



Centraal Planbureau

CPB Notitie | 07 oktober 2015

Een internationale literatuurstudie naar hoe efficiënt gemeenten met hun geld omgaan



CPB Notitie

Aan: Kennisunit Infrastructuur en Ruimtelijke Economie (KIRE)

Centraal Planbureau
Van Stolkweg 14
Postbus 80510
2508 GM Den Haag
T (070)3383 380
I www.cpb.nl

Contactpersonen
Joost van Gemeren
Sander Hoogendoorn
Peter Zwaneveld

Datum: 7 oktober 2015

Betreft: Een internationale literatuurstudie naar hoe efficiënt gemeenten met hun geld omgaan

Samenvatting en conclusies

Inleiding

Op verzoek van de Kennisunit Infrastructuur en Ruimtelijke Economie (KIRE) verkent deze notitie de internationale empirische literatuur rondom de vraag in hoeverre gemeenten op een economisch efficiënte wijze met hun geld omgaan. De focus ligt op twee vormen van efficiëntie onder gemeenten. Enerzijds is de vraag in hoeverre gemeenten de kwantiteit en kwaliteit van diensten leveren naar behoefte van de burger (allocatieve efficiëntie), anderzijds of dit tegen de laagst mogelijke kosten gebeurt (technische efficiëntie of doelmatigheid).

Deze notitie volgt de methoden die in de literatuur het meest worden toegepast. Allocatieve efficiëntie wordt in de literatuur getoetst aan de hand van vastgoedprijzen. Deze weerspiegelen namelijk de aantrekkelijkheid van een gemeente: die is maximaal als de gemeente efficiënte keuzes maakt met betrekking tot het dienstenpakket. Technische efficiëntie wordt gemeten door de zogeheten benchmarkmethode (*frontier analysis*). Benchmarking stelt verschillen in doelmatigheid vast door de kosten van een bepaalde gemeente te vergelijken met die van de goedkoopst producerende gemeenten. Het gaat hier dus om een maat van relatieve (en niet per se absolute) efficiëntie.

De belangrijkste bevindingen zijn als volgt:

1. Er is weinig bekend over allocatieve efficiëntie. Dat wat bekend is, geeft een voorzichtig positief signaal af

De enige studie over Nederland suggereert dat gemeenten het dienstenpakket aanbieden naar behoefte van hun burgers. De buitenlandse literatuur is verdeeld over de allocatieve efficiëntie van gemeenten: ongeveer de helft vindt bewijs dat gemeenten

een te sober of te ruim dienstenpakket aanbieden naar de behoefte van burgers. Gemeenten lijken dus niet in alle situaties allocatief efficiënt.

2. Verschillen in technische efficiëntie tussen gemeenten zijn klein

De literatuur geeft aan dat gemeenten in Noordwest-Europa niet veel zwakker presteren dan de technisch meest efficiënte (meest doelmatige) gemeenten. De gemiddelde gemeente moet circa 12 procent van de kosten terugdringen om ook tot de top te behoren, zonder dat dit ten koste gaat van het dienstenpakket. Dat percentage is redelijk laag te noemen, zeker in internationaal perspectief. De drie bestaande studies over Nederland vinden een nog kleiner verschil in doelmatigheid tussen gemeenten. Deze Nederlandse studies betreffen een specifiek beleidsdomein; er bestaan geen studies over de technische efficiëntie van Nederlandse gemeenten in het algemeen.

3. Op het gebied van technische efficiëntie is vervolgonderzoek over Nederland kansrijk

Deze notitie identificeert twee verder te ontwikkelen onderzoekstakken, beide met betrekking tot technische efficiëntie. Als eerste is onderzoek naar de algehele technische efficiëntie van Nederlandse gemeenten van toegevoegde waarde. De uitkomsten daarvan bieden bijvoorbeeld waardevolle inzichten voor overkoepelende beleidsvraagstukken, zoals de mate van decentralisatie en het lokale belastinggebied. Ook verdienen benchmarkonderzoeken op meer beleidsdomeinen aanbeveling, om de vele benchmarks die in de beleidssfeer circuleren, wetenschappelijk te verifiëren.

1 Inleiding

Deze notitie vloeit voort uit overleg met vertegenwoordigers van de ministeries van Binnenlandse Zaken en Koninkrijksrelaties (BZK) en Infrastructuur en Milieu (IenM), en is ingebed in de Kennisunit Infrastructuur en Ruimtelijke Economie (KIRE). De inzichten uit deze notitie zijn mogelijk ook van belang voor de ministeries van Economische Zaken (EZ) en Financiën (MinFin), alsmede voor de Algemene Rekenkamer en de Vereniging van Nederlandse Gemeenten (VNG).

Aan de hand van een internationale empirische literatuurstudie onderzoekt deze notitie hoe efficiënt gemeenten omgaan met hun geld, door te kijken naar twee vormen van efficiëntie onder gemeenten. Enerzijds dient een gemeente te bepalen welke diensten zij aanbiedt, en per dienst wat de kwaliteit daarvan moet zijn. Conform de wetenschappelijke literatuur noemt deze notitie dit de allocatieve functie van gemeenten. Een gemeente is allocatief efficiënt als zij diensten levert naar de behoefte van haar burgers. Anderzijds is het zaak om het gekozen pakket aan diensten tegen lage kosten te leveren. De literatuur beschrijft dit met de term technische efficiëntie.¹ De focus van deze notitie ligt op de efficiëntie van het gehele gemeentelijke dienstenpakket. Het doel is niet om aan te wijzen op welke specifieke beleidsdomeinen gemeenten stappen kunnen maken, of wat de onderliggende oorzaken zijn van allocatieve of technische efficiëntie.

Allocatieve efficiëntie

Allocatieve efficiëntie wordt in de literatuur gemeten met behulp van de vastgoedwaarde in een gemeente. Het achterliggende idee daarbij is dat de vastgoedwaarde de aantrekkelijkheid van een gemeente voor bewoners weerspiegelt. Als de gemeente voor haar bewoners een betere samenstelling, kwaliteit of omvang van diensten aanbiedt, zal de vastgoedwaarde hoger zijn.

De literatuur bevat voornamelijk onderzoeken uit het buitenland. Er bestaat één studie naar de allocatieve efficiëntie van Nederlandse gemeenten die gebruik maakt van de beschreven methode (Allers en Vermeulen, 2013a;b). Hoofdstuk 2 van deze notitie tracht onder meer in kaart te brengen hoe de resultaten van Allers en Vermeulen zich verhouden tot de internationale literatuur.

Technische efficiëntie

Technische efficiëntieverschillen tussen gemeenten worden in de wetenschappelijke literatuur vooral bepaald op basis van de benchmarkmethode (*frontier analysis*). Deze methode vergelijkt de kosten van elke gemeente met de goedkoopst producerende

¹ Om de notitie inzichtelijker te maken wordt technische efficiëntie bij de analyse van allocatieve efficiëntie als een gegeven verondersteld, en vice versa. Dit doet de complexe werkelijkheid niet volledig recht, maar is wel in lijn met de literatuur. Technische efficiëntie kan bijvoorbeeld ook worden gezien als onderdeel van allocatieve efficiëntie: als dienst A heel erg duur is om te produceren, zal deze dienst geen onderdeel zijn van het allocatief efficiënte dienstenpakket. Het allocatief efficiënte dienstenpakket weegt namelijk zowel kosten als baten mee.

(vergelijkbare) gemeenten. De literatuur toetst dus technische efficiëntieverschillen. Benchmarks kunnen daarom niet uitsluiten dat de best presterende gemeente in de praktijk nog efficiënter kan produceren.

De algehele technische efficiëntieverschillen onder Nederlandse lokale overheden zijn nog nooit onderzocht. Dit constateerde het CPB eerder al (CPB, 2007) en daar is tot op heden nog geen verandering in gekomen. Dat geldt niet voor andere landen. Verder bestaan er vele analyses van individuele beleidsdomeinen en bedrijfsonderdelen van Nederlandse gemeenten. Dit is voornamelijk toegepast onderzoek, waarvan de rapportage niet eenvoudig te vinden is en de kwaliteit moeilijk verifieerbaar. Hoofdstuk 3 van deze notitie focust daarom op onderzoek naar de technische efficiëntie van gemeenten dat is verschenen in de wetenschappelijke literatuur.

Andere prestatiecriteria voor gemeenten

De prestaties van gemeenten kunnen, naast allocatieve en technische efficiëntie, ook op andere manieren worden gemeten. Zo bestaan er verschillende zogeheten productiviteitsonderzoeken, die draaien om de verhouding tussen productie en kosten. Een voorbeeld daarvan is Maten voor Gemeenten van het SCP (2015). Om de kwaliteit van de dienstverlening mee te nemen in het productiviteitsonderzoek, maakt het SCP onder meer gebruik van burgers tevredenheid (zie bijvoorbeeld www.waarstaatjegemeente.nl). Een andere optie is het zogeheten effectiviteitsonderzoek: is een bepaalde interventie effectief of niet? (zie bijvoorbeeld Dur en Vollaard, 2013; Gradus en Dijkgraaf, 2014; Koning, 2014). Productiviteits- en effectiviteitsonderzoek verschillen van allocatieve en technische efficiëntie, maar zijn onderling wel gerelateerd (zie voor een nadere uitleg: CPB, 2007; Blank et al., 2010).

De Atlas voor Gemeenten (Marlet en Van Woerkens, 2014) is nog een voorbeeld om de prestaties van Nederlandse gemeenten te meten. Hierin komt de ontwikkeling en rangschikking van de vijftig grootste gemeenten in Nederland aan bod voor wat betreft bevolking, sociaaleconomische positie en woonaantrekkelijkheid. Een verschil met de twee efficiëntieconcepten in deze notitie is dat de Atlas voor Gemeenten niet naar kosten kijkt, maar wel uitgebreider ingaat op de specifieke baten van gemeentebeleid.

Een (lokale) overheid moet ook meer zaken voor ogen houden dan alleen efficiëntie. De andere welvaartseconomische overheidsdoelen gelijkheid, stabiliteit en groei zijn 'niet ondergeschikt maar nevensgeschikt' aan efficiëntie (Boorsma en Allers, 2006, p.16). Naast het eindresultaat zelf is de gemeente ook verantwoordelijk voor de manier waarop ze die tracht te bereiken.

2 Gemeentelijke allocatieve efficiëntie

De gemeente is de overheid die het dichtst bij de burger staat. Dit brengt verantwoordelijkheid met zich mee. Zij wordt onder meer geacht om de juiste samenstelling, omvang en kwaliteit (hierna aangeduid met pakket) van diensten samen te stellen voor haar bewoners, naast de wettelijke taken die zij van het Rijk krijgt.

Een allocatief efficiënt pakket van diensten laat de behoefte van burgers leidend zijn. Dit betekent dat voor iedere gemeentelijke dienst geldt dat de marginale baten voor inwoners gelijk zijn aan de marginale kosten om de dienst te leveren (Samuelson, 1954; 1955). Dit

hoofdstuk bespreekt de vraag: *in hoeverre leveren gemeenten een allocatief efficiënt dienstenpakket?*

Het antwoord op die vraag komt voort uit de wetenschappelijke internationale empirische literatuur. In de gevonden artikelen wordt gebruik gemaakt van de toets van Brueckner (1979, 1982). De door hem ontwikkelde toets voor allocatieve efficiëntie wordt tot op de dag van vandaag veelvuldig toegepast.

In het overzicht van de literatuur gaat speciale aandacht uit naar het artikel van Allers en Vermeulen (2013a;b). Dit is de enige studie die de allocatieve efficiëntie van Nederlandse gemeenten onderzoekt. De meeste andere studies hebben betrekking op gemeenten uit de Verenigde Staten. Dat is een beperking van dit onderzoek: Nederlandse gemeenten kennen immers een ander takenpakket en andere financiële verhoudingen dan Amerikaanse gemeenten. Desondanks bevat de literatuur ook voor de Nederlandse situatie waardevolle inzichten.

2.1 Vastgoedprijzen meten allocatieve efficiëntie

2.1.1 Theorie

Vastgoedprijzen zijn over het algemeen een goede maatstaf voor de aantrekkelijkheid van een gemeente. Als burgers een gemeentelijke besteding waarderen, slaat deze doorgaans neer in de vastgoedprijzen (Oates, 1969; zie ook Stad en Land: De Groot et al., 2010).² Als een gemeente bijvoorbeeld besluit om de veiligheid te verbeteren, zal de kwaliteit van leven in die gemeente vooruitgaan. Dat resulteert in meer vraag naar woningen aldaar, met als gevolg een hogere vastgoedprijs.

Uiteraard heeft een verruiming van het dienstenpakket ook een keerzijde: het belastingtarief moet omhoog. Althans, zodra de uitkeringen vanuit de Rijksoverheid besteed zijn.³ Belastingen maken de gemeente onaantrekkelijker om te wonen, en daarom volgt een daling van de vastgoedprijs bij een toename van bijvoorbeeld het ozb-tarief. Als de gemeente een uitbreiding van het pakket aan diensten volledig financiert met rijksuitkeringen, gaat deze conclusie niet op. De methode die past bij financiering via uitkeringen vanuit de Rijksoverheid is echter zeer vergelijkbaar met de hieronder beschreven methode (zie paragraaf 2.3).

Er werken dus twee tegenovergestelde effecten bij een investering in diensten voor de burger. Enerzijds stijgt de vastgoedprijs doordat de burger de vruchten plukt van een ruimer pakket aan diensten. Anderzijds wordt de burger in de portemonnee geraakt, waarop een daling van de vastgoedwaarde volgt. Bij elke dienst is het de vraag of een extra

² Chaudry-Shah (1988), Ross en Yinger (1999), Hoyt en Garen (2005) en Hilber (2011) hebben inzichtelijke literatuuroverzichten van de studies op dit gebied gemaakt.

³ De inkomsten van gemeenten bestaan momenteel voor minder dan 10% uit eigen inkomsten (CPB, 2015). Veruit het meeste geld komt uit rijksuitkeringen.

investering de aantrekkelijkheid van een gemeente, gemeten door de vastgoedprijs, per saldo verhoogt of verlaagt.

De toets van Brueckner (1979, 1982) voor allocatieve efficiëntie baseert zich op deze tegenwerkende krachten. Hij stelt dat gemeenten een efficiënt pakket aan diensten leveren als de vastgoedprijzen maximaal zijn.⁴ Dit gebeurt zodra er niet te weinig en niet te veel wordt uitgegeven aan elke gemeentelijke dienst. Als de gemeente te weinig levert, is de aantrekkelijkheid die een extra investering oplevert ruim voldoende om te compenseren voor de hogere belasting. Bij te veel diensten overheerst echter de onaantrekkelijkheid van stijgende belastingen.

Deze redenering kan worden verduidelijkt door het voorbeeld van gemeentelijke afvalverwerking. Als de gemeente nooit afval zou laten ophalen, ligt er overal afval op straat. In dat geval zouden burgers het zeker waarderen als de gemeente een keer per maand de vuilnis komt ophalen. De bijbehorende belastingverhoging nemen zij dan waarschijnlijk voor lief. Maar als de vuilnis vandaag al is opgehaald, wordt de straat niet veel schoner als morgen de vuilniswagen opnieuw langs komt. De belastingen nemen echter wel toe. Het allocatief efficiënte aantal vuilnisrondes zal dus ergens tussen maandelijks en dagelijks liggen, afhankelijk van hoe schoon burgers de straat willen hebben.

