

Kosten van kerndepartementen: een empirische analyse van de productiviteitsontwikkeling tussen 2012 en 2019

Jos Blank, Alex van Heezik en Bas Blank

De Rijksoverheid probeert voortdurend de productiviteit van publieke organisaties te stimuleren. Maar hoe staat het eigenlijk met de productiviteit van die Rijksoverheid zelf? Uit de analyse van de productiviteitsontwikkeling van de kerndepartementen tussen 2012 en 2019 blijkt dat er sprake is van grote verschillen tussen de departementen. De verschillen kunnen oplopen tot tientallen procenten. Opvallend is ook dat de productiviteit van de gezamenlijke kerndepartementen in de loop der tijd niet toeneemt. Onduidelijk is of een gebrek aan innovatief gedrag, vergaande bureaucrativering, steeds complexere dossiers of een combinatie hiervan hieraan debet is. Voor helderheid hierover is een forse verbeterslag in overheidsboekhouding nodig.

1 Inleiding

De publieke sector levert een belangrijke bijdrage aan de maatschappelijke welvaart. Onderwijs, rechtshandhaving en gezondheidszorg zijn belangrijke sectoren voor een goed functionerende economie en dragen bij aan een sociaal rechtvaardige samenleving. Omdat deze voorzieningen veelal uit belastingen en premies worden gefinancierd en de tucht van de markt dikwijls ontbreekt, is inzicht in het presteren van deze sectoren buitengewoon belangrijk (Blank en Lovell, 2000; Blank en Valdmanis, 2019). Het analyseren van de productiviteit, doelmatigheid en effectiviteit van publieke voorzieningen heeft dan ook een hoge vlucht genomen. De wetenschappelijke ontwikkelingen op dit terrein in de afgelopen veertig jaar zijn indrukwekkend geweest, vooral door de ontwikkeling van empirische methoden om doelmatigheid en productiviteit te meten. Deze zogenoemde frontier analyses, bekend onder namen als SFA en DEA, bewijzen hun waarde vooral door toepassingen in de publieke dienstverlening (Fried et al., 2008; Kumbhakar en Lovell, 2000).

De focus van dit productiviteitsonderzoek ligt vooral op organisaties (of sectoren) die verantwoordelijk zijn voor de uitvoering van de publieke dienstverlening. Denk bijvoorbeeld aan scholen, ziekenhuizen en drinkwaterbedrijven, maar ook aan organisaties verantwoordelijk voor de aanleg en het beheer van wegen of de opvang van asielzoekers. Om een indruk te krijgen van deze 'mer à boire' aan onderzoek op dit terrein, verwijzen we graag naar de website www.ipsestudies.nl met rapporten en artikelen die duizenden verwijzingen bevatten naar andere onderzoeken in binnen- en buitenland.

De sterke focus op de uitvoering hangt voor een belangrijk deel samen met de relatieve eenvoud van het meten van de geleverde diensten, of producten van de uitvoerende organisaties. In veel gevallen gaat het om vormen van finale dienstverlening die zich vrij makkelijk in kengetallen laten vatten, zoals het aantal gediplomeerden, het aantal ziekenhuisopnamen

of de hoeveelheid geleverd drinkwater. Er zijn echter ook veel publieke organisaties die activiteiten verrichten die minder eenvoudig in getallen zijn uit te drukken. Dat geldt in het bijzonder voor organisaties die zich bezighouden met de aansturing van de uitvoering en het daarvoor benodigde beleid ontwikkelen. Op decentraal niveau zijn dat voornamelijk de beleidsafdelingen van de gemeenten en provincies. Op nationaal niveau zijn dat de beleidsdirecties van de ministeries, die gezamenlijk de zogenoemde kerndepartementen vormen. In opdracht van de minister dragen zij zorg voor het ontwikkelen van beleid en wet- en regelgeving en het aansturen van de uitvoering daarvan, waaronder het organiseren van de bekostiging. Ook zijn de kerndepartementen verantwoordelijk voor de evaluatie van het gevoerde beleid.

Over het gebrek aan belangstelling bij onderzoekers is een aantal mogelijke verklaringen te geven. Zo levert het meten van prestaties van dit soort intermediaire dienstverlening veel problemen op. Daarbij speelt ook mee dat de financiële verantwoordingen van deze (intermediaire) organisaties vaak weinig transparant zijn over de kosten die zijn gemaakt om de diensten te leveren. De overheidsboekhouding kent grote lacunes en inconsistenties. Ook dat is wellicht voor onderzoekers een belangrijke reden om dit soort onderzoek maar links te laten liggen. Onderzoek naar de organisaties die verantwoordelijk zijn voor het beleid dat aan de basis ligt van de uitvoering (en het leeuwendeel van de bekostiging daarvan) is dus zeldzaam. Er zijn twee uitzonderingen. Een uitzondering is te vinden op het terrein van de lokale overheid (Bikker en Van der Linde, 2016). In die publicatie staan de apparaatskosten van het gemeentebestuur centraal. Daarnaast is er in het Verenigd Koninkrijk wel langer aandacht voor dit onderwerp. Onderzoek dat enigszins vergelijkbaar is, betreft het werk van Hood en Dixon (2015).

