



TPEdigitaal
Jaargang 16 nr. 2
April 2022

Verzamelnummer:
Zorg, sociaal kapitaal,
welvaart

Het belang van sociaal kapitaal in Nederland	1
<i>Céline Odding en Bas ter Weel</i>	
Overwinsten vooral bij kleine zorgbv's	26
<i>Lucy Kok, Anouk Bekker, Frank van der Lee, Céline Odding en Ward Rougoor</i>	
Kosten van kerndepartementen: een empirische analyse van de productiviteitsontwikkeling tussen 2012 en 2019	34
<i>Jos Blank, Alex van Heezik en Bas Blank</i>	
<i>De CoronaMelder door een economische lens.....</i>	48
<i>Joost Poort</i>	
Maatschappelijke kosten-batenanalyse en monitoring van brede welvaart: wat kunnen ze van elkaar leren?	61
<i>Carl Koopmans</i>	
<i>Historicisme en zijn Monsters</i>	72
<i>Marcel Boumans</i>	

ISSN: 1875-8797

TPEdigitaal is een uitgave van de Stichting TPEdigitaal te Amsterdam.
ISSN: 1875-8797

Colofon

Het tijdschrift TPEdigitaal verschijnt vier maal per jaar en wordt uitgegeven door de onafhankelijke stichting TPEdigitaal te Amsterdam onder ISSN 1875-8797.

Redactie

drs. C. Biesenbeek

dr. K. Bruns

drs. A.R.B.J. Houkes-Hommes

dr. G.W. Meijerink

dr. J.M. Pomp

drs. P.R. Teule

dr. L.A.W. Tieben

Redactieadres

redactie@tpedigitaal.nl

Wetenschappelijke adviesraad

prof. dr. C.L.J. Caminada

prof. dr. H.P. van Dalen

prof. dr. P.W.C. Koning

prof. dr. C.C. Koopmans

prof. dr. H. Vollebergh

prof. dr. B. ter Weel

Bureauredactie

J.L. Verbruggen

Onderhoud website

R. Nieuwenhuizen

Foto website

Shutterstock.com. / superjoseph

Uitgever

Stichting TPEdigitaal

p/a Amstelveenseweg 1056

1081 JV Amsterdam

Het belang van sociaal kapitaal in Nederland

Céline Odding en Bas ter Weel

Gemeentes met meer sociaal kapitaal hebben lagere misdaadcijfers en hogere arbeidsparticipatie. Sociaal kapitaal is gemeten met indicatoren over onbaatzuchtigheid, solidariteit en wederkerigheid. Een stijging van het sociaal kapitaal met één standaarddeviatie leidt tot een daling van misdaad van 0,7 standaarddeviaties. Dit is een daling van 1,1 procentpunt bij een gemiddeld misdaadcijfer van 4,2 per 100 inwoners. De modellen wijzen uit dat een stijging van het sociaal kapitaal met één standaarddeviatie leidt tot een toename van de arbeidsparticipatie van 0,5 standaarddeviaties. In percentages is dit gelijk aan een toename van 1,3 procentpunt op een gemiddelde participatie van 70,1 procent.

1 Inleiding

Sociale cohesie is cruciaal voor een samenleving. Het leidt tot wederzijds begrip tussen individuen en tussen groepen met verschillende standpunten en achtergronden (Fukuyama, 1995; Putnam, 2000). Daarnaast is cohesie belangrijk voor het in stand houden van de democratie en het bouwen van effectieve instituties (Tabellini, 2005) en leidt cohesie tot economische voorspoed (Knack en Keefer, 1997). Als mensen elkaar vertrouwen is er meer ruimte voor experimenten die investeringen en innovatie aanwakkeren en bestaan er meer kansen om te participeren, bijvoorbeeld op de arbeidsmarkt. Sociale cohesie is diepgeworteld in de samenleving. Het is door de geschiedenis gevormd tot het huidige niveau en aan verandering onderhevig door de keuzes en het gedrag van mensen en groepen.¹ De huidige sociale cohesie in Nederland is ontstaan door publieke investeringen in scholing en geletterdheid, een tolerante bevolking die bruggen weet te slaan en een zekere homogeniteit van de bevolking als het gaat om normen en waarden.

Het doel van dit onderzoek is om de Nederlandse sociale cohesie in beeld te brengen en de impact ervan op de gemeenteniveau te meten. Om dit doel te bereiken moet een aantal stappen worden gezet.

Ten eerste moet sociale cohesie op een betrouwbare manier gemeten worden en moeten verschillen op een zo laag mogelijk aggregatieniveau en landelijk dekkend niveau inzichtelijk worden gemaakt. Dit wordt uitgewerkt in de volgende paragraaf aan de hand van het concept van sociaal kapitaal dat door economen wordt gebruikt.

Ten tweede is het van belang om de impact van dit gemeten sociaal kapitaal op maatschappelijke uitkomsten te meten. Het gaat om uitkomsten die op gemeenteniveau verschillen.

¹ Dit volgt uit empirisch onderzoek door Hall en Jones (1999), Acemoglu et al. (2005) en Tabellini (2005), die de argumenten toepassen die door North (1981) zijn ontwikkeld dat historisch opgebouwd 'kapitaal' belangrijk is voor de huidige economische resultaten en maatschappelijke uitkomsten en dynamiek.

Een van de uitkomsten betreft misdaad. Waarom is het misdaadniveau zo verschillend tussen gemeentes en wat is de rol van het gemeten sociaal kapitaal om deze verschillen te verklaren? Misdaad meten we aan de hand van het aantal misdrijven per inwoner. De tweede uitkomstmaat is bruto-arbeidsmarktparticipatie. Deze verschilt ook tussen gemeenten en is een indicator van het meer of minder kunnen deelnemen aan het economische proces. Een hogere participatie zorgt voor een diverse, fitte en dynamische economie. Opnieuw is de vraag wat de rol van het gemeten sociaal kapitaal is om verschillen in participatie te kunnen verklaren.

Ten derde moet overtuigend worden aangetoond dat het gemeten sociaal kapitaal impact heeft op deze twee uitkomstmaten. De causaliteit van deze relatie wordt aannemelijk gemaakt door een analyse waarbij historische gemeentekennmerken dienen om het effect van het huidige gemeten sociaal kapitaal op misdaad en arbeidsmarktparticipatie op gemeentenniveau te schatten.

2 Operationaliseren en meten van sociaal kapitaal in Nederlandse gemeentes

Sociale cohesie in een samenleving heeft betrekking op de structuur van sociale relaties tussen mensen en groepen. Voor individuen geeft de gemeenschap waarin ze opgroeien een basis waarop voortgebouwd wordt. Het is mogelijk om te investeren in deze basis. Zo ontstaat een vorm van sociaal kapitaal die voor een deel is gegeven (het startkapitaal) en afhankelijk is van de gemeenschap waarin mensen geboren worden en opgroeien. Door investeringen in, van en door de gemeenschap waartoe mensen behoren en door individuele keuzes en gedrag kan dit sociaal kapitaal verder toenemen. Het kapitaal kan ook afnemen bij gebrek aan investeringen. Om de prikkel om te investeren aan te wakkeren, moeten relaties bestaan of eenvoudig kunnen worden gevormd en moeten er kansrijke mogelijkheden zijn om banden tussen groepen te versterken.

2.1 Sociaal kapitaal

We definiëren sociaal kapitaal als netwerken van relaties tussen mensen die in een samenleving leven en werken, waardoor die samenleving effectief kan functioneren. De huidige voorraad van sociaal kapitaal wordt gevormd door het sociaal kapitaal uit het verleden en de investeringen in dat kapitaal op dit moment.²

Een factor die bijdraagt aan een effectieve opbouw van netwerken van relaties is de homogeniteit van gemeenschappen en samenlevingen. Homogeniteit leidt tot meer solidariteit. Een hogere mate van homogeniteit maakt onderlinge communicatie en wederzijds begrip binnen groepen eenvoudiger waardoor sterkere banden ontstaan en normen en waarden

² Verbindingen zijn kansrijker als ze voor beide partijen profijtelijk zijn. In de literatuur wordt het versterken van relaties gedefinieerd als *bonding* en het leggen van nieuwe relaties als *bridging* (Putnam, 2000). Beide vormen vergroten het sociaal kapitaal in een samenleving.

eenvoudiger verankerd kunnen worden in de samenleving als geheel (Golub en Jackson, 2012). Sociaal kapitaal dat op deze manier wordt gebouwd heeft niet noodzakelijk alleen positieve effecten, omdat het ook kan leiden tot sektevorming en tunnelvisie waaraan samenlevingen ten onder kunnen gaan (Berman, 2000).

Een tweede factor die bijdraagt aan de opbouw van netwerken van relaties is de mate waarin verbindingen zijn gemaakt tussen groepen in de samenleving. Bijvoorbeeld tussen jongeren en ouderen, hoog- en laagopgeleiden en autochtonen en migranten. Naarmate een samenleving beter in staat is geweest bruggen te bouwen, ontstaat een groter sociaal kapitaal en is deze samenleving succesvoller (Beugelsdijk en Van Schaik, 2005). Effectievere bruggen vergroten onbaatzuchtigheid in een samenleving, omdat mensen zich kunnen inleven in anderen en daar ook rekening mee houden.

Ten slotte bestaat er complementariteit tussen investeringen in menselijk en sociaal kapitaal. Als meer wordt geïnvesteerd in menselijk kapitaal vormen mensen van jongs af aan sterkere banden met elkaar die baten opleveren voor de samenleving (Glaeser et al., 1996). Dit is een vorm van wederkerigheid.

2.2 Meten

Het direct meten van sociaal kapitaal is complex, omdat het een latent construct is dat bestaat uit verschillende elementen, zoals onbaatzuchtigheid, solidariteit en wederkerigheid. Om deze dimensies te meten en in beeld te brengen op gemeenteniveau in Nederland, construeren we een aantal indicatoren.

Als indicator voor onbaatzuchtigheid meten we de omvang van vrijwillige bijdrages en schenkingen aan goede doelen op gemeenteniveau. Deze gegevens bestaan uit de collecte-opbrengsten per huishouden, afkomstig van het Centraal Bureau Fondsenwerving (CBF). Mensen geven aan goede doelen zonder dat ze daar direct iets voor terug willen. Een hogere mate van onbaatzuchtigheid zorgt voor sterkere banden in een lokale gemeenschap, omdat mensen elkaar helpen zonder daar direct een dienst voor terug te willen ontvangen of te willen worden gecompenseerd voor hun inspanning of donatie.

Solidariteit meten we aan de hand van de opkomst bij gemeenteraadsverkiezingen. Deze gegevens zijn afkomstig uit cijfers van de Kiesraad. Een hoge opkomst (gemeten als het percentage stemgerechtigden per gemeente) geeft aan dat mensen beseffen dat er een zeker mate van onderlinge afhankelijkheid is en dat instituties nodig zijn om deze onderlinge afhankelijkheid te laten werken.³ Meer betrokkenheid hierbij, zorgt voor sterkere lokale instituties en besluitvorming.

Ten slotte meten we wederkerigheid met een instrumenteel motief, namelijk het aantal bloeddonoren per gemeente. Deze indicator geeft aan dat mensen anderen willen helpen en

³ De opkomst van de kiezers wordt verondersteld burgerbetrokkenheid en participatie bij besluitvorming in de gemeenschap te meten (e.g., Putnam, 1993; Rosenfeld et al., 2001; Gatti et al., 2003).

er tevens op vertrouwen geholpen te worden als zij hulp nodig hebben (Guiso et al., 2004).⁴ Deze gegevens zijn afkomstig uit de Gezondheidsenquête en beschikbaar via het CBS.

2.3 Relatie sociaal kapitaal en misdaad

Misdaad is geconcentreerd in bepaalde geografische gebieden en komt in andere gemeentes veel minder voor. Verschillende niveaus van sociaal kapitaal zouden kunnen helpen verklaren waarom dit het geval is. De rol van sociaal kapitaal bij het verklaren van verschillen in misdaad tussen gemeentes is niet meteen triviaal (Glaeser et al., 1996). Sociaal kapitaal kan immers op verschillende manieren van invloed zijn op misdaad. Bijvoorbeeld doordat de kans op arrestatie hoger is in gebieden met meer sociale verbondenheid, doordat de kosten van misdaad hoger zijn in zulke gemeenschappen of doordat het verschil tussen illegale en legale inkomsten kleiner is in gebieden met meer sociaal kapitaal. Uit de economische en sociologische literatuur komt naar voren dat er waarschijnlijk twee hoofdkanalen bestaan waardoor sociaal kapitaal de mate van misdaad beïnvloedt.

Het eerste kanaal stelt dat door maatschappelijke betrokkenheid en informele sociale controle de kans op arrestatie na een misdaad groter is.⁵ Ten tweede verhogen informele controle (zoals nabuurschap) en de omvang van informele contacten de kosten van misdaad.⁶ Als een persoon een misdaad begaat, riskeert hij verlies van deze gunstige effecten van sociaal kapitaal via sancties (bijvoorbeeld door het verlies van zijn sociale netwerk), waardoor hij buiten zijn gemeenschap valt.

2.4 Relatie sociaal kapitaal en economische dynamiek

De relatie tussen sociaal kapitaal en economische dynamiek is op verschillende niveaus onderzocht (Knack en Keefer, 1997). De consensus is dat investeringen in onderwijs en innovatie vaker tot stand komen in een samenleving waarin mensen elkaar vertrouwen en de cohesie sterker is. Innovatie en onderwijs leiden vervolgens tot economische vernieuwing en groei. Er is een risico-element verbonden aan onderwijs en innovatie dat zich op verschillende manieren manifesteert en waardoor investeringen niet vanzelf van de grond komen. Het rendement op deze investeringen is onzeker, waardoor private partijen

⁴ Hoewel goede doelen en bloeddonaties vergelijkbaar lijken te zijn, is er verschil. Experimenteel onderzoek wijst op het effect van een goed gevoel bij de gever (Andreoni, 1995) en tegelijkertijd een gevoel van wederkerigheid. Sommige liefdadigheidsorganisaties geven kleine geschenken (pennen, ansichtkaarten, etc.) wat ertoe leidt dat de omvang van giften stijgt met de waarde van het geschenk (Falk, 2007). Compensatie voor het doneren van bloed kan echter bloeddonaties juist verdringen, omdat mensen het als onethisch ervaren om hiervoor directe compensatie te ontvangen (Titmuss, 1970; Mellstrom en Johannesson, 2008). In Nederland bestaat in principe geen vergoeding voor het doneren van bloed, waardoor we het effect van wederkerigheid meten.

⁵ In een studie over geweldsmisdrijven in Chicago rapporteren Sampson et al. (1997) lagere misdaadniveaus in buurten waarin burgers vaker samenwerken. Bursik en Grasmick (1993) vinden dat de effectiviteit van wetshandhaving en openbare controle hoger is in gemeenschappen met uitgebreide maatschappelijke betrokkenheid. Sampson en Groves (1989), Land et al. (1990), Rosenfeld et al. (2001) en Lederman et al. (2002) bevestigen deze bevindingen.

⁶ Williams en Sickles (2002) laten dat zien sociaal kapitaal leidt tot reputatievorming en sociale acceptatie genereert en dat het de anonimiteit vermindert.

terughoudend zijn. Ook bestaan er kapitaalmarktbeperkingen aan de vraagkant, waardoor het moeilijk is geld te lenen voor deze investeringen. Niet voor niets is een groot deel van het onderwijs publiek gefinancierd en wordt innovatie met subsidies en belastingkortingen aangewakkerd. Verder zijn monitoringkosten voor private partijen hoog bij onzekere investeringen, waardoor ze afzien van investeringen. Ten slotte is sprake van informatie-asymmetrie en is er de kans op moreel gevaar, omdat degene die wil investeren meer informatie heeft over het mogelijke succes dan de financier van de investering. Hierdoor kunnen investeringen worden belemmerd.

De vraag is op welke wijze sociaal kapitaal deze belemmeringen kan verminderen. Uit de wetenschappelijke literatuur blijkt dat dit op twee manieren mogelijk is. Ten eerste voorkomt sociaal kapitaal slecht gedrag door het afdwingen van informele normen. Hierdoor wordt minder geld verspild en zijn kapitaalmarktbeperkingen minder relevant (Guiso et al., 2004). Ten tweede zijn monitoringkosten laag als mensen elkaar vertrouwen en spelen de effecten van informatie-asymmetrie en moreel gevaar slechts een kleine rol (Akçomak en Ter Weel, 2009).

Het gevolg hiervan is dat sociaal kapitaal investeringen aanwakkert. Die investeringen leiden tot een hogere vraag naar kapitaal en arbeid. Schaarste zorgt ervoor dat de allocatie van kapitaal en arbeid efficiënter wordt, waardoor de productiviteit en daarmee het inkomen stijgt. We verwachten dus dat in gemeentes waarin meer sociaal kapitaal aanwezig is de arbeidsmarktparticipatie hoger is.

3 Model

De theorie is met een empirische analyse van het verband tussen sociaal kapitaal en maatschappelijke uitkomsten te toetsen. Het model voorspelt misdaad en arbeidsmarktparticipatie aan de hand van een aantal variabelen en sociaal kapitaal:

$$y_i = C + \beta_1 K_i + \beta_2 SC_i + \epsilon_i . \quad (1)$$

In deze vergelijking staat y_i voor de twee uitkomstmaten (i) misdaad en (ii) arbeidsparticipatie in gemeente i . We gebruiken de gemeente-indeling uit 2017 wat leidt tot 380 gemeentes waarvoor we data hebben. De set van variabelen in de vector K_i bestaat uit kenmerken van de gemeentes die relevant zijn voor het verklaren van de uitkomstmaten. De variabele SC_i meet het sociaal kapitaal in een gemeente. De laatste term in de vergelijking (ϵ_i) is de foutterm die de afwijkingen van het model ten opzichte van de werkelijke data oppikt.

Dit model kan niet zomaar worden geschat. Endogeniteit en de mogelijkheid van omgekeerde causaliteit kunnen de schattingen van de bovenstaande modellen onbetrouwbaar maken. Putnam (2000) stelt bijvoorbeeld dat een laag niveau van sociaal kapitaal kan resulteren in hogere misdaad, wat kan resulteren in nog lagere niveaus van sociaal kapitaal. Daarnaast is het mogelijk dat een ‘derde’ en onopgemerkt gebleven variabele zowel

misdaad en arbeidsmarktparticipatie als sociaal kapitaal beïnvloedt en niet in het model is opgenomen. Bepaalde beleidsmaatregelen van de lokale overheid kunnen bijvoorbeeld de arbeidsparticipatie stimuleren en tegelijkertijd een impact hebben op het sociaal kapitaal van de gemeente. In het geval van misdaad kan het zo zijn dat de huizenprijzen lager worden als er meer misdaad is, waardoor bepaalde groepen verhuizen en het sociaal kapitaal van een wijk daalt. Het kan ook voorkomen dat de rapportagepercentages over misdaad zijn gecorrigeerd met het niveau van sociaal kapitaal, waarbij inwoners in gebieden met een hoog sociaal kapitaal vaker melding maken van misdrijven.

Om deze problemen op te lossen passen we een strategie toe waarbij we het model in twee stappen schatten. Hiervoor hebben we een aantal instrumenten nodig die correleren met sociaal kapitaal, maar niet met het huidige niveau van misdaad en met de huidige arbeidsparticipatie in een gemeente (zie bijlage B, figuur B.1). We verklaren het sociaal kapitaal aan de hand van deze variabelen en de geschatte relatie wordt gebruikt om het causale effect van sociaal kapitaal op zowel misdaad als op arbeidsparticipatie te schatten.

De instrumenten die we gebruiken zijn het aandeel protestanten in een gemeente in 1859, het aantal scholen in 1859 en het aandeel buitenlandse burgers in 1859. We hebben het jaar 1859 gekozen omdat dit het eerste jaar is waarvoor administratieve en betrouwbare gegevens beschikbaar zijn over deze variabelen. Omdat sociaal kapitaal vaak diepgeworteld is in lokale gemeenschappen worden historische gegevens vaak toegepast in de empirische literatuur (zie Acemoglu et al., 2005).

Voor de opbouw van sociaal kapitaal is het leggen van verbindingen tussen groepen in een lokale gemeenschap van belang. De protestantse traditie bouwt deze bruggen en wordt gebruikt als schatting voor de mate waarin deze bruggen tot stand komen (P_i). Het aantal scholen (S_i) is een schatting van de mate waarin menselijk kapitaal wordt opgebouwd in de lokale gemeenschap. Fragmentatie (F_i) van de bevolking hindert de vorming van sociaal kapitaal, wat we meten met het aandeel buitenlandse burgers in 1859. Dit leidt tot het volgende model voor sociaal kapitaal:

$$SC_i = C + \gamma_1 P_{i,1859} + \gamma_2 S_{i,1859} + \gamma_3 F_{i,1859} + \gamma_4 K_i + \theta_i, \quad (2)$$

waarbij de laatste term in de vergelijking (θ_i) is de foutterm die de afwijkingen van het model ten opzichte van de werkelijke data oppikt. Het resultaat van dit model (\widehat{SC}_i) wordt gebruikt in de tweede stap om het causale effect van sociaal kapitaal op de uitkomstmaten te bepalen:

$$y_i = C + \beta_1 K_i + \beta_2 \widehat{SC}_i + \epsilon_i. \quad (3)$$

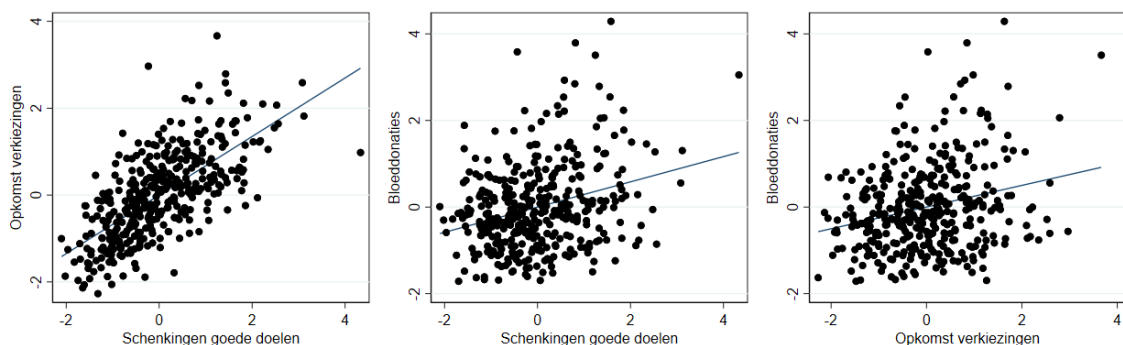
4 Data en beschrijvende statistiek

In deze studie gebruiken we verschillende data-bronnen en bouwen we een database met verschillende indicatoren ten behoeve van de empirische analyse. Bijlage A geeft een overzicht van alle bestanden die we gebruiken en definieert alle variabelen.

De indicator van sociaal kapitaal bestaat uit schenkingen aan goede doelen, de opkomst bij gemeenteraadsverkiezingen en het aantal bloeddonoren. Alle variabelen worden op gemeenteniveau gemeten en gestandaardiseerd, zodat ze kunnen worden vergeleken en gecombineerd in één indicator.

Er zijn 380 gemeentes, wat betekent dat we 380 datapunten in de empirische analyse hebben. De drie variabelen die dienen als input voor de indicator van sociaal kapitaal zijn sterk en positief met elkaar gecorreleerd (figuur 1) en laten een duidelijke spreiding zien.⁷ De correlaties tussen het aantal bloeddonoren en de andere twee variabelen is het minst sterk, wat te zien is aan de wat sterkere spreiding in de twee laatste panels ten opzichte van het eerste panel. De variabelen worden gecombineerd tot één indicator van sociaal kapitaal door middel van een hoofdcomponentenanalyse, op basis van de analyse in Akçomak en Ter Weel (2012).

Figuur 1 De variabelen die worden gebruikt voor een indicator van sociaal kapitaal correleren sterk



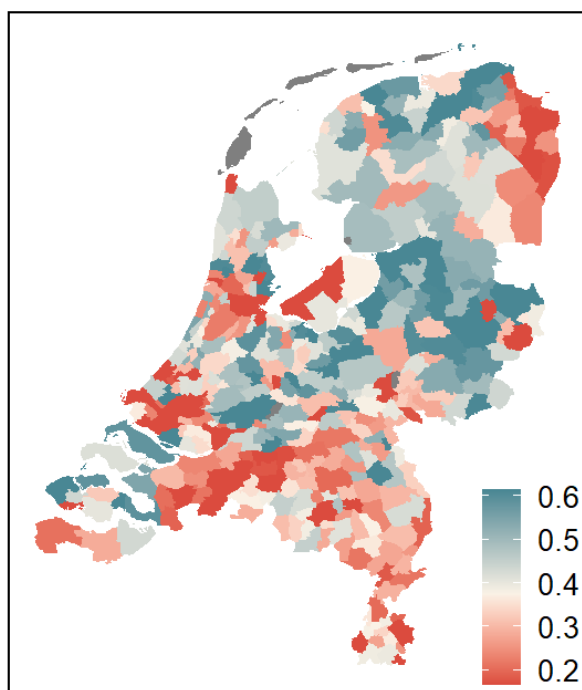
Noot: Alle variabelen zijn gestandaardiseerd. Bijlage A geeft een overzicht van de gebruikte databronnen en de bijbehorende beschrijvende statistiek.

