



Kan Nederlandse overheid Baumol verslaan?

Deze studie onderzoekt de productiviteitsontwikkeling van de Nederlandse overheid. Suggesties worden gedaan voor verbetering van de meting en analyse. De productiviteitsontwikkeling van de overheid is lastig te meten, vooral omdat marktconforme prijzen en omzetten ontbreken. Training, scholing en ICT lijken belangrijke factoren om de productiviteitsgroei te stimuleren. Beleidsexperimenten zijn nodig om de exacte effecten nauwkeuriger te bepalen.

De productiviteitsgroei is geschat voor 2012-2017 voor 33 instituten van de Rijksoverheid. Deze schattingen zijn erg onzeker. De gemiddelde groei ligt tussen de -0,1% en 1,3% per jaar.

CPB Notitie

Martin Mellens
Claudia Deijl
Sander van Veldhuizen

Samenvatting

De Nederlandse overheid staat de volgende kabinetsperiode grote uitdagingen te wachten om de kwaliteit van de publieke dienstverlening op peil te houden. Nadat de afgelopen jaren delen van de publieke sector zoals de zorg, het onderwijs, de politie en defensie te maken hadden met personeelstekorten deed begin dit jaar de coronacrisis haar intrede. De overheidsfinanciën zijn door de crisis fors verslechterd, waardoor extra middelen voor de publieke sector de komende jaren onder druk staan.

Productiviteitsgroei van de overheid is aantrekkelijk: een hogere productie met dezelfde hoeveelheid mankracht en kapitaal betekent dat burgers en bedrijven meer waar krijgen voor hun belastinggeld. Tegelijkertijd is dit makkelijker gezegd dan gedaan. De taken van de overheid bestaan vooral uit dienstverlening en diensten zijn naar hun aard minder makkelijk productiever te maken dan goederen (Baumol-effect). Daarnaast ontbreken prikkels door het gebrek aan concurrentie. Met dit onderzoek¹ willen we onder andere beter zicht bieden op de factoren die van invloed zijn op de productiviteitsgroei van de overheid. Dit kan aanknopingspunten bieden om de productiviteitsgroei te stimuleren.

Productiviteitsgroei is bij de overheid veel lastiger te meten dan bij commerciële bedrijven, omdat marktconforme prijzen en omzetten ontbreken. Normaliter wordt de productiviteitsgroei berekend als de verandering van het productievolume gedeeld door de ontwikkeling van het volume van de productiefactoren arbeid en kapitaal. Het productievolume is gelijk aan de omzet gedeeld door de prijs. Voor publieke goederen en diensten als dijkbewaking en defensie bestaan omzet noch prijzen. Voor andere producten zoals paspoorten zijn er wel tarieven, maar deze worden deels door de politiek vastgesteld en zijn in beperkte mate gerelateerd aan kostenontwikkelingen.

We proberen deze meetproblemen te ondervangen door de productiviteitsontwikkeling op verschillende manieren te bepalen. De productie wordt geschat met volume-indicatoren. Deze methode is in de literatuur het meest gangbaar. Verschillende indicatoren en specificaties worden gebruikt om inzicht te krijgen in de nauwkeurigheid van de schatting. Daarnaast wordt de schatting vergeleken met die van andere instituten zoals IPSE en het SCP. Dit versterkt de betrouwbaarheid van de schattingen. De productiviteitschatting van de overheid blijft echter een benadering, en daarmee onzeker.

De in deze studie geschatte productiviteitsgroei ligt voor de periode 2012-2017 tussen de -0,1% en 1,3% per jaar. Dit is een schatting voor de 33 grootste instellingen die vallen onder het openbaar bestuur van de rijksoverheid. Deze 33 instellingen vormen de basis voor dit onderzoek. De productiviteitsgroei verschilt sterk tussen instellingen en onderdelen van de overheid. De gemiddelde productiviteitsgroei in de commerciële sector ligt in dezelfde periode op 0,7%.

Scholing en training lijken een effectieve manier om de productiviteit te vergroten. Dit blijkt uit de wetenschappelijke literatuur en een empirische analyse door het CPB. ICT en technologie bieden in theorie ook veel mogelijkheden, maar het empirisch bewijs dat we vinden voor de opwaartse invloed op de productiviteitsgroei is mager. Dit kan mede komen door gebreken in de data. Tot op heden worden uitgaven aan ICT door instituten verschillend behandeld en bevatten ze soms ook loonkosten van IT-personeel.

¹ De auteurs bedanken Luka Bastiaans (BZK), Aafke Belterman (BZK), Hugo de Bondt (CBS), Evelien Eggink (SCP), Mark de Haan (CBS), Bart van Hulst (SCP), Brugt Kazemier (CBS), Thomas Niaounakis (IPSE), Sam Peeperkorn (Financiën), Cheyenne Schouws en Remko ter Weijden (BZK) voor hun waardevolle opmerkingen, analyses en dataverzameling ten behoeve van dit onderzoek. Verder bedanken wij een groot aantal CPB-collega's voor hun suggesties voor het onderzoek en het rapport.

Om effecten van maatregelen op het gebied van scholing, training en ICT op productiviteit beter in kaart te brengen, is gestandaardiseerde informatie nodig over de uitgaven aan training en scholing (opleidingsbudget), en ICT. Voor de ICT-uitgaven zal ook moeten worden bepaald hoe moet worden omgegaan met de afschrijvingen. De uitgaven zijn in de kern input-maatstaven en zeggen weinig over de effectiviteit van deze uitgaven. Idealiter wordt ook informatie verzameld over de kwaliteit van de investeringen in ICT en trainingen.

Daarnaast is aan te bevelen een dashboard te ontwikkelen met kwantitatieve en kwalitatieve indicatoren om de productiviteitsgroei van de overheid te analyseren. Deze indicatoren zouden het takenpakket van de overheidsinstelling zo goed mogelijk moeten weerspiegelen. Bij voorkeur zijn die indicatoren gerelateerd aan het productievolume, zoals omzetten of aantallen. Omdat ook deze indicatoren een onvolledig inzicht geven in de dienstverlening van een overheidsinstelling, moeten de kwantitatieve indicatoren worden aangevuld met kwalitatieve indicatoren, zoals klanttevredenheid en andere metingen van de kwaliteit van de dienstverlening.

Om beter grip te krijgen op welke interventies de productiviteit verhogen, kan worden ingezet op beleidsexperimenten. Door bijvoorbeeld deelname van organisaties of afdelingen aan een data-science-training op basis van loting te laten bepalen, kan worden nagegaan wat het effect is van deze training op de productiviteitsgroei. Na afloop van de training worden de uitkomsten van de overheidsorganisaties waarvan medewerkers de training wél hebben gevolgd vergeleken met de soortgelijke overheidsorganisaties waarvan medewerkers de training niet hebben gevolgd. Het verschil tussen de uitkomsten van deze twee groepen meet dan het effect van de training. Eenzelfde soort experiment kan worden opgezet voor het aanbieden van extra ICT-faciliteiten om het effect van ICT-faciliteiten op productiviteitsontwikkeling te onderzoeken. Aan de hand van de uitkomsten van dit type experimenten kunnen beleidsmakers gerichte maatregelen nemen om de productiviteitsgroei te stimuleren.

1 Inleiding

De overheid heeft als belangrijke taak publieke goederen en diensten zoals veiligheid, infrastructuur en basisonderwijs van voldoende kwaliteit en tegen aanvaardbare kosten aan te bieden. De productiviteitsontwikkeling van de overheid is een belangrijke factor om deze doelstelling te bereiken. Met dezelfde hoeveelheid arbeid en kapitaal kunnen bij hogere productiviteit meer goederen en diensten worden geproduceerd. Voor beleidsmakers is het derhalve aantrekkelijk methoden te vinden om de productiviteitsgroei te stimuleren.