Na intensieve raadpleging van verschillende bronnen en een uitgebreide correctie op de beschikbare gegevens is het niettemin mogelijk een analyse uit te voeren naar de doelmatigheid en productiviteit van de kerndepartementen in Nederland in de periode 2012-2019. In dit artikel bespreken we de bevindingen van dit onderzoek op hoofdlijnen. Daarbij staan drie vragen centraal:

1. Wat is de relatieve doelmatigheid van kerndepartementen?
2. Wat zijn de belangrijkste determinanten voor de relatieve doelmatigheid van kerndepartementen?
3. Hoe groot is de generieke productiviteitstrend van kerndepartementen tussen 2012 en 2019?

We gaan eerst in op de onderzoeksmethode. Daarna bespreken we de gegevens die voor dit onderzoek zijn verzameld en hoe deze zijn bewerkt. In de daarop volgende paragraaf komen de resultaten van de analyses aan bod. In de laatste paragraaf trekken we een aantal conclusies en doen we een aantal aanbevelingen.

2 Methode

De totale factor productiviteit (TFP, vanaf nu productiviteit) van een departement is in feite niets anders dan de verhouding tussen het volume van de productie (Y) en het volume aan ingezette middelen (X) (Blank, 2010; Niaounakis en Van Heezik, 2019):

$$Productiviteit = \frac{V_y(Y)}{V_x(X)} \quad (1)$$

Wanneer een instelling meer dan één product levert en bovendien verschillende middelen moet inzetten dan moeten de verschillende producten en ingezette middelen worden gewogen. In de marktsector kunnen prijzen als weegfactoren dienen. De productiviteit is dan gelijk aan de waarde van de productie ($V_y(Y)$) gedeeld door de waarde van de ingezette middelen ($V_x(X)$). Omdat in de publieke dienstverlening doorgaans marktconforme prijzen voor de geleverde diensten ontbreken, is deze weg niet begaanbaar. We veronderstellen daarom dat de productiewaarde gelijk is aan de gestandaardiseerde of schaduwkosten die gemoeid zijn met een gegeven productie. In dit geval hanteren we de kosten die een departement gemiddeld moeten maken om een bepaalde productie te leveren. Eigenlijk wegen we de verschillende producten met schaduwpreizen die een reflectie zijn van de kostprijzen. Maatschappelijk gezien hebben we dit ervoor over, anders hadden burgers via het politieke proces hiertoe niet besloten. De productiviteit is dus niets anders dan de ratio van gestandaardiseerde kosten en feitelijke kosten. De gestandaardiseerde kosten dienen dus als een benchmark. Vergelijking (1) kan nu als volgt worden geschreven.

$$Productiviteit = \frac{c^{sh}(Y)}{c^{obs}(Y)} \quad (2)$$

Waarbij:

$c^{sh}(Y)$ = schaduw (of gestandaardiseerde) kosten om Y te produceren;
 $c^{obs}(Y)$ = feitelijke kosten om Y te produceren.

Hierbij zijn de feitelijke kosten gecorrigeerd voor de prijsontwikkeling. Voor de personeelskosten maken we gebruik van de indexcijfers over contractuele loonkosten per uur (sector openbaar bestuur en overheidsdiensten) en voor de materiaalkosten van de consumentenprijsindex (CPI).

De gestandaardiseerde kosten berekenen we op basis van de resultaten van een regressieanalyse. Daarbij maken we eerst een aantal aanvullende veronderstellingen. Zo zijn de kosten niet alleen afhankelijk van de geleverde diensten, maar ook van het betreffende jaar en betreffende departement. Door technische ontwikkelingen zijn de kosten van de geleverde diensten in 2012 anders dan in 2019. Ook houden we rekening met het feit dat de geleverde diensten bij het ene departement meer of minder kunnen kosten vanwege verschillen in complexiteit van de beleidsdossiers of de kwaliteit van de geleverde diensten. We duiden dit aan met de heterogeniteit van de dienstverlening of de case mix. Daarnaast bevat het

model een component die de relatieve doelmatigheid weerspiegelt. Hierbij gaat het om het verschil in kosten tussen kerndepartementen die samenhangen met de kenmerken van de bedrijfsvoering, zoals het aandeel materiële kosten, de personeelsopbouw of de kwaliteit van arbeidsrelaties. De relatieve doelmatigheid en de technische ontwikkeling bepalen samen de ontwikkeling van de productiviteit.

We kunnen het voorgaande samenvatten in een vergelijking waarin de verschillende componenten zijn verwerkt. De kostenfunctie ziet er dan als volgt uit:

$$\ln(c_{dt}) = a_0 + \sum_m b_m \ln(y_{dtm}) + h \cdot t + het_d + eff_{dt} + err_{dt} \quad (3)$$

Waarbij:

- c_{dt} = kosten departement d op tijdstip t (gecorrigeerd voor prijzen);
- y_{dtm} = productie van dienst m door departement d op tijdstip t ;
- t = trend;
- het_d = percentage afwijkende kosten departement d door heterogeniteit van de productie
- eff_{dt} = percentage extra kosten door ondoelmatigheid departement d op tijdstip t ;
- err_{dt} = meetfout departement d op tijdstip t .