Figuur 2 geeft de verschillen in sociaal kapitaal op gemeenteniveau in Nederland weer. Het valt op dat het sociaal kapitaal lager is in de grote steden, maar ook in Oost-Groningen en in het zuiden van Nederland. Een hoog niveau van sociaal kapitaal is te vinden in Friesland, Drenthe en het oosten van Nederland.

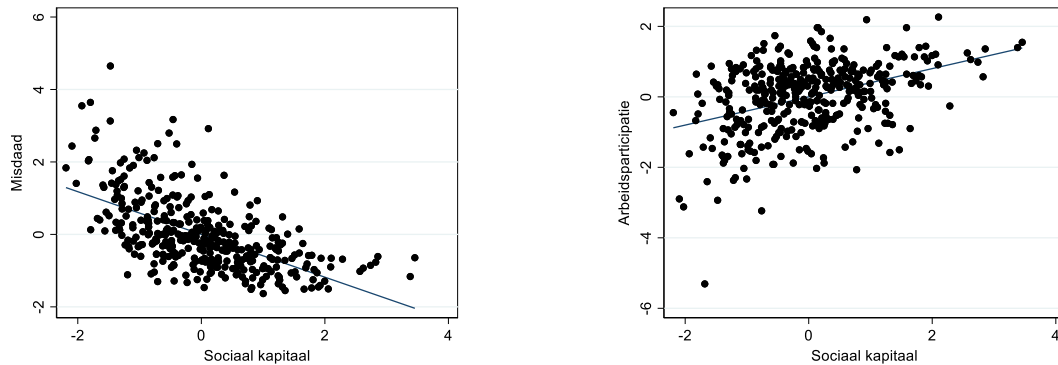
⁷ We testen de relatie van deze componenten met andere indicatoren voor de aan- of afwezigheid van sociaal kapitaal – zie bijlage B, tabel B.1.

De twee uitkomstmaten in dit onderzoek zijn misdaad en arbeidsparticipatie. De eerste wordt gemeten aan de hand van het aantal misdrijven per 100 inwoners, en de arbeidsparticipatie is gedefinieerd als het aandeel van alle volwassenen dat een baan heeft of op zoek is naar een baan (bruto arbeidsparticipatie). Deze variabelen zijn ook gestandaardiseerd. Misdaad is een breed begrip. Een misdrijf kan gaan om vernieling of diefstal, maar ook om moord. Figuur 3 toont de correlatie tussen de indicator van sociaal kapitaal en misdaad (linkerpanel) en arbeidsparticipatie (rechterpanel). Er is een negatieve correlatie tussen het gemeten sociaal kapitaal en de misdaadcijfers. Dit betekent dat in gemeentes waar het sociaal kapitaal relatief laag is, er gemiddeld meer misdaad voorkomt dan in gemeentes waar het sociaal kapitaal relatief hoog is. Het rechterpanel laat zien dat de relatie tussen sociaal kapitaal en arbeidsparticipatie positief is. Gemeentes met een relatief hoog niveau van gemeten sociaal kapitaal hebben gemiddeld een hogere arbeidsparticipatie vergeleken met gemeentes met een relatief laag niveau van gemeten sociaal kapitaal.

Figuur 2 Verschillen in het niveau van sociaal kapitaal tussen Nederlandse gemeentes



Noot: Sociaal kapitaal is genormaliseerd tussen 0 en 1. Bijlage A geeft een overzicht van de gebruikte databronnen en de bijbehorende beschrijvende statistiek.

Figuur 3 Relatie sociaal kapitaal en misdaad (links) en arbeidsparticipatie (rechts)

Noot: Alle variabelen zijn gestandaardiseerd. Bijlage A geeft een overzicht van de gebruikte databronnen en de bijbehorende beschrijvende statistiek.

5 Misdaad

Om de relatie van sociaal kapitaal en misdaad in een gemeente in kaart te brengen, beginnen we met het schatten van een standaard model (OLS) – zie vergelijking (1). Daarna bepalen we de causaliteit van de relatie met behulp van de instrumenten, in een zogenaamde 2SLS analyse – zie vergelijking (2) en (3).

Tabel 1 toont de schattingsresultaten van beide analyses. In kolom (1) is te zien dat misdaad en sociaal kapitaal negatief gecorreleerd zijn – in deze kolom presenteren we de resultaten van vergelijking (1). Wanneer we deze resultaten causaal zouden interpreteren, zorgt een stijging van het sociaal kapitaal in een gemeente met één standaarddeviatie voor een daling van de misdaad van 0,3 standaarddeviaties. Kolom (2) toont de geschatte coëfficiënten van vergelijking (2). Deze stap is nodig voor het schatten van een causaal verband. De instrumenten geven het verwachte teken en zijn statistisch gezien in orde (gezamenlijke F-test is gelijk aan 39,55). Kolom (3) presenteert de geschatte coëfficiënten van het 2SLS-model (vergelijking (3)) waarin een causaal effect waarschijnlijker is. Een stijging van het sociaal kapitaal met één standaarddeviatie leidt tot een daling van de misdaad van 0,7 standaarddeviaties. Omgerekend in percentages betekent dit dat wanneer het sociaal kapitaal met één standaarddeviatie toeneemt, het misdaadcijfer daalt met 1,1 procentpunt (in het niet-causale model is dat 0,5 procentpunt), gegeven dat het gemiddelde misdaadcijfer gelijk is aan 4,2 procent. Dat is een fors effect van sociaal kapitaal op deze uitkomst.

Een eenvoudig voorbeeld geeft duiding aan deze schattingsresultaten. De gemeente Zwolle kent een relatief hoog niveau van sociaal kapitaal. In de gemeente Delft is het sociaal kapitaal een stuk lager. Het verschil tussen beide gemeentes is nagenoeg gelijk aan één standaarddeviatie. Volgens de geschatte coëfficiënten in het model, zou een investering in

het sociaal kapitaal in Delft tot het niveau van het sociaal kapitaal in Zwolle, het misdaadcijfer laten dalen van 6,7 procent naar 5,6 procent.

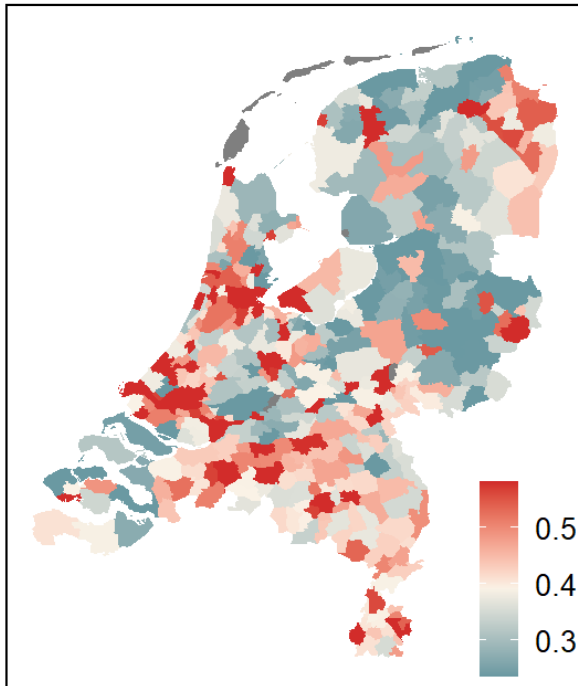
In model (1) en (3) is ook een aantal andere gemeentekennmerken opgenomen die het misdaadcijfer in een gemeente kunnen verklaren. Een hogere bevolkingsdichtheid is positief gecorreleerd met het misdaadcijfer, wat waarschijnlijk het effect van een stedelijke omgeving met meer mogelijkheden voor illegaal gedrag weergeeft. De samenstelling van de bevolking (gemiddeld opleidingsniveau, aandeel werklozen en aandeel jongeren) is ook van belang, al verschilt de mate. Inkomensongelijkheid, gemeten aan de hand van de Gini coëfficiënt, vergroot de misdaad. Het aanbod van winkels doet er niet toe. Wel van belang is het aantal coffeeshops en de recreatiemogelijkheden. Waarschijnlijk hebben beide aantrekkingskracht op criminelen.

Tabel 1 De relatie tussen sociaal kapitaal en misdaad

	(1) OLS	(2) 1st stage	(3) 2SLS
Afhankelijke variabele	Misdaad	Sociaal kapitaal	Misdaad
Sociaal kapitaal	-0,306 (0,035)***		-0,678 (0,080)***
Fragmentatie 1859		-0,066 (0,036)*	
Protestanten 1859		0,363 (0,041)***	
Aantal scholen 1859		0,080 (0,041)*	
Bevolkingsdichtheid	0,203 (0,053)***	-0,095 (0,072)	0,109 (0,063)*
Opleiding	-0,045 (0,043)	0,342 (0,049)***	0,086 (0,054)
Werkloosheid	0,084 (0,035)**	-0,133 (0,044)***	0,052 (0,039)
Jong	0,061 (0,049)	0,071 (0,048)	0,112 (0,053)**
Ongelijkheid (Gini)	0,229 (0,042)***	-0,025 (0,046)	0,212 (0,041)***
Winkels	-0,007 (0,044)	0,058 (0,047)	0,019 (0,048)
Recreatie	0,138 (0,059)**	-0,282 (0,060)***	0,067 (0,064)
Coffeeshops	0,273 (0,044)***	-0,149 (0,034)***	0,220 (0,043)***
Constante	-0,000 (0,029)	-0,000 (0,034)	-0,000 (0,033)
Observaties	380	380	380

Noot: Robuuste standaardfout tussen haakjes. *** $p < 0,01$, ** $p < 0,05$, * $p < 0,1$. Alle variabelen zijn gestandaardiseerd. De F-test, test de gezamenlijke significantie van de instrumenten en is gelijk aan 39,55***. De validiteit van de instrumenten is getest met als nulhypothese dat de instrumenten valide zijn en wordt niet verworpen. De adjusted R-squared voor (1), (2) en (3) is respectievelijk gelijk aan 0,68, 0,55 en 0,60. Bijlage A geeft een overzicht van de gebruikte databronnen en de bijbehorende beschrijvende statistiek.

Op basis van het sociaal kapitaal kan de misdaad per gemeente worden voorspeld. Hiervoor worden de modelschattingen uit kolom (3) van tabel 1 gebruikt. Figuur 4 geeft de voorspelde misdaad op basis van dit model weer. In gemeentes waar het sociaal kapitaal laag is, is de voorspelde misdaad hoog. Terwijl voor de gemeentes met een hoog sociaal kapitaal de voorspelde misdaad laag zijn. Deze resultaten zijn gebaseerd op lineaire modellen en voorspellen de misdaad in gemeentes op basis van verschillen in het gemeten sociaal kapitaal.

Figuur 4 Voorspelde misdaad in Nederland

Noot: De voorspelde misdaad is genormaliseerd tussen 0 en 1. Bijlage A geeft een overzicht van de gebruikte databronnen en de bijbehorende beschrijvende statistiek.

6 Economische dynamiek

Voor het schatten van de relatie tussen sociaal kapitaal en economische dynamiek doorlopen we dezelfde stappen als hierboven. Eerst schatten we het standaard model (OLS), waarna we de 2SLS-analyse uitvoeren. De 2SLS-analyse is deels aangepast, omdat het aandeel buitenlandse burgers in 1859 correleert met de huidige arbeidsparticipatie. Dat maakt dit instrument niet valide voor een dergelijke analyse.

Kolom (1) van tabel 2 toont de schattingsresultaten van het standaard model (vergelijking (1)). Het sociaal kapitaal in een gemeente heeft een positieve relatie met de arbeidsparticipatie. Als we de coëfficiënt causaal zouden interpreteren, leidt een stijging van het sociaal kapitaal in een gemeente met één standaarddeviatie tot een stijging van de arbeidsparticipatie van ongeveer 0,3 standaarddeviaties. Om het causale effect in beeld te brengen, gebruiken we de strategie waarbij we het model in twee stappen schatten. De resultaten van de eerste model stap (vergelijking (2)) zijn gegeven in kolom (2). De instrumenten hebben het verwachte teken en voldoen aan de statistische voorwaarden. De coëfficiënten van vergelijking (3) zijn te vinden in kolom (3) van tabel 2. Deze resultaten mogen we als causaal interpreteren. Een stijging van het sociaal kapitaal met één standaarddeviatie leidt tot een toename van de arbeidsparticipatie van 0,4

standaarddeviaties. In percentages is dit gelijk aan een toename van 1,2 procentpunt (in het niet-causale model is het 0,9 procentpunt).

Neem bijvoorbeeld de gemeentes Zoetermeer en Leiden. Het sociaal kapitaal in Leiden is nagenoeg één standaarddeviatie hoger dan het sociaal kapitaal in Zoetermeer. Volgens de schattingsresultaten leidt een investering in het sociaal kapitaal in Zoetermeer tot het niveau van het sociaal kapitaal in Leiden, tot een stijging van de arbeidsparticipatie van 70,5 naar 71,8 procent. Dit is een fors effect. De gemiddelde bruto-arbeidsparticipatie in Nederland is gelijk aan 70,1 procent met een standaarddeviatie van 2,7 procentpunt.

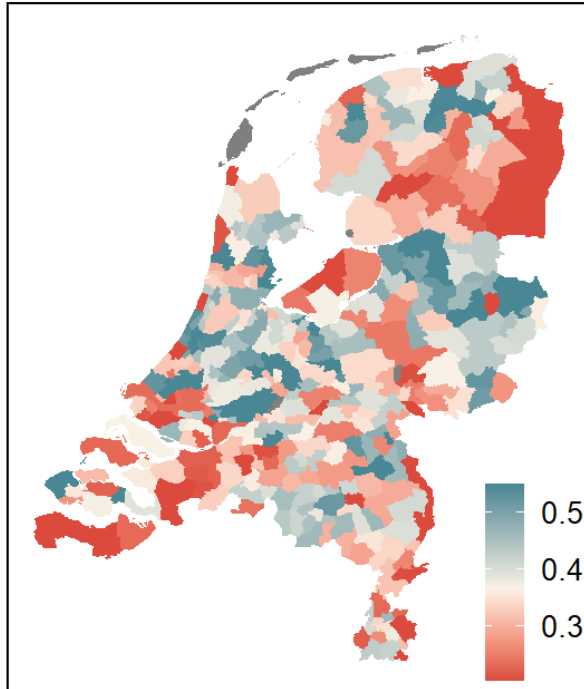
In de modellen in tabel 2 is ook een aantal gemeentekennmerken opgenomen die invloed hebben op de arbeidsparticipatie. De samenstelling van de bevolking heeft invloed op de arbeidsparticipatie. Zo heeft een gemeente met een gemiddeld hoger opgeleide bevolking gemiddeld een hogere arbeidsparticipatie, en in gemeentes waar veel mannen wonen is de participatie hoger. De bevolkingsdichtheid doet er ook toe. Bevolkingsdichtheid heeft een positieve correlatie met de arbeidsparticipatie. De arbeidsparticipatie in een stedelijke omgeving is gemiddeld hoger dan in landelijke omgeving. Inkomensongelijkheid heeft in dit model een negatieve statistische relatie met arbeidsparticipatie.

Tabel 2 De relatie tussen sociaal kapitaal en arbeidsparticipatie

	(1) OLS	(2) 1st stage	(3) 2SLS
Afhankelijke variabele	Bruto-arbeidsparticipatie	Sociaal kapitaal	Bruto-arbeidsparticipatie
Sociaal kapitaal	0,345 (0,044)***		0,435 (0,133)***
Protestanten 1859		0,360 (0,039)***	
Scholen 1859		0,096 (0,049)**	
Opleiding	0,452 (0,056)***	0,506 (0,041)***	0,410 (0,071)***
Bevolkingsdichtheid	0,446 (0,067)***	-0,276 (0,049)***	0,482 (0,093)***
Ongelijkheid (Gini)	-0,285 (0,064)***	-0,059 (0,055)	-0,277 (0,069)
Man	0,297 (0,090)***	0,076 (0,051)	0,300 (0,091)***
Constante	-0,000 (0,039)	-0,000 (0,037)	-0,000 (0,039)
Observaties	380	380	380

Noot: Robuuste standaardfout tussen haakjes. *** $p < 0,01$, ** $p < 0,05$, * $p < 0,1$. Alle variabelen zijn gestandaardiseerd. De F-test, test de gezamenlijke significantie van de instrumenten en is gelijk aan 44,29***. De validiteit van de instrumenten is getest met als nulhypothese dat de instrumenten valide zijn en wordt niet verworpen. De adjusted R-squared voor (1), (2) en (3) is respectievelijk gelijk aan 0,42, 0,49 en 0,41.

Op basis van de modelschattingen van model (3) in tabel 2 voorspellen we de arbeidsparticipatie per gemeente. Figuur 5 laat de voorspelde arbeidsparticipatie per gemeente zien, gegeven het sociaal kapitaal in die gemeente. Over het algemeen is de arbeidsparticipatie hoog op plekken waar het sociaal kapitaal ook hoog is, en laag waar het sociaal kapitaal laag is. Deze resultaten zijn gebaseerd op lineaire modellen en voorspellen de participatie in gemeentes op basis van verschillen in het gemeten sociaal kapitaal.

Figuur 5 Voorspelde arbeidsparticipatie in Nederland

Noot: De voorspelde participatie is genormaliseerd tussen 0 en 1. Bijlage A geeft een overzicht van de gebruikte databronnen en de bijbehorende beschrijvende statistiek.

7 Conclusie en discussie

Gemeentes met een hoger niveau van gemeten sociaal kapitaal hebben minder misdaad en een hogere arbeidsparticipatie. De geschatte coëfficiënten zijn statistisch significant en hebben economisch gezien een fors effect. Dat betekent dat toename van het sociaal kapitaal in gebieden waar dat laag is, omvangrijke positieve effecten kan hebben op het verlagen van misdaad en op het verhogen van arbeidsparticipatie.

Sociale verbanden en sociale cohesie zijn diepgeworteld en langzaam ontstaan in lokale gemeenschappen (North, 1981). We gebruiken de institutionele ontwikkeling in het verleden om de huidige niveaus van sociaal kapitaal te duiden met een regressiemodel waarin we causale effecten van sociaal kapitaal op uitkomsten proberen te kwantificeren. We behandelen sociaal kapitaal in dat model als een indicator waarvan de voorraad gedurende een lange periode is opgebouwd.

Vanuit beleidsoogpunt maakt dit onze studie niet direct van toepassing op het drastisch veranderen van het sociaal kapitaal in gemeentes, omdat dit langetermijninvesteringen vergt. Niettemin komt een aantal patronen naar voren. Fragmentatie van de bevolking

draagt niet bij aan het bouwen van sociaal kapitaal. Dat wil niet zeggen dat strikte homogeniteit van de bevolking in termen van samenstelling optimaal is, maar dat een te gefragmenteerde bevolking leidt tot een lager niveau van sociaal kapitaal en een minder evidente groei ervan. Dat heeft vervolgens een effect op misdaad en arbeidsparticipatie. Maatschappelijke participatie van burgers in lokale gemeenschappen en inburgering van nieuwe burgers in die gemeenschappen is daarom van belang om te bouwen aan gezamenlijk sociaal kapitaal. Vooral onder groepen die af lijken te haken en het maatschappelijk onbehagen voeden lijkt het van belang om ze te blijven bereiken (SCP, 2017). Het gaat hier vaak om lager opgeleide burgers die moeite hebben om inkomen te vergaren en burgers met een niet-westerse migratieachtergrond die moeizaam integreren (aanbod) en minder kansen lijken te krijgen (vraag).

Daarnaast is menselijk kapitaal van belang voor maatschappelijke participatie. De schattingen laten zien dat misdaad hoger is in gemeentes waar meer jongeren wonen en dat de participatie hoger is in gemeentes met een hoger opgeleide beroepsbevolking. Via het onderwijs kan in een vroeg stadium van het leven worden gebouwd aan menselijk kapitaal en is het mogelijk om gezinnen te ondersteunen als zich problemen aandienen. Hierdoor wordt misdaad een minder aantrekkelijke route (met een goede opleiding hebben jongeren meer te verliezen) en zal de arbeidsparticipatie op termijn toenemen. Ook worden in het onderwijs latere sociale netwerken gevormd die bijdragen aan het elkaar ondersteunen en helpen in lokale gemeenschappen (zoals nabuurschap), maar ook aan het vinden van werk.

Ten slotte is het van belang om betere verbindingen tussen groepen tot stand te brengen. Netwerken ontstaan langs lijnen van religie, opleidingsniveau, etniciteit en leeftijd. Dat mechanisme is niet eenvoudig te doorbreken en van alle tijden. Alleen veranderen deze scheidslijnen van bijvoorbeeld de zuilen in Nederland naar scheidslijnen langs etniciteit en opleidingsniveau. Het vergt dat verschillende groepen/netwerken elkaar vaker ontmoeten om wederzijds begrip en belangen tot stand te brengen.

Bijlage A. Data

Deze bijlage beschrijft de data die in dit onderzoek zijn gebruikt. De uitkomsten die op de kaarten worden getoond zijn genormaliseerd (tussen 0 en 1). Voor het schatten van de modellen zijn alle variabelen gestandaardiseerd.

De indicator voor het meten van sociaal kapitaal is opgebouwd uit drie componenten: schenkingen aan goede doelen, de opkomst bij de gemeenteraadsverkiezingen en het aantal bloeddonoren.

- Schenkingen aan goede doelen. Gemiddelde collecteopbrengsten per huishouden in 2017.
Bron: [https://www.cbf.nl/collecte-opbrengsten/filter\(province-12\)](https://www.cbf.nl/collecte-opbrengsten/filter(province-12))
- Opkomst verkiezingen. Het percentage kiesgerechtigden dat heeft gestemd bij de gemeenteraadsverkiezingen in 2014 en 2018. In Leeuwarden en Molenwaard vinden in beide jaren herindelingsverkiezingen plaats. De opkomst bij de verkiezingen in deze twee gemeentes is de opkomst voor de gemeenteraadsverkiezingen in 2010.
Bron: <https://www.verkiezingsuitslagen.nl/verkiezingen/detail/GR20180321>
- Bloeddonoren. Het aantal bloeddonoren per 100 inwoners. Gemiddelde over de jaren 2010 tot en met 2016. Omdat het om enquêteresultaten gaat, is het een gewogen gemiddelde. Voor gemeentes zonder bloeddonoren (34 gemeentes) is het aantal donaties een voorspelling op basis van de overige twee indicatoren voor sociaal kapitaal.
Bron: Gezondheidsenquête (GECON) 2010-2016 CBS Microdata.
- Sociaal kapitaal indicator: Eerste component van de hoofdcomponentenanalyse van drie indicatoren: schenking aan goede doelen, opkomst bij gemeenteraadsverkiezingen en het aantal bloeddonaties.

Met behulp van SC1 en SC2 analyseren we de robuustheid van de resultaten van het gemeten sociaal kapitaal (Zie bijlage B).

- SC1: Eerste component van de hoofdcomponentenanalyse van zes indicatoren: schenkingen aan goede doelen, bloeddonaties, opkomst bij gemeenteraadsverkiezingen, vertrouwen, het aandeel migranten en het aantal scheidingen.
- SC2: Eerste component van de hoofdcomponentenanalyse van vier indicatoren: schenkingen aan goede doelen, bloeddonaties, opkomst bij gemeenteraadsverkiezingen en vertrouwen.

Twee uitkomstmaten staan centraal in dit onderzoek.

- Misdaad. Het totaal aantal misdrijven per 100 inwoners. Gemiddelde over 2013, 2015 en 2017. Bron: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83648NED/table?ts=1610720787038>
- Arbeidsparticipatie. De bruto-arbeidsparticipatie in 2017. Dit is het aandeel van de (werkzame en werkloze) beroepsbevolking in de bevolking (beroeps- en niet-

beroepsbevolking). Deze definitie heeft betrekking op personen die in Nederland wonen (exclusief de institutionele bevolking). De gegevens gelden voor de bevolking van 15 tot en met 75 jaar.

