliggend paradigma vertegenwoordigen) zijn doorgaans niet opgenomen 13);
- Nederlandse bladen zijn niet opgenomen;
- bij meer dan een auteur wordt alleen de eerste geteld.

Meer principiële bezwaren tegen citaten-analyses zijn:

- ze geven geen indruk van recente output, maar van 'past performance'; leeftijd wordt beloond; het prestige van de gevestigde orde wordt gemeten; ze hebben een conserverende werking;
- citaten zijn veelal niet onderling afhankelijk; er blijken citeernetwerken (in een specialisme en/of rond een bepaald blad) te bestaan, waarin men elkaar over en weer citeert en prestige toedeelt (de 'mutual admiration society'). Buitenstaanders lopen de kans genegeerd te worden;
- een aanhaling hoeft niet per se instemming te betekenen, maar kan ook kritiek inhouden; het is niet allemaal vooruitgang in de wetenschap die wordt gemeten;

- er vindt geen weging plaats van de kwaliteit van de bladen waarin een auteur wordt geciteerd;
- bepaalde modes in onderzoekgebieden kunnen onevenredige aandacht krijgen.

Een bruikbare output-maatstaf

Een publikatiemaatstaf is dus een geschikt hulpmiddel bij de beoordeling van onderzoekprogramma's dan een citatenmaatstaf. Wanneer wordt aangesloten bij de bestaande ontwikkelingen kan er gekozen worden tussen de top-40-aanpak, de Groningse methode en de door de Verkenningcommissie gevolgde methode. Zoals gezegd, kan de beoordeling van de 'performance' van wetenschappelijk onderzoek meer doelen dienen. Als evaluatie van voorwaardelijke-financieringsprogramma's het doel 14) is, heeft de benadering van de Verkenningcommissie in beginsel vier voordelen:

- de output van onderzoekgroepen wordt geregistreerd in plaats van die van individuen of grove clusters;
- niet alleen internationale maar ook Nederlandse publikaties tellen mee, zij het met een lager gewicht. Dat doet recht aan de positie van de economische discipline als wetenschap die ook tot taak heeft bij te dragen aan actuele maatschappelijke en beleidsdiscussies;
- niet alleen artikelen, maar ook boeken, bijdragen aan boeken, en dissertaties tellen mee;
- er vindt een vorm van kwaliteitsweging plaats.

Dit overziende lijkt het daarom het overwegen waard wanneer faculteiten intern de evaluatie van voorwaardelijke-financieringsprogramma's aangrijpen om enige vingeroefeningen te doen op maat van de Verkenningcommissie. Daarbij dient er wel rekening mee te worden gehouden dat het rapport van de Verkenningcommissie op onderdelen sterk gekritiseerd is. Een modificatie die b.v. zou kunnen plaatsvinden is om de tijdschriften uit de door de commissie aangemerkte randgebieden (wiskunde, 'operations research') hoger te waarderen. Datzelfde geldt voor een aantal bedrijfs-economische en op deelgebieden gerichte tijdschriften.

De output van de verschillende programma's kan vervolgens gerelateerd worden aan de personele inzet voor onderzoek per programma. Uit een dergelijke procedure resulteert een kengetal per programma. Als norm voor 'evident onder de maat' kan dan bij voorbeeld worden genomen dat een programma lager 'scoort' dan de helft van het gemiddelde. Als dat voorkomt kan dat aanleiding zijn voor een inhoudelijk onderzoek door de faculteit. In gesprekken met de betreffende onderzoeksgroep naar aanleiding van het bibliometrisch materiaal kan dan aan de orde komen of er gegronde redenen zijn die de relatieve 'onder-score' kunnen verklaren. Conclusie: bij zo'n precieze zaak als onderzoek blijven outputmaatstaven altijd hulpmiddelen in een intersubjectief beoordelingsproces; ze treden niet in plaats daarvan.

Ten slotte kan worden opgemerkt dat de voorwaardelijke financiering als zodanig een relatief zwak instrument van onderzoekbeleid van de overheid blijft. In zijn huidige vorm is de toekomst van de voorwaardelijke financiering dan ook twijfelachtig. In het recente rapport van de OECD 15) wordt aanbevolen de voorwaardelijke financiering in het licht te plaatsen van een spiegeling van de Nederlandse universitaire onderzoek-'performance' aan het buitenland. Het is te verwachten dat in een dergelijke procesgang 'excellentie' volledig langs disciplinaire lijnen zal worden gemeten. Interdisciplinair onderzoek zal daarin niet goed tot zijn recht komen. Dat is een reden te meer voor de overheid om ook voor dit soort onderzoek een adequaat instrumentarium te ontwikkelen 16).

Cock Hazeu

12) Het verschil tussen prestatie-indicatoren en effectindicatoren wordt uitgebreid behandeld door H.M. van de Kar, o.a. in: *Beheersen en begroten*, in: D.J. Wolfson (red.), *Naar een beheersbare collectieve sector*, Deventer, 1981, blz. 83-124 en in *Het meten van overheidsvoorzieningen*, Maandblad voor *Accountancy en Bedrijfshuishoudkunde*, 1983, nr. 7/8, blz. 280-292.

13) Vgl. H.L. Theuns, *Over het meten van kwaliteit in de economie en belendende disciplines*, ESB, 12 december 1984, blz. 1181-1182.

14) Dit laat in het midden of de door de Verkenningcommissie gevolgde bibliometrische aanpak ook de meest geschikte is om een inhoudelijke sterkte-zwakte-analyse van het vakgebied te maken.

15) OECD, Directorate for Science, Technology and Industry, *Review of national science and technology policies. The Netherlands*, voorlopige versie, Parijs, maart 1986.

16) Dit punt komt ook aan de orde in: C.A. Hazeu, *De organisatie van het onderzoekbestel; beleidsvoering en besluitvorming in een geïntegreerd systeem*, *Beleid en Maatschappij*, juli/augustus 1986.