Waarbij:

$$eff_{dt} = \exp [-\sum_k \theta_k \ln(z_{atk})] \quad (4)$$

En:

- z_{atk} = kenmerk k van departement d op tijdstip t ;
 - θ_k = te schatten parameters van het model.
- a_0, b_m, h, het_d en θ_k zijn de te schatten parameters van het model. De parameter a_0 is de constante in het model, de parameters b_m zijn elasticiteiten en geven het effect weer van een groei van de productie op de groei van de kosten, en de parameter h geeft de jaarlijkse groei/krimp van de kosten weer door generieke productiviteitstrends of in economenjargon technische ontwikkeling. De parameter het_d geeft het procentuele effect weer van de complexiteit van de dienstverlening op de kosten van een departement. De parameter θ_k geeft het aandeel weer van determinant k in de totale ondoelmatigheid (Alvarez et al., 2006; Blank, 2020; Parmeter, 2018).

Ook leggen we aan het model de voorwaarde op dat een groei van de productie met een bepaald percentage tot een evenredige groei van de kosten leidt (homogeniteitseis). Dus 10 procent meer geleverde diensten leidt automatisch tot 10 procent meer kosten. Deze homogeniteitseis impliceert dat de b_m 's tot 1 moeten sommeren.

Het bovenstaande model is te schatten met een zogenoemd *mixed effects model* (Lindstrom en Bates, 1990). Zoals de naam al aangeeft, combineert een dergelijke aanpak twee typen effecten. Structurele verschillen in de kosten per eenheid product tussen

kerndepartementen worden ‘ingevangen’ door een zogenoemd *random effect* en door ons geïnterpreteerd als een maat voor de heterogeniteit (of case mix). Deze komt tot uitdrukking in de term *het* in vergelijking (3). Daarnaast bestaat vergelijking (4) uit een aantal termen met determinanten voor de doelmatigheid, zoals het ziekteverzuim of de mate van overhead. De effecten van deze determinanten worden tevens geschat. Het gezamenlijk effect van alle determinanten noemen we kosten(on)doelmatigheid en komt tot uitdrukking in de term eff_{at} .

Omdat case mix niet op een directe manier wordt gemeten, is het niet uit te sluiten dat de case mix een deel van de ondoelmatigheid absorbeert. De feitelijke doelmatigheidsverschillen zouden dus nog groter kunnen zijn dan gepresenteerd. Dit treedt op als een kerndepartement structureel ondoelmatig blijkt te zijn.

3 Data

De activiteiten van ambtenaren binnen de kerndepartementen hebben een uiteenlopend karakter, variërend van het beantwoorden van Kamervragen en het afhandelen van moties tot het behandelen van WOB-verzoeken. Eerder zijn al pogingen gedaan deze productie in beeld te brengen, maar dan op het niveau van directoraten-generaal (Ministerie van Binnenlandse Zaken, 2009). Deze onderzoeken bestonden uit een uitgebreide enquêtering van DG'en en hun inspanningen. Meer dan honderd indicatoren werden hier onderscheiden, die uit het oogpunt van analyse nauwelijks hanteerbaar waren. Bovendien zijn deze indicatoren ook niet voor een langere periode beschikbaar. Daarom kiezen we hier voor een andere route, waarbij we met veel minder indicatoren kunnen werken. De gedachte hierachter is dat veel van deze indicatoren een grote mate van samenhang vertonen. Als op enig moment een bepaald dossier veel aandacht krijgt, dan zien we dat tegelijkertijd in veel van de genoemde indicatoren terug. In dit onderzoek hanteren we drie indicatoren die inzicht geven in de ‘beleidsdrukte’ bij een kerndepartement:

- Documenten;
- Kamervragen;
- Programma-uitgaven (in constante prijzen van 2012).

Deze drie indicatoren vertegenwoordigen een grote hoeveelheid daarmee samenhangende activiteiten en dekken samen het grootste deel van de kerndepartementale productie af (Blank et al., 2009). Uit een principale componentenanalyse bleek dat zes indicatoren ruim 90 procent van de totale variatie in de ruim 100 indicatoren dekken. Ook hier zal blijken dat het beperkte aantal indicatoren een zeer groot deel van de variatie in de kosten verklaart. De beleidsdrukte wordt vooral zichtbaar in het aantal documenten en Kamervragen. De documenten betreffen alle door het ministerie gepubliceerde documenten (exclusief niet-verzelfstandige diensten en agentschappen), zoals vermeld op officielebekendmakingen.nl. Daarbij gaat het hoofdzakelijk om wet- en regelgevingsdocumenten, zoals wetten en

wetswijzigingen, Koninklijke Besluiten en Besluiten van algemene strekking. Voor de indicator Kamervragen hebben we het aantal (schriftelijke) antwoorden op de door Kamerleden (schriftelijk) gestelde vragen aan de bewindslieden van de verschillende departementen in kaart gebracht. De programma-uitgaven zijn de totale uitgaven van het departement minus de apparaatsuitgaven van het kerndepartement en gecorrigeerd voor de prijsontwikkeling met het consumentenprijsindexcijfer. Deze uitgaven geven een indicatie van de omvang van de beleidsvelden die door het betreffende kerndepartement worden aangestuurd.

Voor het bepalen van de inzet van middelen van de kerndepartementen hebben we gebruikgemaakt van de cijfers over de gerealiseerde apparaatsuitgaven van de kerndepartementen, zoals vermeld in de jaarverslagen van de ministeries. Daarin zijn de apparaatsuitgaven van de kerndepartementen uitgesplitst naar personele en materiële uitgaven. Binnen de personele uitgaven maken de jaarverslagen een onderscheid tussen uitgaven voor eigen personeel, externe inhuur en overig personeel. De materiële uitgaven zijn uitgesplitst naar bijdragen aan *shared service* organisaties (SSO's), ICT-uitgaven en overige materiële uitgaven, waaronder uitgaven voor huisvesting (Ministerie van Financiën, 2020).