Bron: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83933NED/table?dl=AFE9>

Instrumenten zijn noodzakelijk voor het meten van het causale effect van het gemeten sociaal kapitaal op de uitkomstmaten. De drie instrumenten zijn als volgt gedefinieerd:

- Scholen in 1859. Het aantal scholen per 100 inwoners in 1859. Wanneer het aantal scholen in een gemeente niet bekend is, is het aantal scholen gelijk aan het gemiddelde van de omliggende gemeentes (voor 67 gemeentes).
Bron: Volkstellingen Archief.
- Buitenlandse burgers in 1859. Het percentage buitenlandse burgers in de totale populatie in 1859. Wanneer het percentage buitenlandse burgers in een gemeente niet bekend is, is het percentage buitenlandse burgers gelijk aan het gemiddelde van de omliggende gemeentes (voor twee gemeentes).
Bron: Volkstellingen Archief.
- Protestanten in 1859. Het percentage protestanten in de totale populatie in 1859. Wanneer het percentage protestanten in een gemeente niet bekend is, is het percentage protestanten gelijk aan het gemiddelde van de omliggende gemeentes (voor twee gemeentes).
Bron: Volkstellingen Archief.

De achtergrondvariabelen in de modellen zijn als volgt gedefinieerd (alle op gemeentenniveau).

- Jong. Percentage inwoners tussen de 15 en 24 jaar in 2017.
Bron: Kerncijfers wijken en buurten 2017 – CBS.
- Bevolkingsdichtheid. Logaritme van de bevolkingsdichtheid in 2017.
Bron: Kerncijfers wijken en buurten 2017 – CBS.
- Werkloosheid. Percentage niet-werkende, werkzoekende jongeren van 16 t/m 22 jaar in 2017. Wanneer er geen data beschikbaar is in 2017 zijn de gegevens van 2016 of 2018 gebruikt. De overige missende waarden zijn aangevuld door het gemiddelde van de vijf omliggende gemeentes te nemen (15 gemeentes).
Bron: <https://www.cbs.nl/nl-nl/maatwerk/2019/38/indicatoren-jeugd-naar-gemeente-en-wijk-2016-2018>
- Opleiding. Percentage inwoners tussen de 18 en 50 jaar met een middelbaar of hoog opleidingsniveau in 2017.
Bron CBS microdata.
- Ongelijkheid. Ongelijkheid gestandaardiseerd besteedbaar inkomen: Gini-coëfficiënt in 2017.
Bron: <https://www.cbs.nl/nl-nl/longread/statistische-trends/2019/ongelijkheid-in-inkomen-en-vermogen?onepage=true>
- Recreatie. Percentage landoppervlakte gewijd aan recreatie in 2015.

Bron: <https://opendata.cbs.nl/#/CBS/nl/dataset/70072ned/table?ts=1610463205642>

- Winkels. Percentage landoppervlakte gewijd aan horeca en detailhandel in 2015.
Bron: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/70262ned/table?from-statweb>
- Coffeeshops. Aantal coffeeshops per 10.000 inwoners in 2016.
Bron: https://www.breuerintraval.nl/wp-content/uploads/2018/10/b154_MCN17.pdf
- Man. Het aandeel mannen in 2017.
Bron: Kerncijfers wijken en buurten 2017 – CBS.

De aanvullende maten voor het gemeten sociaal kapitaal zijn gemiddelde enquêteresultaten uit 2002, 2004, 2006 en 2008. Bron: European Social Survey (ESS).

- Vertrouwen medemens. Gebaseerd op de vraag: “De meeste mensen zijn te vertrouwen of je kan niet voorzichtig genoeg zijn.”
- Hulp. Gebaseerd op de vraag: “Mensen zijn meestal behulpzaam of ze denken alleen maar aan zichzelf.”
- Eerlijkheid. Gebaseerd op de vraag: “De meeste mensen proberen alleen maar gebruik van je te maken of proberen eerlijk te zijn.”
- Vertrouwen (algemeen). Het gemiddelde van drie indicatoren: het vertrouwen in de medemens, hulp en eerlijkheid.

De aanvullende maten voor de afwezigheid van sociaal kapitaal zijn:

- Scheidingen. Het percentage gescheiden inwoners in 2017.
Bron: Kerncijfers wijken en buurten 2017 – CBS.
- Migranten. Percentage inwoners met een migratieachtergrond in 2017.
Bron: Kerncijfers wijken en buurten 2017 – CBS.

Misdaad is ingedeeld per soort misdrijf.

- Gewelds- en seksuele misdrijven per 100 inwoners. Gemiddelde 2013, 2015, 2017.
Bron: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83648NED/table?ts=1610720787038>
- Verkrachtingen per 100 inwoners. Gemiddelde 2013, 2015, 2017.
Bron: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83648NED/table?ts=1610720787038>
- Diefstal en inbraak met geweld per 100 inwoners. Gemiddelde 2013, 2015, 2017.
Bron: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83648NED/table?ts=1610720787038>
- Diefstal per 100 inwoners. Gemiddelde 2016 en 2017.
Bron: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83903NED/table?ts=1611050466547>

- Diefstal vervoermiddelen per 100 inwoners. Gemiddelde 2016 en 2017.
Bron: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83903NED/table?ts=1611050466547>
- Totaal diefstal uit woning/schuur/e.d. en diefstal uit niet-residentiële gebouwen per 100 inwoners. Gemiddelde 2016 en 2017.
Bron: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83903NED/table?ts=1611050466547>
- Diefstal uit woning/schuur/e.d. per 100 inwoners. Gemiddelde 2016 en 2017.
Bron: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83903NED/table?ts=1611050466547>
- Misdrijven m.b.t. harddrugs per 100 inwoners. Gemiddelde 2013, 2015, 2017.
Bron: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83648NED/table?ts=1610720787038>

De alternatieve specificaties van inkomen in de robuustheidsanalyse in bijlage B zijn als volgt gedefinieerd.

- Inkomen per persoon. Gemiddeld inkomen per inwoner [x 1.000 euro] in 2017.
Bron: Kerncijfers wijken en buurten 2017 – CBS.
- Inkomen per inkomensontvanger. Gemiddeld inkomen per inkomensontvanger [x 1.000 euro] in 2017.
Bron: Kerncijfers wijken en buurten 2017 – CBS.
- Inkomen westerse migranten. Gemiddeld persoonlijk inkomen van inwoners met een westerse migratieachtergrond [x 1.000] in 2015.
Bron: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83920NED/table?ts=1611064488377>
- Inkomen niet-westerse migranten. Gemiddeld persoonlijk inkomen van inwoners met een niet-westerse migratieachtergrond [x 1.000] in 2015.
Bron: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83920NED/table?ts=1611064488377>
- Inkomensverschil: Inkomen westerse migranten gedeeld door het inkomen van niet-westerse migranten

Tabel A.1 Beschrijvende statistiek

Variabele	Gemiddelde	Standaarddeviatie	Minimum	Maximum
Aantal misdrijven	4,24	1,58	1,67	11,56
Arbeidsparticipatie	70,19	2,66	56,10	76,20
Schenken aan goede doelen	6,93	3,04	0,55	20,09
Bloeddonatie	4,46	2,28	0,55	14,24
Opkomst verkiezingen	57,36	6,13	43,45	79,84
Scheidingen	7,01	1,45	2,52	11,67
Vertrouwen (algemeen)	5,80	0,25	5,21	6,13
Vertrouwen in medemens	5,79	0,30	5,02	6,21
Hulp	5,35	0,30	4,62	5,81
Eerlijkheid	6,26	0,23	5,80	6,70
Migranten	14,89	8,31	3,66	52,71
Sociaal kapitaal indicator	0,00	1,36	-2,98	4,70
Protestanten in 1859	54,04	37,28	0,00	99,98
Buitenlandse burgers in 1859	1,73	2,31	0,00	18,90
Scholen in 1859	0,06	0,04	0,00	0,20
Bevolkingsdichtheid	830,44	1.002,22	59,00	6.347,00
Jong	11,60	1,59	8,62	23,72
Ongelijkheid (Gini-coëfficiënt)	0,27	0,04	0,22	0,52
Werkloosheid	1,64	0,69	0,40	4,26
Opleiding	0,82	0,04	0,71	0,91
Coffeeshops	0,12	0,25	0,00	2,08
Winkels	0,53	0,68	0,00	5,21
Man	0,50	0,01	0,47	0,53
Recreatie	4,41	4,09	0,30	28,97

Bijlage B. Robuustheidsanalyse

We analyseren de kwaliteit van de hoofdanalyse door middel van drie robuustheidsanalyses. Hierbij kijken we naar de specificatie van de indicator van sociaal kapitaal en misdaad en herhalen we de analyses voor verschillende subgroepen.

Indicator van sociaal kapitaal

Ten eerste kijken we naar de invloed van de specificatie van het gemeten sociaal kapitaal. De maatstaf van gemeten sociaal kapitaal bestaat uit schenkingen aan goede doelen, de opkomst bij gemeenteraadsverkiezingen en het aantal bloeddonoren. Eerst testen we de relatie van deze componenten met andere indicatoren voor de aan- of afwezigheid van sociaal kapitaal. Een indicator voor de aanwezigheid van sociaal kapitaal is het algemeen vertrouwen binnen de gemeente. Terwijl het uit elkaar vallen van gezinnen (scheidingen) en een heterogene bevolking juist een afwezigheid van sociaal kapitaal aangeven. Tabel B.1 laat zien dat deze indicatoren allemaal sterk met elkaar zijn gecorreleerd. Het feit dat vertrouwen, heterogeniteit van de bevolking en het aantal scheidingen sterk gecorreleerd zijn met de drie componenten van de indicator van het gemeten sociaal kapitaal is een bevestiging van de kwaliteit van de gezamenlijke maat voor het gemeten sociaal kapitaal.

Tabel B.1 Correlatie tussen de indicatoren van sociaal kapitaal

	(1)	(2)	(3)	(4)	(5)	(6)
(1) Giften aan goede doelen	1,00					
(2) Aantal bloeddonaoren	0,29	1,00				
(3) Opkomst verkiezingen	0,67	0,25	1,00			
(4) Vertrouwen in anderen	0,34	0,22	0,17	1,00		
(5) Aantal migranten	-0,66	-0,14	-0,51	-0,20	1,00	
(6) Echtscheidingen	-0,70	-0,27	-0,58	-0,16	0,66	1,00

Vervolgens construeren we twee alternatieve maatstaven voor sociaal kapitaal: SC1 en SC2. In SC1 zijn, naast schenkingen aan goede doelen, de opkomst bij gemeenteraadsverkiezingen en het aantal bloeddonaoren, ook het algemene vertrouwen, het aantal migranten en het aantal echtscheidingen meegenomen. SC2 is een combinatie van schenking aan goede doelen, bloeddonaaties, opkomst bij verkiezingen en het algemene vertrouwen. Tabel B.2 laat zien dat de 2SLS coëfficiënten voor de twee alternatieven erg dicht bij de coëfficiënt voor de hoofdspecificatie van het gemeten sociaal kapitaal liggen.

Tabel B.2 Resultaten van 2SLS-analyses met verschillende instrumenten

	Misdaad		Arbeidsparticipatie	
	(1) 2SLS	(2) F-test	(3) 2SLS	(4) F-test
SC hoofdspecificatie	-0,678 (0,080)***	39,55***	0,435 (0,133)***	44,29***
SC1	-0,700 (0,080)***	48,94***	0,478 (0,130)***	43,26***
SC2	-0,623 (0,071)***	40,86***	0,403 (0,117)***	46,36***
Giften aan goede doelen	-0,632 (0,073)***	43,50***	0,439 (0,127)***	40,10***
Opkomst verkiezingen	-0,702 (0,094)***	32,42***	0,508 (0,173)**	29,04***
Bloeddonaatie	-2,141 (0,952)*	1,72	0,991(0,411)*	4,73**
Migranten	0,637 (0,147)***	26,09***	-0,929(0,263)***	11,49***
Echtscheidingen	1,552 (0,310)***	9,35***	-0,818 (0,203)***	13,09***
Vertrouwen	-0,559 (0,107)***	12,43***	0,494 (0,162)**	15,31***
Observaties		380		380

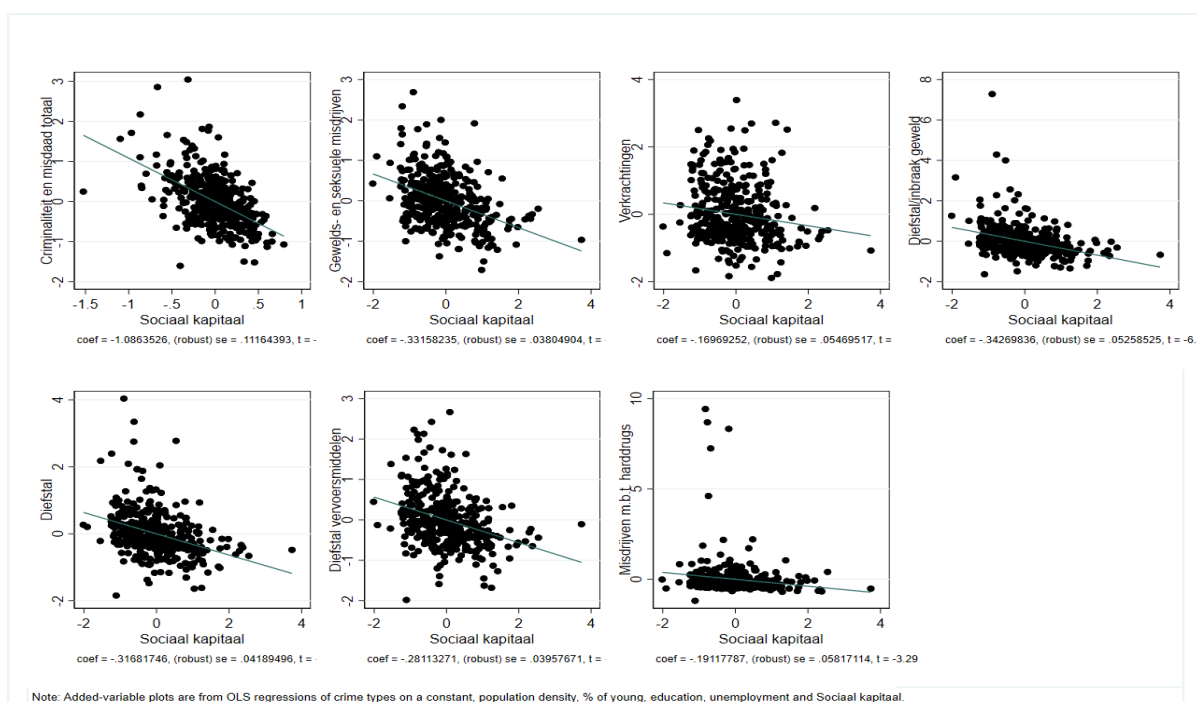
Noot: Kolom (1) en (3) geven de coëfficiënt in de 2SLS analyse (vergelijking (3)) voor verschillende specificaties voor het sociaal kapitaal. *** p<0,01, ** p<0,05, * p<0,1. Alle variabelen zijn gestandaardiseerd. Kolom (2) en (4) geven de resultaten van de test van de gezamenlijke significantie van de instrumenten.

De resultaten zijn dus consistent met onze indicator van het gemeten sociaal kapitaal (zie ook tabel 1 en 2). We schatten dezelfde specificaties ook voor alle bovengenoemde individuele indicatoren voor sociaal kapitaal (schenkingen aan goede doelen, bloeddonaaties etc.). Alle specificaties laten een significant effect van sociaal kapitaal op misdaad zien, met uitzondering van bloeddonaaties. Alle coëfficiënten hebben het verwachte teken. Deze conclusies gelden voor zowel misdaad als arbeidsparticipatie. Al met al, bevestigt dit de kwaliteit van de indicator van het gemeten sociaal kapitaal.

Misdaad

Naast de analyse over het totaal aantal misdrijven in een gemeente, voeren we de analyse ook uit op zes verschillende misdaadcategorieën. Dit doen we om uit te sluiten dat het effect van sociaal kapitaal op het totale misdaadcijfer een vertekende schatting geeft, omdat voor sommige categorieën het misdrijf minder vaak wordt gerapporteerd. De relatie tussen de verschillende categorieën en het gemeten sociaal kapitaal zijn weergegeven in figuur B.1. Alle categorieën hebben een negatieve relatie met het gemeten sociaal kapitaal. De correlatie tussen drugsmisdaad en sociaal kapitaal is minder sterk dan de overige categorieën.

Figuur B.1. Correlatie verschillende vormen van misdaad en sociaal kapitaal



Noot: Alle variabelen zijn gestandaardiseerd. De geschatte lijn en coëfficiënt zijn van de OLS regressie (vergelijking (1))

Tabel B.3 laat zien dat de resultaten voor deze verschillende categorieën overeen komen met de resultaten van het totale misdaadcijfer. Omdat de resultaten van de verschillende categorieën consistent met elkaar zijn, is het aannemelijk dat de hoofdanalyse geen vertekende resultaten weergeeft.

Tabel B.3 OLS-analyses voor verschillende categorieën van misdaad

	Misdaad	Gewelds- en seksuele misdrijven	Verkrachting	Diefstal en inbraak met geweld	Diefstal	Diefstal vervoersmiddelen	Diefstal woning en niet residentieel
Sociaal kapitaal	-0,306*** (0,035)	-0,270*** (0,038)	-0,134* (0,056)	-0,254*** (0,046)	-0,253*** (0,036)	-0,234*** (0,038)	-0,311*** (0,050)
Dichtheid	0,203*** (0,053)	0,0254 (0,055)	0,0824 (0,084)	0,0153 (0,074)	0,297*** (0,055)	0,315*** (0,065)	0,135 (0,081)
Opleiding	-0,0454 (0,043)	-0,207*** (0,041)	-0,125 (0,064)	0,00243 (0,053)	0,0898 (0,047)	0,146*** (0,044)	-0,0268 (0,053)
Werkloosheid	0,0837* (0,035)	0,295*** (0,037)	0,131* (0,066)	-0,0235 (0,054)	0,00553 (0,040)	0,166*** (0,042)	0,169*** (0,051)
Jong	0,0613 (0,049)	0,0949* (0,048)	0,118* (0,057)	0,0768 (0,069)	0,116 (0,059)	0,230*** (0,061)	-0,00662 (0,043)
Ongelijkheid	0,229*** (0,042)	0,131*** (0,037)	0,0521 (0,052)	0,121* (0,047)	0,189*** (0,044)	0,0812* (0,041)	0,283*** (0,047)
Winkels	-0,00670 (0,044)	0,0113 (0,064)	0,0462 (0,089)	0,0533 (0,070)	0,0151 (0,067)	0,0178 (0,073)	-0,0340 (0,058)
Recreatie	0,138* (0,059)	0,126* (0,054)	0,0512 (0,080)	0,313*** (0,085)	0,136 (0,071)	0,112 (0,078)	0,137* (0,064)
Coffeeshops	0,273*** (0,044)	0,257*** (0,044)	0,170* (0,075)	0,303* (0,138)	0,225** (0,071)	0,190*** (0,048)	0,141* (0,055)
Constante	-5,44e-10 (0,029)	7,89e-10 (0,028)	1,14e-08 (0,044)	-3,22e-09 (0,034)	-5,24e-10 (0,030)	-1,04e-09 (0,030)	-4,86e-11 (0,036)
<i>N</i>	380	380	380	380	380	380	380

Noot: Robuuste standaardfout tussen haakjes. De coëfficiënt zijn van de OLS-analyse (vergelijking (1)). *** p<0,01, ** p<0,05, * p<0,1. Alle variabelen zijn gestandaardiseerd.

Subgroepen

We herhalen de schattingen voor verschillende subgroepen - voor gemeentes met meer dan 30.000, 40.000 en 50.000 inwoners. Tabel B.4 laat zien dat er geen grote verschillen zijn ten opzichte van de resultaten in de hoofdtekst (zie tabel 1 en 2) en dat de resultaten dus niet afhankelijk zijn van de sample van gemeentes. Wel wordt arbeidsparticipatie minder precies gemeten, waardoor de statistische significantie daalt bij een kleinere sample.

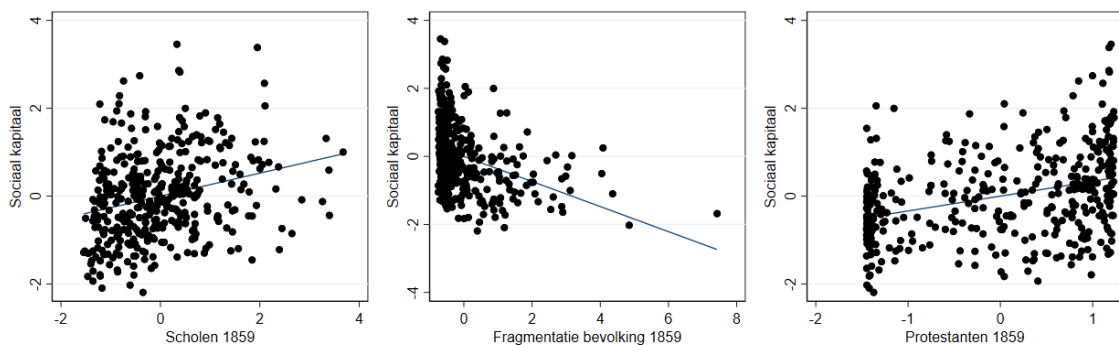
Tabel B.4 Resultaten voor verschillende subgroepen naar omvang

	Misdaad			Arbeidsparticipatie		
	Inwoners >30.000	Inwoners >40.000	Inwoners >50.000	Inwoners >30.000	Inwoners >40.000	Inwoners >50.000
SC	-0,311 (0,054)***	-0,350 (0,064)***	-0,389 (0,098)***	0,189 (0,077)***	0,146 (0,095)	0,141 (0,123)
N	165	118	79	165	118	79

Noot: Robuuste standaardfout tussen haakjes. De coëfficiënt zijn van de OLS-analyse (vergelijking (1)). *** p<0,01, ** p<0,05, * p<0,1. Alle variabelen zijn gestandaardiseerd.

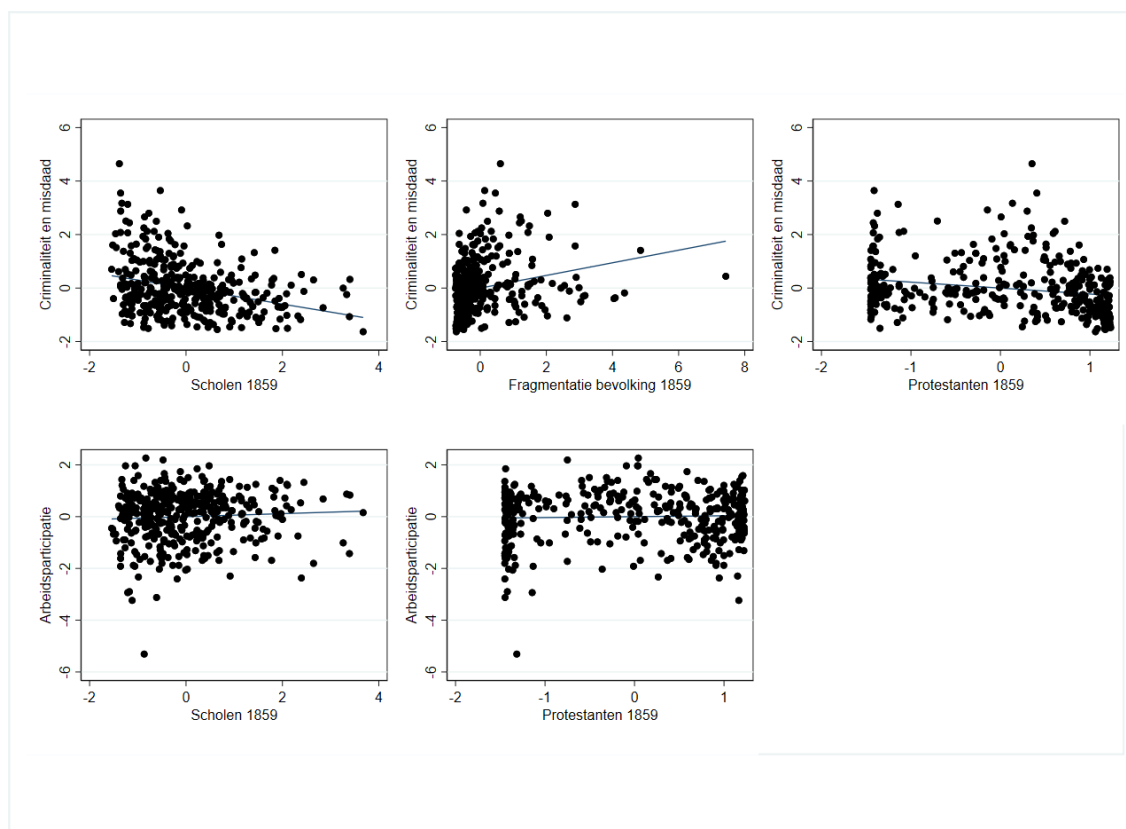
Bijlage C. Instrumenten

Figuur C.1 geeft de correlatie tussen de instrumenten die worden gebruikt in de modellen (model (2) en (3)) en het sociaal kapitaal. Voor de identificatie van de modelschattingen is het van belang dat er een correlatie aanwezig is tussen de instrumenten en het gemeten sociaal kapitaal. Zoals verwacht is het aantal scholen in 1859 en het aantal protestanten in 1859 positief gecorreleerd met het gemeten sociaal kapitaal. De fragmentatie van de bevolking in 1859 is negatief gecorreleerd met het gemeten sociaal kapitaal.