Toch blijft de productiviteitsontwikkeling van de Nederlandse overheid over het algemeen gematigd. De taken van de overheid bestaan vooral uit dienstverlening en diensten zijn naar hun aard minder makkelijk productiever te maken dan goederenproducenten. Voor het onderwijzen van kinderen of surveillancetaken van de politie zijn nog steeds bijna evenveel mensen nodig als dertig jaar geleden. Dit zogenoemde Baumol-effect werd voor het eerst geformuleerd in de jaren zestig van de vorige eeuw (Baumol 1967). Daarnaast ontbreken prikkels door het gebrek aan concurrentie. Dit maakt de vraag pregnant welke factoren wel van invloed zijn op de productiviteitsontwikkeling bij de overheid en hoe de productiviteitsontwikkeling kan worden gestimuleerd.

Deze notitie is als volgt gestructureerd. Sectie 2 geeft de belangrijkste aanbevelingen, terwijl het derde hoofdstuk de problemen bij het meten van de productiviteit, en de achterliggende oorzaken daarvan,

bespreekt. Ondanks deze meetproblemen wordt getracht in sectie 4 een inschatting te maken over de productiviteitsgroei. Hiervoor is onder andere een dataset verzameld met gegevens van de 33 grootste instellingen van de rijksoverheid. Ondanks de onzekerheid over de exacte productiviteitsgroei kunnen we wel iets kunnen zeggen over de factoren die deze groei kunnen stimuleren. De literatuur geeft aanwijzingen welke factoren het kansrijkst zijn (sectie 5). We eindigen met een schets voor een dashboard om de monitoring voor de productiviteitsgroei te verbeteren en suggesties voor verder onderzoek.

2 Aanbevelingen

Dit onderzoek naar de productiviteitsgroei van de overheid leidt tot de volgende aanbevelingen om het inzicht te verbeteren.

- Als overheidsinstellingen een uniforme rapportage aanhouden van de kosten van arbeid en kapitaal, in het bijzonder ICT, dan wordt duidelijker welke interventies tot productiviteitsgroei leiden.
- De productie van de overheid bevat meerdere dimensies. Een dashboard van indicatoren is een zinvolle manier om de productiviteitsontwikkeling te monitoren. Hierbij moet wel bedacht worden dat een dashboard nooit alle aspecten van de dienstverlening van de overheid volledig kan bevatten. De beoordeling van een overheidsinstelling kan derhalve nooit alleen op basis van numerieke criteria plaatsvinden.
- Experimenten kunnen verhelderen op welke wijze productiviteitsgroei kan worden gestimuleerd. Omdat menselijk kapitaal en ICT belangrijke factoren zijn liggen experimenten op dit gebied het meest voor de hand.

3 Productiviteitsgroei moeilijk meetbaar

De productiviteitsontwikkeling is in de publieke sector veel lastiger te meten dan in de marktsector. De productiviteitsgroei is de verhouding tussen de ontwikkeling van het productievolume en die van de productiefactoren arbeid en kapitaal. Hierbij wordt het productievolume berekend als de omzet gedeeld met een relevante prijsontwikkeling.

Bij overheidsdiensten ontbreken meestal marktconforme omzetten en prijzen, waardoor het productievolume niet kan worden berekend. Bij klassieke publieke goederen (zoals defensie of dijkbewaking) bestaat geen relevante marktprijs, omdat mensen niet van het gebruik ervan kunnen worden uitgesloten. Bij andere goederen of diensten van de overheid (denk bijvoorbeeld aan paspoorten) zijn de prijzen door de politiek vastgesteld en niet direct gerelateerd aan de werkelijke kosten.

Naast het cruciale punt van ontbrekende prijzen zijn er nog andere factoren die de meting van het productievolume moeilijker maken. Bij sommige overheidstaken zoals preventie, toezicht en regelgeving is de essentie dat bepaalde gebeurtenissen niet plaatsvinden. Misdrijven of verkeersongevallen die niet hebben plaatsgevonden, worden niet waargenomen; het effect van preventie kan derhalve op zijn best worden geschat. Daarom komen preventie, toezicht en regelgeving niet (of slechts in zeer beperkte mate) tot uiting in

productiecijfers, terwijl het wel belangrijke taken zijn van de overheid. Verder bestaan veel overheidstaken uit diensten en de kwaliteit van dienstverlening is naar haar aard lastig te kwantificeren.

De meetproblemen kunnen deels worden ondervangen door in plaats van omzetten en prijzen andere indicatoren te gebruiken. Deze indicatoren zijn dan een benadering van het productievolume. Voorbeelden van indicatoren zijn het aantal wetenschappelijke publicaties of het aantal afgehandelde belastingaangiften. Dit is een veel gebruikte aanpak om de productiviteitsgroei te meten die onder andere gebruikt wordt door IPSE (zie b.v. IPSE 2012,2013) of het SCP (zie b.v. SCP 2012, 2015a, 2015b).

Bij deze aanpak is de keuze van de indicatoren cruciaal. Deze indicatoren moeten representatief zijn voor het takenpakket van de overheidsinstantie waarvan de productie moet worden gemeten. Dit bepaalt zowel de kwaliteit van de meting als de manier waarop de cijfers moeten worden geïnterpreteerd. Voor een deel van het rijk zijn geschikte indicatoren lastig te vinden. Dat geldt bijvoorbeeld voor defensie, maar ook bijvoorbeeld kernministeries. Ook bij overheidsinstanties die zich bezighouden met toezicht of met regelgeving is het lastig indicatoren te vinden. Vaak worden dan normindicatoren gebruikt zoals het aantal personen/bedrijven waarop toezicht wordt gehouden.

Een verandering van de productie kan het gevolg zijn van exogene schokken die losstaan van de inzet van de productiefactoren. Dit bemoeilijkt de interpretatie van de uitkomsten van de meting van de productiviteitsgroei via de indicatorenmethode. Bij bijvoorbeeld het gevangeniswezen wordt de productiviteitsontwikkeling sterk bepaald door de verandering van het aantal gevangenen. Omdat het aantal geregistreerde misdrijven en daarmee het aantal gevangenen de laatste jaren daalt, waarschijnlijk mede door de vergrijzing, wordt in sommige perioden een productiviteitsdaling gevonden (Niaounakis, Uurlings en van Heezik, 2014). Dit is vooral een reflectie op het feit dat de capaciteit van gevangenen en bewakingspersoneel niet direct kan worden teruggeschroefd. De kosten van de inputfactoren blijven dus gelijk of stijgen, terwijl de productie daalt. De productiviteitsontwikkeling van organisaties die zich bezighouden met de opvang van asielzoekers en vluchtelingen fluctueert sterk als aantallen als indicator voor de productie worden gebruikt. De productiviteitscijfers geven een indicatie van de kosten per eenheid, maar zeggen weinig over de jaar-op-jaar ontwikkeling van de efficiency van een organisatie.

Om de problemen van de indicatorenmethode te ondervangen, kan geprobeerd worden de productiviteitsgroei van een overheidsinstantie te benaderen door een bedrijfstak van de marktsector die qua inputstructuur veel op het overheidsonderdeel lijkt als proxy te nemen. In CPB (2017) wordt dit de benaderingsmethode genoemd. Deze methode schat wat de productiviteitsgroei zou kunnen zijn, als het overheidsonderdeel even productief is als het bedrijf in de marktsector. Het voordeel is dat je deze methode kunt toepassen als geschikte volume-indicatoren ontbreken. Bovendien spelen interpretatiekwesties door externe factoren minder een rol. Een nadeel is dat niet de werkelijke productiviteit wordt gemeten. Het is de productiviteitsgroei die een overheidsinstantie zou hebben als die even productief zou zijn als een evenknie in de marktsector. Dit is nu juist de vraag. Hoewel de methode nuttig is als ijkpunt, biedt die beperkt inzicht in de werkelijke productiviteitsontwikkeling van de onderdelen van de overheid. Daarom is voor deze studie de indicatorenmethode gebruikt.