Naast de gegevens over productie en inzet van middelen zijn er veel gegevens verzameld over (mogelijke) determinanten van kostendoelmatigheid. Daarbij gaat het vooral om hrmerken, zoals ziekteverzuim, werktijdfactor en gemiddelde leeftijd. Om dit soort personeelsgegevens in kaart te brengen, hebben we gebruikgemaakt van via het ministerie van BZK geleverde gegevens uit databestanden van P-direct (het personeelsadministratieprogramma van de Rijksoverheid) en de afdeling Trends, Onderzoek & Statistiek (TOS) van het ministerie van Defensie. Uiteindelijk zijn de gegevens van elf determinanten in de dataset opgenomen. Deze worden in tabel 1 vermeld.

Het voor de analyse gebruikte databestand bestaat uit 88 waarnemingen (8 jaren met 11 kerndepartementen).

Tabel 1 Statistische beschrijving gegevens kerndepartementen, 2012-2019

Variabele	Aantal	Gemiddelde	Std. dev.	Minimum	Maximum
Middelen					
Kosten totaal (x mln.)	88	240,38	134,81	28,92	553,78
Kosten personeel (x mln.)	88	158,01	84,35	16,11	358,78
Kosten materiaal (x mln.)	88	82,36	54,31	8,94	203,46
Productie					
Documenten	88	403,19	226,24	22,00	892,00
Vragen	88	307,16	226,86	5,00	954,00
Programma-uitgaven(x mln.)	88	13507,29	11523,06	27,08	42921,50
Determinanten					
Aandeel vrouwen (%)	88	48,77	9,53	22,41	62,03
Aandeel topfuncties (%)	88	8,09	2,83	2,23	15,83
Aandeel vrouwen in (%)topfuncties	88	31,37	10,79	8,67	50,75
Aandeel ondersteunend personeel (%)	88	31,93	10,32	18,16	57,84
Ziekteverzuimratio (%)	88	4,02	0,78	1,89	5,56
Gemiddelde leeftijd	88	46,14	1,35	42,90	48,56
Instroomratio (%)	88	9,68	4,21	2,08	23,93
Uitstroomratio (%)	88	9,92	2,51	5,30	17,04
Werktijdfactor (%)	88	94,82	1,05	92,61	96,83
Aandeel extern personeel (%)	88	8,38	4,91	1,76	21,12
Kostenaandeel materiaal (%)	88	33,16	6,66	18,84	56,00

4 Resultaten

We hebben verschillende specificaties van het model geschat en ten opzichte van elkaar getoetst. In het uiteindelijke model zijn op basis van statistische significantie en onderlinge relevantie de elf determinanten voor doelmatigheid gereduceerd tot vijf. Tabel 2 geeft de schattingsresultaten van de kostenfunctie weer van de regressie-analyses. Op basis van de schattingen is het mogelijk de marginale kosten te berekenen. Deze geven wat meer houvast over de plausibiliteit van de resultaten. De marginale kosten representeren de extra kosten gemoeid met de productie van één extra eenheid van het betreffende product. Tabel 3 presenteert de schattingen van de marginale kosten in 2019.

Tabel 2 Schattingsresultaten kostenfunctie

Variabelen		Schatting	St. afw.	Signif.
Documenten	b_1	0,383	0,040	0,000
Kamervragen	b_2	0,168	0,030	0,000
Programma	b_3	0,449	0,047	0,000
Constante	b_0	-0,741	0,136	0,000
Ziekteverzuim	θ_5	0,338	0,116	0,004
Aandeel instroom	θ_7	-0,147	0,026	0,000
Werktijdfactor	θ_9	6,943	2,571	0,007
Aandeel externe inhuur	θ_{10}	0,112	0,029	0,000
Kostenaandeel materiaal	θ_{11}	0,458	0,059	0,000

Tabel 3 Schattingen marginale kosten

	Gemiddelde	St. afwijking	Minimum	Maximum
Documenten	26930	11827	11048	44773
Kamervragen	11827	6293	3968	25716
Programma (per miljoen euro)	1477	732	470	2629

De schattingen leveren een betrouwbaar beeld op. Zo zijn de parameters van de productie-indicatoren betrouwbaar. Deze parameters zijn in feite een reflectie van de gewichten die aan de verschillende productie-indicatoren worden toegekend om de productiviteit te kunnen berekenen. Ook de vijf uiteindelijk overgebleven determinanten van doelmatigheid blijven alle significant (op het 5%-niveau).