Figuur C.1 Correlatie historische gemeentekennmerken met sociaal kapitaal

Noot: Alle variabelen zijn gestandaardiseerd. Alle punten geven één van de 380 gemeentes weer.

Daarnaast is het noodzakelijk dat de instrumenten geen direct effect hebben op misdaad en economische dynamiek. Figuur C.2 geeft de correlatie tussen deze variabelen weer. Er is geen sterke correlatie tussen de instrumenten en de uitkomsten. Deze hypothese is ook statistisch getest middels een test voor overidentificatie. De resultaten hiervan bevestigen de conclusie dat de instrumenten valide zijn (zie ook noot tabel 1 en 2). De fragmentatie van de bevolking in 1859 heeft echter wel een effect op de huidige arbeidsparticipatie. Om deze reden gebruiken we voor de arbeidsparticipatie uitkomst maar twee instrumenten: het aantal scholen en het aantal protestanten in 1859.

Figuur C.2 Correlatie historische gemeentekennmerken met de uitkomstmaten

Noot: Alle variabelen zijn gestandaardiseerd. Alle punten geven een van de 380 gemeentes weer.

Auteurs

Céline Odding is promovendus aan de economische faculteit van de Rijksuniversiteit Groningen. Bas ter Weel (e-mail: b.terweel@seo.nl) is Algemeen directeur van SEO Economisch Onderzoek en hoogleraar aan de economische faculteit van de Universiteit van Amsterdam.

Literatuur

- Acemoglu, D., S. Johnson en J. Robinson, 2005, The rise of Europe: Atlantic trade, institutional change, and economic growth, *American Economic Review*, vol. 95(3): 546-579.
- Akçomak, I.S. en B. ter Weel, 2009, Social capital, innovation and growth: Evidence from Europe, *European Economic Review*, vol. 53(5): 544-567.
- Akçomak, I.S., en B. ter Weel, 2012, The impact of social capital on crime: Evidence from the Netherlands, *Regional Science and Urban Economics*, vol. 42(1-2): 323-340.

- Berman, E., 2000, Sect, subsidy, and sacrifice: an economist's view of ultra-orthodox Jews, *Quarterly Journal of Economics*, vol. 115(3): 905-953.
- Beugelsdijk, S. en T. van Schaik, 2005, Differences in social capital between 54 western European regions, *Regional Studies*, vol. 39(8): 1053-1064.
- Bursik, R.J. en H.G. Grasmick, 1993, *Neighborhoods and Crime: The Dimensions of Effective Community Control*. Lexington Books.
- Falk, A., 2007, Gift-exchange in the field, *Econometrica*, vol.75(5): 1501-1511.
- Fukuyama, F., 1995, *Trust. The Social Virtues and the Creation of Prosperity*. Free Press, New York
- Glaeser, E.L., B. Sacerdote en J.A. Scheinkman, 1996, Crime and social interactions, *Quarterly Journal of Economics*, vol. 111(2): 507-548.
- Golub, B. en M.O. Jackson, 2012, How homophily affects the speed of learning and best-response dynamics., *Quarterly Journal of Economics*, vol. 127(3): 1287-1338.
- Guiso, L., P. Sapienza en L. Zingales, 2004, The role of social capital in financial development, *American Economic Review*, vol. 94(3): 526-556.
- Hall, R.E. en C.J. Jones, 1999, Why do some countries produce more output per worker than others?, *Quarterly Journal of Economics*, vol. 114(1): 83-116.
- Knack, S. en P. Keefer, 1997, Does social capital have an economic payoff? A cross-country investigation, *Quarterly Journal of Economics*, vol. 112(4): 1251-1288.
- Land, K.C., P.L. McCall en L.E. Cohen, 1990, Structural covariates of homicide rates: are there any invariance across time and space, *American Journal of Sociology*, vol. 95(4): 922-963.
- Lederman, D., N. Loayza en A.M. Menendez, 2002, Violent crime: does social capital matter?, *Economic Development and Cultural Change*, vol. 50(3): 509-539.
- Mellstrom, C. en M. Johannesson, 2008, Crowding out in blood donation: was Titmuss right?, *Journal of the European Economic Association*, vol. (4): 845-863.
- North, D., 1981, *Structure and Change in Economic History*, Norton, New York.
- Putnam, R.D., 2000, *Bowling Alone: The Collapse and Revival of American Community*, Simon & Schuster, New York.
- Rosenfeld, R., S.F. Messner en E.P. Baumer, 2001, *Social capital and homicide*. *Social Forces*, vol. 79(1): 283-310.
- Sampson, R.J., 1988, Local friendship ties and community attachment in mass society: a multilevel systemic model, *American Sociological Review*, vol. 53(5): 766-779.
- SCP, 2017, Burgerperspectieven 2017/1. SCP: Den Haag. Burgerperspectieven 2017|1 | Monitor | Sociaal en Cultureel Planbureau (scp.nl).
- Tabellini, G., 2005, Culture and institutions: Economic development in the regions of Europe. Working Paper, Bocconi University and IGER.
- Titmuss, R.M., 1970, *The Gift Relationship: From Human Blood to Social Policy*. George Allen and Unwin, Berkley.
- Williams, J., Sickles, R.C., 2002, An analysis of the crime as work model: evidence from the 1958 Philadelphia birth cohort study, *Journal of Human Resources*, vol. 37(3): 479-509.

Overwinsten vooral bij kleine zorgbv's¹

Lucy Kok, Anouk Bekker, Frank van der Lee, Céline Odding en Ward Rougoor

Er is de laatste jaren veel te doen om de grote winsten die sommige zorgaanbieders maken. In dit artikel brengen we dit in kaart voor de extramurale zorg en bespreken de mogelijkheid van het beleidsmatig inperken van de mogelijkheden deze winst uit te keren. Overwinsten blijken vooral voor te komen bij kleine zorgbedrijven, die minder dan 4% van de zorg leveren. Overwinsten dragen daardoor nauwelijks bij aan hoge zorgkosten. Beleid om de winstuitkering te maximaleren kan zowel positieve als negatieve effecten hebben op de toegankelijkheid, betaalbaarheid en kwaliteit van de zorg. Op voorhand is niet te voorspellen of de positieve of negatieve effecten zullen overheersen. Zeker is wel dat dit beleid gepaard gaat met hogere administratieve lasten voor zorgaanbieders en toezichthouders.

1 Inleiding

"Ongebruikelijk hoge winsten bij bijna honderd zorgbedrijven" kopte de NRC op 26 juni 2019 naar aanleiding van onderzoek van Pointer en Follow the Money. De minister van VWS zei vervolgens: "we worden te vaak geconfronteerd met gevallen van fraude, belangenverstrengeling en excessieve winstuitkering." De maatschappelijke doelstelling van de zorgaanbieder, namelijk het leveren van kwalitatief goede, veilige en betaalbare zorg, zou daarbij ondergeschikt gemaakt worden aan zakelijke of privébelangen van individuen (VWS, 2019). Dit was aanleiding voor Kamervragen en de belofte van de minister van VWS om te onderzoeken of het mogelijk is om een norm te stellen voor een maatschappelijk maximaal aanvaardbare winstuitkering.

Maar hoe groot is het probleem nu eigenlijk? Hoe groot zijn de excessieve winsten en hoe vaak komen ze voor? Het is moeilijk om een exact antwoord te geven, omdat de zorgsector sterk gefragmenteerd is en de rapportage die de spelers moeten aanleveren wettelijk beperkt. Op basis van de bij het ministerie ingeleverde jaarverslagen doen we een poging voor de extramurale zorg, die onder de Zorgverzekeringswet en de Jeugdwet valt.

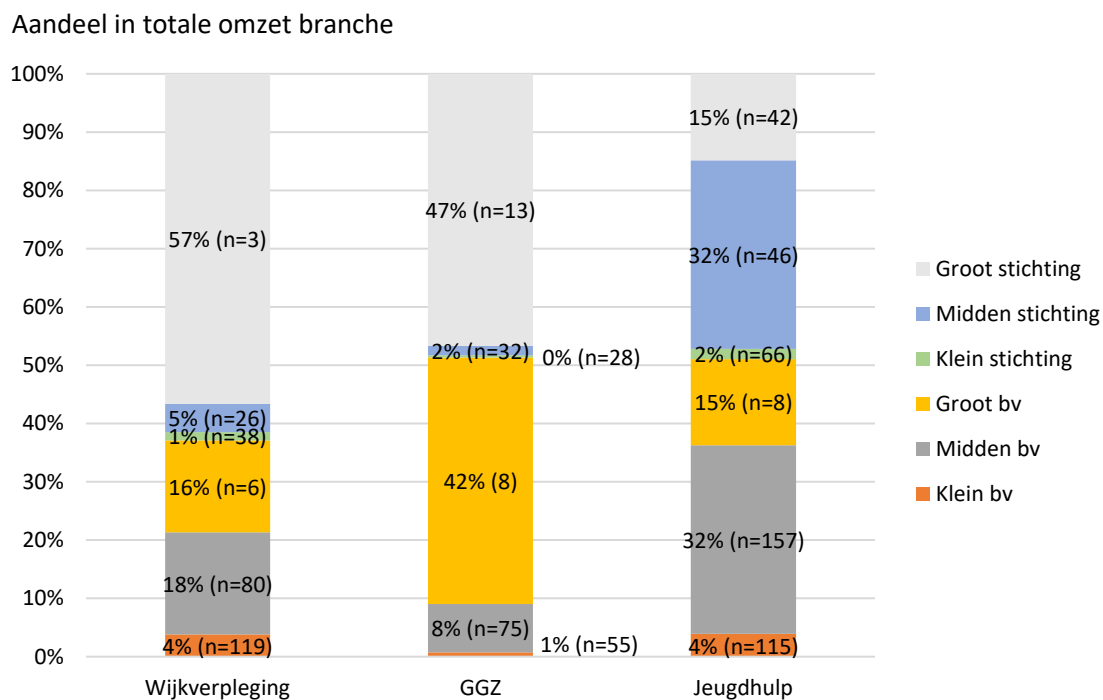
2 Winsten in de extramurale zorg

De extramurale zorg onder de Zorgverzekeringswet en de Jeugdwet omvat de wijkverpleging, de geestelijke gezondheidszorg (GGZ) en de jeugdzorg. In dit deel van de zorgsector zijn talloze kleine aanbieders actief en een aantal grote. Toch leveren de kleine bv's, gemeeten als percentage van de totale omzet, slechts één procent van de GGZ-zorg, en vier procent

¹ Dit artikel is gebaseerd op het onderzoek dat SEO Economisch Onderzoek samen met BDO Advisory heeft verricht in opdracht van het ministerie van VWS, zie Kok et al. 2020.

van de wijkverpleging en jeugdzorg (figuur 1). De bijdrage van middelgrote bv's varieert van 8 tot 32 procent, terwijl de grote bv's 15 tot 42 procent van de omzet voor hun rekening nemen. Naast bv's wordt een groot deel van de zorg geleverd door stichtingen, die formeel geen winst mogen uitkeren. In de Wijkverpleging gaat het daarbij om in totaal 63 procent van de zorg, waarbij 57 procent van de totale omzet geleverd wordt door grote aanbieders. Bij de GGZ en de Jeugdzorg wordt 49 procent van de zorg geleverd door stichtingen, waarvan respectievelijk 47 en 15 procent door grote aanbieders, en maximaal 2 procent door een groot aantal kleine stichtingen.

Figuur 1 Omzet kleine aanbieders maakt klein deel uit van totale omzet in de branche



Bron: Berekeningen gebaseerd op jaarverslagen over 2018, beschikbaar via Digimv. Zie Kok et al. (2020).

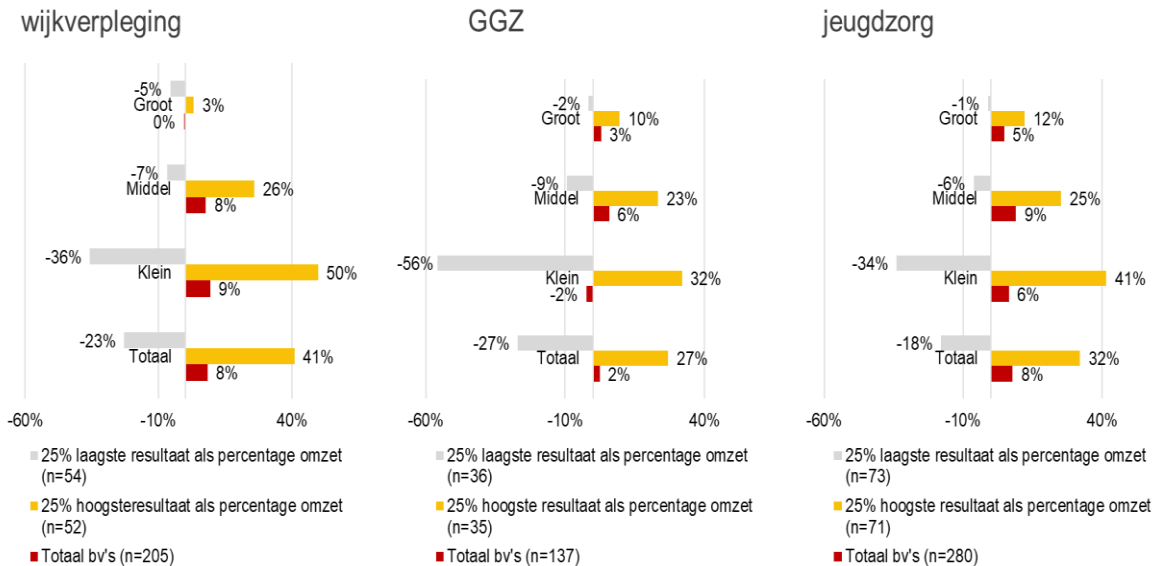
Figuur 2 laat de gemiddelde winst zien voor alle bv's per sector, en per grootteklasse, op basis van de gemiddelde winst van de 25 procent bedrijven met de hoogste en de laagste winst.

De gemiddelde winst van bv's bedraagt in 2018 in de wijkverpleging, de GGZ en de jeugdzorg respectievelijk 8, 2 en 8 procent van de (bedrijfs)omzet, zie figuur 2. Kleine en middelgrote bv's maken gemiddeld meer winst dan grote aanbieders. Bij kleine bv's is de spreiding van de winsten het grootst: zij maken de hoogste winsten, maar ook de grootste verliezen. De gemiddelde winst van de 25 procent succesvolste kleine bedrijven bedraagt voor de wijkverpleging 50 procent, voor de GGZ 32 procent en in de jeugdzorg 41 procent. Voor de succesvolste grote bv's gaat het respectievelijk om 3, 10 en 12 procent. De situatie

voor stichtingen vertoont eenzelfde patroon. De resultaten komen overeen met eerder onderzoek over winsten in de zorg, op basis van cijfers over 2014 en 2015 (Suijs en Verbon 2018).

De winsten voor een deel van de kleine bv's zijn dus aanzienlijk, hoewel vermeld moet worden dat het hier slechts de winsten over 2018 betreft. De grote spreiding van de winsten (zowel naar boven als naar beneden) doet vermoeden dat wanneer over meerdere jaren wordt gekeken de uitschieters kleiner zijn: bedrijven kunnen immers niet langdurig grote verliezen maken. Suijs en Verbon (2018) suggereren dat de grote spreiding in de winsten er ook op zou kunnen duiden dat de zorg door de aantrekkelijke winsten die gemaakt kunnen worden ondernemers met weinig ervaring in de zorg aantrekt. Een groot aantal mislukkingen zou de grote uitschieters naar beneden kunnen verklaren. In het totaal aan geleverde zorg hebben de uitschieters in de winsten van kleine bedrijven echter een klein aandeel, waardoor de invloed op de betaalbaarheid van de zorg beperkt is.

Desondanks is er sprake van maatschappelijk onbehagen over de hoge winsten van sommige aanbieders, mede doordat de zorg collectief wordt gefinancierd. Volgens minister van VWS Hugo de Jonge is "geld voor de zorg bedoeld voor zorg en moet alleen aan de zorg worden besteed." (NOS 2019). Dat er hoge winsten worden gemaakt wil overigens niet per definitie zeggen dat het geld niet aan zorg wordt besteed. Winsten kunnen ook worden gebruikt om te investeren in de zorg, zie box 1.

Figuur 2 Vooral kleine bv's in de extramuraal zorg maken hoge winsten

Bron: Berekeningen gebaseerd op jaarverslagen over 2018, beschikbaar via Digimv. Zie Kok et al. (2020).
 Toelichting: De cijfers betreffen alleen bv's. Kleine bedrijven: omzet minder dan € 700.000 per jaar, middelgrote bedrijven: omzet € 700,000 tot € 12 miljoen per jaar, grote bedrijven: omzet meer dan € 12 miljoen per jaar. Weergegeven zijn de cijfers over het totaal aantal bv's en de 25 procent bv's met de hoogste en de laagste winsten binnen de groep.

Box 1 Winst en winstuitkering

Winst kan worden uitgekeerd of toegevoegd aan het eigen vermogen. Het niet-uitgekeerde deel van de winst komt in het eigen vermogen en blijft eigendom van de zorgaanbieder. Deze kan daarmee investeren in de zorgverlening. Als (een deel van) de winst wordt uitgekeerd kan dat als dividend aan aandeelhouders, of als vergoeding voor arbeid aan het personeel of de eigenaar. Zelfstandige zorgverleners, zoals huisartsen en medisch specialisten, keren hun winst aan zichzelf uit, als vergoeding voor hun arbeid. Stichtingen en bv's kunnen ook winst als resultaatdeling uitkeren aan het personeel. Bv's kunnen daarnaast winst als dividend uitkeren aan aandeelhouders. Dividend is een vergoeding voor het kapitaal dat aandeelhouders hebben ingebracht. In de praktijk wordt de dividenduitkering ook als vergoeding voor arbeid uitgekeerd, als de eigenaar ook directeur is. Dit is fiscaal interessant, omdat dividend anders belast wordt dan inkomen uit arbeid. Bovendien vallen dividenduitkeringen niet onder het inkomensbegrip van de Wet normering topinkomens. De Belastingdienst ziet erop toe dat DGA's niet te veel winst als dividend en te weinig loon aan zichzelf uitkeren door normen te stellen aan het loon dat een DGA zichzelf minimaal moet toekennen. In hoeverre de winst wordt uitgekeerd, en of dat gebeurt in de vorm van een dividenduitkering of als vergoeding voor arbeid is niet te achterhalen uit de jaarverslagen.

3 Oorzaken van overwinsten?

Vanuit economisch perspectief is er sprake van overwinsten als het rendement op het geïnvesteerde kapitaal hoger is dan de kosten van kapitaal. De kosten van kapitaal bestaan uit het verwachte rendement op alternatieve investeringen met een vergelijkbaar risico. Op het niveau van een individueel bedrijf is het niet eenvoudig te bepalen of er sprake is van overwinsten. Winsten fluctueren van jaar op jaar en ex-post zullen volgens deze definitie veel bedrijven soms 'overwinsten' en soms 'onderwinsten' realiseren, ook als gemiddeld over een langere periode sprake is van normale winsten. Van reële overwinsten is daarom pas sprake wanneer de winst voorspelbaar of structureel boven de kosten van kapitaal ligt. Over de vraag wanneer dit het geval is vooralsnog geen consensus (Pomp 2019).

Oorzaken van structurele overwinsten liggen in de marktstructuur. In een volledig competitieve markt zijn er geen overwinsten. Overwinsten ontstaan als er sprake is van een gebrek aan concurrentie of als er sprake is van informatieasymmetrie. Een voorbeeld van informatieasymmetrie is dat consumenten niet altijd de kwaliteit van een product of dienst kunnen beoordelen. De aanbieder met een kwalitatief slecht product kan dan toch hoge prijzen vragen. Oorzaken kunnen ook liggen in overheidsregulering. In de zorgsector zijn patiënten verzekerd waardoor zij niet de werkelijke zorgkosten, maar alleen de verzekeringspremie betalen. Daardoor werkt het prijsmechanisme niet goed, waardoor aanbieders hun prijzen te hoog kunnen zetten.

Er zijn in de praktijk twee manieren waarop regulering voorkomt dat bedrijven te veel winst maken. Namelijk door omzetregulering, waarbij zowel de prijs als het volume bepaald wordt door de toezichthouder, of door tariefregulering. Daarbij bepaalt de toezichthouder de tarieven, waarmee ze indirect een grens stelt aan de omzet van het gereguleerde bedrijf en dus de mogelijkheid om te hoge winsten te maken. Een bedrijf kan dan alleen meer winst maken door de afzet te verhogen of de kosten te verlagen. Beide methodes worden toegepast in Nederland in zogenoemde netwerksectoren, zoals de energie, telecommunicatie en de post, waar vanwege de marktstructuur de concurrentie beperkt is.

In de jeugdzorg is geen sprake van tariefregulering maar onderhandelen zorgaanbieders met gemeenten over tarieven en de hoeveelheid te leveren zorg. In de wijkverpleging en de extramurale GGZ is er sprake van tariefregulering. De NZa stelt de tarieven in de wijkverpleging en de extramurale GGZ vast op basis van de gemiddelde kosten in de sector. In de tarieven zit een opslag voor een redelijk rendement op eigen vermogen. Voor de extramurale zorg gaat het om een vergoeding voor de kosten van kapitaal van zes à zeven procent van een normatief eigen vermogen (Finance Ideas 2017). Het eigen vermogen in de extramurale zorg bedraagt rond de vijftien procent van de omzet, waardoor een redelijk rendement als percentage van de omzet rond de een procent ligt. De NZa doet regelmatig kostenonderzoek om de tarieven te herijken. De toezichthouder, de NZa, stelt de tarieven niet vast per instelling, maar voor de hele sector. Bedrijven met lagere kosten dan gemiddeld kunnen daardoor hogere winsten maken.

Een mogelijke oorzaak voor hoge winsten van kleine aanbieders kan zijn dat zij lagere kosten hebben dan gemiddeld, waardoor er voor hen een hogere marge in het tarief zit. Dit kan bijvoorbeeld komen doordat zij minder complexe patiënten behandelen (zie ARK 2020; IGJ 2021). Ook kan het zijn dat zij minder vastgoed en overhead hebben, onder andere omdat zij niet met 300 gemeenten contracten hoeven af te sluiten. Daarnaast kan ook fraude een rol spelen. Fraude blijkt vooral voor te komen in de wijkverpleging (Zorgverzekeraars Nederland, 2021). Het gaat veelal om kleine zorgaanbieders die werken zonder contract met een verzekeraar, waardoor minder controle mogelijk is. Pointer (2019) noemt een aantal voorbeelden van zorgbedrijven die hoge winsten realiseren door zorg te declareren die niet geleverd is, of slechte zorg leveren met onvoldoende gekwalificeerd personeel. Ook zijn er voorbeelden van belangenverstrengeling, waarbij een zorgaanbieder opdrachten verleende aan een onderaannemer, waar de zorgaanbieder zelf eigenaar van was.

4 Hoe kan de overheid excessieve winstuitkeringen beperken?

Het ministerie van VWS bereidt momenteel regelgeving voor die is gericht op een transparante, integere en professionele bedrijfsvoering bij zorgaanbieders om excessen rond dividenduitkering te voorkomen. Op verzoek van het ministerie van VWS hebben wij onderzocht of er manieren zijn om ex post winstuitkeringen te maximeren. Alle winst boven de norm mag dan niet worden uitgekeerd. Er zijn in theorie twee opties:

1. Norm voor maximale winstuitkering op basis van rendement op geïnvesteerd vermogen. Doordat de extramurale zorg relatief weinig investering behoeft, hebben veel aanbieders een zeer beperkt eigen vermogen. Het gemiddelde eigen vermogen in de wijkverpleging, de extramurale GGZ en jeugdzorg is respectievelijk 11, 19 en 23 procent van de omzet. Er zit daarom een enorme spreiding in het rendement op het eigen vermogen. Voordeel van een dergelijke norm is dat deze aansluit op de economische definitie van overwinst. Nadeel is dat de verantwoordingsgegevens, die vooral kleine zorgaanbieders moeten aanleveren, onvoldoende zijn om de norm toe te passen. Dit betekent dat de norm niet van toepassing wordt op kleine zorgaanbieders, of dat zij meer verantwoordingsgegevens moeten gaan aanleveren.
2. Norm voor maximale winstuitkering op basis van de winst als percentage van de omzet (nettoresultaat/bedrijfsopbrengsten). Voordeel van deze norm is dat deze goed toepasbaar is, ook op kleine aanbieders. Nadeel van deze norm is dat deze niet aansluit bij de economische definitie van overwinst

Voor de hoogte van de norm kan de minister aansluiten bij de berekeningen van de NZa voor het vaststellen van de tarieven. Nadeel van beide normen is dat deze gebaseerd zijn op gemiddelden en voor een specifiek bedrijf niet hoeven aan te sluiten bij de feitelijke kosten van kapitaal. Risicovolle investeringen, bijvoorbeeld in domotica, worden daardoor onvoldoende beloond. Dit geldt sterker voor een norm als percentage van de omzet, omdat deze niet alleen gebaseerd is op een gemiddeld risico van de investering in de sector, maar ook op een normatief eigen vermogen. Om hiermee rekening te houden kan de maximumnorm

worden gesteld op het gemiddelde in de sector plus een opslag. Als er een norm gesteld wordt, zou het moeten gaan om een norm voor structurele winsten over een aantal jaren (bijvoorbeeld een voortschrijdend vijfjaargemiddelde), zodat de aanbieder verliezen kan salderen met winsten.