4 Schatting productiviteitsgroei

Data

Het CPB heeft voor de periode 2012-2017 de productiviteit geschat op basis van data van de 33 grootste instellingen van de rijksoverheid. Het productievolume wordt bepaald via een aantal indicatoren, waarna de

productiviteitsontwikkeling wordt bepaald door het productievolume te delen door de verandering van het volume van de productiefactoren arbeid en kapitaal (Solow-residu). Bijlage 1 gaat dieper in op de methode.

Voor dit onderzoek zijn data verzameld over de uitgaven van 133 organisaties van de Rijksoverheid. Het gaat hier om openbare informatie die door de instanties wordt gepubliceerd. Deze informatie is in de meeste gevallen gevonden in jaarverslagen op www.rijksbegroting.nl. Het gaat om kerndepartementen, agentschappen, ZBO's en uitvoeringsdiensten behorend bij de verschillende ministeries. De totale uitgaven zijn opgesplitst in arbeidskosten, kapitaalkosten en overige uitgaven. Deze overige uitgaven kunnen verschillend van aard zijn en bijvoorbeeld ook uitgaven aan subsidies omvatten. In deze studie spelen de overige uitgaven en de totale uitgaven geen rol. De arbeidskosten zijn verder verdeeld in arbeid intern, externe inhuur, overige arbeid, pensioenlasten en sociale lasten. De kapitaalkosten zijn gespecificeerd in afschrijvingen en materiële kosten en deze materiële kosten vervolgens in huisvesting, materiaal, bureau, ICT, SSO's, exploitatiekosten en overig.

Voor de 33 grootste instellingen van de rijksoverheid (in termen van aantal werknemers) zijn outputindicatoren verzameld. Dit zijn bijvoorbeeld omzetcijfers of cijfers over het aantal verleende producten of diensten. De 33 organisaties bevatten samen ruim 80% van het totale aantal fte's van de oorspronkelijke dataset (uit CPB 2018) van 133 instellingen. Waar mogelijk is het productievolume berekend als de omzet, gedeeld met een interne kostprijs of een tarief. Dit sluit het meest aan bij de conventies van de Nationale Rekeningen. Waar de omzet niet beschikbaar was, zijn aantallen (bijvoorbeeld aantal publicaties of rapporten van de instelling) gebruikt.²

Er is in deze studie geen correctie gemaakt voor de kwaliteit van de dienstverlening. Vanuit theoretisch perspectief zou dit voor de volledigheid van de meting van de dienstverlening wel wenselijk zijn. Kwaliteitsindicatoren worden echter niet op regelmatige basis verzameld en hebben geen gemeenschappelijke noemer. Sommige instituten publiceren het percentage klachten, andere een globale tevredenheid met de dienstverlening. Een éénduidige correctie voor kwaliteit is op basis van deze bronnen niet goed mogelijk.

Deze studie bouwt voort op het onderzoek uit 2018 (CPB 2018), waar het productievolume ook op deze wijze werd geschat. Die studie kijkt keek naar de productiviteit in 2012-2015. In deze studie wordt een langere periode beschouwd (2012-2017). Daarnaast is in deze studie waar mogelijk (voor 15 organisaties) ook informatie uit de personeelsadministratie toegevoegd, zodat verkennend onderzoek kon worden gedaan naar het effect van menselijk kapitaal op productiviteit (zie sectie 5.1).

Resultaten

Op basis van deze beschikbare data wordt de productiviteitsontwikkeling van de overheid in de periode 2012-2017 geschat op gemiddeld jaarlijks tussen de -0,1% en 1,3%. Vanwege de grote spreiding in de uitkomsten is de mediaan genomen en niet het rekenkundige gemiddelde. Gezien de meetproblemen is de schatting van de productiviteitsgroei gevoelig voor de gekozen specificatie. Daarom is ervoor gekozen een interval te presenteren in plaats van een puntschatting, gebaseerd op de uitkomsten van vier verschillende methoden voor de 33 instellingen (Tabel 4.1).

De schattingen zijn niet alleen gevoelig voor de gekozen specificatie, maar ook voor het toevoegen van jaren en instituten. In het algemeen zien we een grote spreiding in de productiviteitsgroei tussen instituten en over verschillende jaren. Ook de gebruikte indicatoren spelen een rol. Vier verschillende methoden zijn gebruikt om de productiviteitsgroei te berekenen. In elk van de methoden wordt op een andere manier een keuze gemaakt uit de gebruikte indicatoren. Bij methode 1 is het gemiddelde van de beschikbare kwantitatieve

² De gebruikte productie-indicatoren per instituut zijn op te vragen bij de auteurs.

indicatoren genomen. Bij methode 2 zijn de relevante kwantitatieve indicatoren bepaald op basis van expert judgement. Dit leidt tot een ietwat hogere schatting van de productiviteitsgroei. Bij methode 3 is de schatting op basis van de kwantitatieve indicator met de laagste ontwikkeling in plaats van een gemiddelde van de beschikbare indicatoren per instituut, om zo een ondergrens van de productiviteitsgroei te schatten. Bij methode 4 is juist de kwantitatieve indicator met de hoogste ontwikkeling genomen om een bovengrens van de productiviteitsontwikkeling te benaderen.

Vanwege de grote spreiding in de indicatoren baseren we ons voornamelijk op de mediaan van de verschillende specificaties. De mediaan ligt tussen de -0,1% en 1,3%. Bij de marktsector ligt de gemiddelde productiviteitsgroei op 0,7% voor de periode 2012-2017.³ De ontwikkeling van de overheid ligt dus, als gekeken wordt naar het midden van het interval, iets lager. Het interval is in lijn met de 0,6% á 0,7% die in de studie van 2018 is gevonden voor de jaren 2012-2015 (CPB 2018). CPB (2017) liet op basis van de benaderingsmethode via de inputstructuur schattingen zien tussen de 0% en 0,7%.

De forse standaarddeviaties wijzen op flinke verschillen in productiviteitsgroei tussen instituten. Dit is in lijn met wat ook bij andere onderzoeken wordt gevonden (Blank en Van Heezik 2019, SCP 2019). Dit komt deels doordat demografische factoren zoals vergrijzing en technologische ontwikkelingen de instituten verschillend beïnvloeden. Ook kunnen plotselinge schokken het beeld behoorlijk beïnvloeden. Een omzet kan sterk dalen doordat een groot onderzoeksproject stopt.

Tabel 4.1 Schattingen productiviteitsgroei (groei per jaar in %, vier methoden)

Specificatie	Gemiddelde	Mediaan	Standaarddeviatie
	%		
Methode 1	0,21	0,06	10,07
Methode 2	0,47	0,26	15,52
Methode 3	-1,35	-0,11	14,13
Methode 4	2,11	1,3	12,85

We proberen de robuustheid van de schattingen verder te verkennen door waar mogelijk een vergelijking te maken met de schatting van andere instituten. Voor zeven instituten kon dit worden gedaan (Moolenaar 2019, Niaounakis et al. 2019, Niaounakis et al. 2020). Omdat het in deze studie niet gaat om de specifieke schattingen van individuele instituten maar het een schatting op macro-niveau betreft, geven we alleen een kwalitatieve beschrijving van de resultaten van dit verkennende onderzoek. In ongeveer de helft van de gevallen komen de gevonden resultaten overeen, maar in de resterende gevallen is de orde van grootte of zelfs het teken verschillend. De mediaan van deze zeven schattingen lijkt wat hoger te liggen dan de mediaan in deze studie, maar ligt binnen het interval. Het gebruik van andere indicatoren en andere methoden om de productiviteit te schatten, verklaren voor een groot deel deze verschillen.

5 Productiviteitsgroei stimuleren

De wetenschappelijke literatuur bevat een groot aantal factoren dat de productiviteitsontwikkeling kan beïnvloeden. Tabel 2.1 bevat een aantal vaak benoemde en belangrijke determinanten, de veronderstelde

³ Zie Statline, bijdrage multifactorproductiviteit op basis van toegevoegde waarde, gemiddelde 2013-2017.

richting van het effect en de mate waarin deze factoren kunnen worden gestuurd door beleidsmakers. Deze tabel is grotendeels gebaseerd op een eerdere literatuurstudie (CPB 2017), aangevuld met recente publicaties. In de secties 5.1. en 5.2 wordt uitgebreid ingegaan op het effect van menselijk kapitaal en ICT. In de volgende alinea's wordt iedere variabele kort toegelicht.