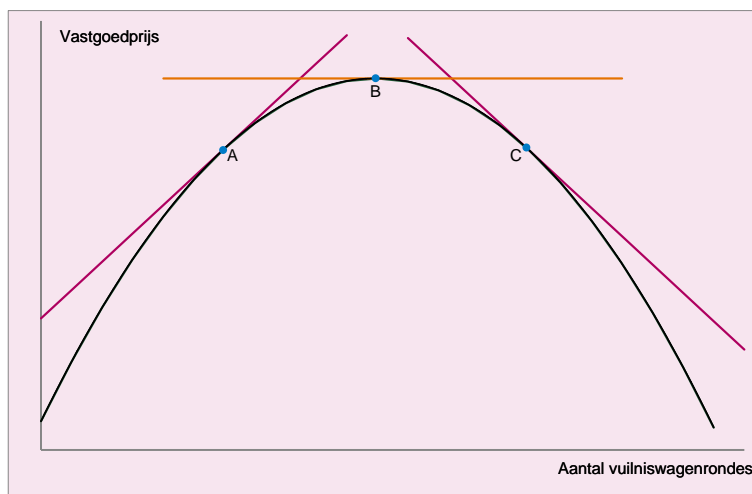
Figuur 2.1 dient ter illustratie van Brueckners stelling. De efficiënte frequentie van de vuilniswagen is aangegeven in punt B. Daar is de waarde van het vastgoed maximaal. In punt A komt de vuilniswagen niet frequent genoeg langs (onderprovisie). In punt C is de vastgoedwaarde even hoog als in punt A, maar in dat geval komt de vuilniswagen juist te vaak langs (overprovisie). Het idee is om alle andere diensten die gemeenten leveren op eenzelfde manier te evalueren.

De stelling van Brueckner kan niet alleen op de kwantiteit maar ook op de kwaliteit van de geleverde diensten worden toegepast. Als de gemeente bijvoorbeeld alleen restafval ophaalt, maar burgers ook hun plastic afval (tegen een verhoging van de belastingen) kwijt willen, is er eveneens sprake van een te sober dienstenpakket. In dat geval bevindt het dienstenniveau zich in punt A, ook al is de frequentie van de vuilnisronde efficiënt.

Ook het volledige dienstenpakket kan op deze manier behandeld worden. Dat is wat dit hoofdstuk doet.

⁴ In een latere publicatie bewijst Brueckner (1983) eveneens het omgekeerde: het streven om de lokale vastgoedwaarde te maximaliseren leidt tot lokale allocatieve efficiëntie. Deze strategie leidt echter niet per se tot een economiebrede allocatieve efficiëntie. Daarvoor is ook een efficiënte verdeling van burgers over de verschillende gemeenten nodig. Ook Stad en Land (De Groot et al., 2010) beschrijft een model waarin maximalisatie van grondwaardesurplus leidt tot allocatieve efficiëntie. Daarnaast brengt vastgoedmaximalisatie in de meeste gevallen herverdelingseffecten met zich mee: huisbezitters profiteren immers van een hogere vastgoedwaarde, maar huurders (in de private sector) bijvoorbeeld niet (Allers en Vermeulen, 2013a;b).

Figuur 2.1 Illustratie van Brueckners toets voor allocatieve efficiëntie



Bron: Deller en Lledo, 2001.

2.1.2 Empirische toetsing

De vraag resteert hoe deze theoretische stelling om te vormen tot een empirische toets. Het middel hiervoor is een regressie. Een regressie is in staat om de helling van de curve in Figuur 2.1 vast te stellen, op het punt van de curve waar gemeenten zich bevinden. De gemeenten vormen de observaties die nodig zijn om de helling, of raaklijn, te schatten. De raaklijn is een rechte lijn langs de curve, zoals de lijnen die horen bij de punten A, B en C in de figuur. Is de lijn stijgend (dalend), dan bevinden de gemeenten zich gemiddeld genomen rond punt A (C). De efficiënte hoeveelheid van een dienst wordt geleverd in punt B, waar de lijn precies horizontaal loopt: daar is namelijk de top van de curve.

Het volledige gemeentelijke dienstenpakket kan als volgt worden geëvalueerd. Als maatstaf voor de ruimhartigheid van het dienstenpakket (de horizontale as van figuur 2.1, en de onafhankelijke variabele in de regressie) geldt het bedrag aan gemeentelijke uitgaven. De vastgoedprijs staat nog steeds op de verticale as als de afhankelijke variabele in de regressie. Als de regressiecoëfficiënt gelijk is aan nul, is de raaklijn horizontaal, en is er gemiddeld genomen dus geen sprake van over- of onderbesteding door gemeenten.

De toets is in staat om alle gemeenten tezamen te evalueren op allocatieve efficiëntie, maar niet om afzonderlijke gemeenten tegen het licht te houden. Dit kan een probleem opleveren, zoals het volgende voorbeeld illustreert. Als de helft van de gemeenten in een studie op punt A ligt, en de andere helft op punt C, dan zal de regressielijn horizontaal lopen. Dat leidt tot de - onjuiste - conclusie dat de onderzochte gemeenten allemaal allocatief efficiënt zijn, terwijl ze dat in de praktijk juist allemaal niet zijn. Wel geldt dan de conclusie dat er geen sprake is van systematische over- of onderprovisie van diensten.

De toets van Brueckner is daarnaast afhankelijk van een aantal aannames. Zo is het gebaseerd op een wereld waarin bevolkingsgroepen niet te veel van elkaar verschillen en burgers eenvoudig kunnen verhuizen. Daarnaast kunnen verschillen tussen gemeenten

van invloed zijn op de regressie, als hiervoor niet of niet voldoende gecorrigeerd is. Ook wordt aangenomen dat gemeenten geen geld verspillen. Dergelijke aannames zijn echter inherent aan een toets die een abstract begrip als allocatieve efficiëntie evalueert. Appendix A1 behandelt deze en andere beperkingen van de toets van Brueckner in meer detail.

2.1.3 Samenvatting

Al met al biedt Brueckner een intuïtieve maar zeker niet onomstreden toets voor allocatieve efficiëntie van gemeenten. De resultaten uit de toets van Brueckner zijn eenvoudig te interpreteren: gemeenten maken wel of niet systematisch een verkeerde allocatiekeuze. Deze evaluatie is wel afhankelijk van een aantal strikte aannames. De literatuur heeft echter nog geen beter alternatief kunnen ontwikkelen.

2.2 Internationale literatuur: een gemengd beeld

2.2.1 Beschrijving van de artikelen

Tabel 2.1 bevat het overzicht van de empirische literatuur op het gebied van allocatieve efficiëntie van het gehele gemeentelijke dienstenpakket.⁵ In totaal betreft het 17 artikelen. De eerste 13 artikelen hebben betrekking op de Verenigde Staten. De overige vier artikelen hebben betrekking op Nederland, Engeland (2) en Canada. Omdat institutionele verschillen tussen landen invloed kunnen hebben op de resultaten, en omdat er maar weinig bekend is over (landen met vergelijkbare instituties als) Nederland, is het belangrijk om terughoudend te zijn bij het interpreteren van de resultaten. De artikelen evalueren soms ook de allocatieve efficiëntie op een of meer individuele beleidsdomeinen. De tekst in het kader op de volgende pagina gaat daar verder op in.

2.2.2 Werkwijze: tellen van artikelen

In deze analyse telt elk artikel even zwaar mee. Dit houdt in dat het aantal positief gestemde artikelen simpelweg wordt opgeteld en vergeleken met het aantal negatieve. Deze procedure is niet helemaal eerlijk, omdat artikel A mogelijk een meer valide onderzoeksmethode hanteert dan artikel B. De ideale situatie zou zijn dat betere artikelen een groter gewicht zouden krijgen. Deze gewichten zijn echter moeilijk vast te stellen.

⁵ De literatuur is gevonden op basis van de "cited by" optie in Google Scholar van de artikelen van Brueckner, in combinatie met de referentielijsten van de op deze manier gevonden papers.

Allocatieve efficiëntie op beleidsdomeinen: Sterke focus op onderwijs

Het gros van de studies uit de hoofdtekst geeft, naast een algehele evaluatie van allocatieve efficiëntie, ook speciale aandacht aan een of meerdere specifieke beleidsdomeinen. Het betreft studies over de Verenigde Staten.

Onderwijs is veruit het meest onderzochte beleidsdomein met veertien studies in totaal. Daaruit blijkt dat de prioriteit van de wetenschappelijke literatuur over allocatieve efficiëntie niet geheel overeenkomt met die van Nederlandse gemeenten. In Nederland staat het onderwijs namelijk onder verantwoordelijkheid van de Rijksoverheid, en niet onder die van gemeenten zoals in de Verenigde Staten. Inzichten uit het Amerikaanse onderwijssysteem kunnen echter wel signalen afgeven over hoe gemeenten in het algemeen presteren.

De evaluaties van de individuele beleidsdomeinen staan vermeld in onderstaande tabel. De beleidsdomeinen zijn gerangschikt naar het aantal studies. Sommige studies komen meerdere keren voor in de tabel. Dit gebeurt als een studie meerdere beleidsdomeinen belicht. Naast de dertien Amerikaanse studies in Tabel 2.1 is ook Barrow en Rouse (2004) onderdeel van dit kader. Dit betreft een onderzoek dat enkel de allocatieve efficiëntie van gemeenten toetst op het terrein van onderwijs.

Onderwijs

50% van de studies over onderwijs kon geen systematische afwijking van de optimale allocatiekeuze aantonen. 36% van de studies, 5 in totaal, toont juist bewijs voor het tegenovergestelde. De afwijkingen van allocatieve efficiëntie betreffen vooral onderprovisie. Onderprovisie wil zeggen dat gemeenten naar de behoefte van burgers te weinig diensten leveren op het gebied van onderwijs. Bij dit beleidsdomein is de aanname van Brueckner, dat bevolkingsgroepen niet veel van elkaar verschillen qua preferenties, echter niet vanzelfsprekend: vooral jonge gezinnen zullen baten ondervinden van investeringen in onderwijs (Hilber en Mayer, 2009).

Andere beleidsdomeinen

De andere beleidsdomeinen lijken consequent te wijzen op onderprovisie door gemeenten. De cijfers zijn enigszins geflatteerd, omdat twee studies de tabel domineren. Deller en Lledo (2001) en Deller en Maher (2009) evalueren gemeenten in Wisconsin op respectievelijk vier en zes beleidsdomeinen. Op alle beleidsdomeinen concluderen zij dat gemeenten te weinig besteden. Als deze studies niet zouden meetellen, zijn de onderste drie rijen in de tabel leeg en valt de evaluatie van het beleidsdomein veiligheid positief uit.

De conclusie is dan ook dat er - zelfs als wereldwijd wordt gekeken - weinig wetenschappelijk onderzoek is uitgevoerd om de allocatieve efficiëntie te toetsen van andere beleidsdomeinen dan het onderwijs, in ieder geval als het gaat om onderzoek dat de toets van Brueckner toepast.

Tabel Allocatieve efficiëntie op individuele beleidsdomeinen

Beleidsdomein	Geen over- of onderprovisie	Overprovisie	Onderprovisie	Niet eenduidig
	Aantal studies	Aantal studies	Aantal studies	Aantal studies
Onderwijs	7	1	4	2
Infrastructuur en mobiliteit	1	0	4	0
Veiligheid	2	0	3	0
Volksgesondheid	0	0	2	1
Overig (Fysieke omgeving, Recreatie en Community Development)	0	0	2	0
Overheidsadministratie	0	0	1	0

Sommige onderzoekers hanteren een nauwere definitie van de beleidsdomeinen. Zij evalueren bijvoorbeeld enkel de politie. De tabel wijst deze toe aan het beleidsdomein veiligheid. In sommige gevallen zijn meerdere evaluaties in dezelfde studie ondergebracht in één beleidsdomein. Zo zijn de brandweer en de politie beide onderdeel van veiligheid.

2.2.3 Bevindingen

De empirische resultaten van de studies naar het totaalpakket aan gemeentelijke voorzieningen blijken wisselend. Acht studies (47%) konden geen systematische onder- of overprovisie door gemeenten aantonen. Twee studies (12%) stuiten op overprovisie en vijf (29%) wijzen erop dat gemeenten te weinig hebben besteed aan het dienstenpakket. Van twee studies in Tabel 2.1 zijn de conclusies afhankelijk van interpretatie: bijvoorbeeld de keuze van het significantieniveau. In dat geval krijgt de studie het label 'niet eenduidig'.

Op basis van deze evaluatie kan worden geconcludeerd dat gemeenten niet in alle situaties allocatief efficiënt zijn. Tegelijkertijd ontstaat er geen schrikbeeld waarin zij systematisch onder de maat presteren. Gezien de resultaten ligt het in ieder geval niet voor de hand dat gemeenten consequent te veel uitgeven. Het is echter niet uit te sluiten dat zij een te sober dienstenpakket aanbieden.

Tabel 2.1 Artikelen over gemeentelijke allocatieve efficiëntie

Artikel	Regio (Verenigde Staten)	Conclusie
Bates en Santerre (2003)	Connecticut	Onderprovisie
Bates en Santerre (2013)	Connecticut	Geen over- of onderprovisie
Bradbury et al. (2001)	Massachusetts	Geen over- of onderprovisie
Brueckner (1979)	New Jersey	Overprovisie
Brueckner (1982)	Massachusetts	Geen over- of onderprovisie
Deller (1990a)	Illinois	Onderprovisie(a)
Deller (1990b)	Maine	Geen over- of onderprovisie
Deller en Chicoine (1993)	Maine	Geen over- of onderprovisie(a)
Deller en Lledo (2001)	Wisconsin	Onderprovisie(a)
Deller en Maher (2009)	Wisconsin	Onderprovisie
Grossman et al. (1999)	VS	Geen over- of onderprovisie
Lee (2009)	New Jersey	Onderprovisie
Taylor (1995)	Hartford, Connecticut	Niet eenduidig(a)
Artikel	Regio (overig)	Conclusie
Allers en Vermeulen (2013a,b)	Nederland	Geen over- of onderprovisie(c)
Bayoumi en Gordon (1991)	Engeland	Overprovisie(b)
Hilber et al. (2011)	Engeland	Geen over- of onderprovisie(c)
Shah (1992)	Edmunton, Canada	Niet eenduidig

Noot: Uitgaven aan onderwijs zijn geen onderdeel van de evaluatie.
(a) Dit oordeel is gebaseerd op meerdere individuele beleidsdomeinen.
(b) Deze studie gebruikt niet de uitgaven aan gemeentelijke diensten als verklarende variabele, maar een geconstrueerde index die de ruimhartigheid van het dienstenpakket weergeeft.
(c) Deze artikelen gebruiken een aangepaste versie van Brueckners toets (zie paragraaf 2.3).

2.3 Literatuur Nederland: een eerste studie stemt positief

De enige relevante studie over de allocatieve efficiëntie van Nederlandse gemeenten is die van Allers en Vermeulen (2013a; b). De resultaten van deze studie stemmen positief. Deze paragraaf zoomt verder in op hun conclusies en de manier waarop die tot stand zijn gekomen.

2.3.1 Methode

Het artikel maakt gebruik van een variant van de toets van Brueckner. Deze methode van Barrow en Rouse (2004) evalueert hoe gemeenten omgaan met uitkeringen vanuit het

Rijk. Let wel: het gaat hier om subsidies, niet te verwarren met bijvoorbeeld bijstandsuitkeringen.⁶ Uitgaven die gefinancierd zijn met subsidies komen niet voor rekening van de lokale burger.^{7,8} Daarom leiden dit soort uitgaven in theorie tot een positief effect op de vastgoedprijs.

De toets van Barrow en Rouse bekijkt in hoeverre het effect van uitkeringen positief is. De gedachtegang is wederom dat vastgoedprijzen de betalingsbereidheid van burgers weerspiegelen. Stel, men besluit om de uitkeringen aan te wenden voor een nieuwe veiligheidscamera op het marktplein. Als de vastgoedwaarde daardoor stijgt met tienduizend euro, dan kan worden gesteld dat deze camera tienduizend euro aan baten oplevert voor de burger.

Een allocatief efficiënte gemeente zet de baten af tegen de investeringskosten. Als de camera vijfduizend euro heeft gekost, verbetert de besteding de allocatieve efficiëntie. De gemeente zou dan nog wel een camera kunnen ophangen. Maar een tweede camera zal de veiligheid minder doen toenemen dan de eerste. Dus kapitaliseert de tweede investering in mindere mate; het levert de burger immers minder baten op. De crux is om te blijven investeren totdat de burger minder dan vijfduizend euro voor een nieuwe camera over heeft.