Geen van beide geschetste opties is ideaal. Beide gaan gepaard met hogere administratieve lasten voor zorgaanbieders en toezichthouders. En beide opties raken de gehele sector, niet alleen de excessen. Er is geen empirische literatuur over het effect van het beperken van de mogelijkheid winst uit te keren in de extramurale zorg door bv's. Op basis van economische theorie en empirisch onderzoek naar verschillen tussen non-profit en for-profit ziekenhuizen kunnen een aantal voorspellingen worden gedaan over de invloed van de maatregel op de betaalbaarheid, de toegankelijkheid en de kwaliteit van de zorg (zie tabel 1; Baeten et al. 2019; Kerste en Kok 2011). Onderstaand overzicht geeft weer wat de mogelijke consequenties zijn van het maximaliseren van de dividenduitkering. De omvang van de effecten hangt af in hoeverre de norm voor de maximale winstuitkering knellend is. Op voorhand is niet te voorspellen of de positieve of negatieve effecten zullen overheersen.

Tabel 1 Mogelijke effecten winstmaximering

	Positieve effecten	Negatieve effecten
Betaalbaarheid	Aanbieders hebben minder belang bij het leveren van onnodige zorg (supplier induced demand)	De maatregel ontmoedigt kostenreductie. De hogere winst door kostenreductie mag immers niet volledig worden uitgekeerd als dividend. Aanbieders kunnen daardoor zelfs kostenverhogende investeringen doen die niet ten goede komen aan de zorg, bijvoorbeeld voor een duur kantoor (gold plating).
Toegankelijkheid	Zorgaanbieders hebben minder prikkels om alleen makkelijke cliënten te selecteren..	<ul style="list-style-type: none"> • Wachlijsten nemen toe, omdat ondernemers minder belang hebben meer cliënten te helpen. • De zorg wordt minder aantrekkelijk voor ondernemers, waardoor minder partijen toetreden. Dit vermindert het aanbod.
Kwaliteit	Meer middelen om te investeren in kwaliteit, bijvoorbeeld in opleiding van het personeel.	Minder prikkel om de kwaliteit te verhogen. Ondernemers hebben een prikkel om hogere kwaliteit te bieden, omdat ze daarmee een hogere marge kunnen realiseren of meer klanten kunnen trekken. Deze prikkel wordt minder als de winstuitkering gemaximeerd wordt.
Innovatie		Kapitaalverschaffers zullen minder geneigd zijn in de sector te investeren. Dat gaat ten koste van innovatie en potentiële productiviteitswinst

Auteurs

Lucy Kok (e-mail: l.kok@seo.nl) is senior projectleider bij SEO Economisch Onderzoek; Anouk Bekker is consultant bij BDO Advisory; Frank van der Lee is partner bij BDO Advisory; Céline Odding is promovenda aan de Rijksuniversiteit Groningen en Ward Rougoor is senior onderzoeker bij SEO Economisch Onderzoek

Literatuur

- Averch, H. en L.L. Johnson, 1962,. Behavior of the firm under regulatory constraint, *The American Economic Review*, vol. 52(5): 1052-1069.
- ARK, 2020, Geen plek voor grote problemen. Aanpak van wachttijden in de specialistische ggz, Den Haag: Algemene Rekenkamer.
- Baeten, S., Diepstraten, P., Heida, J. P., Hoendervanger, J., Willems, L., & Zanen, L., 2019, Uitkering van dividend door zorgaanbieders, Utrecht: Sirm/Finance Ideas.
- Boeger, N., S. Burgess en J. Ellison, 2018, Lessons from the Community Interest Company, in: N. Boeger en C. Villiers (eds), *Shaping the corporate landscape: towards corporate reform and enterprise diversity*, pp. 347-364, Hart Publishing.
- Finance Ideas, 2017, Van nhc naar integraal, een marktconforme vermogenskostenvergoeding voor zorginstellingen, Utrecht: Finance Ideas.
- IGJ, 2021, Factsheet, Onvoldoende tijdige en juiste hulp voor jongeren met ernstige psychische problemen, Inspectie gezondheidszorg en jeugd.
- Kerste, M en L. Kok, 2010, Winst in de eigendomsstructuur. Eigendom, winstbestemming en zeggenschap binnen ziekenhuizen, Amsterdam: SEO Economisch Onderzoek.
- Kok, L., C. Odding en W. Rougoor, 2020, Normering winstuitkering zorg. Amsterdam: SEO Economisch Onderzoek/BDO Advisory.
- NOS, 2019, <https://nos.nl/artikel/2303257-opnieuw-ongebruikelijk-hoge-winsten-bij-85-zorgbedrijven>. 25-9-2019.
- Pointer, 2019, [Slechte zorg leveren, maar toch enorme winst maken: hoe gaan zorgcowboys te werk? | KRO-NCRV](#).
- Pomp, M., 2019, Maken farmaceutische bedrijven excessieve winsten. *TPedigitaal*, vol. 13(3).
- Suijs, J. en H. Verbon, 2018, De winsten van zorginstellingen, *Beleid & maatschappij*, vol. 45(1): 46-77.
- VWS, 2019, Brief minister VWS aan Tweede Kamer, 25 november 2019 Kamerstuk 34 767, Nr. 19 Zorgverzekeraars Nederland (2021), Fraudebeheersing zorgverzekeraars in 2020.

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

Jos Blank, Alex van Heezik en Bas Blank

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

1 Inleiding

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

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

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

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

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

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

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

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

2 Methode

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

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

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

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

Waarbij:

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

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

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

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

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

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

Waarbij:

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

Waarbij:

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

En:

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

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

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

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

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

3 Data

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

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

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

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

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

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

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

Tabel 1 Statistische beschrijving gegevens kerndepartementen, 2012-2019

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

4 Resultaten

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

Tabel 2 Schattingsresultaten kostenfunctie

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

Tabel 3 Schattingen marginale kosten

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

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

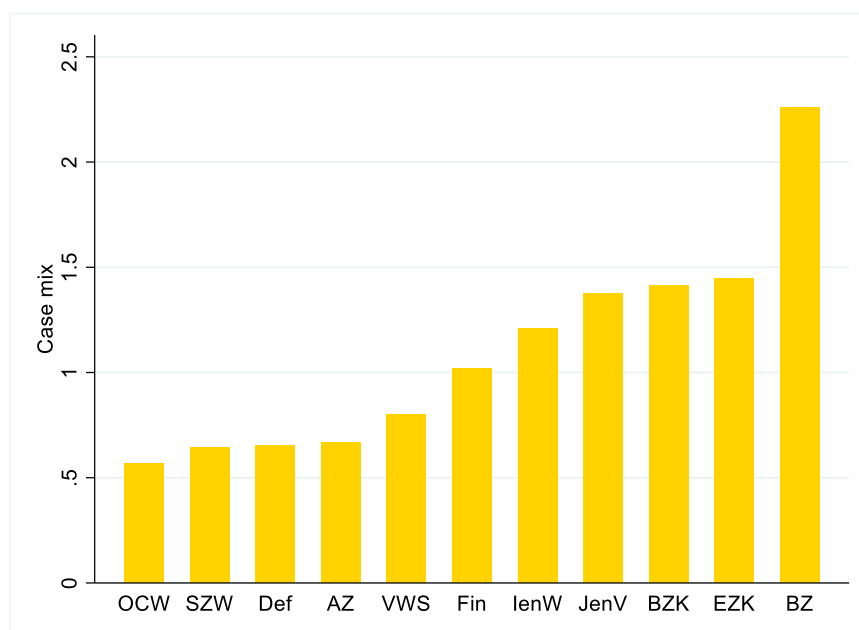
De parameterwaarden van de productie-indicatoren hebben plausibele waarden. Zo blijkt 38 procent van de ingezette middelen gemoeid met het aantal documenten, 17 procent met de behandeling van Kamervragen en 45 procent met het programma. De daarbij geschatte kostprijzen (marginale kosten) bedragen gemiddeld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

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

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

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

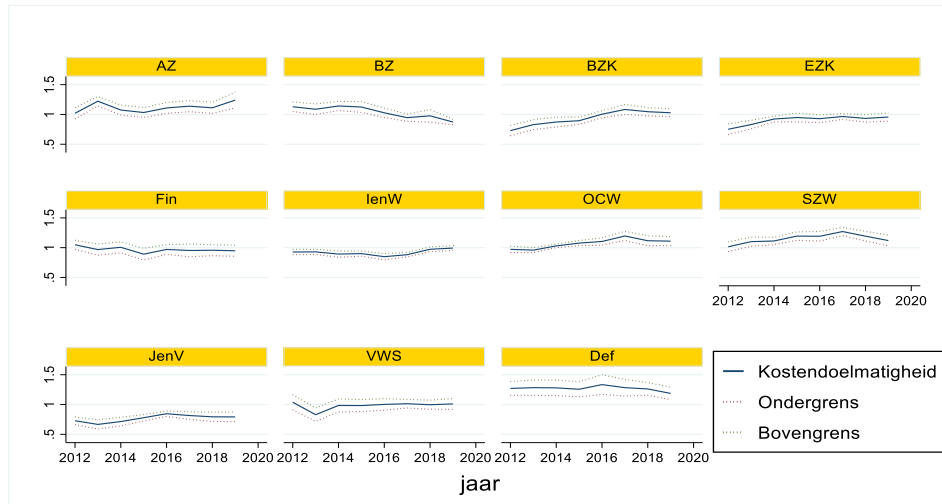
Figuur 1 Case mix per kerndepartement



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

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

Figuur 2 Kostendoelmatigheid kerndepartementen, 2012-2019



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

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

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

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

5 Conclusies en aanbevelingen

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

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

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

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

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

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

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

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

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

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

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

Auteurs

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

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

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

Literatuur

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

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

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

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

De CoronaMelder door een economische lens

Joost Poort

In oktober 2020 lanceerde Nederland de CoronaMelder-app voor geautomatiseerd onderzoek naar blootstelling aan het Covid-19 virus. Een tentatieve analyse van de maatschappelijke kosten en baten van de app suggereert dat de baten in evenwicht zijn met de kosten, zelfs als alleen wordt gekeken naar de bespaarde levensjaren door het voorkomen van sterfgevallen. Het percentage van de positieve tests dat door de app teweeg wordt gebracht, is de sleutel tot de positieve bijdrage ervan aan de maatschappelijke welvaart. Dit percentage is aanzienlijk lager voor de app dan voor handmatig contactonderzoek door de GGD, wat erop wijst dat de app efficiënt kan zijn als aanvulling op handmatig contactonderzoek, maar niet als vervanging daarvan.

1 Inleiding

De Covid-19 pandemie leidde tot tal van maatregelen in de hele wereld, die ingrijpende gevolgen hadden voor het dagelijks leven, de burgerlijke vrijheden en de economie. In Nederland werden scholen gesloten, evenals restaurants, bioscopen, theaters, bibliotheken, sportfaciliteiten en de meeste winkels. Een avondklok en een mondkapjesplicht werden ingevoerd. Evenementen werden afgelast. Reizen naar het buitenland was nog amper mogelijk. Wat nog wel kon en mocht, moest op anderhalve meter afstand van iedereen van buiten je huishouden. Al deze maatregelen hebben gemeen dat zij de verspreiding van het virus en de gevolgen daarvan voor de volksgezondheid en de gezondheidszorg willen voorkomen door het aantal contacten waardoor het virus van de ene gastheer op de andere kan overspringen, *generiek* te beperken.

Naast deze generieke maatregelen doet de GGD bron- en contactonderzoek (hierna: contactonderzoek) om personen op te sporen die in de buurt zijn geweest van iemand met corona en dus een *specifiek* risico lopen om te worden besmet. In aanvulling daarop lanceerde Nederland in oktober 2020 een applicatie (hierna: app) voor digitaal contactonderzoek: de CoronaMelder. Het doel van zowel handmatig contactonderzoek als geautomatiseerd onderzoek via de app is om besmette personen in een vroeg stadium op te sporen en te voorkomen dat zij andere mensen besmetten. Potentieel leidt dit tot een veel gerichtere vermindering van het aantal contacten, namelijk contacten die een grotere kans geven op nieuwe infecties.

In economische termen waren de effecten van de generieke maatregelen immens. In prepandemisch december 2019 voorspelde het CPB (2019) dat de Nederlandse economie in 2020 reëel met 1,3 procent zou groeien, met als grootste bedreigingen voor de groei vraagstukken rond stikstof- en PFAS-normen, Brexit en het Amerikaanse handelsbeleid. Een half jaar later voorspelde het een ongekende daling van 6 procent (Centraal Planbureau 2020). De meest recente prognose uit september 2021 is alweer optimistischer. De krimp kwam in

2020 uit op 3,8 procent terwijl voor 2021 een groei van 3,9 procent werd voorzien (Centraal Planbureau 2021). Ondertussen rezen de overheidsuitgaven de pan uit. Miljarden werden uitgegeven aan medische zorg en aan steun voor de sectoren in de economie die noodgedwongen hun deuren moesten sluiten, waardoor de staatsschuld snel toenam.

En dit zijn nog maar de financiële en kortetermijneffecten. Het effect van uitgestelde medische zorg tijdens de eerste Covid-golf in 2020 in de twaalf specialismen met de hoogste uitgaven in 2018, wordt geschat op 34 tot 50 duizend verloren gezonde levensjaren (van Gieszen et al. 2020, p. 3). Dit is exclusief vele effecten, zoals het negatieve effect van het tijdelijk stopzetten van bevolkingsonderzoek op de vroege diagnose van kanker. Ook hebben economen gewaarschuwd voor langetermijneffecten, bijvoorbeeld van het tijdelijk sluiten van scholen (Teulings 2021), en gesteld dat de huidige interventies meer kosten dan baten voor de samenleving veroorzaken (bijvoorbeeld Baarsma et al. 2021).

Aan de andere kant blijken de eerste kosten-batenanalyses van beperkende maatregelen in de Verenigde Staten positief uit te pakken. Doti (2021) schat de reductie van het aantal doden in de VS op 358 duizend in 2020. Bij een voor leeftijd gecorrigeerde 'waarde van statistisch leven' van 4,2 miljoen dollar weegt dit op tegen de geschatte kosten van verloren baten en negatieve inkomenseffecten met een factor 3,7. Broughel en Kotrous (2021) komen tot soortgelijke positieve nettoresultaten van beperkende maatregelen tijdens de eerste golf van de pandemie. Eerste aanzetten tot kosten-batenanalyses van de beperkende maatregelen in Nederland (Frijters 2021; Kolen 2021) geven sterk uiteenlopende beelden en zijn vatbaar voor fundamentele kritiek (Koopmans 2021). Hoe dat ook zij, geen van deze analyses gaat in op de specifieke kosten en baten van apps voor contactonderzoek. Tegen deze achtergrond bekijkt dit artikel de CoronaMelder door een economische lens en geeft het een eerste ruwe inschatting van de maatschappelijke kosten en baten van de app.

2 De CoronaMelder

Een van de eerste voorstellen vanuit de wetenschap voor het gebruik van apps bij de bestrijding van Covid-19 kwam van Ferretti et al. (2020). Op basis van modellen van de verspreiding van de epidemie en de vaststelling dat 46 procent van de nieuwe besmettingen afkomstig was van pre-symptomatische personen, concludeerden de auteurs dat handmatig contactonderzoek niet zou volstaan de epidemie te beteugelen. Een app die de contacten van een positief bevonden persoon onmiddellijk op de hoogte zou stellen, zou meer restrictieve maatregelen kunnen helpen voorkomen, aldus de onderzoekers. Dit idee won snel aan populariteit. Singapore, China, Zuid-Korea en Israël waren de eerste landen die dergelijke apps introduceerden, in Europa gevolgd door Duitsland en Zwitserland en vele landen daarna (Rehse en Tremöhlen 2020, p. 36-37).

In Nederland is de CoronaMelder-app officieel gelanceerd op 10 oktober 2020. De doelstellingen waren (1) de GGD te helpen bij het opsporen van recente contacten van iemand die positief getest is die mogelijk tot nieuwe infecties hebben geleid, en (2) de verspreiding van

het virus te vertragen door er bij deze contacten op aan te dringen zich te laten testen en in quarantaine te gaan totdat duidelijk is of ze besmet zijn, om te voorkomen dat ze anderen besmetten (Ebbers et al. 2021, p.3). Gebruik van de app is vrijwillig en ook na een positieve test moeten gebruikers nog toestemming geven om hun contacten via de app te waarschuwen.¹ De app registreert een contact wanneer een andere app-gebruiker naar schatting ten minste vijftien minuten binnen een straal van anderhalve meter is geweest.

Op 28 mei 2021 is een evaluatie van CoronaMelder gepubliceerd, met daarin veel feitelijke informatie (Ebbers et al. 2021). Per 23 mei 2021 hadden 4,9 miljoen mensen CoronaMelder gedownload, maar naar schatting 2,9 miljoen mensen (ongeveer 17% van de bevolking) maken daadwerkelijk *gebruik van* de app. Als app-gebruikers gemiddeld evenveel contacten hebben als niet-gebruikers, wordt dus slechts $17\% \times 17\% = 3\%$ van de alle contacten geregistreerd. Het percentage app-gebruikers is redelijk stabiel in alle leeftijdsgroepen, maar neemt toe met het opleidingsniveau. Via CoronaMelder hadden 174 duizend besmette personen hun contacten op de hoogte gebracht van een besmetting, oftewel zo'n 6 procent van de gebruikers. Vervolgens hebben 189 duizend mensen zich aangemeld voor een test na een melding in de app; 77% van hen was niet of nog niet bereikt via handmatige contactopsporing en 14 duizend mensen zijn positief getest bij een test na een melding.

Ebbers et al. stellen vast dat ruim de helft van de mensen die na een melding in de app een test aanvroegen, *nooit* is benaderd door de GGD. Zonder de app zou deze groep niet, of pas nadat zij symptomen ontwikkelden, zijn gesignaleerd. Een andere groep werd door de app op de hoogte gebracht *voordat* zij door de GGD waren bereikt, waardoor zij verdere contacten konden vermijden en zich eerder konden aanmelden voor een test. Ongeveer 3 tot 5 procent van de mensen die na een melding in de app een test aanvroegen maar geen symptomen hadden, testte positief, tegenover ongeveer 1 procent in de algehele bevolking. Al met al is in totaal ongeveer 1 op de 10 testaanvragen en 1 op de 20 positieve tests door de CoronaMelder teweeggebracht. Tussen 26 september 2020 en 18 april 2021 komt dit neer op zo'n 11 duizend positieve testen als gevolg van de app. Uit modellering van het RIVM volgt dat de app tussen december 2020 en maart 2021 ruim 15 duizend besmettingen en ruim 200 ziekenhuisopnames heeft voorkomen. Tegelijk leidde de app tot ongeveer 128 duizend negatieve testen.

Van de app-gebruikers zegt 97 procent bereid te zijn thuis te blijven na een melding en zegt 95 procent bereid te zijn een test te doen. In de praktijk liggen deze cijfers ruim de helft lager: 45% van de mensen die een melding ontvingen, bleef daadwerkelijk thuis, terwijl 41% zich liet testen. Over het geheel genomen concluderen de auteurs dat de app een klein maar merkbaar effect had. Zij stellen dat het begrijpelijk is dat het effect klein is, gezien alle sociale beperkingen die ook golden in de onderzoeksperiode. Naarmate er minder beperkingen zijn, kan de waarde van CoronaMelder daarom toenemen. Die kan ook verder toene-

¹ Ongeveer 75 procent van de positief geteste app-gebruikers deelde inderdaad hun sleutel. Dit is waarschijnlijk een onderschatting van de bereidheid daartoe, aangezien GGD-medewerkers niet altijd vragen naar het gebruik van Coronamelder (Ebbers et al. 2021, p. 19).

men als meer mensen de app gebruiken, als de tijd tussen contacten en notificaties kan worden verkort en als de naleving van de instructies bij notificaties wordt verbeterd. De evaluatie besteedt ook aandacht aan onbedoelde effecten van de app. Zo zijn er aanwijzingen voor het feit dat sommige mensen maatschappelijke druk voelen om de app te installeren en te gebruiken.² Ook noemen Ebbers et al. de kwestie van valse positieven, bijvoorbeeld wanneer signalen door muren gaan.

3 Prikkel voor het gebruik van CoronaMelder

Naleving van de meeste maatregelen die in de inleiding zijn genoemd, komt zowel de individuele als de volksgezondheid ten goede, mits zij inderdaad effectief zijn om de verspreiding van het virus te voorkomen. Door bijvoorbeeld *social distancing* vermindert iemand zowel het risico om het virus op te lopen als dat om anderen te besmetten. Vaccinatie heeft ook een wederzijds effect: zowel de persoon die de vaccinatie krijgt als haar toekomstige contacten hebben er baat bij.³

Digitale apps voor het traceren van contacten zijn in dat opzicht fundamenteel anders. De app voorkomt niet dat de app-gebruiker wordt besmet. Ervan uitgaande dat een vroege diagnose voor de behandeling en het herstel van een patiënt geen verschil maakt, zijn er dus *geen persoonlijke gezondheidsbaten voor een app-gebruiker*.

Wat de app doet, is de kans op vroegtijdige diagnose vergroten en daardoor iemand in staat stellen actie te ondernemen om het virus niet verder te verspreiden. Aldus is het installeren van de app iets wat men doet voor de gezondheid van anderen. De potentiële gezondheidsbaten komen ten goede aan de contacten die een persoon zou hebben gehad en via hen aan anderen, terwijl de kosten ten laste komen van de persoon die de app installeert. Dit is zeer vergelijkbaar met het klassieke publieke-goederenprobleem, waarbij mensen niet bereid zijn te investeren in collectieve goederen en de voorkeur geven aan *free-riding*.

Natuurlijk kunnen de potentieel positieve gezondheidseffecten voor de contacten en voor de samenleving in haar geheel nog steeds een geldige stimulans zijn voor iemand om de app te installeren. In economische termen: een persoon kan nut ontleen aan het bijdragen aan de gezondheid van zijn familieleden en anderen en aan 'een goede burger zijn'.⁴ Daar komt nog bij dat iemand die dankzij de app zijn gezinsleden niet besmet, mogelijk voorkomt hen

² Dit is in lijn met onderzoeksresultaten van Helberger et al. (2021, hoofdstuk 9) die vinden dat veel mensen, met name jongeren, zich moreel verplicht voelen om de app te installeren.

³ De omvang van deze effecten in beide richtingen kan verschillen, afhankelijk van de leeftijd en de algemene kwetsbaarheid van een persoon. Jonge, gezonde mensen hebben relatief weinig te vrezen van besmetting in vergelijking met ouderen, terwijl hun nutsverlies door *social distancing* waarschijnlijk groter is. Dit impliceert dat de netto individuele prikkel om algemene maatregelen na te leven en de prikkel om zich te vaccineren aanzienlijk kleiner is voor jongere leeftijdsgroepen in goede gezondheid: vandaar de clandestiene 'coronafeestjes'.

⁴ In de gedragseconomische literatuur wordt dit de 'warme gloed' van het geven genoemd (bedacht door Andreoni, 1989).

te moeten verzorgen of huishoudelijke taken over te moeten nemen. Een ander indirect effect is dat men door anderen niet te besmetten bijdraagt aan de heropening van de samenleving, wat voor de meeste mensen ook privé-voordelen oplevert.