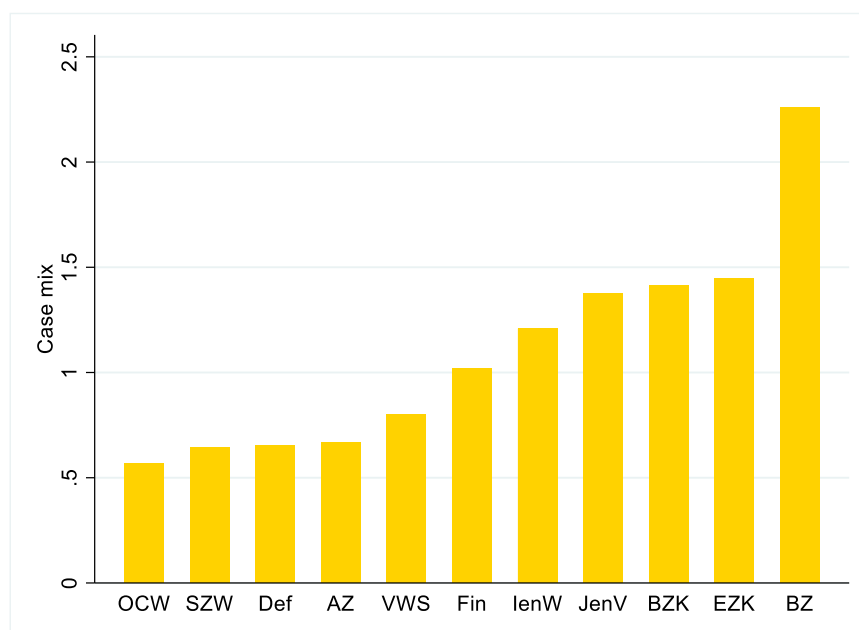
De parameterwaarden van de productie-indicatoren hebben plausibele waarden. Zo blijkt 38 procent van de ingezette middelen gemoeid met het aantal documenten, 17 procent met de behandeling van Kamervragen en 45 procent met het programma. De daarbij geschatte kostprijzen (marginale kosten) bedragen gemiddeld 27.000 euro voor een document, 12.000 euro voor een Kamervraag en 1.500 euro per 1 miljoen euro programmakosten. De correlatie tussen feitelijke kosten en de met het model voorspelde kosten is gelijk aan 99%.

Het kerndepartement AZ is op te vatten als een vreemde eend in de bijt vanwege zijn geringe omvang en specifieke taken. Een dergelijke uitbijter zou de schattingsresultaten substantieel kunnen beïnvloeden. We hebben daarom de schattingen nogmaals gedaan op basis van een gegevensverzameling zonder AZ. Het effect hiervan op de schattingsresultaten is echter beperkt.

Zoals in de modelbeschrijving aangegeven schatten we ook per kerndepartement een effect mee, dat te interpreteren is als de case mix. De case mix geeft aan hoeveel meer (of minder) kosten een kerndepartement maakt door een afwijkende zwaarte in de verrichte

activiteiten. Figuur 1 geeft de resultaten van de case mix weer. Per departement is een cijfer weergegeven rondom 1. Een cijfer kleiner dan 1 impliceert dat de case mix lager is dan gemiddeld, bij een cijfer groter dan 1 is de case mix hoger dan gemiddeld. Een waarde van 1,5 geeft aan dat met een gemiddelde zaak bij een specifiek kerndepartement 50 procent meer kosten gemoeid is dan bij het gemiddelde kerndepartement. Zoals al eerder werd betoogd is het niet uit te sluiten dat deze variabele een deel van de ondoelmatigheid absorbeert. De case mix wordt daarmee overschat en de kostenondoelmatigheid onderschat.

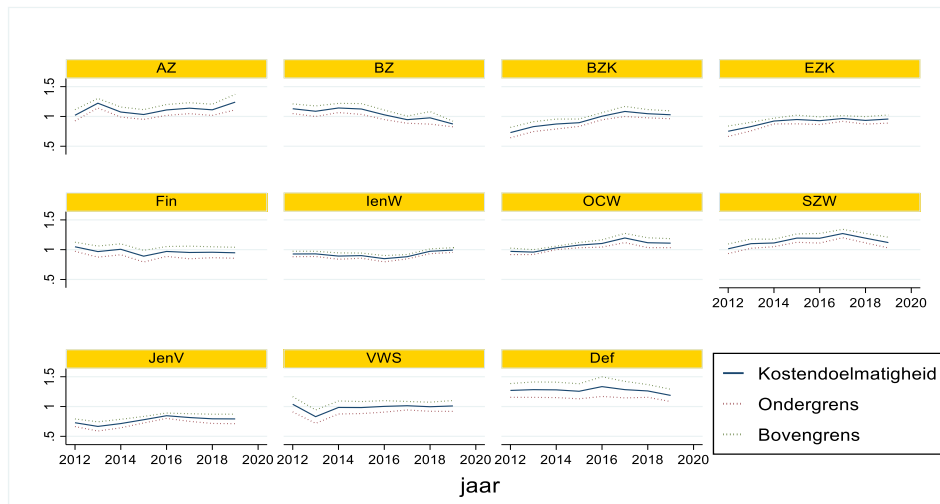
Figuur 1 Case mix per kerndepartement



Uit figuur 1 blijkt dat de kerndepartementen van OCW, SZW, Def en AZ de laagste case mix hebben. De kosten per eenheid product zijn hier ongeveer 60 procent van de gemiddelde case mix. Absolute koploper qua case mix is het kerndepartement van BZ. De kosten per eenheid product zijn hier 120 procent hoger dan bij het gemiddelde kerndepartement. De gemiddelde case mix verschilt dus aanzienlijk per kerndepartement. Wel kan hier sprake zijn van enige over- of onderschatting.