Hoe kan worden getoetst of gemeenten inderdaad zo handelen? Bijvoorbeeld door opnieuw een regressie uit te voeren. Een regressie is immers in staat om te schatten in hoeverre een uitkering van de Rijksoverheid neerslaat in de vastgoedprijzen. Als de kapitalisatie gelijk is aan het bedrag aan uitkeringen (de 'investeringskosten'), besteden gemeenten hun uitkeringen op een allocatief efficiënte manier. Deze toets is afhankelijk van vergelijkbare aannames als de toets van Brueckner.

Allers en Vermeulen passen de methode van Barrow en Rouse toe op veranderingen in de algemene uitkering van de Rijksoverheid aan gemeenten. De hervormingen van de financiële-verhoudingswet in 1997 en 2001 bieden een uitgelezen kans om het effect van gemeentelijke middelen op de vastgoedprijs zuiver te schatten.

2.3.2 Bevindingen

De schattingen van Allers en Vermeulen wijzen erop dat de uitkeringen van het Rijk aan gemeenten vrijwel volledig kapitaliseren in de vastgoedprijzen van die gemeenten. Elke euro per jaar aan uitkeringen is nagenoeg één-op-één terug te zien in de huizenprijzen.⁹ Volgens Allers en Vermeulen impliceren de resultaten dat de uitkeringen goed besteed zijn

⁶ Barrow en Rouse (2004) passen de toets toe op het beleidsdomein onderwijs. Hilber et al. (2011) zijn de eersten die deze gedachtegang voor Engeland toepassen op het geheel aan gemeentelijke financiën.

⁷ De uitkering komt voort uit belastingheffing door het Rijk. Uiteraard dragen burgers in gemeenten daar ook aan bij. Deze bijdrage is echter minimaal, omdat de belasting gespreid wordt over burgers in het hele land.

⁸ Gemeenten kunnen dergelijke uitkeringen ook inzetten om de lokale belastingen te verlagen.

⁹ Vastgoedprijzen bestaan uit huizenprijzen en de prijs van commercieel vastgoed. De toets van Brueckner geldt ook voor de huizenprijzen zolang de rekening van uitbreidingen in het dienstenpakket niet (deels) bij bedrijven ligt. In het geval van Barrow en Rouse (2004) en Allers en Vermeulen (2013a; b) is dat inderdaad geen probleem, omdat de rekening volledig bij het Rijk ligt.

in de ogen van de marginale huizenkoper. Ze zijn echter voorzichtig om dit te interpreteren als allocatieve efficiëntie, met het oog op de restrictieve aannames van de gebruikte toets.

De conclusie die Allers en Vermeulen vooral trekken is dat een nieuwe uitkering vanuit de Rijksoverheid kan leiden tot herverdelingseffecten. Woningeigenaren in een gemeente die meer uitkeringen ontvangt, profiteren doordat hun woning meer waard wordt. Dit effect geldt echter alleen voor woningeigenaren die al in de gemeente woonden.

Woningeigenaren die later in de gemeente komen wonen profiteren niet, omdat de aanschafprijs van de woningen één-op-één meestijgt met de aantrekkelijkheid van het dienstenpakket. Huurders in de private sector profiteren evenmin, omdat ook de huur stijgt. Huurders in de sociale sector gaan er alleen op vooruit als woningcorporaties de toename in woningwaarde niet volledig doorberekenen.

2.4 Conclusie en discussie

Er is weinig bekend over de allocatieve efficiëntie van Nederlandse gemeenten. De enige studie naar Nederlandse gemeenten (Allers en Vermeulen, 2013a; b) concludeert wel dat Nederlandse gemeenten hun algemene uitkering besteden naar wens van de marginale huizenkoper. Dat is een signaal dat gemeenten een efficiënt pakket van diensten leveren. Desondanks wordt dit bewijs slechts gevormd door één studie.

De aannames in de literatuur geven een aanvullende reden voor terughoudendheid. Zo is de toets gebaseerd op een wereld waarin bevolkingsgroepen niet te veel van elkaar verschillen en burgers eenvoudig kunnen verhuizen. Dergelijke aannames zijn echter inherent aan een toets die een abstract begrip als allocatieve efficiëntie evalueert. Het is dan ook goed om deze toets niet los te zien van meer concrete en specifiekere evaluatiemethoden zoals tevredenheidsonderzoek onder burgers (zie bijvoorbeeld SCP, 2015).

Vanuit de internationale literatuur is er ook geen eenduidig oordeel mogelijk over de allocatieve efficiëntie van gemeenten. Er zijn grofweg evenveel onderzoeken die erop wijzen dat lokale overheden een inefficiënt dienstenpakket leveren, als dat er studies zijn die geen inefficiëntie aantonen. Wel blijkt daaruit dat gemeenten niet in alle omstandigheden allocatief efficiënt zijn. Tot slot lijkt het waarschijnlijker dat gemeenten naar de behoefte van burgers een te sober dienstenpakket leveren, dan dat zij te veel diensten aanbieden.

3 Gemeentelijke technische efficiëntie

Naast de keuze voor de samenstelling, omvang en kwaliteit van diensten, is het zaak voor gemeenten om hun diensten zo goedkoop mogelijk te leveren. In dat geval profiteren burgers van een lage gemeentelijke belasting, waardoor zij bijvoorbeeld meer andere producten kunnen aanschaffen.

Om vast te stellen of gemeenten inderdaad goedkoop leveren, moet worden gekeken naar het productieproces van gemeenten. In de literatuur wordt een goed draaiend productieproces technisch efficiënt genoemd.¹⁰ Factoren die bijdragen aan technische efficiëntie zijn een goede organisatie, gespecialiseerd personeel en geschikt management.¹¹

De literatuur beoordeelt het productieproces van gemeenten met name vanuit de benchmarkmethode.¹² Deze methode toont technische efficiëntieverschillen tussen gemeenten. Het gaat hier dus om een maat van relatieve (en niet per se absolute) efficiëntie. Tot dusver kent de literatuur nog geen techniek die kan vaststellen of gemeenten daadwerkelijk op het kostenminimum zitten.¹³ Dit hoofdstuk onderzoekt daarom de volgende vraag: *in hoeverre verschillen de kosten om het dienstenpakket te produceren tussen gemeenten?*^{14,15}

De focus ligt op de technische efficiëntie van het volledige gemeentelijke productieproces. Voor Nederland is over een dergelijk vraagstuk nog geen bewijs geleverd (CPB, 2007). Daarom zal bewijs hierover uit het buitenland komen. Wel zijn drie wetenschappelijke benchmarks bekend die Nederlandse gemeenten op een specifiek beleidsdomein evalueren.¹⁶ Deze komen in paragraaf 0 aan bod.

¹⁰ In de literatuur wordt naast technische efficiëntie ook gesproken van kostenefficiëntie, een gerelateerde term. Voor het exacte verschil, zie Coelli et al. (2005) of Knox Lovell (2000). In dit onderzoek wordt omwille van het overzicht geen onderscheid gemaakt. Dit is in lijn met de meeste literatuurstudies over dit onderwerp (De Borger en Kerstens, 2000; Worthington en Dollery, 2000; Afonso en Fernandes, 2008; Blank et al., 2010; Kalb et al., 2012; Kutlar et al., 2012). Nieswand en Seifert (2011) en Štastná en Gregor (2011) kiezen er daarentegen voor om te focussen op kostenefficiëntie. Een deel van deze studies bevat ook evaluaties van specifieke beleidsdomeinen.

¹¹ Voor een volledig overzicht van de oorzaken van technische efficiëntie, zie Blank et al. (2010).

¹² Studies die een andere methode hanteren vallen niet binnen de focus van dit onderzoek.

¹³ Door voortdurende technologische ontwikkelingen en arbeidsproductiviteitsverbeteringen is een dergelijk absoluut minimum tevens tijdgebonden.

¹⁴ De vraag doet wellicht vermoeden dat schaalvoordelen en -nadelen ook onderdeel zijn van dit hoofdstuk. Dit is niet het geval; daar wordt in de cijfers over technische efficiëntie namelijk dikwijls voor gecorrigeerd. Zie Allers en Geertsema (2014) voor een bespreking van schaalvoordelen bij Nederlandse gemeenten.

¹⁵ Een vergelijkbare vraag is: behalen gemeenten een zo groot mogelijke productie, gegeven het budget? Het verschil draait om welk van de twee posten vaststaat: het budget om te produceren of het aantal te produceren diensten. Beide vragen kennen sterke overeenkomsten, maar zijn niet noodzakelijk hetzelfde.

¹⁶ In de beleids sfeer circuleert nog een scala aan andere benchmarks voor gemeenten. Deze geschieden soms op vrijwillige basis en de kwaliteit ervan is niet altijd verifieerbaar. Deze notitie kiest er daarom voor om alleen de wetenschappelijke benchmarkmethode de revue te laten passeren.

3.1 Benchmarking meet technische efficiëntieverschillen

3.1.1 Theorie

Benchmarking is een veelvuldig toegepaste methode in de literatuur om technische efficiëntieverschillen te meten.¹⁷ Benchmarking is niets anders dan het productieproces van gemeenten te vergelijken met een norm. Wat die norm is, kan een onderzoeker of beleidsmaker zelf vaststellen. De norm kan bijvoorbeeld worden gesteld op het gemiddelde. In dat geval wordt elke gemeente vergeleken met de gemiddelde productieprestatie van alle gemeenten bij elkaar.

In de literatuur is de standaard om de best presterende gemeenten als norm te stellen: de zogeheten 'best practice', of *efficient frontier*. Deze gemeenten krijgen een efficiëntiescore van 100% toebedeeld. Voor gemeenten die de best practice niet halen, kan vervolgens worden vastgesteld in hoeverre zij tekortschieten. Scoort een gemeente bijvoorbeeld 90%, dan moet zij de kosten met 10% terugdringen om tot de top te behoren. De kwaliteit van de diensten mag dan echter niet onder deze kostenreductie lijden. Is dit wel het geval, dan scoort de gemeente alsnog onder de 100%. Dit komt doordat technische efficiëntie zowel om kwantiteit, kwaliteit als om kosten draait. In de literatuur staat deze manier van benchmarken bekend als '*frontier analysis*'.

In navolging van de literatuur gebruikt dit hoofdstuk als maatstaf de gemiddelde technische efficiëntiescore. De gemiddelde efficiëntiescore is op dezelfde wijze te interpreteren als de score per gemeente: een score van 90% impliceert dat er een kostenverschil van 10% zit tussen de gemiddelde gemeente en de best practice.

Als de best practice tegen het kostenminimum aanzit, dan is de efficiëntiescore ook een absolute maat: een score van 100% geeft dan aan dat de betreffende gemeente maximaal presteert als het gaat om technische efficiëntie. Gemeenten met een lagere score weten dan ook aan de hand van hun score welk deel van de kosten teruggedrongen moet worden om ook technisch efficiënt te zijn. Als de top zelf echter nog aanzienlijke stappen kan maken, is het alleen juist om de scores te interpreteren als efficiëntieverschillen.

Doorgaans beoordeelt de literatuur hoge gemiddelde scores (boven de 85%) als 'goed', maar dat is niet altijd terecht.¹⁸ Hoge scores kunnen bijvoorbeeld ook duiden op een weinig innovatieve instelling van gemeenten: ieder volgt de gebaande paden, met als gevolg weinig efficiëntieverschillen. Deze notitie houdt daarom een slag om de arm als er in de literatuur een waardeoordeel wordt gekoppeld aan efficiëntiescores.

Aanvullend op de gemiddelde efficiëntiescore beschrijft de literatuur twee andere maatstaven. De laagste efficiëntiescore geeft aan hoe ver de zwakste gemeente van de top

¹⁷ Het leeuwendeel van deze paragraaf is gebaseerd op Blank (2000) en Coelli et al. (2005). De Algemene Rekenkamer (2003a; b) biedt een andere, toegankelijke uitleg van benchmarking.

¹⁸ De grens van 85% is vastgesteld op basis van Van Hulst en De Groot (2011), die een score van 86% karakteriseren als een 'redelijke score'.

af zit. Het aandeel gemeenten dat tot de best practice hoort, meet hoe breed die top is. Deze indicatoren krijgen minder gewicht, enerzijds omdat ze meer afhankelijk zijn van de gebruikte methode en uitbijters, anderzijds omdat deze maatstaven niet altijd door onderzoekers gerapporteerd worden.

Desondanks geven deze maatstaven wel een beeld van de verdeling van gemiddelde efficiëntiescores, wat helpt bij de interpretatie. Een gemiddelde score van 90% kan bijvoorbeeld ontstaan doordat de helft van de gemeenten 80% haalt en de andere helft tot de best practice hoort. Een andere mogelijkheid is dat een kwart van de gemeenten 60% scoort en de rest 100%. Een hoog aandeel best practicegemeenten wijst eerder op de tweede situatie, net als een lage score voor de zwakste gemeente.

3.1.2 Empirische schatting

De empirische schatting draait om het vaststellen van de best practice. In de literatuur worden twee verschillende methoden voorgesteld om dit te doen. De zogeheten parametrische methode maakt gebruik van regressieanalyse. De niet-parametrische methode neemt simpelweg aan dat gemeenten met de meest gunstige verhouding tussen productie en kosten de best practice vormen. De niet-parametrische methode vergelijkt een gemeente dus met werkelijk bestaande best practice gemeenten, terwijl de parametrische methode een hypothetische gemeente creëert om als norm te fungeren. Zowel de parametrische als de niet-parametrische methode houdt rekening met de kwaliteit van de diensten, eventuele schaalvoordelen en gemeentekennmerken.

Beide methoden zijn op hun beurt weer onderverdeeld in verschillende stromingen. Onder de parametrische methode vallen de zogeheten Deterministic Frontier Analysis (DFA) en de Stochastic Frontier Analysis (SFA). Aangezien SFA geavanceerder is en DFA zelden wordt gebruikt voor de analyse van algehele technische efficiëntie, is DFA niet meegenomen in deze notitie. Onder de niet-parametrische methode vallen de zogeheten Data Envelopment Analysis (DEA) en *Free Disposable Hull* (FDH).¹⁹ Het kader op de volgende pagina gaat verder in op deze methoden en illustreert ze aan de hand van een voorbeeld.

¹⁹ De econometrische voordelen en beperkingen van DEA en SFA zijn te vinden in Knox Lovell (2000) en Coelli et al. (2005); die van FDH zijn te vinden in Vanden Eeckaut et al. (1993).

Empirische methoden voor technische efficiëntie

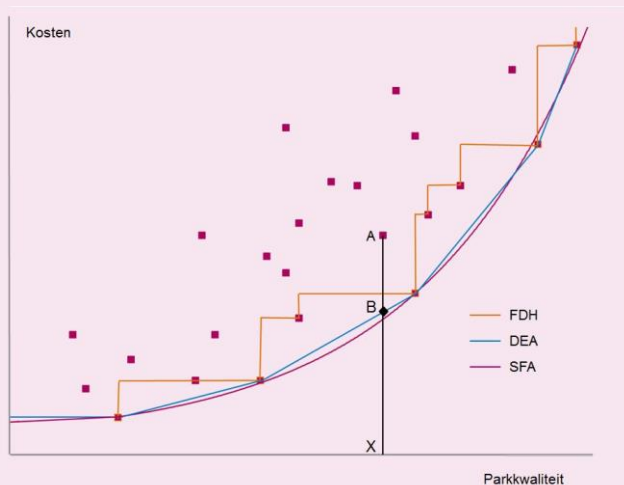
Zoals gezegd draait het berekenen van technische efficiëntieverschillen om het bepalen van de 'best practice'. Daartoe beschikt een onderzoeker over meerdere opties.

Om de empirische methoden voor technische efficiëntie te illustreren, kijkt dit kader naar een versimpelde versie van de werkelijkheid die illustratief is voor de meer complexe realiteit. Stel dat gemeenten enkel verantwoordelijk zijn voor één dienst: het onderhoud van parken. Gemeenten zijn daarnaast allemaal even groot, schaalvoordelen spelen dus geen rol. De vraag is: welke gemeenten lukt het om parkdiensten volgens de standaard van de best practice te leveren?

In de figuur zijn de kwaliteit van de gemeentelijke parken en bijbehorende kosten voor enkele gemeenten afgebeeld (elke punt in de wolk is een gemeente). Het gaat hier niet om werkelijke data. Aangenomen wordt dat de gemeente de kwaliteit van de parken heeft vastgesteld volgens de behoefte van haar burgers.