Tabel 2.1 Determinanten voor productiviteitsontwikkeling overheid

Determinant	Richting effect	Mate van stuurbaarheid
Algemene trend		
Aard overheidstaken	-	-
Menselijk kapitaal	+	+
ICT	+	+
Budgetkortingen	?	+
Schaalvergroting	?	+
Organisatiecultuur	+	+
Kwaliteit managers	+	+
Transparantie	+	+
Accountability	+	+
Kwaliteit fysieke infrastructuur	+	-
Kwaliteit institutionele infrastructuur	+	-

De productiviteitsgroei wordt allereerst beïnvloed door de aard van de overheidstaken. Dit kan niet door beleidsmakers worden beïnvloed. Overheidstaken bestaan vooral uit dienstverlening. Productiviteitswinsten zijn bij diensten minder makkelijk te behalen dan bij goederen (Baumol 1967). Voor verzorging van ouderen, lesgeven aan kinderen en surveillance bij de politie zijn nog steeds ongeveer evenveel mensen nodig als veertig jaar geleden. Natuurlijk zorgen ook hier innovaties voor productiviteitswinsten, maar de effecten daarvan zijn veel kleiner dan bij de productie van goederen.

Het effect van budgetkortingen op de productiviteitsgroei is lastig te bepalen. Budgetkortingen kunnen op korte termijn de productiviteitsgroei stimuleren, omdat de kosten lager worden, terwijl het takenpakket (en dus de productie) gelijk blijft (zie b.v. Blank en van Heezik 2019). De budgetprijkkels kunnen zo, zeker op de korte termijn, ook een opwaarts effect hebben op de productiviteitsgroei. Indien budgetkortingen duurzaam leiden tot lagere lonen binnen de organisatie, kan dat echter de motivatie van overheidspersoneel verlagen. Zo kunnen budgetkortingen, zeker op lange termijn, de kwaliteit van de overheid verlagen en het bovendien lastiger maken goed gekwalificeerd personeel te vinden. Productiviteitsgroei op korte termijn kan dus op de langere termijn gevolgd worden door een lagere productiviteitsontwikkeling.

Het effect van schaalvergrotingen is gemengd. Door schaalvergroting dalen de vaste kosten per eenheid product, hebben instanties meer marktmacht bij de inkoop van goederen en diensten en kunnen kapitaal en arbeid mogelijk efficiënter en flexibeler worden aangewend. Tegenover deze schaalvoordelen staan ook schaalnadelen zoals toenemende bureaucratie en coördinatiekosten. Een optimale schaalgrootte kan worden afgeleid op basis van de geschatte kosten en een veronderstelde productiefunctie. Blank (2015) noemt een groot aantal studies voor de optimale schaal voor bijvoorbeeld zorg en onderwijsinstellingen. Op basis van deze analyses komt Blank tot de conclusie dat de schaalvergroting bij zorginstellingen en bij het voortgezet onderwijs is doorgeschoten. Het ankerpunt van deze analyse wordt echter wel bepaald door theoretische veronderstellingen. Aan de andere kant moeten naast economische aspecten ook de maatschappelijke kosten worden genoemd. Hieronder vallen bijvoorbeeld de bereikbaarheid en toegankelijkheid van de voorziening.

De perceptie van burgers speelt hier ook een rol. Mogelijk zijn grote instellingen efficiënter, maar ervaren burgers toch minder dienstverlening door afnemende persoonlijke betrokkenheid.

De organisatiecultuur heeft ook invloed op de productiviteitsgroei. Waardering voor het werk van werknemers en gedeelde waarden en normen verhogen de motivatie en de productiviteit (zie b.v. Matko en Takacs 2017). Er is ook een positief verband tussen transparantie van de overheid en productiviteit (Douglas en Meijer 2016). Verder lijkt er een positief verband tussen de kwaliteit van de managers en productiviteit (Syverson 2011). Het is uiteraard niet eenvoudig om deze factoren beleidsmatig aan te sturen. Investeren in cultuur en managementkwaliteit loont, maar de beheersbaarheid is begrensd.

De productiviteitsontwikkeling van overheidsinstellingen hangt ook samen met de omgeving waarin ze opereren. De productiviteit van bedrijven is hoger wanneer de instituties en de fysieke infrastructuur (b.v. wegen en andere netwerken) van hogere kwaliteit is (Kim en Loayza 2017). Hoewel deze factoren in beperkte mate ook op het niveau van organisaties kunnen worden beïnvloed, wordt het in de literatuur meer gezien als een factor die betrekking heeft op de structuur van een land. De geschiedenis van een land speelt hierbij ook een rol die naar haar aard minder te beïnvloeden is. Het opleidingsniveau van bewoners kan een positief effect hebben op de efficiëntie van gemeentediensten (Alfonso en Fernandes 2008; Kalb et al. 2012).

5.1 Training en scholing belangrijk

Investeren in menselijk kapitaal door training en scholing lijkt een effectieve manier om de productiviteitsontwikkeling te vergroten. Geschoolde en getrainde medewerkers kunnen bij een gegeven stand van de techniek hun taken sneller en efficiënter uitvoeren en zo een hogere productie behalen. Daarnaast kunnen geschoolde werknemers makkelijker nieuwe technieken leren. Er is derhalve een positief verband tussen innovatie en scholing. Er is uitgebreid onderzoek gedaan naar het verband tussen training en scholing en productiviteit in de economische literatuur, op macro- bedrijfstak en bedrijfsniveau (zie o.a. Ahmed en Bhatti 2020, Blundell et al. 1999, Sasso en Ritzen 2020). Over het algemeen bevestigt de economische literatuur op bedrijfsniveau het verband tussen training en scholing en productiviteit (e.g. Konings en Vanormelingen 2015). Op macro-economisch niveau zijn de schattingsresultaten echter gemengd (zie b.v. Botev et al. 2019). Specifiek onderzoek voor de publieke sector is schaars.

Hoewel een positief effect van training in theorie evident lijkt, is dit empirisch niet makkelijk aan te tonen. Dit komt in de eerste plaats doordat een éénduidige maatstaf voor training en scholing niet evident is. Vaak worden aantal scholingsjaren of het hoogste behaalde eindexamen als maatstaf genomen. Voor de productiviteit op de werkvloer zijn echter ook taak-specifieke kennis en zogenoemde zachte kennis belangrijk. Het is kortom lastig alle aspecten van menselijk kapitaal in cijfers te vangen.

Daarnaast kan er samenhang zijn tussen scholing en de maatstaf van de productiviteitsgroei. Geschoold personeel is duurder. Sterker nog: voor de werknemer worden de opbrengsten van training en scholing gemeten in termen van hoger loon. De stijging van de productie moet deze kostenstijging overtreffen, wil er sprake zijn van productiviteitsgroei. Als dit niet het geval is, dan wordt geen verband gevonden tussen hogere productiviteit en opleidingsniveau.

Training en scholing hangt bovendien samen met andere factoren die productiviteitsgroei beïnvloeden. Dit geldt vooral voor ICT-gebruik en R&D-uitgaven. Deze samenhang en de meetproblemen maken het moeilijk het exacte effect van een individuele factor te vinden.

Voor deze studie is onderzoek verricht naar het effect van training en scholing op de productiviteitsgroei, dat wijst op een positief verband. Het databestand is hiervoor voor een aantal instituten aangevuld met informatie uit de personeelsadministratie (P-direkt). Deze informatie bevat onder andere het schaalniveau van het personeel. Deze data zijn echter alleen beschikbaar voor instellingen die vallen onder het rijk. Dit is slechts voor 15 onderzochte instituten het geval. De verandering van het percentage werknemers in een bepaald schaalniveau is gebruikt als verklarende variabele voor de productiviteitsgroei (zie bijlage 1.3 voor meer gedetailleerde schattingen). Meer werknemers in een hoger schaalniveau (ten opzichte van het laagste schaalniveau) is positief gerelateerd aan productiviteitsgroei. Dit duidt erop dat scholing en training een rol kan spelen in het stimuleren van productiviteitsgroei. Overigens is er niet voor alle schaalniveaus een statistisch significant effect.