Op basis van de schattingen en toepassing van vergelijking (4) berekenen we per kerndepartement de kostendoelmatigheid. Figuur 2 toont de kostendoelmatigheid van de kerndepartement in de periode 2012-2019. De kostendoelmatigheid is de verhouding tussen de kosten van de gemiddelde praktijk en de feitelijke kosten. Een waarde van bijvoorbeeld 0,90 betekent dat dezelfde productie kan worden gerealiseerd tegen 90 procent van de feitelijke kosten (ten opzichte van de gemiddelde praktijk). Met andere woorden: er is een doelmatigheidswinst van 10 procent te realiseren ten opzichte van de gemiddelde praktijk.

Figuur 2 Kostendoelmatigheid kerndepartementen, 2012-2019



Figuur 2 laat zien dat er forse verschillen bestaan tussen de kostendoelmatigheid van kerndepartementen. Zo blijkt de doelmatigheid van de kerndepartementen BZK, EZK en JenV gemiddeld over de jaren 2012-2019 aanvankelijk maar 70 à 80 procent te bedragen ten opzichte van de gemiddelde praktijk. Vooral bij BZK is de laatste jaren een aanzienlijke verbetering te zien. Het kerndepartement van Defensie steekt qua kostendoelmatigheid ruimschoots boven de andere kerndepartementen uit. Een negatieve trend is waar te nemen bij het kerndepartement van BZ.

De resultaten kennen een zekere mate van statistische onzekerheid. Daarom is behalve de puntschattingen ook het gebied aangegeven waarbinnen de productiviteit met een zekerheid van 95 procent valt. Hieruit wordt duidelijk dat alleen Def in alle jaren een hogere kostendoelmatigheid kent dan het gemiddelde kerndepartement. Voor AZ en SZW geldt dit voor zeven van de acht onderzochte jaren. Het kerndepartement van JenV scoort voor alle acht jaren significant lager dan het gemiddelde kerndepartement.

Verder kunnen we de kostendoelmatigheid nog nader analyseren aan de hand van de schattingsresultaten in tabel 2. Een positief teken in tabel 2 impliceert een opwaarts effect op de kosten en dus een negatief effect op de productiviteit. Het negatieve effect van ziekteverzuim op de kostendoelmatigheid ligt voor de hand. Het positieve doelmatigheidseffect van de instroomratio is minder evident. Een hoge instroomratio leidt wellicht in eerste instantie juist tot extra kosten in verband met werving en inwerken. De instroomratio kan ook een indicatie zijn voor 'verse' krachten met een hogere arbeidsproductiviteit of lagere loonkosten. Een andere hypothese is dat door de instroom de organisatie weer beter aansluit bij de feitelijke behoefte. Denk hierbij aan mensen met een iets ander profiel dan de mensen die de dienst hebben verlaten. Het effect van de werktijdfactor is een opvallend resultaat, al is

bekend dat ook in andere sectoren een hoog percentage deeltijders positief kan bijdragen aan de productiviteit (Künn-Nelen et al., 2013; ROA, 2011). Een hypothese is dat mensen in deeltijd productiever zijn, omdat zij juist de laagproductieve uren van de dag of de week niet werken (Collewet en Sauermann, 2017). Daar staat tegenover dat er per gewerkt uur meer overheadkosten worden gemaakt voor bijvoorbeeld kantoorruimte, hrm-diensten en de salarisadministratie. Niet uit te sluiten is dat beide theorieën juist zijn, maar de gekozen modelspecificatie dit niet toelaat. Er is immers gekozen voor een lineair verband. Externe inhuur kan theoretisch twee effecten hebben. Externe inhuur is doorgaans flexibeler en daardoor doelmatiger. Daar staat tegenover dat de loonkosten per gewerkt uur waarschijnlijk aanzienlijk hoger zijn, omdat de marges voor de intermediair en de leegloop voor dit type personeel ook betaald moeten worden. Het negatieve effect van een hoog kostenaandeel materiaal kan wijzen op een al te 'uitbundige' aankoop van goederen en diensten. Bekend is het fenomeen dat overschotten in budgetten aan het einde van het jaar alsnog worden aangewend voor allerlei aankopen en inhuur. Materiële uitgaven lenen zich daar beter voor dan personele.

5 Conclusies en aanbevelingen

Het onderzoek naar de productiviteit en doelmatigheid van de publieke dienstverlening richt zich vrijwel altijd op de uitvoering. De productiviteit van publieke organisaties die zich bezighouden met beleidsontwikkeling en aansturing van de uitvoering van beleid wordt zelden onderzocht. De analyse van de productiviteit en doelmatigheid van de Nederlandse kerndepartementen vult deze leemte.

Het voor de analyse gebruikte databestand bestaat uit 88 waarnemingen (8 jaren met 11 kerndepartementen). Het databestand bevat voor ieder kerndepartement een aantal productindicatoren, kostencategorieën en doelmatigheidsdeterminanten. Op basis van de gegevens en een geavanceerde regressiemethode is een kostenfunctie geschat waaruit de onderzoeksresultaten zijn afgeleid. Op grond daarvan kunnen we de volgende conclusies trekken.