De technische efficiëntiescore van een gemeente is als volgt af te leiden. De score is berekend als de verticale afstand van de best practice tot de horizontale as gedeeld door de afstand tussen de horizontale as en de gemeente in kwestie. De verschillende lijnen weerspiegelen de best practice volgens de drie onderzochte methoden. Volgens DEA is de efficiëntiescore van gemeente A bijvoorbeeld dus $\frac{BX}{AX}$ gedeeld door AX .

Verschillende best practice methoden



Bij de niet-parametrische methoden ontstaat de best practice door de punten van gemeenten met de laagste kosten-kwaliteitverhouding te verbinden. Bij DEA wordt een rechte doorgetrokken lijn gebruikt. FDH verbindt best practicegemeenten door een tweetal lijnen in de vorm van een trap. De best practice grens die FDH levert, ligt hoger dan de DEA-lijn. Dit betekent dat de technische efficiëntiescores bij FDH altijd hoger zijn dan bij DEA. Ook liggen er meer gemeenten op de best practice lijn bij FDH.

Bij SFA is een vloeiende kromme getrokken door de puntenwolk. Hierdoor liggen maar weinig gemeenten exact op de best practice lijn. Gemeenten kunnen zich in een enkel geval ook onder de best practice lijn bevinden. Deze gemeenten zijn zo goedkoop (deels) vanwege factoren buiten technische efficiëntie om. In dergelijke gemeenten zijn burgers bijvoorbeeld zelf actief in het onderhouden van parken. Deze 'geluksfactor' wordt door SFA uit de best practice gefilterd, om te voorkomen dat appels met peren worden vergeleken. Hoe SFA zich exact verhoudt tot DEA en FDH, is te zien in Appendix A4: de scores zijn bij SFA doorgaans lager dan bij FDH en vergelijkbaar aan DEA.

De Algemene Rekenkamer (2003b) biedt een overzicht van hoe de drie methodes zich tot elkaar verhouden in verschillende situaties, en biedt aanknopingspunten voor praktische toepassing.

De literatuur wijst op twee beperkingen bij de interpretatie van benchmarking. Enerzijds kan het lastig zijn om te meten wat de productie van de gemeente nu precies is. Met name is het lastig om de kwaliteit van gemeentelijke diensten te vangen in een cijfer. Anderzijds kunnen de cijfers vertekend zijn als er niet voldoende is gecorrigeerd voor verschillen tussen gemeenten. In dat geval vergelijkt de benchmarkmethode appels met peren. Deze en andere beperkingen van de benchmarkmethode worden uitgebreider besproken in Appendix A2.

3.1.3 Samenvatting

De benchmarkmethode biedt een handzame manier om technische efficiëntieverschillen tussen gemeenten te analyseren. Maar ook deze methode is volgens de literatuur niet zonder valkuilen. Daarom is het ook hier zaak om de wetenschappelijke resultaten bij een vertaalslag naar de praktijk met de nodige voorzichtigheid te interpreteren.

3.2 Internationale literatuur: efficiëntie tamelijk homogeen

3.2.1 Beschrijving van de studies

Tabel 3.1 toont de 41 studies die de algehele technische efficiëntieverschillen onder gemeenten evalueren, gesorteerd naar landencluster.²⁰

De landenclusters zijn gerangschikt op basis van vergelijkbaarheid met Nederland. Het cluster Noordwest-Europa en dan met name studies over België en Duitsland geven wellicht de beste indicatie voor technische efficiëntie in Nederland, omdat er geen evaluatie van de algehele technische efficiëntie van Nederlandse gemeenten bekend is. Wel bestaan er drie studies over specifieke beleidsdomeinen van Nederlandse gemeenten. Paragraaf 0 bespreekt deze.

3.2.2 Werkwijze: gemiddelde per landencluster

Om de meet- en vergelijkbaarheidsproblemen van de benchmarkmethode te beperken, bekijkt deze notitie het gemiddelde over een aantal vergelijkbare studies. De gedachte hierachter is dat sommige studies de efficiëntiescores overschatten (of de best practice onderschatten) en andere de scores onderschatten (of de best practice overschatten). Het gemiddelde levert dan een betere indicatie.

²⁰ De wetenschappelijke studies zijn gevonden op basis van zoektermen in Google Scholar als "local government cost efficiency", "municipality technical efficiency", "DEA municipalities", en dergelijke. Ook zijn deze zoektermen gebruikt in de standaard Google zoekmachine om onder meer beleidsrapporten te detecteren die van de benchmarkmethode gebruikmaken. Ook de referentielijsten zijn een bron geweest om studies te vinden. Studies die oudere frontiemethoden als COLS toepassen, laat deze notitie buiten beschouwing.

Tabel 3.1 Studies over gemeentelijke technische efficiëntie

Studie naar landencluster	Land	Studie naar landencluster	Land
Noordwest-Europa		Centraal- en Zuid-Europa (incl. Turkije)	
De Borger et al. (1994)	België	Athanassopoulos en Triantis (1998)	Griekenland
De Borger en Kerstens (1996a)	België	Boetti et al. (2012)	Italië
De Borger en Kerstens (1996b)	België	Lo Storto (2013)	Italië
Vanden Eeckaut et al. (1993)	België	Karbowaniki en Kulaii (2011)	Polen
Geys (2006)	België	Afonso en Fernandes (2006)	Portugal
Geys en Moesen (2009)	België	Afonso en Fernandes (2008)	Portugal
Bönisch et al. (2011)	Duitsland	Jorge et al. (2006)	Portugal
Geys et al. (2013)	Duitsland	Arcelus et al. (2013)	Spanje
Kalb et al. (2012)	Duitsland	Balaguer-Coll et al. (2007)	Spanje
Kriese (2008)	Duitsland	Balaguer-Coll en Prior (2009)	Spanje
Loikkanen en Susiluoto (2005)	Finland	Balaguer-Coll et al. (2010)	Spanje
Nieswand en Seifert (2013)	Frankrijk	Giménez en Prior (2007)	Spanje
Widmer en Zweifel (2008)	Zwitserland	Zafra-Gómez et al. (2010)	Spanje
		Štastná en Gregor (2011)	Tsjechië
		Kutlar et al. (2012)	Turkije
Angelsaksische landen			
Worthington (2000)	Australië		
Grossman et al. (1999)	Verenigde Staten		
Hayes en Chang (1990)	Verenigde Staten		
Lee (2009)	Verenigde Staten		
Nold Hughes en Edwards (2000)	Verenigde Staten		
Ontwikkeld Azië		BRICS-landen	
Nakazawa (2013)	Japan	De Sousa et al. (2005))	Brazilië
Nijkamp en Suzuki (2009)	Japan	De Sousa en Stošić (2005)	Brazilië
Lim (2007)	Zuid-Korea	Charnes et al. (1989)	China
Sung (2007)	Zuid-Korea	Monkam (2014)	Zuid-Afrika

Een enkele studie evalueert een andere lokale overheid dan gemeenten.

Deze notitie kiest ervoor om ongewogen gemiddelden te nemen per landencluster voor de drie efficiëntie maatstaven.²¹ Vervolgens wordt vooral gekeken naar de met Nederland meest vergelijkbare landenclusters. Dan gaat het vooral om de ontwikkelde landen in Noordwest-Europa en, in het bijzonder, onze buurlanden.

In paragraaf 3.2.4 volgt een robuustheidscontrole, die de gemiddelde efficiëntiescores corrigeert voor heterogeniteit tussen studies. Op deze manier spelen keuzevrijheid van de onderzoekers en andere versturende factoren een kleinere rol. Hierdoor wordt de vergelijking tussen studies, en daarmee landenclusters, betrouwbaarder.

²¹ Idealiter zou men de studies indelen naar land. Over het algemeen zijn er echter onvoldoende studies per land beschikbaar. België en Duitsland vormen daarop een uitzondering: daar zijn wel meerdere onderzoeken uitgevoerd. Om die reden zijn de gemiddelde maatstaven voor beide landen ook afzonderlijk in de resultaten opgenomen. De studies over deze landen zijn ook deel van het cluster Noordwest-Europa.

3.2.3 Bevindingen

In Tabel 3.2 staan de maatstaven voor technische efficiëntieverschillen zoals gevonden in de internationale literatuur. Het betreft ongewogen gemiddelden van de studies per landencluster.²²

Tabel 3.2 Indicatoren voor gemeentelijke technische efficiëntieverschillen naar landencluster

Landencluster	Gemiddelde score (%)	Laagste score (%)	Aandeel met 100% score (%)	Aantal studies
Noordwest-Europa	88	47	37	13
Angelsaksische landen	80	55	38	5
Ontwikkeld Azië	85	51	29	4
Centraal- en Zuid-Europa	73	31	25	15
BRICS-landen	55	21	14	4
Buurlanden				
België	89	48	55	6
Duitsland	88	38	36	4
België en Duitsland	88	44	52	10
Totalen				
Europese Unie	79	39	29	27
Alle landen	78	39	27	41

Appendix A3 bespreekt de indicatoren per studie (voor Noordwest-Europa) en hoe deze zijn samengebracht in de tabel.

De gemiddelde gemeente in Noordwest-Europa moet de kosten met 12% terugdringen om tot de beste practice te horen. Dit is te zien aan de gemiddelde efficiëntiescore van 88% voor deze landencluster. Dit geldt overigens alleen bij een gelijkblijvend niveau van het dienstenpakket. België en Duitsland presteren vergelijkbaar met het volledige cluster Noordwest-Europa. Dat is niet heel verrassend aangezien deze landen 10 van de 13 studies voor hun landencluster leveren.

Vergeleken met andere benchmarkstudies in andere sectoren is de score tussen 88 en 89% tamelijk hoog. Dat houdt in dat efficiëntieverschillen onder Noordwest-Europese gemeenten gemiddeld relatief klein zijn. Desondanks moet de zwakste gemeente 53% van de kosten zien te reduceren om tot de top te behoren. Daar zit dus nog aanzienlijke ruimte tot verbetering. De top is breed en bestaat uit 37% van de gemeenten.

Opvallend is dat de gemiddelde efficiëntiescore bij de individuele studies in het cluster Noordwest-Europa geen uitschieters kent. Voor alle 13 studies ligt de score boven de 81%.²³ In andere landenclusters liggen uitschieters vaak beduidend lager. In alle landenclusters, dus ook in Noordwest-Europa, verschillen de laagste score en het aandeel best practicegemeenten wel sterk tussen individuele studies. Dit komt doordat deze maatstaven minder robuust zijn dan de gemiddelde efficiëntiescore.

²² Scores van DEA-methoden die gebruikmaken van constante schaalopbrengsten zijn alleen meegenomen als er geen variant met variabele schaalopbrengsten gerapporteerd is.

²³ Als een correctie wordt uitgevoerd voor vergelijkbaarheid (zie paragraaf 3.2.4 en Appendix A4), dan komt de laagste gemiddelde efficiëntiescore op 79%.

Angelsaksische gemeenten moeten de kosten met gemiddeld 20% terugdringen om technisch efficiënt te worden. De verschillen met de best practice zijn in Angelsaksische landen gemiddeld dus groter dan in Noordwest-Europa. Toch lijkt het verschil in scores niet heel groot. De minimale efficiëntiescores en het aandeel van de best practice zijn in beide landenclusters nagenoeg gelijk.

De Aziatische gemeenten scoren qua gemiddelde en minimumscore ongeveer gelijk met Noordwest-Europese gemeenten. Het percentage best practicegemeenten ligt echter 8 procentpunt lager. Bij deze vergelijking moet worden opgemerkt dat er slechts vier studies uit Azië zijn meegenomen, die enkel betrekking hebben op Japan en Zuid-Korea: twee ontwikkelde landen. De gemeenten in de clusters Centraal- en Zuid-Europa en in de BRICS-landen scoren opvallend lager op alle maatstaven dan de andere clusters.

3.2.4 Robuustheidscontrole

Een onderlinge vergelijking tussen studies is niet altijd perfect, omdat efficiëntiescores bijvoorbeeld afhankelijk zijn van de onderzoeksopzet. Zo blijkt uit het kader dat de keuze voor een bepaalde methode de prestatie-indicatoren beïnvloedt. Daarom is het van belang om de invloed van deze en andere factoren uit de scores te filteren. Deze subparagraaf doet dit voor de belangrijkste indicator: de gemiddelde technische efficiëntiescore.²⁴

Deze notitie gebruikt twee verschillende manieren om de scores beter vergelijkbaar te maken via regressieanalyse. Beide correcties komen neer op een herverdeling van de scores. De gemiddelde efficiëntiescore van alle studies bij elkaar verandert daardoor dus niet tot nauwelijks. De correctie werkt als volgt. De keuze voor meerdere indicatoren om de productie van gemeenten te meten, levert bijvoorbeeld een hogere score. Een regressie toont aan hoe groot die 'score-inflatie' precies is. De scores die volgen uit een studie met veel productie-indicatoren worden vervolgens met deze factor gecorrigeerd, zodat ze vergelijkbaar zijn met andere studies. Appendix A4 beschrijft de correctiemethodes in verder detail.

Tabel 3.3 beschrijft het resultaat van de opschoningsprocedure aan de hand van de twee correcties. Het Angelsaksische cluster blijkt het meest veranderlijk: dit cluster scoort na opschoning onder de 65%, terwijl de ongewogen gemiddelde efficiëntiescore nog 80% was. De reden is dat twee van de vijf Angelsaksische studies een aanpak kiezen die scores sterk verhoogt. De lage score in deze landen is mogelijk te wijten aan de institutionele context: in de Verenigde Staten staat men bijvoorbeeld over het algemeen toleranter tegenover regionale verschillen dan in Europa.

²⁴ Deze procedure is ook toepasbaar op de andere maatstaven, maar hiervoor zijn minder gegevens beschikbaar.

Tabel 3.3 Geschoonde gemiddelde technische efficiëntiescores per landencluster

Landencluster	Ongewogen gemiddelde	Correctie 1	Correctie 2	Aantal Studies
	Score in %	Score in %	Score in %	
Noordwest-Europa	88	91	92	13
Angelsaksische landen	80	63	64	5
Ontwikkeld Azië	85	88	89	4
Centraal- en Zuid-Europa	73	72	73	15
BRICS-landen	55	53	54	4
Buurlanden				
België	89	88	90	6
Duitsland	88	92	93	4
België en Duitsland	88	90	91	10
Totalen				
Europese Unie	79	80	82	27
Alle landen	78	77	78	41

Wellicht belangrijker is dat de hoge scores van de clusters Noordwest-Europa niet het gevolg zijn van een score-opdrijvende onderzoeks aanpak. De correcties leiden zelfs tot een toename van de scores tot boven de 90%. Ook Ontwikkeld Azië scoort hoger. Dit heeft met herverdeling te maken: als een cluster (Angelsaksische landen) te hoog zat, zit de rest automatisch te laag. De efficiëntieverschillen binnen de clusters Noordwest-Europa en Ontwikkeld Azië zijn na correcties nog kleiner dan uit de ongewogen scores bleek. In het tekstkader op de volgende pagina verkennen we enkele mogelijke oorzaken van het feit dat de scores per landencluster en per studie ook na opschoning nog flinke verschillen laten zien.

De conclusie van deze robuustheidsparagraaf is dat de resultaten uit Tabel 3.2 over Noordwest-Europa en Duitsland en België relatief ongevoelig zijn voor de opschoningsprocedure.

Wat hangt samen met technische efficiëntiescores?