De resultaten van deze studie zijn indicatief. Dit komt vooral door beperkingen van de data. Het aantal waarnemingen waarop de schatting is verricht en het aantal instituten (15) is klein. Bovendien is het schaalniveau niet perfect als indicator van het menselijk kapitaal. Het niveau waarop iemand wordt ingeschaald, hangt samen met het opleidingsniveau en ervaring in de organisatie en zegt dus iets over het menselijk kapitaal. Het hangt ook samen met onderhandelingsvaardigheid, ambitieniveau en vele andere factoren die niet noodzakelijk iets zeggen over de productiviteit van een werknemer. Ook het type organisatie heeft invloed op het aandeel hoog en laag geschaalde medewerkers. Dit maakt het lastig te bepalen of de productiviteitsgroei komt door menselijk kapitaal of door het type taken dat moet worden uitgevoerd.

5.2 Potentieel ICT niet volledig benut

Het gebruik van ICT kan de productiviteit van de overheid vergroten en kan het contact tussen de overheid en de burgers vergemakkelijken. Door ICT zijn sommige taken van de overheid zoals administratieve procedures en het schrijven van rapporten makkelijker en sneller uit te voeren. Ook de dienstverlening van uitvoeringsinstanties wordt vergemakkelijkt. Denk aan de digitale aangifte van de belastingen. Burgers hebben makkelijker toegang tot informatie en de kwaliteit van de data wordt vergroot door ICT (Palvalin et al. 2013). De communicatie met de burger kan bijvoorbeeld door de sociale media worden vergemakkelijkt. Derhalve lijken investeringen in ICT nuttig voor het verhogen van productiviteit.

Het gebruik van grote datasets (big data) en zogenoemd e-government kunnen de productiviteitsgroei van de overheid stimuleren. E-government is het gebruik van ICT om overheidsdiensten te verbeteren. Hieronder vallen bijvoorbeeld applicaties om vragen van burgers te beantwoorden. Voor zowel big-data als voor e-government geldt dat er in potentie grote mogelijkheden zijn voor productiviteitswinsten, maar dat tot nog toe onduidelijk is hoe groot het effect is (Bright en Margetts 2016, Severo et al. 2016, Twizeyimana en Andersson 2019).

Recente economische literatuur wijst op een positief verband tussen investeringen in ICT en productiviteit (zie b.v. Syverson 2011, Hüseyin en Gönel 2020) voor bedrijven in de commerciële sector. Een hogere IT-intensiteit werd onder andere genoemd als oorzaak voor de hogere productiviteitsgroei van bedrijven in de VS ten opzichte van bedrijven in de EU. Palvalin et al (2013) vindt een positief effect voor kennisinstituten. Dit artikel wijst er echter op dat de positieve effecten niet gegarandeerd zijn. Het effect hangt sterk af van de manier waarop de IT-investeringen worden geïmplementeerd. De literatuur wijst er bovendien op dat het effect van IT samenhangt met andere factoren zoals de organisatiestructuur en menselijk kapitaal.

Het onderzoek naar het effect van ICT-investeringen op de productiviteitsontwikkeling van de publieke sector is schaarser en de resultaten zijn gemengd (van Reenen 2010). Ook een verkennende empirische studie voor deze notitie vindt geen sterk positief verband. De magere empirische resultaten kunnen mede komen door

meetproblemen. In de overheidsadministratie worden alleen ICT-uitgaven bijgehouden. Dit is mogelijk geen goede maatstaf voor de ICT-investeringen, omdat uitgaven aan ICT door instituten verschillend worden behandeld. Een consistente rapportage tussen jaren en instituten ontbreekt. De reeks zou bijvoorbeeld ook de personeelskosten van intern IT-personeel moeten bevatten. Doordat overheidsorganisaties soms gebruik maken van shared-services-organisaties (SSO), wordt het steeds moeilijker de impact van ICT te meten (Moolenaar 2019).

Bovendien worden kwaliteitsverbeteringen door ICT-gebruik onvoldoende meegenomen. Beleidsnota's kunnen sneller worden gemaakt dan in het verleden, maar hierdoor worden ook de eisen in termen van presentatie (figuren, lay-out, uitgevoerde analyses) vergroot. Het aantal nota's is misschien nog steeds hetzelfde, maar de kwaliteit van het rapport is verbeterd. Als alleen het aantal rapporten wordt beschouwd, komt dit niet tot uiting in de productiviteitscijfers.

Ook economische factoren dragen bij aan de magere empirische resultaten, in het bijzonder voor de overheid. Het duurt mogelijk enige tijd voordat de investeringen gaan renderen. Bij ICT-projecten bij de overheid zijn er extra hindernissen, die het effect op de productiviteitsgroei dempen. Dunleavy et al. (2006) benoemt een aantal factoren. Door de complexiteit van de projecten bij de overheid is de organisatie vaak inflexibel. De technische flexibiliteit wordt verminderd door uitbesteding van projecten. Verder zijn er relatief weinig bedrijven die de complexe projecten kunnen uitvoeren. Dit gebrek aan competitie zou de kwaliteit van de ICT-investeringen kunnen verminderen. Daarnaast speelt een rol dat efficiencyoverwegingen moeten worden afgewogen tegen bijvoorbeeld juridische en politieke factoren. De geleverde producten hangen samen met regelgeving die (vaak) kan veranderen. Bij dienstverlening aan burgers gaat het om grote aantallen en de behoeften verschillen. Uit democratische overwegingen kan de overheid niet zoals een commercieel bedrijf groepen burgers uitsluiten, wat een commercieel bedrijf wel kan doen.

Een parlementair onderzoek naar ICT-projecten bij het rijk laat zien dat ook in Nederland de potentie van ICT niet volledig wordt waargemaakt voor de overheid (Kamerstukken 2014). In het rapport worden gebrek aan ICT-kennis, een gebrekkige verantwoordings- en besluitvormingsstructuur bij ICT-projecten en onvoldoende inzicht in de kosten en baten van de projecten als oorzaken voor de soms moeizame implementatie van ICT bij de overheid genoemd.

Voor deze studie is eigen empirisch onderzoek verricht naar het effect van ICT op de productiviteitsgroei. Daaruit blijkt dat het lastig is een positief effect van ICT empirisch te staven. Het gevonden verband tussen ICT-uitgaven en productiviteitsgroei is negatief, maar niet statistisch significant, hieruit kunnen dus geen conclusies worden getrokken. Bovengenoemde meetproblemen bij de werkelijke omvang van de ICT-voorraad en het relatief geringe aantal datapunten zijn een belangrijke oorzaak voor het resultaat van dit onderzoek.

6 Betere indicatoren noodzakelijk

Betere maatstaven zijn noodzakelijk om meer inzicht te krijgen in de productiviteitsontwikkeling en de factoren die eraan ten grondslag liggen. Dit geldt zowel voor het productievolume als voor de inputfactoren arbeid en kapitaal. Er is geen uniforme maatstaf voor alle onderdelen van de overheid. Mede daarom lijkt een

dashboard van indicatoren de beste weg om meer inzicht te krijgen. Het instituut IPSE geeft een voorbeeld van hoe een dergelijk dashboard eruit zou kunnen zien voor individuele organisaties in de publieke sector.⁴

De analyses over productiviteitsgroei worden versterkt als de kwaliteit van de productiefactor arbeid beter in kaart wordt gebracht. In de analyses van deze notitie (en in veel andere studies) zijn de loonuitgaven gebruikt om de productiviteitsontwikkeling te meten, omdat dit cijfer goed beschikbaar is. Dit cijfer geeft echter geen volledig inzicht in de kwaliteit van de productiefactor arbeid. In deze studie is schaalniveau als nadere indicator toegevoegd, maar voor een volledige analyse zal dit cijfer moeten worden aangevuld met gegevens over het opleidingsniveau en het niveau van investeringen in menselijk kapitaal door middel van trainingen en cursussen.