De belangrijkste conclusie is dat de doelmatigheid tussen kerndepartementen sterk verschilt. Het meest doelmatig zijn de kerndepartementen van AZ, SZW en Def. Het kerndepartement van AZ dankt zijn hoge score aan de gunstige werktijdfactor en het lage ziekteverzuim, het kerndepartement van SZW vooral aan de lage inzet van materiaal en het kerndepartement van Def vooral aan het lage ziekteverzuim. Het kerndepartement van JenV heeft de laagste kostendoelmatigheid, vooral veroorzaakt door een hoog ziekteverzuim en een relatief hoge inzet van materiaal. Hier valt dus nog winst te boeken, al is er de laatste jaren sprake van een verbetering, welke vooral tot stand komt door een relatief lagere inzet van materiaal. De doelmatigheidsverschillen staan los van de zwaarte van de beleidsdossiers, aangezien bij het vaststellen van de doelmatigheid al rekening gehouden met de (grote) verschillen in case mix. Omdat de case mix niet op een directe manier wordt

gemeten, is het niet uit te sluiten dat de case mix een deel van de ondoelmatigheid absorbeert. De feitelijke doelmatigheidsverschillen zouden dus nog groter kunnen zijn dan gepresenteerd. Dit treedt op als een kerndepartement structureel ondoelmatig blijkt te zijn.

Uit de analyse van de effecten van een aantal doelmatigheidsdeterminanten blijkt hoge ziekteverzuimpercentages, hoge werktijdfactoren, hoge aandelen externe inhuur en materiaal-kosten tot lage kostendoelmatigheid leiden. Een hoge instroomratio zorgt voor een hoge kostendoelmatigheid. Deze uitkomsten geven belangrijke aanwijzingen voor mogelijkheden om de doelmatigheid te verbeteren. Daarbij is de grootste winst te boeken bij het verminderen van het ziekteverzuim, het vergroten van het aantal deeltijders en het verminderen van de inzet van materiaal. Dit kan wel per kerndepartement verschillen. Ook is het zaak rekening te houden met samenhangende effecten.

Op basis van de onderzoeksresultaten blijkt geen generieke productiviteitstrend voor de kerndepartementen vast te stellen. Hiermee bedoelen we dat er bijvoorbeeld geen technische of institutionele ontwikkelingen zijn die de productiviteit van alle kerndepartementen in min of meer dezelfde mate beïnvloeden. Denk hierbij aan nieuwe IT-systemen en veranderde werkprocessen maar ook aan nieuwe regelgeving op het terrein van veiligheid of milieu. Extra kosten om aan milieueisen te voldoen zouden zelfs kunnen bijdragen aan een lagere generieke productiviteit. Een andere oorzaak kan de groeiende complexiteit van de uit te voeren taken zijn. Dat is een moeilijk beïnvloedbaar fenomeen, al zijn er wellicht wel mogelijkheden om de toenemende bureaucrativering te verminderen. De cijfers laten bijvoorbeeld ook een buitengewoon hoge overhead zien.

Uit de analyse blijkt verder dat er sprake is van aanzienlijke verschillen in de gemiddelde case mix. Zo kost bij BZ het afhandelen van bijvoorbeeld een document of Kamervraag ruim 120 procent meer dan gemiddeld. Voor de kerndepartementen van OCW, SZW, Def en AZ geldt juist dat de case mix maar 60 procent is van het gemiddelde kerndepartement. Volledigheidshalve benadrukken we dat de gepresenteerde kostendoelmatigheden al zijn gecorrigeerd voor deze case mix verschillen. Zonder deze correctie zouden de verschillen alleen nog maar groter worden. Op grond van deze bevindingen komen we tot een drietal aanbevelingen.

1. Inkrimpen budgetten. Gezien de grote verschillen in kostendoelmatigheid tussen de kerndepartementen, lijkt er bij verscheidene kerndepartementen nog flink wat ruimte voor verbetering. Vanwege de permanent intrinsieke druk om de bureaucratie uit te breiden (Wet van Niskanen, zie Niskanen, 1968) en beschikbare budgetten op te maken (Wet van Bowen, zie Archibald en Feldman, 2008; Bowen, 1980), zijn er weinig prikkels voor de ambtelijke leiding om die ruimte te benutten. Het moet daarom door de politiek worden afgedwongen om vervolgens het management hierop aan te spreken of af te rekenen. Zoals in eerder productiviteitsonderzoek is aangetoond, geldt daarbij de krimp van budgetten als een effectief instrument. Uiteraard moet het management van de kerndepartementen dan wel duidelijk zijn dat er ook ruimte is om met minder geld toe te kunnen. Daartoe kunnen de inzichten uit dit onderzoek behulpzaam zijn. Vooral het terugdringen van het

ziekteverzuim en het stimuleren van deeltijdwerk kunnen daaraan bijdragen. Ook een kritische blik op de externe inhuur en de materiële kosten kan doelmatigheidswinst opleveren. Op termijn kan dit bij sommige kerndepartementen resulteren in een doelmatigheidswinst van tientallen procenten.

2. Gericht onderzoek naar oorzaken gebrek aan productiviteitgroei. Verder bevelen wij aan om nader onderzoek te verrichten naar de oorzaak van het ontbreken van een generieke productiviteitsverbetering bij de kerndepartementen gedurende de onderzoeksperiode. Op basis van beschikbare gegevens is niet af te leiden of de min of meer constante productiviteit het gevolg is van een gebrek aan focus op productiviteitsverhogende innovaties of dat de kerndepartementen steeds meer te maken krijgen met complexere taken en wet- en regelgeving en met zwaardere eisen ten aanzien van personeelsbeleid, duurzaamheid en kwaliteitszorg. Om daar zicht op te krijgen zijn meer gedetailleerde gegevens nodig over de bedrijfsvoering.