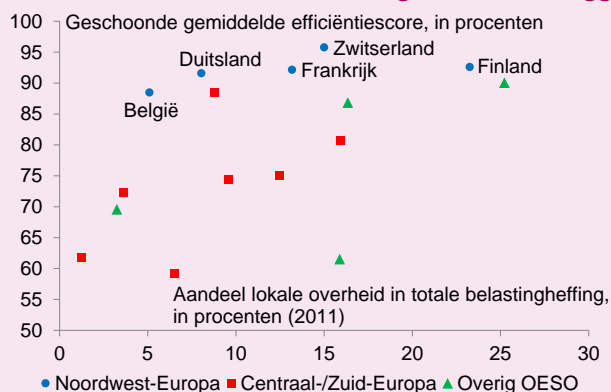
Dit kader koppelt een aantal mogelijk invloedrijke factoren aan de geschoonde gemiddelde efficiëntiescore per land. Het gaat in dit geval om de zestien OESO-landen die onderdeel zijn van deze studie. De analyse draait om het plotten van de (slechts) zestien efficiëntiescores tegen telkens een van de factoren. De analyse is daarom exploratief. Het gaat om een statistisch verband. Of de verbanden ook causaal zijn, is op basis van deze analyse niet te zeggen.

Decentralisatie en lokaal belastinggebied: positief verband?

Hoe groter het aandeel dat de lokale overheid heeft in de belastingheffing, hoe hoger de gemiddelde efficiëntiescores, zo illustreert de figuur links. De puntenwolk bevindt zich namelijk vooral linksonder en rechtsboven in de grafiek. Maar heel sterk is het verband zeker niet: de correlatie is krap 0,25. Het is goed om op te merken dat dit verband niet zozeer opgaat voor landen in Noordwest-Europa. Als men een denkbeeldige lijn trekt tussen de (ronde) punten die deze landen voorstellen, is deze lijn nagenoeg horizontaal.

Aan het statistische verband in de niet-Noordwest-Europese OESO landen kunnen meerdere oorzaken ten grondslag liggen. Het zou kunnen dat efficiëntie (onder zwakkere gemeenten) groeit als gevolg van een groter lokaal belastinggebied. Het aandeel van de lokale overheid in de totale overheidsbestedingen en de hoogte van de onroerendgoedbelasting correleert immers positief met de technische efficiëntiescores, zij het zwak. Anderzijds kan het zijn dat de omvang van het lokale belastinggebied pas wordt vergroot als gemeenten onderling weinig verschillen qua efficiëntie.

Technische efficiëntie naar omvang lokaal belastinggebied



Bron: Eigen berekening en gegevens van de OESO.

Nationaal inkomen en bevolkingsdichtheid: positief verband

Naast het aandeel van de lokale overheid in de belastingheffing, blijken twee andere kenmerken een positieve statistische samenhang te hebben met de efficiëntiescores, zij het minder sterk. Als het nationaal inkomen toeneemt, nemen technische efficiëntieverschillen bijvoorbeeld af. Daarnaast is de bevolkingsdichtheid hoger in landen waarin gemeenten een hogere score laten zien.

Andere factoren: geen verband

De andere variabelen die op deze manier zijn onderzocht, blijken geen statistisch verband te hebben met de efficiëntiescores. Het gaat dan om opleidingsniveau, vergrijzing, uitgaven aan sociale doeleinden (een indicator voor politieke oriëntatie), verstedelijking, inkomensongelijkheid en de mate van democratie.

3.3 Literatuur over Nederland: weinig efficiëntieverschillen

Er bestaan drie wetenschappelijke studies naar de technische efficiëntieverschillen onder Nederlandse gemeenten op een specifiek beleidsdomein. De bevindingen zijn samengevat in Tabel 3.4. De resultaten wijzen nog sterker dan de vorige paragraaf richting de conclusie dat gemeenten niet veel van elkaar verschillen wat betreft technische efficiëntie.

De scores op het gebied van de bijstand en het afvalbeheer zijn hoog tot zeer hoog te noemen. Met scores van respectievelijk 94 en 97% ligt de gemiddelde gemeente vlak achter de best practice.²⁵ De technische efficiëntie is dus sterk homogeen op deze vlakken. Dat is ook te zien aan de hoge minimumscores die rond de 85% liggen.

Tabel 3.4 Indicatoren voor gemeentelijke technische efficiëntieverschillen naar beleidsdomein

Studie	Beleidsdomein	Datajaar	Gemiddelde score (%)	Laagste score (%)	Aandeel best practice (%)
Broersma et al. (2013) (a)	Bijstand	2007	94	85	30
Felsö et al. (2011)	Afvalbeheer	2009	97	84	32
Van Hulst en De Groot (2011)	Burgerzaken	2008	86	52	25

(a) Schatting op basis van een figuur met de verdeling van de scores per gemeente. De scores betreffen een gemiddelde van verschillende rekenmethodes.

De verschillen op het gebied van burgerzaken zijn aanmerkelijk groter. Op dat terrein wordt gemiddeld een score van 86% behaald. De laagste score ligt net boven de 50%. Voor de minder presterende gemeenten lijkt er in ieder geval dus ruimte voor verbetering. Van Hulst en De Groot (2011) stellen dat de 25% die tot de best practice behoren, de zwakker presterende gemeenten kunnen helpen hun achterstand weg te werken.

Desalniettemin suggereren de drie studies dat de efficiëntieverschillen onder Nederlandse gemeenten over het algemeen klein zijn. Hoe groot de verschillen in totaal zijn, is niet te zeggen omdat de resultaten van de drie specifieke studies niet per se opgaan voor het volledige scala aan gemeentelijke diensten.

3.4 Conclusie en discussie

Op basis van de literatuurstudie ontstaat een eenduidig beeld dat technische efficiëntieverschillen tussen gemeenten in Noordwest-Europa redelijk klein zijn. De gemiddelde technische efficiëntiescore ligt hier tussen de 88 en 89%. Dit betekent dat de gemiddelde gemeente de kosten met 11 à 12% moet terugdringen om de prestatie van de meest efficiënte gemeenten te evenaren. Die kostenreductie mag dan niet ten koste gaan van de omvang of de kwaliteit van het dienstenpakket. Individuele studies over Noordwest-Europa laten daarnaast opvallend weinig uitschieters zien wat betreft de gemiddelde efficiëntiescore.

²⁵ De efficiëntiescore is van 2003 op 2004 aanzienlijk gestegen als gevolg van de invoering van de Wet Werk en Bijstand (WWB).

Omdat de efficiëntieverschillen ook na correctie voor studiekekenmerken blijven staan, kunnen ze ook voor Nederlandse gemeenten worden verwacht. Drie studies over Nederlandse gemeenten op individuele beleidsterreinen bevestigen dit beeld. De efficiëntieverschillen blijken in deze studies zelfs nog kleiner. Desondanks zijn er aan de onderkant van de prestatieladder nog een aantal gemeenten die aanzienlijk achter liggen ten opzichte van de top. Hier liggen kansen voor efficiëntieverbetering. Ook is het niet uit te sluiten dat de goedkoopst producerende gemeenten nog stappen kunnen maken. Dat is op basis van onderzoek naar efficiëntieverschillen niet te zeggen.

Een volgende stap is dat onderzoek naar de algehele technische efficiëntie van Nederlandse gemeenten plaatsvindt. In dat kader is een doordachte onderzoeksaanpak van belang om te zorgen dat appels niet met peren vergeleken worden. Ook moet worden gezorgd voor accurate maatstaven om de gemeentelijke productie te meten. Deze twee beperkingen van benchmarkonderzoek zijn tot dusver reden geweest om de algehele technische efficiëntie van gemeenten in Nederland niet te toetsen.

Dit betekent niet dat dergelijk onderzoek vruchteloos is. Het bestaan van vele buitenlandse, wetenschappelijke studies suggereert dit. Een algehele evaluatie biedt waardevolle informatie voor overkoepelende beleidsvraagstukken zoals de mate van decentralisatie en het lokale belastinggebied (De Borger en Kerstens, 2000). Ook helpt het om te achterhalen in welke gemeenten er in het algemeen verbetering te realiseren is. Tot slot kunnen de resulterende scores gebruikt worden als uitkomstmaat in regressies die de impact van beleid op algehele technische efficiëntie toetsen.

Ook is wetenschappelijk gefundeerd onderzoek naar de technische efficiëntie in Nederland op meer beleidsdomeinen van toegevoegde waarde. Dergelijk specifiek onderzoek vult enkele hiaten in de literatuur op en geeft middelen in handen om de reeds vele bestaande, beleidsgerichte benchmarks wetenschappelijk te verifiëren.

Referenties

Geciteerde bronnen

Afonso, A., en S. Fernandes,, 2008, Assessing and explaining the relative efficiency of local government, *The Journal of Socio-Economics*, vol. 37(5): 1946-79.

Algemene Rekenkamer, 2003a, Handreiking meten van doelmatigheid, Den Haag: Algemene Rekenkamer.

Algemene Rekenkamer, 2003b, Technieken voor het meten van doelmatigheid, Internetbijlage bij rapport 'Handreiking meten van doelmatigheid', Beschikbaar op: www.rekenkamer.nl/Publicaties/Handreikingen/Meten_van_doelmatigheid.

Allers, M.A. en J.B. Geertsema, 2014, Geen grotere doelmatigheid door herindeling gemeenten, *ESB*, vol. 99(4688): 406-9.

Allers, M.A. en W. Vermeulen, 2013a, Fiscal Equalization and capitalization: evidence from a policy reform, CPB Discussion Paper 245.

Allers, M.A. en W. Vermeulen, 2013b, Kapitalisatie van de algemene uitkering uit het gemeentefonds in woningprijzen, Groningen: Centrum voor Onderzoek van de Economie van Lagere Overheden.

Barrow, L. en C.E. Rouse, 2004, Using market valuation to assess public school spending, *Journal of Public Economics*, vol. 88: 1747-69.

Blank, J.L.T. (red.), 2000, *Public provision and performance: contributions from efficiency and productivity measurement*, Amsterdam: Elsevier.

Blank, J.L.T., A.C.M. Dumaij, C.M.G. Haelermans, A.A.S. van Heezik, B.L. van Hulst en P.M. Koot, 2010, Is de ziekte van Baumol te behandelen? Een verkennend onderzoek naar de mogelijkheden om de productiviteit in de publieke sector te vergroten. Delft: Centrum voor Innovatie en Publieke Sector Efficiëntie Studies.

Boorsma, P.B. en M.A. Allers, 2006, De financiële verhouding onder de loep, Den Haag: Vereniging van Nederlandse Gemeenten.

Borger, B. de en K. Kerstens, 2000, What is known about municipal efficiency? The Belgian case and beyond, in: J.L.T. Blank (red.), *Public provision and performance: contributions from efficiency and productivity measurement*, 299-330, Amsterdam: Elsevier.

- Brueckner, J.K., 1979, Property values, local public expenditure and economic efficiency, *Journal of Public Economics*, vol. 11(2): 223-45.
- Brueckner, J.K., 1982, A test for allocative efficiency in the local public sector, *Journal of Public Economics*, vol. 19(3): 311-31.
- Brueckner, J.K., 1983, Property value maximization and public sector efficiency, *Journal of Urban Economics*, vol. 14(1): 1-15.
- Chaudry-Shah, A., 1988, Capitalization and the theory of local public finance: an interpretive essay, *Journal of Economic Surveys*, vol. 2(3): 209-43.
- Coelli, T.J., D.S. Prasada Rao, C.J. O'Donnell en G.E. Battese, 2005, *An introduction to efficiency and productivity analysis* (2^{de} ed.), New York: Springer.
- CPB, 2007, Productiviteit en effectiviteit van de collectieve sector: kansen voor beleidsrenderend onderzoek, CPB Notitie, 18 december, Den Haag: Centraal Planbureau.
- CPB, 2015, Een ruimer lokaal belastinggebied, CPB Policy Brief, 9 april, Den Haag: Centraal Planbureau.
- Deller, S.C., en V. Lledo, 2001, Local public sector performance: are Wisconsin city and village taxes too high? Wisconsin: Madison Agricultural and Applied Economics Department Staff Paper 440.
- Deller, S.C. en C. Maher, 2009, Government, effectiveness, performance, and local property values, *International Journal of Public Administration*, vol. 32(13): 1182-1212.
- Dur, A.J. en B.A. Vollaard, 2013, Slecht voorbeeld doet slecht volgen in de buitenruimte, *ESB*, vol. 98(4672S): 65-69.
- Eeckaut, P. vanden, H. Tulkens en M.A. Jamar, 1993, Cost efficiency in Belgian municipalities, in: H.O. Fried, C.A. Knox Lovell en S.S. Schmidt (red.), *The Measurement of Productive Efficiency—Techniques and Applications*, 300-334, New York: Oxford University Press.
- Ewijk, C., van, H.L. de Groot en A.J. Santing, 2012, A meta-analysis of the equity premium, *Journal of Empirical Finance*, vol. 19(5): 819-30.
- Folmer, K., 2009, Why do macro wage elasticities diverge? A meta analysis, CPB Research Memorandum 224.
- Geys, B., F. Heinemann en A. Kalb, 2013, Local government efficiency in German municipalities, *Raumforschung und Raumordnung*, vol. 71(4): 283-93.

- Gradus, R.H.J.M. en E. Dijkgraaf, 2014, Kostenbesparingen door tariefsystemen voor huisafval, *ESB*, vol. 99(4686): 333-35.
- Groot, H.L.F. de, G. Marlet, C.N. Teulings en W. Vermeulen, 2010, *Stad en land*, Den Haag: Centraal Planbureau.
- Haelermans, C.M.G., 2012, On the productivity and efficiency of education, The role of innovations in Dutch secondary education, Proefschrift: Universiteit Maastricht.
- Hedges, L.V., E. Tipton en M.C. Johnson, 2010, Robust variance estimation in meta-regression with dependent effect size estimates, *Research Synthesis Methods*, vol. 1(1): 39-65.
- Hilber, C.A.L., 2011, The economic implications of house price capitalization: a survey of an emerging literature, SERC Discussion Paper 91.
- Hilber, C.A.L., T. Lyytikäinen en W. Vermeulen, 2011, Capitalization of central government grants into local house prices: panel data evidence from England, *Regional Science and Urban Economics*, vol. 41(4): 394-406.
- Hilber, C.A.L. en C. Mayer, 2009, Why do households without children support local public schools? Linking house price capitalization to school spending, *Journal of Urban Economics* vol. 65(1): 74-90.
- Hoyt en Garen, 2005, Fiscal policy and property values, CBER Research Report, 10.
- Hulst, B.L. van, en H. de Groot, 2011, Benchmarking Burgerzaken: een empirisch onderzoek naar de kostendoelmatigheid van burgerzaken, Delft: Centrum voor Innovatie en Publieke Sector Efficiëntie Studies.
- Kalb, A., B. Geys en F. Heinemann, 2012, Value for money? German local government efficiency in a comparative perspective, *Applied Economics*, vol. 44(2): 201-18.
- Knox Lovell, C.A., 2000, Measuring efficiency in the public sector, in: J.L.T. Blank (red.), *Public provision and performance: contributions from efficiency and productivity measurement*, 23-53, Amsterdam: Elsevier.
- Koning, P.W.C., 2014, Door schuldhulpverlening uit de bijstand, *ESB*, vol. 99(4677): 38-41.
- Kutlar, A., F. Bakirci en F. Yüksel, 2012, An analysis on the economic effectiveness of municipalities in Turkey, *African Journal of Marketing Management*, vol. 4(3): 80-98.
- Marlet, G. en C. van Woerkens, 2014, *Atlas voor gemeenten 2014. Economie en arbeidsmarkt*, Nijmegen: VOC uitgevers.

Nieswand, M. en S. Seifert, 2013, What drives intermediate local governments' spending efficiency: the case of French départements, *Local Government Studies*, vol. 40(5): 766-90.

Nguyen, K.-H. en T.J. Coelli, 2009. Quantifying the effects of modelling choices on hospital efficiency measures: a meta-regression analysis, *Centre for Efficiency and Productivity Analysis Working Paper*, 07.

Oates, W.E., 1969, The effects of property taxes and local public spending on property values: An empirical study of tax capitalization and the Tiebout hypothesis, *The Journal of Political Economy*, vol. 77(6): 957-71.

Oates, W.E., 1999, An essay on fiscal federalism. *Journal of Economic Literature*, vol. 37(3): 1120-49.

Raad voor de Financiële Verhoudingen, 2011, Uitbreiding lokaal belastinggebied, Advies aan staatssecretaris Wiebes, 26 maart 2015, Den Haag: Raad voor de Financiële Verhoudingen.