Uit een consumentenonderzoek blijkt dat 47 procent van de CoronaMelder-gebruikers het eens is met de stelling dat men door het gebruik van de app een goede burger is, tegenover 14 procent van de mensen die gestopt zijn met het gebruik van de app en 9 procent van de niet-gebruikers (Ebbers et al. 2021; p. 53). Een vergelijkbaar patroon wordt gevonden voor de stelling dat het gebruik van de app de economie helpt. Het instemmingsniveau met de stelling dat de app helpt bij het beschermen van mensen met een kwetsbare gezondheid is met 78 procent substantieel hoger onder gebruikers van de app. Interessant is dat 34 procent van de niet-gebruikers en 41 procent van de voormalige gebruikers het ook eens zijn met deze laatste stelling. Dit roept de vraag op of deze groepen er bewust voor kiezen om *niet* bij te dragen aan die bescherming, de app niet gebruiken omdat ze (menen) zo weinig contacten te hebben dat het geen zin heeft, of niet begrijpen dat een dergelijke bescherming wordt geboden als *anderen* dan deze kwetsbare groepen de app installeren. Er zijn sterke netwerkeffecten verbonden aan het installeren van de app: naar mate meer mensen de app installeren, neemt de kans dat contacten van een besmet persoon inderdaad worden opgespoord ongeveer kwadratisch toe.

Terwijl deze potentiële voordelen van de app voornamelijk indirect zijn, komen de kosten voor rekening van de gebruiker. Wat zijn de kosten voor de persoon die de app installeert? In geld uitgedrukt is de app gratis. Het installeren kost ongeveer een minuut voor wie goed met zijn smartphone overweg kan en een paar MB, het gebruik daarna wat energie en batterijduur. Dit zijn over het algemeen zeer kleine belemmeringen, maar bij gebrek aan private gezondheidsbaten kunnen ze toch bijdragen aan de relatief lage adoptiegraad van de app. Bovendien lijken deze 'kosten' door niet-gebruikers te worden overschat, al kan het zo zijn dan niet-gebruikers gemiddeld minder digitaal vaardig zijn: 22 procent van hen denkt dat het veel tijd en energie zou kosten om de app te installeren (tegenover 3 procent van de gebruikers) en 49 procent van de niet-gebruikers verwacht dat de app gebruiksvriendelijk is (tegenover 99 procent van de gebruikers). Gebruikers van de app zien vaak persoonlijke voordelen (67 procent) en zelden nadelen (7 procent), terwijl niet-gebruikers zelden persoonlijke voordelen zien (9 procent) en vaker nadelen (24 procent) (Ebbers et al. 2021, p. 16-17).

Zorgen over privacy vormen een meer diepgaande categorie van private kosten van de app. Uit de enquêteresultaten blijkt dat dergelijke zorgen sterk correleren met app-gebruik, terwijl er wijdverspreide misvattingen zijn over de privacyaspecten van de app zelf. Van de app-gebruikers gelooft 85 procent dat persoonlijke informatie strikt vertrouwelijk wordt gehouden, terwijl slechts 55 procent van de niet-gebruikers dat denkt. Niettemin denkt 57 procent van de app-gebruikers dat de apps de locatie van de gebruiker vastlegt; binnen de groep niet-gebruikers is dit 68 procent. Opvallend is dat deze misvatting vaker voorkomt bij hoger opgeleide groepen. In dezelfde lijn denkt 35 procent van de app-gebruikers en 55

procent van de niet-gebruikers ten onrechte dat de app hun naam en persoonlijke gegevens vastlegt.

Het bestrijden van deze misvattingen door betere informatievoorziening kan helpen de adoptiegraad van de app te verbeteren. Tot op zekere hoogte zal hier echter zelfrechtvaardiging een rol spelen: mensen die de app om wat voor reden dan ook niet hebben geïnstalleerd, sussen hun geweten door te stellen dat het erg ingewikkeld en tijdrovend zou zijn om dat wel te doen. In dergelijke gevallen zal het verstrekken van betere informatie weinig soelaas bieden. Overigens worden ook in andere landen, zoals Duitsland en Zwitserland, adoptiecijfers waargenomen die vergelijkbaar zijn met die in Nederland (Rehse en Tremöhlen 2020, p. 2, 38).

De asymmetrie tussen de kosten en baten van app-gebruik blijft dus een fundamenteel obstakel voor grootschalige vrijwillige toepassing. Het is het klassieke probleem van het publieke goed. Daarom blijft het essentieel om de werkelijke en de vermeende private kosten van het gebruik van de app te verminderen. Met name op het vlak van privacy kunnen misvattingen over de werking van de app worden verminderd door betere informatie te verstrekken, wat waarschijnlijk zal helpen om de adoptiegraad van de app te verbeteren. Vanuit economisch perspectief kan zelfs het subsidiëren van app-gebruikers om hen te compenseren voor de positieve externe effecten van het installeren en gebruiken van de app gerechtvaardigd zijn, hoewel ervoor moet worden gewaakt dat de intrinsieke motivatie op deze manier wordt geschaad (zie ook: Rehse en Tremöhlen 2020, p. 24-27). Een manier om dit te doen zou een loterij onder actieve app-gebruikers kunnen zijn, zoals de staat Ohio deed onder mensen die gevaccineerd waren, wat leidde tot een aanzienlijke stijging van de vaccinatiegraad.⁵ Een meer radicale manier om de adoptie te verhogen is integratie van de app met apps voor vaccinatie- en testbewijzen (in Nederland de CoronaCheck-app), zoals in Singapore is gebeurd.

4 Naar een maatschappelijke kosten-batenanalyse

Hoewel een volledige maatschappelijke kosten-batenanalyse niet binnen het bestek van dit artikel valt, worden in deze paragraaf een paar rudimentaire stappen gezet om de directe en indirecte maatschappelijke kosten en baten van de CoronaMelder in kaart te brengen. Vertrekpunt voor een dergelijke analyse is het formuleren van het nul-alternatief: wat is het alternatieve beleid ten opzichte waarvan de kosten en baten worden beoordeeld? Zoals hierboven vermeld, hebben Ferretti et al. (2020) digitale apps voor het traceren van contacten voorgesteld als alternatief voor meer restrictieve maatregelen. Een dergelijk nul-alternatief zou echter niet alleen een analyse van de effecten van de app vereisen, maar ook van de andere maatregelen waar het hier om gaat. Bovendien werd in de inleiding literatuur aangehaald waarin wordt gesteld dat voor verschillende van deze generieke maatregelen de kosten hoger zijn dan de baten. Als dat zo is, zouden dergelijke nul-alternatieven het

⁵ Zie: <https://odh.ohio.gov/wps/portal/gov/odh/media-center/odh-news-releases/odh-news-release-05-20-21>.

beeld voor de CoronaMelder flatteren. Anderzijds kwamen de twee genoemde kosten-batenanalyses voor de VS op een zeer gunstige kosten-batenverhouding voor het generieke pakket van beperkende maatregelen aldaar. Dat zou de lat voor de CoronaMelder juist weer heel hoog leggen. Daarom verdient het de voorkeur bij wijze van nul-alternatief ofwel te kijken naar intensiever handmatig contactonderzoek door de GGD, ofwel naar 'niets doen'.

Kosten en baten ten opzichte van niets doen. Met 'niets doen' als nul-alternatief, staan aan de kostenkant van de balans de kosten van het ontwikkelen en onderhouden van de app en van de publiciteitscampagne die is gehouden om de app te promoten. Deze kosten zijn over 2020 en 2021 alles bij elkaar ongeveer 23 miljoen euro (Tweede Kamer 2021; p. 30). In een volledige kosten-batenanalyse moeten deze kostencijfers nader worden onderzocht. Zo kunnen er bijvoorbeeld extra kosten van de app zijn aan de kant van de GGD. Anderzijds leiden sommige van deze kosten rechtstreeks tot baten voor andere economische actoren, met name de kosten van reclame die grotendeels bestaan uit schaarste-opbrengsten. Merk verder op dat de kosten betrekking hebben op heel 2020 en 2021, terwijl de baten hieronder worden berekend over de periode tot 18 april 2021.

Bijkomende kosten zijn de onbedoelde 'zachte' effecten zoals genoemd door Ebbers et al. (2021) – de publieke druk die sommige mensen voelen en de zorgen over de privacy – evenals het energie- en dataverbruik van het installeren en gebruiken van de app. Ook leiden vals-positieve meldingen tot aanzienlijke tijd- en kostenverspilling aan testen en zelfisolatie.

De meeste van deze effecten zijn moeilijk te kwantificeren, laat staan in euro's uit te drukken, zonder uitgebreid onderzoek. Wel kan een poging worden gedaan om een voorzichtige schatting te maken van de maatschappelijke kosten van vals-positieve meldingen, ervan uitgaande dat alle negatieve testen die uitsluitend door de app worden uitgelokt, zonder de app niet zouden hebben plaatsgevonden. Het gaat dan om 128 duizend testen tussen 26 september 2020 en 18 april 2021.

Wat zijn dan de totale maatschappelijke kosten van een negatieve test? Ten eerste zijn er de werkelijke kosten van de test zelf, inclusief de analyse ervan. Aangezien geen gedetailleerde informatie over de kosten van officiële GGD-tests (en handmatige contractopsporing) beschikbaar is, moeten ruwe schattingen worden gemaakt. Momenteel wordt in Nederland geadverteerd met commerciële PCR-tests voor prijzen vanaf ongeveer 75 euro. Uitgaande van een winstmarge van 15% geeft dit een kostenraming van 65 euro per PCR-test. GGD-teststraten hebben mogelijk wat lagere kosten door schaalvoordelen.

Daarbij komen de *tijdskosten* van het ondergaan van de test en het verlies van productiviteit of consumentensurplus door zelfisolatie, totdat de negatieve uitslag beschikbaar is. Voor deze schatting wordt hier uitgegaan van een gemiddeld verlies van een werkdag op deze manier, gewaardeerd tegen het gemiddelde brutoloon in Nederland, dat in 2020 24 euro

bedroeg.⁶ De totale welvaartskosten van een door CoronaMelder ingegeven negatieve test zullen dus in de orde zijn van $8 \times 24 + 65 = 257$ euro. De totale kosten van 128 duizend van dergelijke negatieve testen tussen 26 september 2020 en 18 april 2021 bedragen 33 miljoen euro. De totale kosten van de app en het testen als gevolg van valse waarschuwingen bedragen dan samen ongeveer 56 miljoen euro. Inclusief de kosten voor het testen van positieve gevallen zijn deze ongeveer 59 miljoen euro.

Aan de batenkant staan de welvaartswinsten van het voorkomen van infecties. Eerder werd de schatting van het RIVM genoemd dat de app tussen december 2020 en maart 2021 ruim 15 duizend infecties en ruim 200 ziekenhuisopnames heeft voorkomen. Extrapolatie hiervan naar september-april om kosten en baten over dezelfde tijdspanne te vergelijken, zou leiden tot ruim 22 duizend vermeden infecties en ongeveer 300 vermeden ziekenhuisopnames.⁷

Een alternatieve manier om tot een schatting van het aantal vermeden besmettingen te komen is via de 11 duizend positieve testen die de app heeft opgeleverd. Veronderstel een reproductiegetal R_0 dat in de tijd constant is op 0,9 (feitelijk is het tussen september 2020 en april 2021 gedurende aanzienlijke perioden groter geweest dan 1). En stel dat een besmette persoon die door de app wordt gewaarschuwd erin slaagt dit tot 0,45 te halveren door zijn contacten te reduceren⁸, waarna de besmettingsketen doorgaat met de 'gewone' reproductiesnelheid van 0,9. Dan voorkomt elke positieve test die door de app wordt geactiveerd $0,45 \times (1 + 0,9 + 0,9^2 + \dots) = 4,5$ volgende besmettingen. Bij een R_0 van 0,8 is dat opgeteld 2. Zo voorkomen 11 duizend positieve testen die door de app worden getriggerd dus een keten van ongeveer 22 tot 50 duizend volgende infecties. Volgens de door het RIVM gehanteerde verhouding tussen besmettingen en ziekenhuisopnamen zou het aantal voorkomen ziekenhuisopnamen tussen de 300 en 680 liggen.

Het sterftecijfer van Covid-infecties in een land hangt sterk af van de samenstelling van een bevolking (leeftijd, de prevalentie van obesitas en diabetes). In een meta-analyse van verschillende studies schatten Brazeau et al. (2020) het sterftecijfer in hoge-inkomenslanden (met een grotere concentratie van risicogroepen) op 1,15 procent (0,78 tot 1,79 procent). Dit ligt vrij dicht bij de ratio van 1,06 procent op basis van de officiële statistieken voor Nederland per 8 juni 2021⁹, en ligt ook binnen de bandbreedte van 0,5 tot 1,4 procent voor de eerste golf in Europese landen, genoemd in een meer recente CBS-publicatie (Stoeldraijer

⁶ Zie: <https://www.cbs.nl/nl-nl/visualisaties/dashboard-arbeidsmarkt/ontwikkeling-cao-lonen/uurloon>. Ook dit is een benadering. Enerzijds is de gemiddelde productiviteit van werknemers hoger dan deze brutolonen, dus als een volledige dag werk wordt verzuimd, zullen de kosten groter zijn. Anderzijds zullen de kosten bij mensen die met pensioen zijn, naar school gaan of om andere redenen niet werken, kleiner zijn.

⁷ De schatting van het RIVM was gebaseerd op 7,5 duizend positieve testen die door de app werden getriggerd, terwijl er in de periode september-april 11 duizend van dergelijke testen waren, dus 47% meer.

⁸ Eerder werd genoemd dat 45% van de mensen die een melding ontvingen thuisbleven en 41% zich liet testen. Het is plausibel dat die groepen sterk overlappen en dat degenen die dankzij een melding van de app positief bevonden zijn (en dus tot die 41% behoorden), zich in grote meerderheid aan het voorschrift hielden thuis te blijven en minder contacten hadden.

⁹ 17.695 sterfgevallen en 1,67 miljoen infecties; beide hebben waarschijnlijk te lijden van onderrapportage.

et al. 2021). Omdat in de tweede helft van het tijdvak september 2020 - april 2021 het vaccinatieprogramma op stoom begon te komen waardoor kwetsbare groepen vanaf januari werden beschermd, wordt in onderstaande schattingen een sterfte van 0,5 procent – aan de onderkant van deze schattingen – gehanteerd. Dit impliceert dat tussen de 110 en 250 sterfgevallen zijn voorkomen door CoronaMelder.

Hoe verhouden de gezondheidsvoordelen van de CoronaMelder zich tot de geraamde kosten van de app en de maatschappelijke kostenraming van het testen? Aangezien de nadruk hier ligt op gezondheidsvoordelen als gevolg van besmettingen die via de app worden opgespoord, moeten ook de kosten van het testen van deze gevallen worden meegerekend, en wordt een totale kostenraming van 59 miljoen euro gebruikt. Als alleen naar de voorkomen infecties wordt gekeken, liggen deze totale kosten tussen 1180 en 2670 euro per infectie; als alleen naar de voorkomen sterfgevallen wordt gekeken, liggen ze op 235 à 534 duizend euro per voorkomen sterfgeval. Bij een waarde van 80 duizend euro per verloren gezond levensjaar (Zorginstituut Nederland, 2018) zouden deze kosten per voorkomen sterfgeval gelijk zijn aan gemiddeld 2,9 tot 6,7 verloren gezonde levensjaren. Bij de voor maatschappelijke kosten-batenanalyses voorgeschreven bandbreedte van 50 tot 100 duizend euro per verloren levensjaar (Tweede Kamer 2016 en bijlagen) corresponderen de kosten met die van 2,4 tot 10,7 verloren jaren. Op het eerste gezicht lijken deze getallen niet onrealistisch. Frijters (2020) gaat uit van gemiddeld 2,9 verloren gezonde levensjaren, net boven de onderkant van deze bandbreedte, en Broughel en Kotrous (2021) komen voor de VS op een waarde per gered leven van rond de 1,0 miljoen dollar.

Deze getallen impliceren dat de maatschappelijke baten van de CoronaMelder – puur op basis van vermeden sterfgevallen – ongeveer opwegen tegen de kosten. Een meer verfijnde berekening van de verwachte geredde levensjaren door voorkomen sterfgevallen zou dit nauwkeuriger kunnen maken. In zo'n berekening moeten ook de maatschappelijke baten worden meegenomen van het voorkomen van symptomatische kortdurende Covid-gevallen (effecten in de orde van enkele dagen ziekteverzuim en tijdelijk verlies van kwaliteit van leven voor de patiënt en zijn naasten); het voorkomen van *long Covid* met ernstig verlies van levenskwaliteit en productiviteit gedurende verscheidene maanden; en de vermeden kosten van ziekenhuisopname.

Kosten en baten in vergelijking met handmatig contactonderzoek. Tot besluit wordt een vergelijking gemaakt van kosten en baten ten opzichte van die van handmatige contactonderzoek door de GGD. In Ebbers et al. (2021, p. 22) worden 1,4 miljoen testaanvragen toegeschreven aan handmatig contactonderzoek. In vergelijking met de CoronaMelder is een groter deel van 18 procent ofwel 226 duizend van deze testen positief uitgevallen. Uitgaande van dezelfde schatting voor de maatschappelijke kosten per test als voor de app (257 euro), vertaalt zich dit naar een maatschappelijke kostenpost van 354 miljoen euro aan testen die door handmatig contactonderzoek zijn veroorzaakt.

Om het beeld compleet te maken is een schatting nodig van de kosten van handmatig contactonderzoek. Nauwkeurige cijfers daarover zijn vooralsnog niet beschikbaar, dus wordt

een vrij grove inschatting gemaakt. GGD-directeur Sjaak de Gouw schatte in een kranteninterview de tijd die nodig is voor een goed onderzoek op ongeveer 8 uur, maar door de grote aantallen is het vaak niet mogelijk zoveel tijd te investeren.¹⁰ Uitgaand van gemiddeld 2 tot 4 uur contactonderzoek per besmetting tegen kosten per uur die gelijk zijn aan het eerder genoemde landelijk gemiddelde bruto uurloon van 24 euro, zouden de kosten van handmatig contactonderzoek voor de opvolging van alle 1,1 miljoen besmettingen in de periode september 2020 - april 2021 in de orde van 54 tot 108 miljoen euro liggen. Dit zou de totale kosten van het handmatige contactonderzoek en de daarbij behorende tests op ongeveer 410 tot 464 miljoen euro brengen. Per positieve test die handmatig contactonderzoek te weegbrengt, komt dit overeen met 1800 tot 2050 euro. Dit cijfer is gebaseerd op een ruwe schatting van het aantal uren dat per besmetting in contactonderzoek wordt geïnvesteerd en gaat ervan uit dat deze contacten zich niet zouden aanmelden voor een test zonder het contactonderzoek.

Voor de CoronaMelder-app is deze verhouding van de kosten per positieve test ongeveer 5400 euro (11 duizend positieve tests die door de app worden uitgelokt, tegen totale maatschappelijke kosten van ongeveer 59 miljoen euro). Ondanks het feit dat deze getallen noodzakelijkerwijs gebaseerd zijn op vele aannames, waarvan sommige nogal grof zijn, suggereert dit dat de maatschappelijke kosten van het opsporen van besmettingen via de CoronaMelder-app hoger zijn dan via handmatig contactonderzoek. Dit hangt niet zozeer af van de kosten van de ontwikkeling van de app, maar vooral van het lagere percentage positieve tests dat door de app wordt uitgelokt: 10,4 procent voor de app, tegenover 18,1 procent voor handmatig contactonderzoek. Als gevolg daarvan drukken de maatschappelijke kosten van negatieve tests veel zwaarder op de maatschappelijke efficiëntie van de app.

5 Conclusies en discussie

In dit artikel is een economisch perspectief ontwikkeld op CononaMelder, de Nederlandse app voor contactonderzoek die op 10 oktober 2020 is gelanceerd. Doel van de app is de GGD te helpen bij het opsporen van recente contacten van iemand die positief is getest op Covid-19, die mogelijk tot nieuwe infecties hebben geleid, en de verspreiding van het virus te vertragen door er bij deze contacten op aan te dringen zich te laten testen en in quarantaine te gaan totdat duidelijk is of zij zijn besmet.

De app biedt geen directe gezondheidsvoordelen voor een gebruiker ervan. De app vergroot de kans op vroegtijdige ontdekking van besmetting en stelt iemand daardoor in staat actie te ondernemen om het virus niet verder te verspreiden. De potentiële gezondheidsvoordelen komen dus ten goede aan de contacten die een persoon zou hebben gehad (en via hen aan anderen), terwijl de kosten ten laste komen van de persoon die de app installeert. Deze asymmetrie blijft een fundamenteel obstakel voor grootschalig vrijwillig gebruik. Daarom

¹⁰<https://www.parool.nl/nederland/ggd-baas-over-contactonderzoek-liever-zelf-bellen-omdat-er-schaamte-is~b590d7cb/>.

blijft het van groot belang om de werkelijke en de vermeende persoonlijke kosten van het gebruik van de app te verminderen als men het gebruik wil vergroten. Met name in verband met privacy-kwesties zouden misvattingen over de werking van de app kunnen worden verminderd door betere informatie te verstrekken. Vanuit economisch oogpunt kan het zelfs gerechtvaardigd zijn app-gebruikers te subsidiëren om hen te compenseren voor de positieve externe effecten van het installeren en gebruiken van de app. Een manier om dit te doen zou een loterij onder actieve app-gebruikers kunnen zijn, zoals de staat Ohio heeft gedaan onder mensen die werden gevaccineerd, wat er heeft geleid tot een aanzienlijke stijging van de vaccinatiegraad.

Een tentatieve analyse van de maatschappelijke kosten en baten van de CoronaMelder suggereert dat de baten in evenwicht zijn met de kosten, zelfs als alleen wordt gekeken naar de bespaarde levensjaren door het voorkomen van sterfgevallen. In een meer verfijnde berekening zouden extra maatschappelijke baten voortvloeien uit het voorkomen van symptomatische kortdurende Covid-gevallen met effecten in de orde van enkele dagen ziekteverzuim en tijdelijk verlies van kwaliteit van leven voor de patiënt en zijn naasten; het voorkomen van *long Covid*-gevallen met ernstige verliezen in levenskwaliteit en productiviteit gedurende meerdere maanden; en de vermeden kosten van ziekenhuisopname.

Het percentage tests dat door de app wordt geactiveerd en positief uitvalt is de sleutel tot de positieve bijdrage ervan aan de maatschappelijke welvaart. Dit percentage is aanzienlijk lager voor de app dan voor handmatig contactonderzoek door de GGD, wat erop wijst dat de app efficiënt kan zijn als aanvulling op handmatig contactonderzoek, maar niet als vervanging daarvan. Dit wordt onderstreept door de relatief lage adoptiegraad van de app.

Sinds de onderzoeksperiode die ten grondslag ligt aan dit artikel (26 september 2020 en 18 april 2021) is de vaccinatiegraad opgelopen tot 86,3 procent van iedereen van 18 jaar en ouder. Bovendien heeft 60,6 procent van de volwassenen een derde 'booster'-prik gehad. Hoe hoger de leeftijdsgroep – en daarmee het risico op ernstig beloop van de ziekte – hoe hoger de vaccinatiegraad bovendien is.¹¹ Inmiddels is bekend dat vaccinaties het risico op een ernstig beloop met zo'n 95 procent verkleinen, maar dat het risico op een besmetting minder afneemt en ook gevaccineerde mensen na een infectie anderen kunnen besmetten. De zogeheten omikron-variant van het virus – eind januari 2022 goed voor 99 procent van alle besmettingen – blijkt daarbij aanzienlijk besmettelijker dan eerdere varianten en breekt gemakkelijker door een vaccinatie heen, maar leidt ook minder vaak tot opname in het ziekenhuis of overlijden. Ondanks ongekend hoge besmettingscijfers van soms meer dan 100 duizend per dag, was dat voor het kabinet reden om de beperkende maatregelen vanaf januari 2022 stapsgewijs af te schalen.

Al deze ontwikkelingen hebben gevolgen voor zowel de kosten als de baten van de CoronaMelder. Enerzijds biedt de app potentieel meer toegevoegde waarde naast handmatig contactonderzoek, nu de maatschappij geleidelijk weer open is gegaan, waardoor mensen meer

¹¹ <https://coronadashboard.rijksoverheid.nl/>, geraadpleegd 14 februari 2022.

contacten hebben met personen die niet privé of via handmatig contactonderzoek op de hoogte kunnen worden gebracht: denk aan de passagier naast je in de trein of de persoon aan de volgende tafel in een restaurant. Daar komt bij dat de app potentieel meer schaalvoordelen heeft dan handmatig contactonderzoek waarvoor bij 100 duizend besmettingen per dag nog nauwelijks tijd is. De app zou dus naar verhouding meer contacten kunnen identificeren die niet langs andere weg in beeld komen. Gelet op de grote besmettelijkheid van de omikron-variant zal een dergelijk contact bovendien vaker tot een besmetting kunnen leiden, waarbij vooral de vraag is of de anderhalve meter die de app beoogt aan te houden bij omikron niet te krap is voor een risicovol contact.

Anderzijds neemt door de hoge vaccinatiegraad de kans op *false positives* toe en nemen de maatschappelijke baten van een *true positive* af. Om met het eerste te beginnen: steeds vaker zullen volledig gevaccineerde en geboosterde app-gebruikers een melding krijgen. Dat kan het percentage negatieve tests als gevolg van een melding verhogen en (achteraf) onnodige quarantaine tot gevolg hebben, wat de kosten-batenverhouding verslechtert. Die verhouding wordt verder verslechterd doordat de baten van een vermeden besmetting van een gevaccineerde gebruiker beduidend lager zijn door de bescherming die de vaccinatie biedt tegen ernstig beloop. Maar ook ongevacineerde gebruikers hebben bij een besmetting met omikron betere kaarten.