Om het belang van scholing en training voor de overheid goed te kunnen analyseren, zullen de uitgaven aan cursussen en opleidingen ten minste op een uniforme wijze moeten worden geboekt. De vraag is dan nog in hoeverre deze uitgaven een compleet beeld geven van de investeringen in training en scholing. Deze uitgaven zullen over het algemeen geen rekening houden met de interne training en scholing. Bovendien houden deze cijfers geen rekening met de zogenoemde “zachte kennis” die werknemers opdoen, bijvoorbeeld door samen te werken en te leren van collega’s.

Een betere en binnen de overheid éénduidige manier om de ICT-uitgaven te boeken, is een belangrijke voorwaarde om goed onderzoek te doen naar de effectiviteit van ICT. Hierbij zou op zijn minst moeten worden aangegeven hoe wordt omgegaan met outsourcing en welk deel is uitgegeven aan personeel en welk deel aan (het onderhouden van) apparatuur.

⁴ Zie Trends in Publieke Sector (TiPS) ([link](#)).

Schets voor een dashboard

Om de productiviteitsgroei te monitoren, kan het beste een dashboard van indicatoren worden gebruikt. In zo'n dashboard kan rekening worden gehouden met de verschillende dimensies van de productiviteitsontwikkeling van de overheid. De exacte invulling van het dashboard hangt af van het soort instituten dat wordt gemonitord. Hier wordt een schets gegeven van de elementen die een dashboard kan bevatten.

Een dashboard moet allereerst volume-indicatoren bevatten. Deze indicatoren moeten representatief zijn voor de taken van de overheid. Verschillende taken vragen om verschillende indicatoren. Voor onderzoekinstellingen kunnen aantallen publicaties worden gebruikt, voor uitvoeringsinstellingen zoals de Belastingdienst aantallen aangiften. Voor toezichthoudende instanties is het lastiger indicatoren te vinden. In zulke gevallen zouden normindicatoren zoals de omvang van de potentiële doelgroep kunnen worden gebruikt. De inputfactoren arbeid, kapitaal, training en scholing en ICT zullen eveneens in het dashboard vertegenwoordigd moeten zijn.

Het is belangrijk dat het dashboard ook maatstaven bevat over de tevredenheid met de dienstverlening van de overheid en de kwaliteit van de dienstverlening. Kwaliteit is een onderdeel van het productievolume. Verder gaat het bij veel overheidsdiensten niet om de productie, maar om het nut dat aan deze diensten wordt toegeschreven. Bij de politie gaat het niet alleen om het aantal inbrekers dat wordt gepakt, maar ook om de vraag of burgers zich veilig voelen.

Een dashboard kan gebruikt worden om de witte vlekken in de meting te identificeren. Uit pragmatische overwegingen zal het dashboard eerst gevuld worden met indicatoren die beschikbaar zijn. De huidige situatie kan vervolgens aangevuld worden met de theoretisch optimale situatie. Op deze wijze kan worden nagegaan wat het meest kansrijk is om de meting te verbeteren.

Om de indicatoren van verschillende instituten te kunnen aggregeren in een dashboard is het noodzakelijk dat het productievolume en de inputfactoren waar mogelijk op éénduidige wijze worden gemonitord. Om recht te doen aan de heterogeniteit van de overheid zullen wel een aantal typen instellingen moeten worden onderscheiden.

Voor een analyse op macroniveau is het raamwerk van de Nationale Rekeningen zinvol. Dit biedt een handvat om aan te sluiten bij andere macro-economische grootheden. Daarom is het zinvol op microniveau in eerste instantie te zoeken naar omzetten en prijzen. Er kunnen ook indicatoren op bedrijfstakniveau worden verzameld of gekeken worden naar de verschillende functies van de overheid. Deze indicatoren kunnen vervolgens worden gekoppeld aan relevante macro-economische begrippen (lonen, investeringen) die door het CBS eveneens op dit niveau worden verzameld.

Een dashboard is vooral een analyse- en monitoringsinstrument en is geen vervanging van een meting van de productiviteitsgroei. Bij een dashboard is het nadrukkelijk niet de bedoeling de dimensies op te tellen tot één cijfer. Als een dashboard wordt opgesteld, moet het zoeken naar een goede manier om productiviteitsgroei te meten derhalve onverminderd doorgaan.

7 Beleidsexperimenten nodig

Gerandomiseerde beleidsexperimenten zijn de beste manier om meer inzicht te krijgen in wat de invloed is van intensiveringen in training of ICT-faciliteiten op de productiviteitsgroei van overheidsdiensten. Experimenten kunnen een oorzakelijk verband aantonen tussen een verklarende variabele (door middel van een interventie) en de productiviteitsgroei van de overheid. In principe kunnen experimenten gebruikt worden voor alle in deze notitie genoemde factoren, dus ook voor bijvoorbeeld managementkwaliteit en organisatiecultuur.⁵ Omdat training en ICT als belangrijke factoren zijn geïdentificeerd, is het zinvol om in eerste instantie hierop de aandacht te richten bij het opzetten van experimenten.

Hoe werkt een gerandomiseerd beleidsexperiment? Kernpunt is dat de medewerkers van sommige overheidsinstellingen bijvoorbeeld deelnemen aan een data-science training en medewerkers van andere soortgelijke overheidsorganisaties niet. Deelnemen aan de training wordt dan een treatment genoemd. Welke organisatie een treatment krijgt en welke de controlegroep is, wordt door loting bepaald. Na afloop van de training worden de uitkomsten van de organisaties waarvan de medewerkers een training gevolgd hebben, vergeleken met de organisaties waarvan de medewerkers die de training niet hebben gevolgd. Het verschil tussen de uitkomsten meet dan het effect van de training. Eenzelfde soort experiment kan worden opgezet voor het aanbieden van extra ICT-faciliteiten om het effect van ICT-faciliteiten op productiviteitsontwikkeling te onderzoeken.

Het ligt het meest voor de hand een beleidsexperiment op organisatieniveau uit te voeren. Te denken valt aan onderdelen van verschillende ministeries of gemeenten die dezelfde taken uitvoeren of kantoren van uitvoeringsinstanties als belastingen, douane of UWV. Zowel de controlegroep als de treatmentgroep moeten vanuit statistisch oogpunt voldoende waarnemingen hebben. Beleidsexperimenten kunnen ook op medewerkersniveau worden uitgevoerd, maar het is zeker voor het organisaties bij het openbaar bestuur lastig op individueel niveau een goede productiviteitsmaatstaf te vinden.

Voorwaarde voor het laten slagen van een beleidsexperiment is dat er een goede maatstaf voor productiviteit van de respectievelijke organisaties beschikbaar is. Daarnaast is het belangrijk dat de interventiegroep (de groep die bijvoorbeeld een training ondergaat) en de controlegroep statistisch gezien in alle aspecten vergelijkbaar zijn, met als enige uitzondering dat de ene groep wél deelneemt aan de interventie en de andere groep niet (Duflo et al. 2006). Deze vergelijkbaarheid wordt bewerkstelligd door de gerandomiseerde toewijzing (loting) van de interventie. Het is dus niet mogelijk dat medewerkers zelf bepalen of ze deelnemen aan de training. Zo zou er immers door zelfselectie een verschil kunnen ontstaan tussen de interventiegroep en de controlegroep die mogelijk ook de uitkomsten beïnvloedt.

Gerandomiseerde experimenten kennen overigens ook nadelen. De externe validiteit bij 'kleine' experimenten is beperkt. Dat wil zeggen dat de resultaten op basis van enkele organisaties mogelijk niet representatief zijn voor de gehele overheid.