Ross, S., en J. Yinger, 1999, Sorting and voting: a review of the literature on urban public finance, In: Cheshire, P. en E.S. Mills (red.), *Handbook of Regional and Urban Economics*, vol. 3: 2001-60.

Rousseeuw, P.J. en A.J. Leroy, 1987, Robust regression and outlier detection, *Wiley Series in Probability and Mathematical Statistics*, New York: Wiley and Sons.

Samuelson, P.A., 1954, The pure theory of public expenditure, *The Review of Economics and Statistics*, vol. 36(4): 387-89.

Samuelson, P.A., 1955, Diagrammatic exposition of a theory of public expenditure, *The Review of Economics and Statistics*, vol. 37(4): 350-56.

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

Simar, L. en P.W. Wilson, 2008, Statistical inference in nonparametric frontier models: recent developments and perspectives, in: H.O. Fried, C.A. Knox Lovell en S.S. Schmidt (red.), *The measurement of productive efficiency and productivity growth*, Oxford: Oxford University Press, 421-521.

Štastná, L. en M. Gregor, 2011, Local government efficiency: evidence from the Czech municipalities, IES Working Paper 14.

Worthington, A.C. en B.E. Dollery, 2000, An empirical survey of frontier efficiency measurement techniques in local government, *Local Government Studies*, vol. 26(2): 23-52.

Studies in de literatuurstudie over allocatieve efficiëntie

Allers, M.A. en W. Vermeulen, 2013a, Fiscal Equalization and capitalization: evidence from a policy reform, CPB Discussion Paper 245.

Allers, M.A. en W. Vermeulen, 2013b, Kapitalisatie van de algemene uitkering uit het gemeentefonds in woningprijzen, Groningen: Centrum voor Onderzoek van de Economie van Lagere Overheden.

Barrow, L. en C.E. Rouse, 2004, Using market valuation to assess public school spending, *Journal of Public Economics*, vol. 88: 1747-69.

Bates, L.J. en R.E. Santerre, 2003, The impact of a state mandated expenditure floor on aggregate property values, *Journal of Urban Economics*, vol. 53(3): 531-40.

Bates, L.J. en R.E. Santerre, 2013, Does the US health care sector suffer from Baumol's cost disease? Evidence from the 50 states, *Journal of Health Economics*, vol. 32(2): 386-91.

Bayoumi, T. en J. Gordon, 1991, The determinants and efficiency of local authority spending in England, IMF Working Paper 9.

Bradbury, K.L., C.J. Mayer en K.E. Case, 2001, Property tax limits, local fiscal behavior, and property values: evidence from Massachusetts under Proposition 2½, *Journal of Public Economics*, vol. 80(2): 287-311.

Brueckner, J.K., 1979, Property values, local public expenditure and economic efficiency, *Journal of Public Economics*, vol. 11(2): 223-45.

Brueckner, J.K., 1982, A test for allocative efficiency in the local public sector, *Journal of Public Economics*, vol. 19(3): 311-31.

Deller, S.C., 1990a, An application of a test for allocative efficiency in the local public sector, *Regional Science and Urban Economics*, vol. 20(3): 395-406.

Deller, S.C., 1990b, Pareto-Efficiency and the Provision of Public Goods Within a Rural Setting, *Growth and Change*, vol. 21(1): 30-39.

Deller, S.C., en D.L. Chicoine, 1993, Representative versus direct democracy: a test of allocative efficiency in local government expenditures, *Public Finance Review*, vol. 21(1): 100-114.

Deller, S.C., en V. Lledo, 2001, Local public sector performance: are wisconsin city and village taxes too high? Wisconsin: Madison Agricultural and Applied Economics Department Staff Paper 440.

Deller, S.C. en C. Maher, 2009, Government, effectiveness, performance, and local property values, *International Journal of Public Administration*, vol. 32(13): 1182-1212.

Grossman, P.J., P. Mavros, en R.W. Wassmer, 1999, Public sector technical inefficiency in large US cities, *Journal of Urban Economics*, vol. 46(2): 278-99.

Hilber, C.A.L., T. Lyytikäinen en W. Vermeulen, 2011, Capitalization of central government grants into local house prices: panel data evidence from England, *Regional Science and Urban Economics*, vol. 41(4): 394-406.

Lee, S.H., 2009, Tax Competition among governments and the effects on government performance: empirical evidence from local governments in New Jersey, Proefschrift: State University of New Jersey.

Shah, A.M., 1992, Empirical tests for allocative efficiency in the local public sector, *Public Finance Review*, vol. 20(3): 359-77.

Taylor, L.L., 1995, Allocative inefficiency and local government, *Journal of Urban Economics*, vol. 37(2): 201-11.

Studies in de literatuurstudie over technische efficiëntie

Afonso, A., en Fernandes, S., 2006, Measuring local government spending efficiency: evidence for the Lisbon region, *Regional Studies*, vol. 40(1): 39-53.

Afonso, A., en Fernandes, S., 2008, Assessing and explaining the relative efficiency of local government, *The Journal of Socio-Economics*, vol. 37(5): 1946-79.

Arcelus, F.J., P. Arocena, F. Cabasés en P. Pascual, 2013, On the cost-efficiency of service delivery in small municipalities, *Regional Studies*, DOI: 10.1080/00343404.2013.837872.

Athanassopoulos, A.D. en K.P. Triantis, 1998, Assessing aggregate cost efficiency and the related policy implications for Greek local municipalities, *Infor*, vol. 36(3): 66-83.

Balaguer-Coll, M.T., D. Prior en E. Tortosa-Ausina, 2007, On the determinants of local government performance: A two-stage nonparametric approach, *European Economic Review*, vol. 51(2): 425-451.

Balaguer-Coll, M.T. en D. Prior, 2009, Short-and long-term evaluation of efficiency and quality, An application to Spanish municipalities, *Applied Economics*, vol. 41(23): 2991-3002.

Balaguer-Coll, M.T., D. Prior en E. Tortosa-Ausina, 2010, Decentralization and efficiency of local government, *The Annals of regional science*, vol. 45(3): 571-601.

Boetti, L., M. Piacenza en G. Turati, 2012, Decentralization and local governments' performance: how does fiscal autonomy affect spending efficiency? *FinanzArchiv: Public Finance Analysis*, vol. 68(3): 269-302.

Bönisch, P., P. Haug, A. Illy en L. Schreier, 2011, Municipality size and efficiency of local public services: Does size matter? IWH-Diskussionspapiere 18.

Borger, B. de, K. Kerstens, W. Moesen, en J. Vanneste, 1994, Explaining differences in productive efficiency: an application to Belgian municipalities, *Public Choice*, vol. 80(3-4): 339-58.

Borger, B. de en K. Kerstens, 1996a, Cost efficiency of Belgian local governments: A comparative analysis of FDH, DEA, and econometric approaches, *Regional Science and Urban Economics*, vol. 26(2): 145-70.

Borger, B. de en K. Kerstens, 1996b, Radial and nonradial measures of technical efficiency: an empirical illustration for Belgian local governments using an FDH reference technology, *Journal of Productivity Analysis*, vol. 7(1): 41-62.

Broersma, L, A.J.E. Edzes en J. van Dijk, 2013, Have Dutch municipalities become more efficient in managing the costs of social assistance dependency?, *Journal of Regional Science*, vol. 53(2): 274-91.

Charnes, A., W.W. Cooper en S. Li, 1989, Using data envelopment analysis to evaluate efficiency in the economic performance of Chinese cities, *Socio-Economic Planning Sciences*, vol. 23(6): 325-44.

Eeckaut, P. vanden, H. Tulkens en M.A. Jamar, 1993, Cost efficiency in Belgian municipalities, In: Fried, H.O., C.A. Knox Lovell and S.S. Schmidt (red.), *The Measurement of Productive Efficiency-Techniques and Applications*, 300-334, New York: Oxford University Press.

Felsö, F.A., H. de Groot, A.A.S. van Heezik, 2011, Benchmark gemeentelijk afvalbeheer. Een empirisch onderzoek naar de productiviteit en kostendoelmatigheid, Delft: Centrum voor Innovatie en Publieke Sector Efficiëntie Studies.

Geys, B., 2006, Looking across borders: A test of spatial policy interdependence using local government efficiency ratings, *Journal of Urban Economics*, vol. 60(3): 443-62.

Geys, B., en W. Moesen, 2009, Measuring local government technical (in)efficiency, *Public Performance and Management Review*, vol. 32(4): 499-513.

- Geys, B., F. Heinemann en A. Kalb, 2013, Local government efficiency in German municipalities, *Raumforschung und Raumordnung*, vol. 71(4): 283-293.
- Giménez, V.M. en D. Prior, 2007, Long-and short-term cost efficiency frontier evaluation: evidence from Spanish local governments, *Fiscal Studies*, vol. 28(1): 121-39.
- Grossman, P.J., P. Mavros, en R.W. Wassmer, 1999, Public sector technical inefficiency in large US cities, *Journal of Urban Economics*, vol. 46(2): 278-99.
- Hayes, K. en S. Chang, 1990, The relative efficiency of city manager and mayor-council forms of government, *Southern Economic Journal*, vol. 57(1): 167-77.
- Hulst, B.L. van, en H. de Groot, 2011, Benchmarking burgerzaken: een empirisch onderzoek naar de kostendoelmatigheid van burgerzaken, Delft: Centrum voor Innovatie en Publieke Sector Efficiëntie Studies.
- Jorge, S.M., P.J. Camões, J.B.D.C. Carvalho en M.J. Fernandes, 2006, Portuguese Local Government relative efficiency: A dea approach, Artikel gepresenteerd bij de 8^e CIGAR workshop over performance measurement en output based budgeting in de publieke sector, Hamburg, September 14-15.
- Kalb, A., B. Geys en F. Heinemann, 2012, Value for money? German local government efficiency in a comparative perspective, *Applied Economics*, vol. 44(2): 201-18.
- Karbowaniki, B. en G. Kulaii, 2011, Efficiency of public sector at the level of local governments in Poland, Warschau: National Bank of Poland.
- Kriese, M., 2008, Effizienzanalyse der Sächsischen Gemeinden (in het Duits), *Ifo Dresden Berichtet*, vol. 5: 3-13.
- Kutlar, A., F. Bakirci en F. Yüksel, 2012, An analysis on the economic effectiveness of municipalities in Turkey, *African Journal of Marketing Management*, vol. 4(3): 80-98.
- Lee, S.H., 2009, Tax Competition among governments and the effects on government performance: empirical evidence from local governments in New Jersey, Proefschrift: State University of New Jersey.
- Lim, D.J., 2007, A comparative study of performance measurement in Korean local governments using data envelopment analysis and stochastic frontier analysis, Proefschrift: University of Texas at Arlington.
- Lo Storto, C., 2013, Evaluating technical efficiency of Italian major municipalities: a data envelopment analysis model, *Procedia - Social and Behavioral Sciences*, vol. 81: 346-50.

Loikkanen, H.A. en I. Susiluoto, 2005, Cost efficiency of Finnish municipalities in basic service provision 1994-2002, *Urban Public Economics Review*, vol. 4: 39-64.

Monkam, N.F., 2014, Local municipality productive efficiency and its determinants in South Africa, *Development Southern Africa*, vol. 31(2): 1-24.

Nakazawa, K., 2013, Cost inefficiency of municipalities after amalgamation, *Procedia Economics and Finance*, vol. 5: 581-88.

Nieswand, M. en S. Seifert, 2013, What drives intermediate local governments' spending efficiency: the case of French départements, *Local Government Studies*, vol. 40(5): 766-90.

Nijkamp, P. en S. Suzuki, 2009, A generalized goals-achievement model in data envelopment analysis: an application to efficiency improvement in local government finance in Japan, *Spatial Economic Analysis*, vol. 4(3): 249-74.

Nold Hughes, P.A en M.E. Edwards, 2000, Leviathan vs. Lilliputian: a data envelopment analysis of government efficiency, *Journal of Regional Science*, vol. 40(4): 649-69.

Sousa, M.D.C.S. de, F. Cribari-Neto en B.D. Stošić, 2005, Explaining DEA technical efficiency scores in an outlier corrected environment: the case of public services in Brazilian municipalities, *Brazilian Review of Econometrics*, vol. 25(2): 287-313.

Sousa, M.D.C.S. de en B.D. Stošić, 2005, Technical efficiency of the Brazilian municipalities: correcting nonparametric frontier measurements for outliers, *Journal of Productivity analysis*, vol. 24(2): 157-81.

Šťastná, L. en M. Gregor, 2011, Local government efficiency: evidence from the Czech municipalities, IES Working Paper 14.

Sung, N., 2007, Information technology, efficiency and productivity: evidence from Korean local governments, *Applied Economics*, vol. 39(13): 1691-1703.

Widmer, P. en P. Zweifel, 2008, Provision of public goods in a federalist country: Tiebout competition, fiscal equalization, and incentives for efficiency in Switzerland, University of Zurich Socioeconomic Institute Working Paper 0804.

Worthington, A.C., 2000, Cost efficiency in Australian local government: a comparative analysis of mathematical programming and econometrical approaches, *Financial Accountability & Management*, vol. 16(3): 201-23.

Zafra-Gómez, J. L., M. Antonio en M. Pérez, 2010, Overcoming cost-inefficiencies within small municipalities: improve financial condition or reduce the quality of public services?, *Environment and planning C: Government and policy*, vol. 28(4): 609-29.

Appendix A

A1 Allocatieve efficiëntie: beperkingen van Brueckners toets

Een beperking van Brueckners toets is dat deze enkel een beeld geeft over de gemiddelde gemeente.²⁶ Individuele gemeenten zijn met deze toets niet te beoordelen.

Daarnaast ontstaan econometrische problemen wanneer de helft van de gemeenten te veel diensten aanbiedt, en de andere helft te weinig. De conclusie zou dan luiden dat de gemiddelde gemeente allocatief efficiënt is, terwijl er in werkelijkheid geen enkele gemeente voldoet aan de eisen voor efficiëntie. Wel geldt in zulke gevallen dat gemeenten niet systematisch te veel of te weinig diensten leveren.

Een vergelijkbare tekortkoming geldt bij een evaluatie van het volledige pakket aan diensten. Als gemeenten op een beleidsdomein te veel diensten leveren en op een ander beleidsdomein te weinig, dan streept de toets van Brueckner deze soms tegen elkaar weg. In dat geval kan niet meer over het efficiënte pakket aan diensten worden gesproken.

Elke gemeente geldt als een observatie, maar elke gemeente heeft ook andere kenmerken. Deze kenmerken kunnen zowel de vastgoedprijs als het dienstenpakket bepalen. Als hiervoor niet of niet voldoende gecorrigeerd wordt, zijn de resultaten van de regressie vertekend, en zegt de kapitalisatiegraad weinig meer over allocatieve efficiëntie.

Datzelfde geldt als gemeenten (een deel van) hun geld verspillen aan ongewenste zaken. Deze 'diensten' kapitaliseren namelijk niet in de vastgoedprijzen. In dat geval kan Brueckners toets wijzen op allocatieve efficiëntie als er juist sprake is van onderprovisie aan gewenste diensten (en dus overprovisie aan niet gewenste zaken).

Dit heeft te maken met het feit dat de literatuur de omvang en de kwaliteit van het gemeentelijke dienstenpakket meet aan de hand van de bestedingen in geldwaarde. Deze indicator is niet perfect. Een toename in de bestedingen kan bijvoorbeeld ook duiden op een toename van de kosten, terwijl bestedingsverschillen worden geïnterpreteerd als andere keuzes met betrekking tot het dienstenpakket.

Een belangrijke aanname van de toets van Brueckner is verder dat vastgoedprijzen de aantrekkelijkheid van een gemeente voor burgers volledig weerspiegelen. Dan neemt de vastgoedprijs alle relevante sociale baten en kosten van elke dienst mee. Een situatie waarin deze aanname niet opgaat, is wanneer de behoeften van bevolkingsgroepen onderling te veel verschillen. In dat geval zegt kapitalisatie in vastgoedprijzen alleen iets over hoe de marginale huizenkoper diensten waardeert, maar niet de gehele bevolking.

²⁶ Zie bijvoorbeeld Hilber (2011) voor een uitgebreider overzicht van de beperkingen van Brueckners toets.