3. Boekhouding op orde. Om dit soort analyses te kunnen uitvoeren, is het belangrijk te beschikken over een goede overheidsboekhouding. Tijdens ons onderzoek bleek dat deze veel te wensen overlaat. Zo blijkt dat niet alle departementen hetzelfde verstaan onder apparaatskosten, dat er slordig wordt omgegaan met verrekeningen in de sfeer van shared services en dat er over de geleverde prestaties nagenoeg helemaal geen verantwoording wordt afgelegd. Dat laatste is vooral opmerkelijk omdat de kerndepartementen hun uitvoeringsorganisaties op dit punt wel de maat nemen. Een verbeteringsslag wordt hier dan ook dringend aanbevolen.

Auteurs

Jos L.T. Blank (j.blank@ipsestudies.nl) is wetenschappelijk directeur stichting IPSE Studies en verbonden aan de TU Delft.

Alex A.S. van Heezik (a.vanheezik@ipsestudies.nl) is algemeen directeur stichting IPSE Studies.

Bas Blank (b.blank@ipsestudies.nl) is onderzoeker stichting IPSE Studies en promovendus aan de TU Delft.

Literatuur

Alvarez, A., C. Amsler, L. Orea en P. Schmidt, 2006, Interpreting and testing the scaling property in models where inefficiency depends on firm characteristics, *Journal of Productivity Analysis*, vol. 25(3): 201-212.

Archibald, R.B. en D.H.Feldman, 2008, Explaining Increases in Higher Education Costs, *The Journal of Higher Education*, vol. 79(3): 268-295. <http://doi.org/10.1353/jhe.0.0004>

Bikker, J. en D. van der Linde, 2016, Scale economies in local public administration, *Local Government Studies*, vol. 42(3): 441-463. <http://doi.org/10.1080/03003930.2016.1146139>

- Blank, J.L.T., 2010, *Principes van productiviteitsmeting. Elementaire handleiding voor kwantitatief onderzoek naar de productiviteit, doelmatigheid, effectiviteit en kwaliteit van de publieke sector*. Maastricht: Shaker Publishing B.V.
- Blank, J.L.T., 2020, The use of the scaling property in a frontier analysis of a system of equations. *Applied Economics*, 00(00): 1–11. <http://doi.org/10.1080/00036846.2020.1763246>
- Blank, J.L.T., B. Enserink en A.A.S. van Heezik, 2019. Policy Reforms and Productivity Change in the Dutch Drinking Water Industry: A Time Series Analysis 1980–2015, *Sustainability*, vol. 11(12): 3463. <http://doi.org/10.3390/su11123463>
- Blank, J.L.T., P.M. Koot, P. van Aken en C. Koolhaas, 2009, 2009, Beleidsdruk gemeten, Delft: TU Delft, IPSE Studies.
- Blank, J.L.T. en C. Lovell, 2000, Performance assessment in the public sector: contributions from efficiency and productivity measurement, in: J.L.T. Blank (ed.), *Public provision and performance*. Amsterdam: Elsevier.
- Bowen, H.R., 1980, *The Costs of Higher Education: How much do colleges and universities spend per student and how much should they spend?* San Francisco: Jossey-Bass.
- Collewet, M. en J. Sauermaann, 2017, Working hours and productivity, *Labour Economics*, vol. 47: 96-106.
- Fried, H.O., C.A.K. Lovell en S.S. Schmidt, 2008, *The measurement of productive efficiency and productivity growth*, New York: Oxford University Press.
- Hood, C. en R. Dixon, 2015, *A Government that Worked Better and Cost Less?: Evaluating Three Decades of Reform and Change in UK Central Government*, Oxford, UK: Oxford University Press.
- Kumbhakar, C., en C. Lovell, 2000, *Stochastic frontier analysis*, New York: Cambridge University Press.
- Künn-Nelen, A., A. de Grip en D. Fouarge, 2013, IS PART-TIME EMPLOYMENT BENEFICIAL FOR FIRM PRODUCTIVITY?, *ILR Review*, vol. 66(5): 1172-1191.
- Lindstrom, M.J. en D.M. Bates, 1990, Nonlinear Mixed Effects Models for Repeated Measures Data. *Biometrics*, vol. 46(3). <http://doi.org/10.2307/2532087>
- Ministerie van Binnenlandse Zaken, 2009, Beleidsdruk in beeld: een kwantitatieve vergelijking van directoraten-generaal, Den Haag: Ministerie van Binnenlandse Zaken.
- Ministerie van Financiën, 2020, Rijksbrede kostensoorten tabel, Den Haag: Ministerie van Financiën.
- Niaounakis, T.K. en A.A.S. van Heezik, 2019, Op afstand de beste? Een analyse van de productiviteitsontwikkeling bij IND, CJIB, SVB, RDW en het Kadaster, Delft: IPSE Studies.
- Niskanen, W.A., 1968, The Peculiar Economics of Bureaucracy, *American Economic Review*, vol. 57(2): 293-321.
- Parmeter, C.F., 2018, Estimation of the two-tiered stochastic frontier model with the scaling property. *J Prod Anal*, vol. 49: 37-47. <http://doi.org/10.1007/s11123-017-0520-8>
- ROA, 2011, De arbeidsmarkt naar opleiding en beroep tot 2016, Reports. Maastricht: Researchcentrum voor Onderwijs en Arbeidsmarkt/Maastricht University.