Al met al zijn beleidsexperimenten een veelbelovende weg om te bewandelen. Beleidsexperimenten geven de overheid de kans om verschillende, uit de literatuur succesvolle, interventies uit te testen in de praktijk en

⁵ Er is een aanzienlijke literatuur over beleidsexperimenten in de context van organisatiecultuur en managementkwaliteit (zie b.v. Bloom en van Reenen 2011), waarbij ook variabelen worden bestudeerd die samenhangen met productiviteit van de medewerkers zoals ziekteverzuim. Lastig punt bij deze studies is wel dat twee verklarende factoren worden opgevoerd (organisatiecultuur en kwaliteit management) die zelf ook weer lastig zijn te meten. Dit bemoeilijkt het opzetten van experimenten.

daaruit de meest effectieve interventies te destilleren. De interventies die worden gekenmerkt door de grootste kosteneffectiviteit kunnen dan worden opgeschaald naar andere overheidsorganisaties. Op deze manier kan de productiviteitsgroei van de overheid worden gestimuleerd en krijgt de Nederlandse overheid handvatten aangereikt om Baumol te verslaan.

Literatuur

- Ahmed, T. en A.A. Bhatti, 2020. Measurement and determinants of multifactor productivity: a survey of literature, *Journal of Economic Surveys*, vol. 34 (2): 293-319.
- Alfonso, A. en S. Fernandes, 2008, Assessing and explaining the relative efficiency of local government, *The Journal of Socio-Economics*, vol. 37: 1946-1979.
- Baumol, W.J., 1967, Macroeconomics of unbalanced growth: the anatomy of urban crisis, *The American Economic Review*, vol. 57(3): 415-426.
- Blank, J., 2015. Illusies over fusies: een kritische beschouwing over de schaalvergroting in de Nederlandse publieke sector 1985-2012. IPSE studies.
- Blank, J.L.T. en A.S. van Heezik, 2019, Een essay over de relatie tussen beleid en productiviteit in onderwijs, zorg, veiligheid & justitie en netwerksectoren, IPSE studies.
- Bloom, N. en J. van Reenen, 2011, Human resource management and productivity, in: Orley Ashenfelter en David Card (eds), *Handbook of Labor Economics*, vol. 4B, chapter 19, 1697-1769.
- Blundell, R., L. Dearden, C. Meghir en B. Sianesi, 1999, Human capital investment: The returns from education and training to the individual, the firm and the economy, *Fiscal Studies*, vol. 20 (1) : 1-23.
- Botev, J., B. Egert, Z. Smidova en D. Turner, 2019, A new macroeconomic measure of human capital with strong empirical links to productivity, OECD Economics Department Working Papers 1575.
- Bright, J. en H. Margetts, 2016, Big data and public policy: can it succeed where e-participation has failed, *Policy & Internet*, vol. 8(3): 218-223.
- CPB, 2017, Productiviteitsontwikkeling van de Nederlandse overheid. CPB notitie, 8 november 2017.
- CPB, 2018, Productiviteitsgroei Nederlandse overheid: twee methodes vergeleken, CPB notitie, 7 december 2018.
- Douglas, S. en A. Meijer, 2016. Transparency and public value – analyzing the transparency practices and value creation of public utilities, *International Journal of Public Administration*, vol. 39 (12): 940-951.
- Duflo, E., R. Glennerster en M. Kremer, 2006, Using Randomization in Development Economics Research: A Toolkit, NBER Technical Working Papers 0333, National Bureau of Economic Research, Inc.
- Dunleavy, P., H. Margetts, S. Bastow, en J. Tinkler, 2006, *Digital era Governance*, Oxford University Press.
- Hüseyin, T. en F. Gönel, 2020, ICT labor, software usage and productivity: firm-level evidence from Turkey, *Journal of Productivity Analysis*, vol. 53: 265-285.

IPSE, 2012, Dumaij, A.C.M. en J.A. Wilschut, Waterzuivering transparant: een empirisch onderzoek naar de productiviteit en doelmatigheid van de Nederlandse rioolwaterzuiveringsinstallaties, 2002-2009, IPSE Studies, november 2012.

IPSE, 2013, Blank, J.L.T., Arbeidsproductiviteit publieke sector in historisch perspectief: de aard van technologische ontwikkelingen, *Tijdschrift voor Openbare Financiën*, vol. 45(3): 135-163.

Kalb, A., B. Geys en F. Heinemann, 2012, Value for money? German local government efficiency in a comparative perspective, *Applied Economics*, vol. 44(2): 201-218.

Kamerstukken, 2014. Parlementair onderzoek naar ICT-projecten bij de overheid, Kamerstukken, 15 oktober, <https://www.tweedekamer.nl/kamerstukken/detail?id=2014Z17985&did=2014D36603>.

Konings, J. en S. Vanormelingen, 2015, The impact of training on productivity and wages: firm-level evidence, *The Review of Economics and Statistics*, vol. 97 (2): 485-497.

Lucas, R.E., 1988. On the Mechanics of Economic Development, *Journal of Monetary Economics*, vol. 22(1): 3-42.

Matko, A. en T. Takacs, 2017, Examination of the relationship between organizational culture and performance, *International Review of Applied Sciences and Engineering*, vol. 8(1): 99-105.

Moolenaar, D.E.G., 2019. Productiviteitsontwikkeling in het justitieveld: een verkenning van de mogelijkheden. WODC Cahier 2019-5.

Niaounakis, T.K., A.S. van Heezik en J. Blank, 2020, Productiviteit uitvoering sociale zekerheid: een analyse van de productiviteitsontwikkeling van de overheid bij UWV en SVB tussen 2002 en 2018. IPSE studies, juni 2020.

Niaounakis, T.K., A.S. van Heezik en D. Deurloo, 2019, Op afstand de beste? Een analyse van de productiviteitsontwikkeling bij IND, CJIB, SVB, RDW en het Kadaster. IPSE studies, december 2019.

Niaounakis, T.K., T. Urlings en A.S. van Heezik, 2014, Productiviteitstrends in het gevangeniswezen, Een empirisch onderzoek naar het effect van regulering op de productiviteitsontwikkeling tussen 1970 en 2012, IPSE studies, oktober 2014.

Palvalin, M., A. Lönnqvist en M. Vuolle, 2013, Analysing the impacts of ICT on knowledge work productivity, *Journal of Knowledge Management*, vol. 17(4): 545-557, <https://doi.org/10.1108/IKM-03-2013-0113>.

Reenen, J. van, 2010, The economic impact of ICT, Center for Economic Performance.

Sasso, S. en J. Ritzen, 2019, Sectoral cognitive skills, R&D, and productivity: a cross-country cross-sector analysis, *Education Economics*, vol. 27(1): 35-51.

SCP, 2012, Waar voor ons belastinggeld? Prijs en kwaliteit van publieke diensten, Sociaal en Cultureel Planbureau, Den Haag, januari 2012.

SCP, 2015a, Maten voor gemeenten 2014: Prestaties en uitgaven van de lokale overheid in de periode 2007-2012, Sociaal en Cultureel Planbureau, Den Haag, maart 2015.

SCP, 2015b, Public sector achievement in 36 countries: A comparative assessment of inputs, outputs and outcomes, Sociaal en Cultureel Planbureau, Den Haag, december 2015.

SCP, 2019, Publiek Voorzien. <https://digitaal.scp.nl/publiekvoorzien/> .

Severo, M., A. Feredj en A. Romele, 2016, Soft data and public policy: can social media offer alternatives to official statistics in urban policymaking, *Policy & Internet*, vol. 8(3): 354-372.

Syverson, C., 2011. What determines productivity, *Journal of Economic Literature*, vol. 49(2): 326-365.

Twizeyimana, J.D. en A. Andersson, 2019. The public value of E-government – a literature review, *Government Information Quarterly*, vol. 36: 167-178.