Beperkte of selectieve verhuismobiliteit is een andere factor waardoor de vastgoedmarkt de betalingsbereidheid van burgers minder goed weergeeft. Als mensen niet snel willen of kunnen verhuizen, zullen gemeentelijke diensten minder sterk kapitaliseren in vastgoedprijzen dan Brueckners toets aanneemt. Anderzijds is de kapitalisatie te sterk als diensten vooral rijke inwoners naar gemeenten toe trekken. Zij drijven de vastgoedprijzen namelijk op vanwege hun grotere budget. Ten slotte mogen aanbodreacties op de verandering in de vraag (huizenbouw- en sloop) niet te veel verschillen per gemeente.

Ook is belangrijk op te merken dat lokale belastingen versturende effecten kunnen hebben. Een gemeente die gebruik maakt van de ozb om diensten te financieren, verstoort bijvoorbeeld vraag en aanbod op de woningmarkt (Brueckner, 1983). De toets van Brueckner heeft dus alleen betrekking op de economische efficiëntie van de gehele lokale economie, gegeven dat de woningmarkt niet volledig efficiënt functioneert.

Vanwege de bovenstaande beperkingen van Brueckners toets is het niet gepast om te stellig te zijn over de allocatieve efficiëntie van gemeenten. Daarom betekent een nuleffect enkel dat gemeenten niet systematisch de plank misslaan. Dat betekent dus niet dat gemeenten allocatief efficiënt zijn bij een nuleffect, maar het kan wel gezien worden als een randvoorwaarde voor allocatieve efficiëntie.

A2 Technische efficiëntie: beperkingen van de benchmarkmethode

De belangrijkste beperking is dat benchmarking vooral is bedoeld om verschillen in een steekproef vast te stellen. Het gaat om *relatieve* scores ten opzichte van de best practice. Het is niet valide om scores van individuele studies als *absolute* efficiëntiecijfers te interpreteren.

Een voorbeeld illustreert dit probleem. In studie A produceren alle gemeenten hun diensten voor 300 euro. Omdat alle gemeenten even goed presteren, krijgen alle gemeenten een efficiëntiescore van 100% toebedeeld. In studie B levert de helft van de gemeenten hetzelfde pakket diensten tegen 100 euro en de andere helft doet dit tegen 200 euro. De gemiddelde efficiëntiescore ligt dan op 75%, ook al zijn alle gemeenten in studie B efficiënter dan die in studie A.

Een tweede probleem is dat er mogelijk onbeïnvloedbare factoren meespelen bij efficiëntiescores. Een voorbeeld van dergelijke factoren voor Nederland is de ligging op veengrond ten opzichte van zandgrond. Een gemeente kan niet beïnvloeden dat binnen haar grenzen meer of minder grond verzakt dan elders. Desondanks zijn gemeenten op veen duurder uit dan die op zand. Als hiervoor niet gecorrigeerd wordt, lijken veengemeenten ten onrechte inefficiënt. Gemeenten kunnen door dergelijke factoren uiteraard ook efficiënter lijken dan ze in werkelijkheid zijn.

Ook is het moeilijk om de output van gemeenten correct te meten (De Borger en Kerstens, 2000). Meestal gebruikt men daarom onder andere gegevens over de bevolkingssamenstelling en -omvang als globale maatstaf voor productie. Een van de problemen van een dergelijk globale maatstaf is dat de kwaliteit van de productie niet altijd onderdeel uitmaakt van de evaluatie. Zo zou het kunnen zijn dat gemeenten die vooral bezuinigen op de kwaliteit, positief uit de benchmark komen. Er zijn enkele studies die dit als maatstaf meenemen, maar gemiddeld genomen laat de literatuur kwaliteit buiten beschouwing, vermoedelijk vanwege de problemen met de meetbaarheid.

Tot slot geldt dat niet eenduidig is wanneer een score hoog is of niet. Deze notitie volgt Van Hulst en De Groot (2011), die andere studies in andere sectoren gebruiken als vergelijkingsmateriaal om de scores te interpreteren. Als de scores hoger liggen dan wat andere studies doorgaans rapporteren, dan is er sprake van een hoge score.²⁷

A3 Technische efficiëntie: berekening landenclusterscore

De maatstaven voor technische efficiëntie (gemiddelde score, laagste score, efficiënt in %) in Tabel 3.2 zijn als volgt berekend. De studies zijn ingedeeld in landenclusters. Het landencluster Noordwest-Europa bestaat bijvoorbeeld uit dertien studies. De score (op elke maatstaf) van dit landencluster betreft een ongewogen gemiddelde van de scores per studie. Daarmee mogen er in dat cluster dus maximaal dertien scores per maatstaf zijn.

Feit is echter dat dit er in de praktijk soms meer zijn, omdat sommige studies de efficiëntiescore middels meerdere modellen berekenen.²⁸ De gekozen oplossing is om elk model een gelijk gewicht te geven in de berekening van de score per studie. Ook hier gaat het dus om een ongewogen gemiddelde.

Ongewogen gemiddelden zijn gebruikt omdat er onvoldoende informatie bestaat over bijvoorbeeld kwaliteitsverschillen tussen studies om onderbouwd een gewogen gemiddelde te kunnen uitrekenen. Een onderbouwde weging geeft mogelijk een beter beeld van de werkelijke scores.

Tabel A1 en Tabel A2 illustreren de procedure aan de hand van de dertien studies over Noordwest-Europa. In Tabel A1 Tabel A5 stelt binnen elke studie elke rij een ander model voor. In sommige gevallen (zoals Geys et al., 2013) worden er op een regel meerdere waarden weergegeven. In dat geval hebben de auteurs meerdere schattingen uitgevoerd waartussen het verschil niet eenduidig is. In de verwerking worden die modelvarianten

²⁷ Andere, meer technische beperkingen van de benchmarkmethode zijn te vinden in Coelli et al. (2005).

²⁸ Sommige studies werken meerdere DEA-modellen uit. Indien een studie zowel geavanceerde DEA-modellen toepast (waarbij rekening wordt gehouden met schaalvoordelen of -nadelen) als minder geavanceerde DEA-modellen, dan gebruikt deze notitie alleen de resultaten van de geavanceerde DEA-modellen.

eerst samengenomen tot één model door middel van een ongewogen gemiddelde. Tabel A6 toont de gemiddelde cijfers per studie en voor het hele cluster.²⁹

Tabel A5 Efficiëntie maatstaven per model voor Noordwest-Europa

Studie	Land	Model	Gemiddelde score (%)	Laagste score (%)	Efficiënt (%)
De Borger et al. (1994)	België	FDH	86,30	16,30	55,69
		FDH	96,70	36,30	82,34
		FDH	96,90	72,40	55,69
		FDH	99,10	76,60	82,34
		FDH	97,40	72,40	55,69
		FDH	99,40	86,30	82,34
De Borger en Kerstens (1996a)	België	FDH	93,70	44,10	66,38
		DEA	72,70	31,80	10,87
		SFA	80,90	34,70	12,22
		SFA	78,10	34,70	0,00 (a)
De Borger en Kerstens (1996b)	België	FDH	95,40	44,10	74,19
		FDH	97,20	61,90	74,19
		FDH	98,10	70,90	74,19
Geys (2006)	België	SFA	84,35	51,13	0,00 (a)
Geys en Moesen (2009)	België	FDH	95,04	28,81	73,03
		DEA	64,32	24,51	12,83
		SFA	85,76	57,18	0,00 (a)
Vanden Eeckaut et al. (1993)	België	FDH	97,94	50,00 (e)	80,43
		DEA	83,96	40,00 (e)	19,57
Geys et al. (2013)	Duitsland	SFA	81,37 87,87 87,64 (b)	22,96 21,90 22,23 (b)	0,10 (a)
		SFA	83,19 89,29 89,13 (b)	28,57 26,79 27,28 (b)	0,10 (a)
Kalb et al. (2012)	Duitsland	SFA	83,02 88,18 (b)	21,17 22,47 (b)	0,10 (a)
		SFA	85,31 89,59 (b)	21,70 24,66 (b)	0,10 (a)
Kriese (2008)	Duitsland	SFA	79,58	39,50	0,00 (a)
		SFA	90,06	53,66	0,00 (a)
Bönisch et al. (2011)	Duitsland	DEA	93,46 (c)	58,48	36,45
Loikkanen en Susiluoto (2005)	Finland	DEA (d)	85,80 (c)	57,10	3,40
		DEA (d)	85,60 (c)	59,90	1,98
		DEA (d)	89,80 (c)	61,70	7,37
		DEA (d)	87,10 (c)	58,70	5,10
Nieswand en Seifert (2013)	Frankrijk	DEA	87,90	65,30	29,50
		DEA	78,20	45,80	6,30
Widmer en Zweifel (2008)	Zwitserland	DEA (d)	85,00	52,00	3,85

(a) Het aantal efficiënte gemeenten is hier 0 of 1. Dat gebeurt vaak bij de SFA-methode.

(b) Het gaat hier om varianten van hetzelfde model. Daarom zijn twee á drie cijfers in deze cel opgenomen.

(c) Voor dit model is geen gemiddelde gerapporteerd, wel de mediaan. De mediaan is behandeld alsof het ging om een gemiddelde.

(d) De gebruikte methode gaat uit van constante schaalopbrengsten. De score op technische efficiëntie is dus deels het gevolg van schaalvoordelen- en nadelen.

(e) Schattingen op basis van een figuur.

²⁹ Een andere benadering is om eerst de gemiddelden per land te nemen, alvorens naar landencluster te aggregeren. Dit beïnvloedt de gemiddelde efficiëntiescores nauwelijks (maximaal 2 procentpunt, in het geval van de BRICS-landen). De andere twee maatstaven zijn wel gevoelig voor de aggregatiemethode.

Tabel A6 Efficiëntiemaatstaven per studie voor Noordwest-Europa

Studie	Land	Gemiddelde score (%)	Laagste score (%)	Efficiënt (%)
De Borger et al. (1994)	België	95,97	60,05	69,02
De Borger en Kerstens (1996a)	België	81,35	36,33	38,62
De Borger en Kerstens (1996b)	België	96,90	58,97	74,19
Geys (2006)	België	84,35	51,13	(b)
Geys en Moesen (2009)	België	81,71	36,83	42,93
Vanden Eeckaut et al. (1993)	België	90,95	45,00	50,00
Bönisch et al. (2011)	Duitsland	93,46	58,48	36,45
Geys et al. (2013)	Duitsland	86,42	24,96	(b)
Kalb et al. (2012)	Duitsland	86,53	22,50	(b)
Kriese (2008)	Duitsland	84,82	46,58	(b)
Loikkanen en Susiluoto (2005)	Finland	87,08	59,35	4,46
Nieswand en Seifert (2013)	Frankrijk	83,05	55,55	17,90
Widmer en Zweifel (2008)	Zwitserland	85,00	52,00	3,85
Noordwest-Europa		87,51	46,75	37,49

(a) Het aantal efficiënte gemeenten bij de SFA-methode is vaak 0 of 1. Deze cijfers zijn uit de berekening van het gemiddelde per studie gelaten, vanwege de beperkte vergelijkbaarheid met de scores op basis van de andere methoden.
(b) Deze studie betreft enkel scores op basis van de SFA-methode; daarom is deze cel leeg.

A4 Berekening van geschoonde efficiëntiescores

De resultaten van individuele studies en de daarin gebruikte modellen worden beïnvloed door de opzet van de studie. De scores voor technische efficiëntie hangen bijvoorbeeld af van de methode, zoals het kader 'Empirische methoden voor technische efficiëntie' beschrijft. De in Tabel 3.2 gerapporteerde ongewogen gemiddelden houden echter geen rekening met deze afhankelijkheden. De robuustheidscontrole in deze appendix toetst in hoeverre dit de resultaten uit de hoofdtekst beïnvloedt.

Het gaat in deze appendix enkel om de gemiddelde technische efficiëntiescore, omdat de literatuur deze maatstaf beschouwt als het meest bruikbaar van de drie maatstaven voor technische efficiëntie. Ook is hiervoor veruit het meeste informatie beschikbaar.

De robuustheidscontrole draait om een tweetal correcties, die elk gebruikmaken van een regressie. De twee regressies trachten aan te tonen wat de invloed is van studiekekenmerken op de gemiddelde technische efficiëntiescore. Vervolgens wordt deze informatie gebruikt om de ongewogen gemiddelde scores voor technische efficiëntie te corrigeren tot een 'geschoond' gemiddelde. De correcties hebben als doel de scores uit verschillende studies meer vergelijkbaar te maken.

Het verschil tussen beide correcties is dat correctie 1 de efficiëntiescores in een eerder stadium opschoont dan correctie 2. Correctie 1 tracht de scores per model op te schonen, dus nog voordat de score per studie berekend wordt. In correctie 2 is de regressieopzet hetzelfde, maar die wordt pas toegepast nadat de ruwe scores zijn samengevoegd tot een gemiddelde score per studie. Correctie 1 bevat in totaal 87 observaties of scores op modelniveau; correctie 2 bevat 41 observaties, gelijk aan het totaal aantal studies.

Een simpele OLS-regressie volstaat niet om het effect van de studieopzet te schatten. Er zijn namelijk een aantal problemen die zorgen dat de t-waarden van een OLS-regressie onbetrouwbaar worden. Dit komt onder andere doordat de efficiëntiescores niet onafhankelijk van elkaar zijn (Hedges et al., 2010). Zo is het denkbaar dat een onderzoeker meerdere studies uitvoert met vergelijkbare opzet of databestanden, wat resulteert in onderling gecorreleerde scores. Deze notitie corrigeert de OLS-regressie hiervoor door middel van Huber-White robuuste standaardfouten (zie Van Ewijk et al., 2012).

Daarnaast kunnen uitbijters van invloed zijn op de regressieschatting (Rousseeuw en Leroy, 1987). Als dit het geval is, levert een mediane regressie andere schattingen op dan een OLS-regressie. Voor beide correcties is een dergelijke methode ook toegepast, inclusief Huber-Sandwich robuuste standaardfouten (zie Folmer, 2009).

De opbouw van het vervolg van deze appendix is als volgt. Als eerste worden de verklarende variabelen behandeld, die de studieopzet typeren en daarom onderdeel zijn van de correctie. Daarna volgt regressie 1, waarna de ongewogen scores worden omgerekend tot geschoonde scores. Deze twee stappen heten samen 'correctie 1'. Correctie 2, die daarna komt, bestaat eveneens uit twee delen: regressie 2 en de bijbehorende berekening van geschoonde scores. Het uiteindelijke resultaat van de opschoning is te zien in Tabel 3.3.

Verklarende variabelen

De selectie van de verklarende variabelen volgt grotendeels Haelermans' metaregressie (2012) van technische efficiëntie in het onderwijs.³⁰ De volgende negen kenmerken worden in deze notitie gebruikt om de technische efficiëntiescore te verklaren, mits deze een significante invloed op de scores blijken te hebben:

1.	Landencluster	Zie Tabel 3.1
2.	Methode	DEA-CRS, DEA-VRS, SFA of FDH
3.	Dataset vanaf 2000	Studie bevat gegevens over de periode vanaf 2000
4.	Gemeenteselectie	Analyse bevat alleen een specifiek type gemeente
5.	Peer reviewed	Studie staat in een peer reviewed tijdschrift of boek
6.	Aantal kostenindicatoren	Aantal variabelen die de kosten meten
7.	Aantal productie-indicatoren	Aantal variabelen die de productie meten
8.	Aantal gemeenten	Aantal gemeenten in de analyse (in logaritmen)
9.	Aantal modellen in studie	(Enkel bij correctie 2)

³⁰ Haelermans (2012) maakt gebruik van de STATA-optie *metareg*. Deze methode weegt elke gemiddelde efficiëntiescore met de eigen standaardfout, om meer verklarende kracht te geven aan betrouwbaar gemeten scores. Deze werkwijze is echter niet goed toepasbaar in dit onderzoek, omdat de gemiddelde efficiëntiescores sterk negatief gecorreleerd zijn met de eigen standaarddeviatie. Zo levert een gemiddelde score van 100% per definitie een standaarddeviatie op van 0%. Daarom kiest deze notitie voor de ongewogen benadering, zoals Nguyen en Coelli (2009) ook doen bij hun meta-analyse van efficiëntiescores van ziekenhuizen.