Een gezonde verhouding tussen kosten en baten vergt dan andere adviezen bij een melding van de app. Het advies voor volledig gevaccineerde en recent herstelde app-gebruikers om na een melding in quarantaine te gaan en zich te laten testen is al enige tijd geleden verval-¹² Een melding is voor hen vooral een extra waarschuwing geworden om een zelftest te doen en in quarantaine te gaan wanneer ze toch milde symptomen ontwikkelen. Met name in kringen van niet-gevaccineerde gebruikers kan de app toegevoegde waarde houden, als ze bereid zijn de app te installeren en zich in acht te nemen als ze een melding krijgen.

Auteur

Joost Poort (poort@uva.nl) is Universitair hoofddocent en co-directeur aan het Instituut voor Informatierecht (IViR), Universiteit van Amsterdam.

Literatuur

- Andreoni, J., 1989, Geven met onzuiver Altruïsme: Toepassingen op liefdadigheid en Ricardiaanse Gelijkwaardigheid, *Tijdschrift voor Politieke Economie*, vol. 97-6 (Dec 1989): 1447-1457.
- Baarsma, B., E. van den Broek-Altenburg, G. van den Berg en C. Teulings, 2021, Langetermijnbelangen worden bij de aanpak van corona veronachtzaamd, *ESB 16* (april 2021).
- Brazeau, N.F., R. Verity en S. Jenks et al., 2020, [Report 34: COVID-19 Infection Fatality Ratio: Estimates from](#)

¹² <https://www.coronamelder.nl/nl/faq/3-wat-als/>

- [Seroprevalence](#). Imperial College COVID-19 response team, 29 oktober 2020.
- Broughel, J. en M. Kotrous, 2021, [De voordelen van de onderdrukking van het Coronavirus: A cost-benefit analysis of the response to the first wave of COVID-19 in the United States](#), *Covid Economics* 67, 4 februari 2021: 128-171.
- Centraal Planbureau, 2019, [Decemberraming: Economisch Vooruitzicht 2020](#). Den Haag.
- Centraal Planbureau, 2020, [Juniraming 2020](#). Den Haag.
- Centraal Planbureau, 2021, [Macro Economische Verkenning 2022](#). Den Haag.
- Doti, J.L., 2021, [Benefit-cost analysis of COVID-19 policy intervention at the state and national level](#), *Covid Economics* 67, 4 februari 2021: 94-127.
- Ebbers, W., L. Hooft, N. van der Laan en E. Metting, 2021, [Evaluatie CoronaMelder. Een overzicht na 9 maanden](#). Erasmus Universiteit Rotterdam/UMC Utrecht/Tilburg University/Rijksuniversiteit Groningen.
- Ferretti, L. C. Wymant en M. Kendall et al., 2020, [Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing](#), *Wetenschap* 368, 619. (8 mei 2020).
- Frijters, P., 2020, Vanuit geluk bezien zijn de kosten van een lockdown veel hoger dan de baten, *ESB* (12 november 2020).
- Giessen, A. van, A. de Wit en C. van den Brink et al., 2020, [Impact van de eerste COVID-19 golf op de reguliere zorg en gezondheid. Inventarisatie van de omvang van het probleem en eerste schatting van gezondheidseffecten](#), RIVM-rapport 2020-0183. Bilthoven.
- Helberger, N., S. Eskens, J. Strycharz, G. Bouchè, J. van Hoboken, J. van Mil, J. Toh, with N. Appelman, J. van Apeldoorn, M. van Eechoud, N. van Doorn, M. Sax en C. de Vreese, *Conditions for technological solutions in a COVID-19 exit strategy, with particular focus on the legal and societal conditions*. Report for ZonMw (<https://www.ivir.nl/conditions-for-technological-solutions-covid-19-exit-strategy/>).
- Kolen, B., 2020, Een eerste kwantitatieve analyse van de Nederlandse coronamaatregelen, *ESB* (12 november 2020).
- Koopmans, C., 2020, Kosten en baten van lockdowns blijven onzeker, *ESB*, (12 november 2020).
- Rehse, D. en F. Tremöhlen, 2020, Bevordering van participatie in digitale volksgezondheidsinterventies: The Case of Digital Contact Tracing (2020, ZEW - Centrum voor Europees Economisch Onderzoek Discussion Paper No. 20-076, beschikbaar op SSRN: <https://ssrn.com/abstract=3761710>).
- Stoeldraijer, L., T. Traag en C. Harmsen, 2021, [Oversterfte tijdens eerste golf corona-epidemie bijna dubbel zo hoog als tijdens griepepidemie](#), CBS 21 mei 2021.
- Teulings, C.N., 2021, [Schoolsluiting is contraproductief en zelfvernietigend](#), *Covid Economics* 69, 18 februari 2021: 166-175.
- Tweede Kamer, 2016, *Brief van de Minister van Volksgezondheid welzijn en sport* (Sociaal domein), 34 477 (4).
- Tweede Kamer, 2021, *Brief van de Minister van Volksgezondheid* (Infectieziektenbestrijding), 25 295(1422).
- Zorginstituut Nederland, 2018, [Ziektelast in de praktijk. De theorie en praktijk van het berekenen van ziektelast bij pakketbeoordelingen](#).

Maatschappelijke kosten-batenanalyse en monitoring van brede welvaart: wat kunnen ze van elkaar leren?

Carl Koopmans

Maatschappelijke kosten-batenanalyses (MKBA's), de Monitor Brede Welvaart en de Brede Welvaartsindicator zijn alle drie gebaseerd op een breed welvaartsbegrip, maar de invulling daarvan verschilt sterk. De instrumenten kunnen leren van elkaars sterke punten. De MKBA kan worden verbeterd door een breder palet aan welvaartseffecten zichtbaar te maken, zoals onderzocht door CPB en PBL. Ook kan de discontovoet worden aangepast om beter aan te sluiten bij het streven naar duurzaamheid. De Monitor Brede Welvaart en de Brede Welvaartsindicator kunnen worden verbeterd door het belang van indicatoren in te schatten met behulp van de MKBA-methodiek. Het verdient aanbeveling om ook dat te onderzoeken.

1 Meten van brede welvaart

Brede welvaart is in de mode. Van verschillende kanten wordt gepleit om beleid te baseren op welvaart in brede zin, inclusief bijvoorbeeld het milieu en sociale omstandigheden (ESB, 2021; Dietz, 2021). Daarbij is er ook groeiende aandacht voor het meten van welvaart (ESB, 2019). Het CBS publiceert sinds 2018 jaarlijks de Monitor Brede Welvaart, waarin een groot aantal cijfers en trends over de samenleving wordt gepresenteerd (CBS, 2021). Het CBS plaatst deze ontwikkelingen in de context van de Sustainable Development Goals (SDG's) van de Verenigde Naties (2020) waarbij internationale vergelijking een belangrijk aspect is. De Universiteit Utrecht en de Rabobank ontwikkelden de Brede Welvaartsindicator die bedoeld is om alle aspecten van brede welvaart samen te vatten in één integraal cijfer (Stegeman et al., 2017; Van Bavel et al., 2019). De Monitor Brede Welvaart en de Brede Welvaartsindicator worden ook ingezet om de brede welvaart van regio's in beeld te brengen (CBS, 2020; Badir et al., 2017).

De Tweede Kamer is bezorgd dat maatschappelijke kosten-batenanalyses (MKBA's) onvoldoende rekening houden met brede welvaart. In 2019 nam de Kamer een motie aan (Sneller et al., 2019) waarin de regering wordt gevraagd om de Algemene leidraad voor maatschappelijke kosten-batenanalyses (Romijn en Renes, 2013) te evalueren vanuit het concept van brede welvaart. De motie stelt dat de interpretatie van de 'maatschappelijke component' de laatste jaren is verbreed. Dit zou volgens de motie beter verankerd moeten worden in MKBA's. De motie veronderstelt impliciet dat brede welvaart onvoldoende is weerspiegeld in MKBA's. Dat roept de vraag op of dit inderdaad het geval is. Deze vraag is opgepakt door het Centraal Planbureau en het Planbureau voor de Leefomgeving (Bos et al., 2021) en komt ook aan de orde in dit artikel. Een vraag die de Tweede Kamer niet stelt is of de Monitor Brede Welvaart en de Brede Welvaartsindicator de brede welvaart wel goed weergeven. Er is veel kritiek op deze instrumenten, waarbij het de vraag is of zij de welvaart goed meten

(Hardeman, 2021; Jacobs, 2021). Daarnaast is van belang hoe relevant informatie over brede welvaart is bij het vormgeven van beleid.

In dit artikel gaan we na in hoeverre de instrumenten MKBA, Monitor Brede Welvaart en Brede Welvaartsindicator van elkaar kunnen leren. Welke zwakke punten van elk instrument kunnen worden verbeterd door gebruik te maken van de sterke punten van andere instrumenten? Om de sterke en zwakke punten in beeld te brengen beoordelen we in welke mate de instrumenten voldoen aan criteria met betrekking tot inhoudelijke kwaliteit en bruikbaarheid voor beleid. Daarbij baseren we ons deels op een overzicht van overeenkomsten en verschillen tussen MKBA en de Monitor Brede Welvaart van Filak et al. (2020). Visser en Wortelboer-van Donselaar (2021) onderzochten specifiek voor het mobiliteitsbeleid hoe MKBA's en de Monitor Brede Welvaart kunnen worden verbeterd.

Eerst wordt kort beschreven welke rol brede welvaart in economisch onderzoek speelt. Daarna worden de criteria geïntroduceerd. Vervolgens analyseren we in hoeverre de drie instrumenten voldoen aan de criteria en welke mogelijkheden er zijn om de instrumenten te verbeteren. Tot slot worden conclusies getrokken.

2 Brede welvaart in economisch onderzoek

Welvaartseconomie is een belangrijk onderdeel van economisch onderzoek (zie bijvoorbeeld Boadway en Bruce, 1984). Daarbij wordt uitgegaan van een breed of 'open' welvaartsbegrip, waarin alles in de economie ('allocatie') dat het nut van mensen beïnvloedt, onderdeel is van de welvaart. Dat omvat dus ook bijvoorbeeld de hoeveelheid vrije tijd en het behoud van het milieu (Wolfson, 1988, p. 22, 26). Dit brede welvaartbegrip lijkt dus meer op welzijn dan op welvaart in enge, financiële zin. Een belangrijk instrument van de welvaartseconomie is de sociale welvaartsfunctie, een soort 'nutsfunctie voor een heel land', die aangeeft hoe de sociale welvaart afhangt van de omvang en verdeling van schaarse middelen. Ook woeden er al tientallen jaren discussies over een 'groen bbp' dat rekening zou moeten houden met het milieu (Van den Bergh, 2019). Er wordt gepleit voor het presenteren van meerdere indicatoren naast het bbp, ook in de Tweede Kamer (Tijdelijke Commissie Breed Welvaartsbegrip, 2016). Deze wens van de Tweede Kamer was de aanleiding voor het ontwikkelen van de Monitor Brede Welvaart (CBS, 2021).

MKBA's zijn gebaseerd op de welvaartstheorie, en daarmee ook op een brede interpretatie van het begrip welvaart (Boadway & Bruce, 1984). Dat geldt ook voor MKBA's in Nederland (Eijgenraam et al., 2000). Daarbij wordt het Kaldor-Hicks-criterium toegepast (Kaldor, 1939; Hicks, 1939). Er wordt geschat hoe groot een financiële overdracht zou zijn die een positief dan wel negatief effect op de welvaart precies zou compenseren. Dit wordt aangeduid als de 'betalingsbereidheid' van burgers en bedrijven voor effecten van beleid. Bij gewenste effecten is de betalingsbereidheid positief (baten), bij ongewenste effecten negatief (kosten). Daarbij gaat het om alle beleidseffecten waar mensen belang aan hechten, dus

inclusief effecten op het milieu en andere niet-financiële effecten. De kosten en baten wordt opgeteld over alle burgers en bedrijven in Nederland (of een regio, Europa, of de wereld). Het wordt in MKBA-richtlijnen aanbevolen om ook de kosten en baten van specifieke groepen (verdelingseffecten) te laten zien (Romijn en Renes, 2013).

Het meten van (brede) welvaart op basis van betalingsbereidheid is niet onomstreden. Deze aanpak is gebaseerd op utilitarisme: het maximaliseren van tevredenheid en welzijn voor iedereen (Bentham, 1789). Sen (1999) stelt dat de utilitaristische benadering te weinig aandacht geeft aan rechtvaardigheid en aan essentiële rechten die mensen moeten hebben. Hij stelt voor om uit te gaan van een ‘capabilities’ benadering, waarin de mogelijkheden van mensen om te functioneren centraal staan. Aspecten van capabilities zijn een lang leven, gezondheid, veiligheid, opleiding, een (zinvolle) baan, zelfrespect, tevredenheid en kalmte. De capabilities-benadering is sterk van invloed op het Human Development Report (UNDP, 2020) en de Sustainable Development Goals van de Verenigde Naties, en op de Better Life Index van de OESO (OECD, 2020). Ook de Monitor Brede Welvaart en de Brede Welvaarts-indicator baseren zich (ten dele) op capabilities.

MKBA's worden in de praktijk vooral ex ante ingezet om beleidsopties te vergelijken, maar kunnen ook ex post worden ingezet. Indicatoren van brede welvaart worden gebruikt om ex post na te gaan hoe de welvaart zich heeft ontwikkeld.

3 Criteria en vergelijking

We passen inhoudelijke criteria toe die de kwaliteit van het instrument weerspiegelen. Het eerste inhoudelijke criterium is *completeid*. Het idee achter brede welvaart is dat alle aspecten die van belang zijn voor de welvaart (in brede zin) door het instrument worden meegenomen. Het tweede criterium is *wetenschappelijke onderbouwing*. Het instrument moet zich zoveel mogelijk baseren op wetenschappelijke methoden en feitelijke gegevens. Ook is de selectie en presentatie van informatie bij voorkeur niet gebaseerd op subjectieve keuzes of toevalligheden.

Daarnaast is de bruikbaarheid van een instrument van groot belang. Het is voor beleidsmakers en politici prettig als er *helderheid* is over de uitkomsten en de wijze waarop deze zijn bereikt. Een belangrijke functie van beleidsvoorbereiding is verder dat *trends zichtbaar* worden waar ‘het beleid’ eventueel op kan reageren. Daarnaast is het nuttig als instrumenten de voor- en nadelen van *beleidsopties vergelijken* en *nieuwe beleidsopties genereren*. En tot slot is het belangrijk dat de *onderzoekskosten* niet te hoog zijn.

De Monitor Brede Welvaart is het meest compleet van de drie instrumenten omdat honderden indicatoren zijn opgenomen die een breed spectrum van dimensies van (brede) welvaart omvatten. De Brede Welvaartsindicator is gebaseerd op een beperkt aantal indicatoren vergeleken met de Monitor. Een MKBA streeft naar completeid, maar in de praktijk worden soms relevante effecten niet opgenomen of niet gemonetariseerd, terwijl dit wel

mogelijk was (met name effecten op de natuur; zie Annema en Koopmans, 2015). Verdelingseffecten worden in MKBA's in beperkte mate weergegeven (Annema et al., 2007).

Wat betreft de wetenschappelijke onderbouwing scoort de MKBA vrij goed. De MKBA maakt onderscheid tussen belangrijke en minder belangrijke zaken door effecten te waarderen in geld (monetariseren). Effecten waarvoor een hoge betalingsbereidheid bestaat, wegen relatief zwaar mee in het kosten-batensaldo. Een MKBA volgt strakke regels die aangeven hoe de berekeningen dienen te worden uitgevoerd (Romijn en Renes, 2013; en diverse werkwijzers). Er is echter ook enige ruimte voor eigen keuzes door onderzoekers die in de praktijk mogelijk minder goed onderbouwd zijn, bijvoorbeeld bij de invulling van toekomstscenario's of de keuze van methoden om effecten monetariseren.

De Monitor Brede Welvaart en de Brede Welvaartsindicator bevatten indicatoren die volgens wetenschappelijke standaarden zijn gemeten, maar de keuze van deze indicatoren is niet wetenschappelijk onderbouwd. Alkire (2002) laat zien dat 'human development' op sterk uiteenlopende manieren kan worden gemeten. Jacobs (2021) stelt vast dat de indicatoren niet zijn gebaseerd op schaarste of economische theorie. Kalshoven (2018) noemt de keuze van indicatoren in de Monitor "willekeurig". De keuze van indicatoren lijkt in belangrijke mate gebaseerd op de beschikbaarheid van statistieken. Ook overlappen de indicatoren elkaar. In de Monitor Brede Welvaart is de indicator 'bbp per hoofd' sterk bepalend voor 'Besteedbaar inkomen'. 'Tevredenheid met het leven' wordt waarschijnlijk sterk beïnvloed door 'Tevredenheid met vrije tijd', 'Tevredenheid met woning' en 'Ervaren regie over het eigen leven'. Bij de Brede Welvaartsindicator (Rijpma et al., 2017) meten 'Happiness' en 'Life satisfaction' ongeveer hetzelfde, en worden deze indicatoren waarschijnlijk sterk bepaald door bijvoorbeeld 'Long term unemployment', 'Housing satisfaction' en 'Household income'.

De Monitor Brede Welvaart presenteert verschillende indicatoren als ongeveer gelijkwaardig, maar in werkelijkheid kunnen sommige indicatoren (veel) sterker bijdragen aan de welvaart dan andere indicatoren. Zo is het waarschijnlijk dat de indicator 'groei tijdverlies door files en vertraging' veel minder belangrijk is voor de totale (brede) welvaart dan de indicator 'bbp-groei', want het bbp (circa 800 miljard euro per jaar) is honderden malen groter dan de kosten van files en vertraging (3 tot 4 miljard euro per jaar; KiM, 2019). Bij de Brede Welvaartsindicator wordt (na schaling) een gewogen gemiddelde van indicatoren berekend. De gewichten zijn gebaseerd op de waardering die Nederlanders in enquêtes geven aan de indicatoren (Stegeman et al., 2017; Van Bavel et al., 2019). Het is echter bekend dat enquêtes vaak aanleiding geven tot sociaal wenselijke en strategische antwoorden. Daarnaast is het de vraag of de antwoorden gaan over het gepercipieerde belang van onderwerpen (economie, milieu, rechtvaardige verdeling) in het algemeen, of over de specifieke indicatoren in de Brede Welvaartsindicator.

Beleidskeuzes worden uiteindelijk gemaakt door 'de politiek'. Daarbij kunnen politici beleidseffecten verschillend wegen. Zo zal een politicus van GroenLinks waarschijnlijk meer belang hechten aan de CO₂-uitstoot dan een politicus van de PVV. Een MKBA weegt effecten op basis van de gemiddelde voorkeuren (gemeten als betalingsbereidheid) van alle inwoners van Nederland. Dit kan worden gezien als een wetenschappelijk onderbouwde manier om politieke keuzes te ondersteunen, waarbij politici het laatste woord hebben en daarbij (impliciet) eigen gewichten kunnen toepassen. Dat impliceert dat een MKBA afzonderlijke

Nu, later en de discountvoet

Indicatoren die de toekomst betreffen worden in de Monitor Brede Welvaart als even belangrijk gepresenteerd als de indicatoren over het heden. De Monitor maakt onderscheid tussen indicatoren 'hier en nu' en 'later' (daarnaast zijn er indicatoren 'elders', over het buitenland). De indicatoren 'later' geven aan hoeveel economisch, natuurlijk, menselijk en sociaal kapitaal Nederland nalaat aan toekomstige generaties. Daarmee zijn het geen indicatoren van toekomstige ontwikkelingen, maar zijn deze indicatoren wel relevant voor de toekomst van Nederland.

In MKBA's worden welvaartseffecten ingeschat die in de toekomst optreden, die vervolgens worden gereduceerd door te disconteren. De discountvoet weerspiegelt enerzijds een negatief risicovrij (reeël) rendement op kapitaalmarkten van min 1%. Daarnaast is er een risico-opslag van 3,25% die voortkomt uit de (macro-economische) risico's die gepaard gaan met investeringen. Per saldo is de standaard discountvoet 2,25% (Werkgroep discountvoet, 2020).

Op het eerste gezicht lijkt het of de MKBA de toekomst onderwaardeert door met een positieve discountvoet te disconteren, maar dat ligt anders. Uit de onderbouwing van deze discountvoet blijkt dat de discountvoet positief is vanwege risico's die verbonden zijn aan investeren. De risicovrije discountvoet is negatief en weerspiegelt dus juist (tijds)voorkeur voor de toekomst vergeleken met het heden. Dit neemt niet weg dat de verre toekomst in MKBA's een beperkte rol speelt. Bijvoorbeeld een effect over 50 jaar wordt bij een discountvoet van 2,25% vermenigvuldigd met $(1/(1+0,0225)^{50})$ 0,33, en dus met twee derde verminderd.

In diverse andere landen wordt in MKBA's gewerkt met een discountvoet die niet constant is, maar daalt naarmate kosten en baten verder in de toekomst liggen (Koopmans en Rietveld, 2013). Dit kan niet alleen worden onderbouwd vanuit normatieve duurzaamheids-overwegingen, maar ook met de onzekerheid die inherent is aan langetermijnbeslissingen (Gollier et al., 2008). In Nederland is niet gekozen voor een dalende discountvoet, want om verschillende redenen zouden gevoeligheidsanalyses met een (over de gehele periode) lagere discountvoet volgens de Werkgroep discountvoet (2020) beter zijn. Wel beveelt de Werkgroep nader onderzoek aan. Wellicht kan de toenemende aandacht voor duurzaamheid en in het bijzonder klimaatverandering, aanleiding zijn om dit onderzoek snel te starten.

* De Brede Welvaartsindicator richt zich minder op de toekomst dan de Monitor Brede Welvaart en blijft daarom in deze box buiten beschouwing.

effecten die politiek relevant zijn expliciet moet laten zien. Daarbij gaat het niet alleen over soorten effecten zoals CO₂-uitstoot of reductie van files, maar ook over de verdelingseffecten voor groepen in de samenleving die voor politici belangrijk zijn.

Wat betreft helderheid hebben alle drie de instrumenten zowel pluspunten als beperkingen. De uitkomst van een MKBA is doorgaans helder: de baten overtreffen de kosten, of juist niet, of baten en kosten zijn ongeveer even groot. De wijze waarop kosten en baten zijn berekend wordt door beleidsmakers echter vaak ervaren als een 'black box' (Dopheide et al., 2011). De indicatoren in de Monitor Brede Welvaart zijn duidelijk en worden visueel aantrekkelijk gepresenteerd met kleuren, maar de veelheid aan indicatoren leidt tot een minder goed overzicht. Sommige indicatoren ontwikkelen zich gunstig, andere minder gunstig, maar er is geen duidelijk totaalbeeld of de welvaart toeneemt dan wel afneemt. Voor de Monitor Brede Welvaart en de Brede Welvaartsindicator geldt dat de hierboven beschreven onevenwichtigheid in de presentatie respectievelijk weging van indicatoren en dubbeltellingen tussen indicatoren leiden tot een minder helder beeld. De indicatoren waaruit de Brede Welvaartsindicator is opgebouwd zijn helder, maar de betekenis van de samengestelde index is vrij abstract.

Beleidsrelevante trends worden het best zichtbaar gemaakt door de Monitor Brede Welvaart en de Brede Welvaartsindicator. De Monitor maakt de ontwikkeling van een groot aantal indicatoren zichtbaar. Voor diverse indicatoren zijn er concrete beleidsdoelen; dit onderstreept het belang van deze indicatoren. Een MKBA heeft niet als primair doel om trends zichtbaar te maken, maar laat wel *en passant* - in een knelpuntenanalyse en in het nulalternatief - zien welke knelpunten in de toekomst worden verwacht. Echter, dit betreft alleen knelpunten gerelateerd aan de onderzochte beleidsmaatregelen. De Brede Welvaartsindicator brengt in beeld of de brede welvaart toeneemt of afneemt (of de brede welvaart daarbij correct wordt gemeten, hoort bij het criterium wetenschappelijke onderbouwing).

De MKBA geeft van de drie instrumenten als enige ook duidelijke informatie over de merites van beleidsopties. Een MKBA geeft aan of beleidsopties de welvaart verhogen of verlagen, en in welke mate. De Monitor Brede Welvaart en de Brede Welvaartsindicator geven alleen informatie over ontwikkelingen tot nu toe. Daar kunnen wel doelen op worden gebaseerd, maar geen concrete beleidsmaatregelen. Wel worden de indicatoren in deze instrumenten beïnvloed door eerder gevoerd beleid; dit zou lessen kunnen opleveren voor toekomstig beleid. De indicatoren worden echter mede bepaald door andere ontwikkelingen zoals de groei van de wereldeconomie of EU-beleid, waardoor de effecten van Nederlands beleid niet afzonderlijk zichtbaar zijn.