Bijlage 1 Empirische studie

Bijlage 1.1 Data en methode

Van 133 instellingen van de rijksoverheid zijn voor de periode 2010-2017 gegevens verzameld voor de inputfactoren arbeid en kapitaal. De gegevens komen uit begrotingen en jaarverslagen (openbare bronnen). De dataset bevat kerndepartementen, dienstonderdelen van kerndepartementen, zelfstandige bestuursorganen (ZBO), agentschappen en rechtspersonen met een wettelijke taak (RWT). Beschikbaar zijn onder andere de arbeidskosten, de kapitaalkosten en de ICT-kosten.

Voor de 33 grootste instellingen in termen van aantal werknemers zijn outputindicatoren verzameld. Het productievolume van de overheid kan niet op de gangbare manier worden gemeten, maar moet worden geschat met relevante indicatoren. Er zijn geen standaard regels over welke indicatoren moeten worden opgenomen en hoe deze moeten worden gepresenteerd. Alleen voor de 33 grootste organisaties zijn data verzameld, waarbij gebruik is gemaakt van de indicatoren die openbaar zijn. Voor kleinere instituten is het lastiger bruikbare openbare indicatoren te vinden. Bovendien is voor een conclusie op macroniveau voldoende dat de instellingen gezamenlijk een substantieel deel van de dataset vormen. De 33 organisaties bevatten samen ruim 80% van het totale aantal fte's van de oorspronkelijke dataset van de 133 instellingen. De dataset met outputindicatoren bevat omzetcijfers, tarieven, (kost)prijsindices, cijfers over het aantal verleende producten of diensten en/of kwaliteitsindicatoren.⁶

De dataset is voor de studie van 2019 verder aangevuld met hr-gegevens van de instituten. Deze gegevens worden beschikbaar gesteld door organisatie P-direkt. Uit P-direkt halen we geaggregeerde gegevens over de leeftijdsopbouw, het schaalniveau en het geslacht van de werknemers van een instituut. Alleen voor instellingen die vallen onder de cao-rijk zijn de data uit P-direct beschikbaar. Dit zijn 15 instellingen.

We nemen de productiviteitsschatting op basis van productie-indicatoren als uitgangspunt, omdat dit het beste aansluit bij wat in de empirische literatuur gangbaar is. We kijken naar de ontwikkeling van de totale factor productiviteit. De productiviteitsgroei wordt berekend als een Solow-residu.

Om de determinanten van productiviteitsgroei vast te kunnen stellen wordt een regressieanalyse gedaan met fixed-effects op organisatieniveau. De fixed effects dienen om te corrigeren voor organisatie-specifieke factoren die de productiviteitsgroei mogelijk beïnvloeden. De geschatte productiviteitsgroei is hierbij de afhankelijke variabele. De verklarende variabelen zijn onder andere de reële groei van de ICT uitgaven, de leeftijd van de werknemers en het niveau van scholing en training. Schaalniveau van werknemers wordt gebruikt als een proxy voor scholing en training. Hierbij worden vier categorieën onderscheiden en wordt het laagste niveau in de regressieanalyse als referentiecategorie gebruikt.

Vanwege de beperkte beschikbaarheid van de hr-gegevens, slinkt het aantal observaties in de dataset. De resultaten van de regressieanalyse dienen derhalve slechts als indicatief bewijs.

⁶ De gebruikte productie indicatoren per instituut zijn op te vragen bij de auteurs.

Bijlage 1.2 : achtergronden bij de variabelen

Inputfactoren

- We gebruiken de gegevens van de totale arbeidskosten en kapitaalkosten uit de jaarverslagen. Gegevens over de prijzen van de inputfactoren zijn niet beschikbaar. Om de volumecijfers te bepalen, worden deze cijfers gedeïflecteerd met de prijs van de beloning van werknemers van de overheid (prijs overheidsconsumptie, beloning werknemers van het CBS) en de prijs van de overheidsinvesteringen (iboi) van het CBS.
- Het percentage werknemers van een instituut dat in een bepaalde leeftijdscategorie valt wordt weergegeven. Er worden vier leeftijdscategorieën onderscheiden: < 35 jaar, 35-45 jaar, 45-55 jaar, > 55 jaar.
- Het percentage werknemers van een instituut dat in een bepaald schaalniveau valt. Er zijn 4 categorieën: < schaal 7, schaal 8-9, schaal 10-13, > schaal 14.
- De reële groei van de ICT-uitgaven.

Outputindicatoren

De volgende algemene regels zijn gebruikt voor het bepalen van het productievolume:

- Idealiter willen we zo dicht mogelijk aansluiten bij de conventies van de Nationale Rekeningen. Dat houdt in dat we in eerste instantie zoeken naar omzetten en kostprijzen.
- Soms zijn voor een organisatie omzetten en tarieven beschikbaar. In dat geval wordt het volume bepaald door de omzet te defleren met een prijsindex van het CBS voor de specifieke bedrijfstak. Daarnaast worden de tarieven zoals gerapporteerd door de organisatie zelf, gebruikt om de omzet te defleren.
- Soms zijn er geen omzetten, maar wel aantallen beschikbaar. Bijvoorbeeld aantallen publicaties of aantal belastingaangiften. In dat geval is het productievolume bepaald op basis van deze aantallen.
- Als het mogelijk is, is het productievolume bepaald met twee methoden (zowel gedeïflecteerd omzetten als aantallen). Het productievolume is gelijkgesteld aan het gemiddelde van de geïndexeerde variabelen.

Kanttekeningen bij de data

- Vanwege de transparantie is het gebruik van openbare bronnen wenselijk. Daarom is er voor gekozen data te gebruiken uit jaarverslagen en andere openbare bronnen. Verder kan op deze wijze een groot aantal verschillende instellingen worden meegenomen.
- Bij de keuze van de productie-indicatoren is geprobeerd zoveel mogelijk aan te sluiten bij de conventies van de Nationale Rekeningen en wat gangbaar is bij studies naar de productiviteitsontwikkeling van de overheid. Deze indicatoren moesten snel en makkelijk beschikbaar zijn. We hebben op basis van expert judgement een selectie gemaakt uit beschikbare indicatoren, maar geen additionele correcties gemaakt voor exogene schokken zoals bijvoorbeeld een vluchtelingen crisis (wat leidt tot grote instroom van asielzoekers) of de stand van de conjunctuur.
- We hebben de beperkingen van de data proberen te ondervangen door als dat mogelijk is met verschillende indicatoren het productievolume te bepalen, nauwkeurig te onderzoeken wat de oorzaak is van sterke fluctuaties en verschillende methoden te gebruiken om met uitbijters (sterk afwijkende datapunten) om te gaan.

Bijlage 1.3: Determinanten productiviteitsgroei overheid

Specificatie	(1)	(2)	(3)	(4)	(5)	(6)
Leeftijd 35-45	-0.230 (0.775)			0.489 (0.896)	-0.197 (0.804)	0.640 (0.800)
Leeftijd 45-55	0.512 (0.778)			0.638 (0.990)	-0.478 (0.973)	0.990 (0.982)
Leeftijd 55 +	0.325 (1.106)			0.611 (1.307)	-0.082 (1.145)	0.877 (1.170)
Schaal 8-9		2.814 (1.758)		2.516* (1.405)	2.619* (1.348)	2.598* (1.366)
Schaal 10-13		1.011 (0.869)		0.972 (0.855)	1.099 (0.882)	1.156 (0.922)
Schaal 14 +		-6.969 (5.492)		-7.569 (5.935)	-7.362 (5.558)	-7.187 (5.288)
ICT groei			-0.066 (0.069)		-0.095 (0.089)	
ICT medium						-4.167 (2.830)
ICT hoog						-5.857 (3.770)
Constant	-19.549 (62.075)	-68.596 (54.555)	0.408 (0.254)	-108.204 (104.538)	-129.337 (104.308)	-140.776 (109.063)
Observaties	60	60	60	60	60	60
R-squared	0.030	0.178	0.018	0.193	0.226	0.257
Panelvar	15	15	15	15	15	15
(i) Panel regressies met fixed effects op organisatieniveau. (ii) Robuuste standaard errors tussen haakjes. (iii) *** p<0.01, ** p<0.05, * p<0.1						