We bespreken kort de keuzes voor de verklarende variabelen. Allereerst wordt er een onderscheid gemaakt naar landenclusters gezien de evidente verschillen in technische efficiëntie tussen groepen van landen in Tabel 3.2.

De potentiële invloed van de gekozen methode op de uitkomsten is grotendeels al besproken in het kader 'Empirische methoden voor technische efficiëntie', waarbij DEA-VRS gelijkstaat aan 'DEA' in het kader. DEA-CRS is niet behandeld in het kader. Deze methode neemt aan dat de verhouding tussen productie en kosten niet veranderen. In het kader zou dit betekenen dat de best practicelijn een rechte lijn is door de oorsprong. DEA-CRS scores liggen doorgaans lager dan DEA-VRS scores.³¹ De gekozen methode is als categorische variabele opgenomen in de regressie van correctie 1. Bij correctie 2 neemt de regressie het aandeel van elk van de methodes in de studie mee als variabele.

Het is mogelijk ook relevant welke jaartallen de databases beslaan. Oudere datasets zijn bijvoorbeeld van lagere kwaliteit en oudere technieken zijn ook minder geavanceerd. Om hiervoor te corrigeren is een dummy opgenomen, die de waarde 1 aanneemt als de dataset informatie bevat over de periode na 2000. Deze dummy geeft mogelijk ook informatie over de vraag of efficiëntiescores door de jaren heen zijn toe- of afgenomen.

Soms kiezen de onderzoekers ervoor om maar een deel van de gemeenten te bestuderen. Zo nemen Vanden Eeckaut et al. (1993) alleen Waalse gemeenten op in hun studie, in plaats van alle Belgische gemeenten. Dit kan ervoor zorgen dat de steekproef van gemeenten niet meer representatief is, wat in een hogere of lagere score resulteert.

In de literatuur blijken studies in een peer-reviewed tijdschrift of boek soms anders van aard dan andere publicaties zoals onderzoeksrapporten. In peer-reviewed studies is de technische efficiëntiescore op zichzelf eigenlijk weinig interessant. Het gaat meer om de methode om de score te schatten. In onderzoeksrapporten zijn de scores zelf veel meer van belang. Dit kan er bijvoorbeeld toe leiden dat artikelen in tijdschriften minder moeite doen om goede datasets te vinden, wat gevolgen heeft voor de efficiëntiescore. Een ander mogelijk verschil is dat peer-reviewed artikelen doorgaans een hogere kwaliteit hebben, waardoor scores anders uitvallen.

Daarnaast speelt het aantal kosten- en productie-indicatoren een rol. Dit bepaalt namelijk hoe de best practice wordt geschat. Naarmate er meer indicatoren onderdeel zijn van de schatting, vergelijkt men gemeenten namelijk op meer dimensies. Een vergelijkbaar verhaal geldt voor het aantal gemeenten: hoe meer er worden toegevoegd, hoe meer gegevens er zijn voor de onderzoekers om de best practice te schatten.³²

³¹ Modellen die constante schaalopbrengsten hanteren als aanname zijn alleen opgenomen in deze literatuurstudie als er in dezelfde studie geen variant beschikbaar was die uitgaat van variabele schaalopbrengsten.

³² Het effect van steekproefgrootte en het aantal kosten- en productie-indicatoren kan per methode verschillen (Simar en Wilson, 2008). In navolging van Haelermans (2012) kiest deze notitie er echter voor om deze verschillen niet mee te nemen in de regressies. Toevoeging van interactietermen zou de schattingen daarnaast te veel vrijheidsgraden ontnemen.

De schattingen starten met alle verklarende variabelen, waarna statistisch insignificante variabelen een voor een uit de regressie worden verwijderd. Voor variabelen met betrekking tot de methode geldt daarbij de gezamenlijke significantie van de vier categorieën. Deze notitie hanteert een significantieniveau van 10%. De uiteindelijke schatting waarin alleen variabelen zijn opgenomen die een significant effect uitoefenen op de efficiëntiescore, vormt de basis om de ruwe scores op te schonen.

Correctie 1: Regressie op modelniveau

Tabel A3 toont de uitkomsten van de regressie op modelniveau. De onderstaande interpretatie heeft betrekking op de OLS-schatting met alleen significante variabelen.

De coëfficiënten zijn als procentpunten te interpreteren. Bij de categorische en dummyvariabelen gaat het om procentpunten ten opzichte van de referentiecategorie. De waarde van -10,47 bijvoorbeeld houdt daarmee in dat modellen die de methode DEA-VRS gebruiken, gemiddeld genomen 10,47% procentpunt lagere efficiëntiescores kennen dan FDH-modellen. De variabele 'methode' is als geheel overigens insignificant.

Tabel A7 Correctie 1: technische efficiëntiescores verklaard, regressies op modelniveau

		OLS-regressie		Mediane regressie
		Alle variabelen	Significante variabelen	Alle variabelen
Variabele	Categorie	Coëfficiënt (%)	Coëfficiënt (%)	Coëfficiënt (%)
Cluster	Noordwest-Europa	(referentie)	(referentie)	(referentie)
	Angelsaksische landen	-21,48 **	-25,23 ***	-23,55 ***
	Ontwikkeld Azië	-2,94	-2,63	-6,51
	Centraal- en Zuid-Europa	-17,08 ***	-18,31 ***	-14,93 ***
	BRICS-landen	-31,66 ***	-33,46 ***	-35,40 ***
Methode	FDH	(referentie)	(referentie)	(referentie)
	DEA-CRS	-0,64		-3,17
	DEA-VRS	-10,47 **		-10,03 **
	SFA	-8,45 *		-11,18 ***
Dataset vanaf 2000		-9,83 **	-10,09 ***	-5,37
Gemeenteselectie		5,12		7,85 **
Peer reviewed		-2,28		-0,62
Aantal kostenindicatoren		2,10 ***	2,23 ***	2,25 ***
Aantal productie-indicatoren		1,08 **	0,84 **	1,14 *
Aantal gemeenten (a)		0,41		0,52
Constante		86,30 ***	85,67 ***	82,25 ***
R ²		0,52	0,48	nvt

De OLS-schattingen maken gebruik van Huber-White robuuste standaardfouten. De mediane regressie gebruikt Huber-Sandwich robuuste standaardfouten. Lege cellen geven weer dat de betreffende variabele is verwijderd op grond van statistische insignificantie. Deze verwijderde variabelen voegen gezamenlijk ook geen significante verklaringskracht toe aan het model met alleen de significante variabelen.
(a) In logaritmen.
*/**/** Significant op respectievelijk tien-, vijf- en eenprocentniveau. Aantal observaties: 87.

De coëfficiënt bij schaalvariabelen interpreteert men net iets anders. Het aantal kostenindicatoren heeft bijvoorbeeld een coëfficiënt van 2,23 (in het model met alleen significante variabelen). Dit betekent dat de gemiddelde efficiëntiescore toeneemt met 2,23 procentpunt per kostenindicator. Het aantal productie-indicatoren is van kleinere invloed, zo blijkt uit de kleinere coëfficiënt van 0,84.

Opvallend is dat modellen over Angelsaksische landen een flink lagere efficiëntiescore laten zien dan die over Noordwest-Europa: 25 procentpunt. De coëfficiënt is zelfs groter dan die voor Centraal- en Zuid-Europa. De ongecorrigeerde score van 80% voor Angelsaksische landen (zie Tabel 3.2) komt vooral door de keuze van de onderzoekers voor erg veel kostenindicatoren. Dit was het geval in twee van de vijf studies voor deze landen. Bij de correctie die hieronder volgt zal de score dus beduidend lager uitvallen. De geschoonde efficiëntiescore zal niet veel gaan verschillen tussen Noordwest-Europa en Azië, gegeven de kleine en insignificante coëfficiënt van -2,63.³³

Meer recente databestanden leveren lagere scores op. Wellicht zijn nieuwere studies beter in staat om inefficiënties aan te tonen dan oudere, bijvoorbeeld door betere technieken of data. Een andere mogelijke verklaring is dat de efficiëntieverschillen door de jaren heen zijn toegenomen.

Tot slot blijken de schattingen nauwelijks bepaald te zijn door uitbijters. De coëfficiënten in de laatste kolom van de tabel stemmen vrijwel overeen met die van de OLS-regressie met alle variabelen. Dat betekent dat de OLS-coëfficiënten bruikbaar zijn als middel om de ruwe scores op te schonen.

Omrekening naar geschoonde scores

De individuele score per model moet worden gecorrigeerd naar een door de onderzoeker zelf te specificeren referentiewaarde, zodat de modellen onderling vergelijkbaar worden. Men kan er bijvoorbeeld voor kiezen om de scores zo te wijzigen dat alle modellen als het ware 1 kostenindicator hebben. Modellen met 2 kostenindicatoren komen dan 2,23 procentpunt lager te liggen: de mate van correctie is de coëfficiënt voor 'aantal kostenindicatoren' in Tabel A3. Hetzelfde geldt voor aantal productie-indicatoren en de vraag of de dataset van voor of na 2000 is.

Deze notitie kiest voor het gemiddelde als de referentiewaarde. Dat is een intuïtieve keuze, omdat het gemiddelde van alle 87 scores zowel voor als na de opschoning dan gelijk is. De keuze voor de referentiewaarde heeft geen effect op de *verschillen* tussen de net berekende geschoonde individuele scores voor technische efficiëntie. De *absolute* grootte van de geschoonde waarden hangt daarentegen wel af van de gekozen referentiewaarde voor de verklarende variabelen.

³³ De variabele 'cluster' heeft als geheel wel een significante invloed op de technische efficiëntiescore.

De enige variabele waarvoor geen correctie is gewenst is 'cluster'. Op deze manier worden de scores voor alle 'externe invloeden' geschoond, behalve voor dat waar we in geïnteresseerd zijn: de efficiëntieverschillen tussen clusters. Als ook hiervoor gecorrigeerd zou zijn, zouden de clusters qua score bij definitie (los van wat ruis) gelijk worden en zegt de gemiddelde score per cluster niets meer.

Het effect van de opschoning op individuele scores blijkt aanzienlijk: de technische efficiëntiescores in een van de (Angelsaksische) artikelen daalde met 37%. Eén artikel krijgt door de opschoning een score vlak boven de 100%. Deze is teruggezet naar 100% alvorens door te gaan met de procedure. De gemiddelde verandering per groep van landen is beperkt, behalve in de Angelsaksische cluster.

Na de correctie naar referentiewaarden en de toevoeging van de constante, volgt de aggregatieprocedure zoals beschreven in Appendix A3. De geschoonde scores per model voegen we eerst samen tot een gemiddelde geschoonde score per artikel, en daarna per landencluster. De uiteindelijk verkregen geschoonde gemiddelde scores staan weergegeven in Tabel 3.3.

Correctie 2: Regressie op studieniveau

De correctie kan ook op studieniveau plaatsvinden, in plaats van op modelniveau. In een regressie is de gemiddelde efficiëntiescore per studie dan de afhankelijke variabele. Heeft een studie bijvoorbeeld vier modellen toegepast, dan is het gemiddelde van de vier bijbehorende scores de nieuwe afhankelijke variabele. De verklarende variabelen worden op dezelfde manier naar studieniveau gebracht. Zo geldt een gemiddelde voor de schaalvariabelen. De variabele 'methode' wordt als volgt meegenomen. De regressie bevat vier variabelen: het aandeel van elk van de vier methodes in de studie. Als een studie vier modellen gebruikt, waarvan drie DEA-VRS en een SFA, dan is de variabele fractie DEA-VRS voor die studie gelijk aan $\frac{3}{4}$, en de fractie SFA $\frac{1}{4}$. De rest van de methodes staan in dat geval op 0.

Tabel A4 geeft de output van de regressie weer. Qua resultaten wijken de schattingen niet veel af van de OLS-regressie op modelniveau, zeker niet als alleen de (gezamenlijk) significante variabelen – dat zijn in beide correcties dezelfde – overblijven. De mediane regressie wijkt ook hier niet zoveel af van de OLS-regressie. De effecten wijzen overal dezelfde kant op, dus uitbijters hebben wederom geen grote invloed op de schattingen. Wel komt de significantie van de variabelen wat lager uit, waarschijnlijk veroorzaakt door het kleinere aantal observaties in correctie 2.

Omrekening naar geschoonde scores

Eenzelfde procedure als bij correctie 1 wordt toegepast om de individuele scores op te schonen. De grootste wijziging is opnieuw in het Angelsaksische cluster te vinden: ook daar werd een studie met 37 procentpunt naar beneden bijgesteld. Twee studies komen net boven de 100% uit, die zijn teruggezet naar 100%.

Tabel A8 Correctie 2: technische efficiëntiescores verklaard, regressies op studieniveau


		OLS-regressie		Mediane regressie
		Alle variabelen	Significante variabelen	Alle variabelen
Variabele	Categorie	Coëfficiënt (%)	Coëfficiënt (%)	Coëfficiënt (%)
Cluster	Noordwest-Europa	(referentie)	(referentie)	(referentie)
	Angelsaksische landen	-27,13 ***	-28,22 ***	-22,24 *
	Ontwikkeld Azië	-4,32	-2,94	-4,87
	Centraal- en Zuid-Europa	-19,18 ***	-19,15 ***	-13,89 **
	BRICS-landen	-42,10 ***	-38,66 ***	-46,60 **
Methode	Fractie FDH	(referentie)		(referentie)
	Fractie DEA-CRS	6,15		4,76
	Fractie DEA-VRS	-2,52		-2,05
	Fractie SFA	-4,26		-6,36
Dataset vanaf 2000		-12,65 **	-11,65 ***	-12,07 *
Gemeenteselectie		2,32		3,63
Peer reviewed		-0,28		-0,28
Aantal kostenindicatoren		2,27 ***	2,37 ***	1,54 **
Aantal productie-indicatoren		1,57 *	1,33 *	0,85
Aantal gemeenten (a)		1,06		1,88
Aantal modellen		-1,26		-2,61
Constante		82,23 ***	85,27 ***	86,42 ***
R ²		0,63	0,61	nvt

De OLS-schattingen maken gebruik van Huber-White robuuste standaardfouten. De mediane regressie gebruikt Huber-Sandwich robuuste standaardfouten. Lege cellen geven weer dat de betreffende variabele is verwijderd op grond van statistische insignificantie. Deze verwijderde variabelen voegen gezamenlijk ook geen significante verklaringskracht toe aan het model met alleen de significante variabelen,
(a) In logaritmen.
*/**/** Significant op respectievelijk tien-, vijf- en eenprocentniveau. Aantal observaties: 41.

De resultaten van correctie 2 staan weergegeven in Tabel 3.3. Er is weinig verschil met de uitkomsten aan de hand van correctie 1. Op basis van het feit dat de resultaten van beide robuustheidscontroles vrijwel gelijk zijn, gaat de voorkeur uit naar correctie 1. Deze controle maakt gebruik van meer observaties en zal daarom per artikel een helderder beeld geven van de scores per land. Dat is ook de reden waarom het kader 'Empirische methoden voor technische efficiëntie' gebruik maakt van de scores op basis van correctie 1.

Samenvatting

Over het algemeen verandert er in elk van de robuustheidsanalyses niet zo veel aan de gemiddelde scores (behalve voor het cluster van Angelsaksische landen). Zo presteren de clusters Noordwest-Europa, België en Duitsland overal boven de 87. De correcties doen de scores zelfs verder stijgen, wat suggereert dat ze zeker niet door onderzoeksanpak of achtergrondkenmerken kunstmatig zijn opgehoogd. Voor Nederland, als onderdeel van het cluster Noordwest-Europa, is dus met enige voorzichtigheid te concluderen dat gemeenten op het gebied van technische efficiëntie relatief weinig van elkaar verschillen.



Dit is een uitgave van:

Centraal Planbureau
Van Stolkweg 14
Postbus 80510 | 2508 GM Den Haag
T (070) 3383 380

info@cpb.nl | www.cpb.nl

Oktober 2015