De drie instrumenten zijn niet goed in staat om nieuwe beleidsopties te genereren. De Monitor Brede Welvaart en de Brede Welvaartsindicator bevatten hierover geen enkele informatie. Een MKBA levert geen geheel nieuwe beleidsmogelijkheden op, maar kan wel suggesties doen om bestaande beleidsvarianten aan te passen om tot een betere kosten/baten

uitkomst te komen. Het lijkt niet goed mogelijk om de instrumenten geschikter te maken voor het bedenken van nieuw beleid.

De onderzoekskosten van sommige instrumenten zijn hoog, maar veel kleiner dan de kosten van verkeerde beleidskeuzes. De kosten van de Monitor Brede Welvaart zijn hoog door het grote aantal indicatoren, maar worden tegelijk beperkt doordat wordt voortgebouwd op bestaande statistieken. De kosten van de Brede Welvaartsindicator zijn minder hoog omdat het aantal indicatoren kleiner is, en ook hierbij bestaande statistieken worden gebruikt. Een MKBA kost meestal twintigduizend tot honderdduizend euro. Alle drie instrumenten kunnen beleidskeuzes beïnvloeden die over vele miljoenen euro's gaan. Bij de Monitor Brede Welvaart en Brede Welvaartsindicator is die invloed indirect, terwijl een MKBA over concrete beleidskeuzes gaat.

Onderstaande tabel laat zien dat elk van de instrumenten zowel voordelen als nadelen heeft. Dit hangt samen met de doelen van de instrumenten. De Monitor Brede Welvaart en de Brede Welvaartsindicator hebben als doel om trends te laten zien terwijl MKBA's zich richten op beleidskeuzes. Ook blijkt dat diverse voordelen keerzijdes hebben. De relatief goede wetenschappelijke onderbouwing van de MKBA komt voort uit het volgen van een technisch-economische methode, waardoor rapportages minder helder worden. De Monitor scoort goed bij compleetheid omdat een groot aantal indicatoren wordt gepresenteerd, maar dit gaat ten koste van de overzichtelijkheid en daarmee van de helderheid.

Tabel 1 Plussen en minnen van de drie instrumenten

criterium	MKBA	Monitor Brede Welvaart	Brede Welvaartsindicator
Inhoudelijke kwaliteit			
Compleetheid	+	++	0
Wetenschappelijke onderbouwing	+	0	0
Bruikbaarheid voor beleid			
Helderheid	0	0	0
Beleidsrelevante trends zichtbaar	-	++	++
Beleidsopties vergelijken	++	-	-
Nieuwe beleidsopties genereren	-	--	--
Onderzoekskosten	+	+	++

4 Instrumenten verbeteren

De analyse van voor- en nadelen van de instrumenten biedt ook aanknopingspunten voor verbeteringen (zie ook Bos et al., 2022; Visser en Wortelboer-van Donselaar, 2021):

- De MKBA kan afzonderlijke effecten, kosten en baten beter presenteren, om duidelijk te maken hoe het kosten-batensaldo is berekend. Hiervoor zijn eerder diverse suggesties

gedaan (Lijn 43 en Studio Barten, 2013; Renes et al., 2011; Koopmans, 2004), maar deze worden in de praktijk vrij weinig toegepast. Naast deze eerdere voorstellen kan de Monitor Brede Welvaart een inspiratiebron zijn. Met name het indelen van effecten in heldere categorieën en het weergeven van effecten met kleuren in visueel aantrekkelijke figuren kan bijdragen aan de helderheid van MKBA's.

- De MKBA kan nagaan of alle relevante effecten zijn meegenomen. De Monitor Brede Welvaart kan daarbij dienen als 'checklist'. Het gaat hierbij niet alleen om effecten die naar verwachting tot substantiële kosten en baten leiden, maar ook om effecten die veel maatschappelijke aandacht krijgen, zoals effecten op natuur en milieu. Door deze effecten ook op te nemen als ze weinig kosten of baten genereren, wordt voorkomen dat een beeld ontstaat dat deze effecten geen aandacht krijgen in de MKBA. Ook wordt hierdoor helderder hoe de kosten en baten zijn bepaald. De checklist kan tevens leiden tot meer aandacht voor verdelingseffecten in MKBA's.
- De discontovoet in MKBA's kan beter rekening houden met duurzaamheid. Een manier om dat te bereiken is de discontovoet verlagen voor effecten die in de verre toekomst liggen.
- De Monitor Brede Welvaart en de Brede Welvaartsindicator kunnen het relatieve belang van indicatoren voor de totale welvaart beter laten zien. Een hulpmiddel hierbij is het in geld uitdrukken van indicatoren (zowel voor niveaus als veranderingen in de tijd). Onderzoeken naar de financiële waarde van milieuschade, files, arbeidsparticipatie en levensjaren, die in MKBA's worden gebruikt om effecten te monetariseren, kunnen worden benut om bij een deel van de indicatoren in de Monitor en in de Brede Welvaartsindicator gewichten te schatten. Dat lukt niet bij alle indicatoren; bijvoorbeeld verdelingseffecten zijn lastig in geld uit te drukken. De gewichten kunnen in de Brede Welvaartsindicator worden gebruikt om indicatoren te wegen. In de Monitor kunnen verschillen tussen gewichten aanleiding zijn om indicatoren meer dan wel minder prominent te presenteren. Ook kunnen indicatoren met een laag gewicht worden weggelaten of samengevoegd. Dat geeft de lezer meer overzicht door alleen belangrijke indicatoren te laten zien.
- De Monitor Brede Welvaart en de Brede Welvaartsindicator kunnen de selectie van indicatoren beter onderbouwen, met name door overlap tussen indicatoren te verkleinen. Aangezien allerlei aspecten van tevredenheid (met gezondheid, wonen etc.) zijn opgenomen, is het de vraag wat een indicator van algemene tevredenheid nog toevoegt. En evenzo, als inkomen en arbeidsparticipatie zijn opgenomen, is het de vraag wat de meerwaarde is van het ook nog opnemen van het bbp per hoofd.

5 Conclusies

Er bestaat geen ideale methode om welvaart te meten. De drie beschreven instrumenten zijn alle drie gebaseerd op een breed welvaartsbegrip, maar hebben verschillende doelen en operationaliseren het begrip welvaart sterk verschillend. De instrumenten kunnen leren van elkaars sterke punten. De MKBA kan worden verbeterd door een breder palet aan

welvaartseffecten zichtbaar te maken, zoals onderzocht door CPB en PBL (Bos et al., 2022). Ook kan de discontovoet worden aangepast om beter aan te sluiten bij het streven naar duurzaamheid. De Monitor Brede Welvaart en de Brede Welvaartsindicator kunnen worden verbeterd door het belang van indicatoren in te schatten met behulp van de MKBA-methodiek. Het verdient aanbeveling om ook dat te onderzoeken.

Auteur

Carl Koopmans (e-mail: c.koopmans@seo.nl) is onderzoeksdirecteur bij SEO Economisch Onderzoek en hoogleraar beleidsevaluatie aan de Vrije Universiteit.

De auteur dankt Frits Bos (CPB), Bert Hof (PBL), Theo Roelandt (ministerie EZK), Bert Tieben (SEO), Henry van der Wiel (ministerie EZK) en een anonieme reviewer voor nuttig commentaar.

Literatuur

- Alkire, S., 2002, Dimensions of human development. *World development*, vol. 30(2): 181-205.
- Annema, J.A. en C. Koopmans, 2015, The practice of valuing the environment in cost-benefit analyses in transport and spatial projects, *Journal of Environmental Planning and Management*, vol. 58(9): 1635-1648.
- Annema, J.A., C. Koopmans en B. van Wee, 2007, Evaluating transport infrastructure investments: The Dutch experience with a standardized approach, *Transport Reviews*, vol. 27(2): 125-150.
- Badir, M., B. van Bavel, S. Hardeman en A. Rijpma, 2017, Brede Welvaartsindicator 2017. Brede welvaart in Nederland: nationaal en regionaal, Utrecht: Universiteit Utrecht en Rabobank.
- Bavel, B.J.P. van, A. Rijpma en S. Hardeman, 2019, Vervolgstappen voor integrale welvaartsmeting. *Economisch Statistische Berichten*, vol. 104(4772s): 22-25.
- Bentham, J., 1789, An introduction to the principles of morals and legislation. <https://www.earlymoderntexts.com/assets/pdfs/bentham1780.pdf>
- Bergh, J. van den, 2019, BBP vervangen door brede welvaartsmaat is moeilijk en onnodig, *Economisch Statistische Berichten*, vol. 104(4772): 66-71.
- Boadway, R.W. en N. Bruce, 1984, *Welfare economics*. New York: B. Blackwell.
- Boardman, A.E., D.H. Greenberg, A.R. Vining en D.L. Weimer, 2018, *Cost-benefit analysis: concepts and practice*, Fifth edition. Cambridge University Press.
- Bos, F., B. Hof en J. Tijn, 2022, Maatschappelijke kosten-batenanalyse en brede welvaart; een aanvulling op de Algemene MKBA-Leidraad. Den Haag: Centraal Planbureau en Planbureau voor de Leefomgeving. (te verschijnen).
- CBS, 2020, Regionale Monitor Brede Welvaart toont grote verschillen stad en platteland. <https://www.cbs.nl/nl-nl/nieuws/2020/49/regionale-monitor-brede-welvaart-toont-grote-verschillen-stad-en-platteland>
- CBS, 2021, Monitor Brede Welvaart & de SDG's 2021, Den Haag: Centraal Bureau voor de Statistiek.
- Dietz, 2021, Economisch herstelbeleid vraagt om discussie over welvaart, *Economisch Statistische Berichten*, vol. 106(4798S).
- Dopheide, E.J.M., F. Sijtsma en A. van der Veen, 2011, Begrip van MKBA is beperkt onder gebruikers. *Economisch Statistische Berichten*, vol. 96(4614-4615): 460-463.

- Ederveen, S. en M. Stoel, 2021, Veranker sturen op 'brede welvaart' in het regeerakkoord, *Economisch Statistische Berichten*, vol. 106(4800s).
- Eijgenraam, C.J.J., C.C.Koopmans, P.J.G. Tang en A.C.P. Verster, 2000, Evaluatie van infrastructuurprojecten; Leidraad voor kosten-batenanalyse, Sdu, Den Haag.
- ESB, 2019, Dossier Meten van welvaart, *Economisch Statistische Berichten*, vol. 104(4772s).
- ESB, 2021, Dossier Sturen op welvaart. *Economisch Statistische Berichten*, vol. 106(4800s).
- Filak, M., A. Ouarraki, S. Gerritsen en M. Stoel, 2020, Maatschappelijke Kosten Baten Analyse en de Monitor Brede Welvaart & SDG's, Versie 1.0. Bureau Strategische Analyse, Inspectie der Rijksfinanciën, Ministerie van Financiën, 3 november.
- Gollier, C., P. Koundouri en T. Pantelidis, 2008, Declining discount rates: Economic justifications and implications for long-run policy. *Economic Policy*, vol. 23(56): 758-795.
- Hardeman, S., 2021, Aandacht voor brede welvaart vraagt nieuwe kijk op beleidsindicatoren, *Me Judice*.
- Hicks, J., 1939, The Foundations of Welfare Economics, *Economic Journal*, vol. 49(196): 696-712.
- Hoekstra, R.S., 2021, Integreer welvaart in de beleidscyclus, *Economisch Statistische Berichten*, vol. 106(4800s): 46-50.
- Jacobs, B., 2021, Het beleidsdebat over welvaart mist theoretische fundering, *Economisch Statistische Berichten*, vol. 106(4800S): 35-39.
- Kaldor, N., 1939, Welfare Propositions of Economics and Interpersonal Comparisons of Utility, *Economic Journal*, vol. 49(195): 549-552.
- Kalshoven, F., 2018, De nieuwe indicator is nog een rommeltje. Dat kan beter, CBS, *de Volkskrant*, 18 mei.
- KiM, 2019, Mobiliteitsbeeld 2019. Den Haag: Kennisinstituut voor Mobiliteitsbeleid.
- Koopmans, C.C., 2004, Heldere Presentatie OEI – Aanvulling op de Leidraad OEI. Den Haag: Ministerie van Verkeer en Waterstaat en Ministerie van Economische Zaken.
- Koopmans, C. en P. Rietveld, 2013, Long-term impacts of mega-projects: the discount rate, in: *International handbook on mega-projects*, Edward Elgar Publishing.
- Lijn 43 en Studio Barten, 2013, Handreiking visualiseren van MKBA-resultaten, Tien tips voor een heldere presentatie.
- OECD, 2018, What matters the most to people? Evidence from the OECD Better Life Index users' responses. Working Paper 90, Parijs: OECD.
- OECD, 2020, *How's Life? 2020, Measuring Well-being*, Parijs: OECD Publishing.
- Renes, G., A. van Hoorn en D. Hamers, 2011, Verbetering van de communicatie en presentatie rondom de MKBA bij verstedelijkingsprojecten. Discussiepaper, Den Haag: Planbureau voor de Leefomgeving.
- Romijn, G. en G. Renes, 2013, Algemene leidraad voor maatschappelijke kosten-batenanalyse. Den Haag: Centraal Planbureau en Planbureau voor de Leefomgeving.
- Rijpma, A., M. Moatsos, M. Badir en H. Stegeman, 2017, Netherlands beyond GDP: A Wellbeing Index.
- Sen, A., 1999, Commodities and capabilities, *OUP Catalogue*.
- Sneller, Bruins, Slootweg, Snels, 2019, Motie van het lid Sneller c.s., Tweede Kamer 35300 nr. 54, 3 oktober 2019.
- Stegeman, H., M. Badir, A. Rijpma en M. Moatsos, 2017, Een indicator voor bredere welvaart voor Nederland, *ESB*, vol. 102(4747): 132-135.
- Tijdelijke Commissie Breed Welvaartsbegrip, 2016, Rapport. Tweede Kamer, 34 298 nr.3.
- UNDP, 2020, Human Development Report 2020, The next frontier. Human development and the Anthropocene, New York: United Nations Development Programme.
- Verenigde Naties, 2020, The Sustainable Development Goals Report 2020.

- Visser, J. en P. Wortelboer-van Donselaar, 2021, *Uitwerking van brede welvaart voor de monitoring en evaluatie van mobiliteitsbeleid*, Den Haag: Kennisinstituut voor Mobiliteitsbeleid.
- Werkgroep Discontovoet, 2020, *Rapport Werkgroep discontovoet 2020*.
- Wolfson, D.J., 1988, *Publieke sector en economische orde*, Wolters-Noordhoff.

Historicisme en zijn Monsters

Marcel Boumans

Dit artikel is een Nederlandse vertaling van de rede die Marcel Boumans heeft uitgesproken bij de aanvaarding van het ambt van hoogleraar Geschiedenis van de Economische Wetenschap aan de Universiteit Utrecht op 9 december 2021. De oratie duikt in de wetenschapsgeschiedenis voor een verklaring van hoe de sociale wetenschappen bepalen wat normaal is. Boumans laat zien hoe statistisch onderzoek (inclusief machine learning) haar eigen monsters creëert. In een 'historicistische' benadering van de wetenschap worden deze monsters als gevaarlijk gezien, omdat ze de orde verstoren. Boumans pleit voor het omgekeerde: verwelkom de afwijking van de norm. Dat past bij de open houding van de wetenschap.

“Doe normaal of ga weg”

Deze oproep was de kernboodschap van een brief van de minister-president Mark Rutte, gepubliceerd in de twee grootste Nederlandse ochtendkranten in januari 2017.¹

Het probleem is dat om aan deze oproep gevolg te kunnen geven, men moet weten wat normaal is. In de brief werd dit alleen aangegeven door enkele voorbeelden van sociaal gedrag te noemen, zoals “dat je elkaar de hand schudt” en te benadrukken dat dit soort gedrag typisch Nederlands is. Om erachter te komen wat in Nederland normaal is, wordt dit in de sociale wetenschappen meestal onderzocht met statistische methoden.

De leeropdracht die ik met deze oratie officieel aanvaard, is geschiedenis van de economische wetenschap, die ik beschouw als een onderdeel van wetenschapsgeschiedenis. Door een wetenschapshistorische analyse van Mark Ruttes oproep hoop ik de relevantie van wetenschapsgeschiedenis aan te tonen voor het begrijpen van de invloed van die moderne sociale wetenschappen op het politieke debat.² Ik zal laten zien dat deze roep om normaliteit, of eigenlijk elke roep om normaliteit, in tegenspraak is met de kernwaarden van de Universiteit Utrecht, namelijk die van “open blik, open houding, open wetenschap.”³

Deze analyse bestaat uit drie delen. Het eerste deel is de historische analyse van de term *normaal*, die zijn oorsprong vindt in de geschiedenis van de statistiek. In de statistiek wordt de normaal gedefinieerd als het midden van de normale verdeling (zie figuur 1). Om te

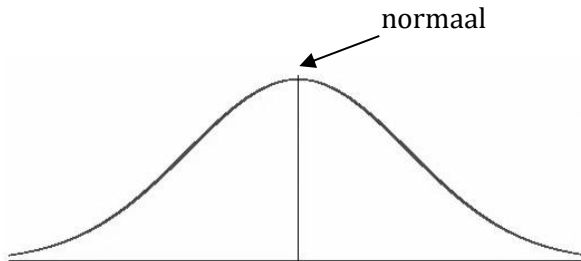
¹ Deze brief is geplaatst op de website van de VVD: <https://vvd.nl/content/uploads/2017/01/briefvanmark.pdf>. Laatst ingezien op 15 oktober 2021.

² Voor een bredere beschouwing verwijs ik graag naar Daston (2020).

³ <https://www.uu.nl/organisatie/over-ons/wie-wij-zijn>. Laatst ingezien op 17 december 2021.

bepalen wat in Nederland normaal is, onderzoeken sociaalwetenschappers wat de statistische gemiddelden zijn van verschillende soorten van sociaal gedrag.

Figuur 1 Normale verdeling



Het tweede deel laat zien dat een term, die verwijst naar een statistisch gemiddelde, een representant is geworden van nationale identiteit en attitude. En het laatste deel zal de consequentie van deze naturalisatie laten zien: waarom niet normaal doen betekent dat men weg moet, dat wil zeggen, Nederland moet verlaten.

1 Hoe bepalen we in de sociale wetenschappen wat normaal is?

We leven in het tijdperk van *Big Data* en *Machine Learning*, een tijd die wordt gekenmerkt door een sterk geloof in de mogelijkheid om over sociale groepen objectieve feiten te verkrijgen dankzij de beschikbaarheid van bijna onbeperkte aantallen data en *unsupervised* statistische analyse, waarbij *unsupervised* impliceert dat de *machine*, dat wil zeggen het algoritme, het alleen aankan, zonder menselijke tussenkomst.

Om te begrijpen hoe de normaal wordt bepaald, moeten we daarom eerst bekijken wat *machine learning* inhoudt. *Machine learning* is de moderne term voor wat vroeger statistische inductie werd genoemd. Het probleem van inductie is hoe we op grond van een bepaalde dataset tot algemene feiten komen over, bijvoorbeeld, een sociale groep. De vraag is: vertelt een statistisch gemiddelde ons alleen iets over de gebruikte dataset of vertelt het ons iets meer algemeen over de betreffende sociale groep?

De eerste wetenschapper die benadrukte dat inductie meer is dan alleen een samenvatting van de data, is de Britse negentiende-eeuwse wetenschapper William Whewell.⁴ Volgens Whewell is het kenmerkende van een inductie het vinden van een concept dat de data met elkaar verbindt. Whewell gebruikte daarvoor de term *colligation*. Als een voorbeeld voor de methoden voor de constructie van een dergelijke *colligation* noemde hij de kleinste kwadratenmethode.

⁴ Whewell 1858.

Dit idee dat inductie meer is dan een samenvatting van de data, werd ook uitgebreid besproken door John Maynard Keynes in zijn *Treatise on Probability*.⁵ Keynes maakte een onderscheid tussen de beschrijvende functie van statistiek, dat wil zeggen het samenvatten van data, en de inductieve functie. Met deze laatste functie probeert men aan de hand van de beschrijvingen van waargenomen gebeurtenissen iets te zeggen over gebeurtenissen die niet zijn waargenomen.⁶ Om met een verzameling gegevens iets te zeggen over een fenomeen buiten die verzameling, moeten er aannames moeten worden gemaakt over de relatie tussen die gegevens en dat fenomeen.

Volgens Keynes was inductie gebaseerd op wat hij de methode van analogie noemde. De kenmerken die door een reeks waarnemingen worden gedeeld, vormen een positieve analogie van de reeks, terwijl verschillen in kenmerken tussen de waarnemingen een negatieve analogie vormen. Keynes betoogde dat voor inductie een zorgvuldige weging van de positieve en negatieve analogieën, de overeenkomsten en verschillen, nodig is. Voordat men tot een gevolgtrekking kan komen, moet men de overeenkomsten in de gegevens vaststellen die daarmee een positieve analogie creëren. Een voorbeeld van een dergelijke analogie is de veronderstelling dat de natuur *uniform* is. De natuur wordt uniform geacht als positieverschillen in tijd en ruimte als irrelevant kunnen worden beschouwd. Keynes noemde deze veronderstelling het principe van uniformiteit.

Dit uniformiteitsprincipe was een belangrijk onderdeel van de kritiek van Keynes op Jan Tinbergens gebruik van regressieanalyse. In de jaren dertig had Tinbergen deze methode in de economie geïntroduceerd om tijdreeksen te analyseren om zodoende de causale verbanden te vinden die ten grondslag liggen aan de conjunctuurcyclus. De tijdreeksen die Tinbergen gebruikte betroffen allerlei economische waarnemingen over een tijdvak dat liep van 1919 tot 1939, een tijdvak van enorme economische turbulentie, een periode die dus nauwelijks uniform in tijd te noemen is. Deze kritiek van Keynes en Tinbergens reactie daarop, werd bekend als het Keynes-Tinbergen-debat. Keynes belangrijkste punt van kritiek was dat voor de regressieanalyse gegevens van een economie gebruikt werden waarvan duidelijk was dat in die periode deze economie niet constant was. Volgens hem was regressieanalyse alleen van toepassing op gegevens van een systeem dat uniform is in de tijd.⁷

Kort samengevat, het probleem dat Whewell en Keynes benoemen is dat een statistisch gemiddelde alleen meer is dan een samenvatting van een dataset als de individuele leden van een sociale groep iets gemeenschappelijks hebben, iets uniforms. Alleen dan geeft het gemiddelde een karakterisering van die groep.

⁵ Keynes 1973a.

⁶ Keynes 1973a, p. 359.

⁷ Keynes 1973b, p. 286.

2 De naturalisatie van het normale

Regressieanalyse is de statistische methode om relaties tussen variabelen te schatten. Deze statistische relaties worden correlaties genoemd. De gevonden relatie kan toevallig zijn of een echt bestaand structureel verband weergeven. Dat hangt af van de uniformiteit.

Regressieanalyse is afkomstig uit de biologie. De wetenschapshistoricus Judy Klein heeft laten zien dat uniformiteit een gegeven was in de eerste biometrische studies.⁸ Het ging in deze studies om correlaties tussen *hetzelfde* orgaan van verschillende generaties of tussen verschillende organen van *hetzelfde* organisme. De variabelen werden dus gemeten in vergelijkbare eenheden. En als het metingen van verschillende populaties waren, waren ze in ieder geval *organisch verwant*. Bovendien vertoonden de metingen meestal een normale frequentieverdeling met de kenmerkende klokvormige curve (zie figuur 1).

Klein liet zien dat het gebruik van regressieanalyse in de sociale wetenschappen de vraag opriep welke van de verschillende componenten van de sociale en economische statistieken gecorreleerd moesten worden. Keynes had aangegeven dat deze componenten uniform moesten zijn.

Een van de grondleggers van de moderne statistiek is de negentiende-eeuwse Britse wetenschapper Francis Galton. Hij benadrukte ook dat uniformiteit een voorwaarde was voor statistische inductie. Volgens hem werd aan deze voorwaarde voldaan als de data samenklonteren in een klokvormige curve. En hij benadrukte dat als de gegevens niet zo'n klokvormige kromme zouden laten zien, het gemiddelde dan "monsterlijk" en betekenisloos is.⁹

Op de vraag: "Wat hebben de leden van een sociale groep gemeen, wat is wat ze delen?," antwoordde Galton dat zij tot dezelfde *natuurlijke klasse* behoren, dat ze van hetzelfde *type* zijn, en dat de gemeenschappelijke kenmerken *typerend* zijn voor de natuurlijke klasse waartoe ze behoren. Dat wil zeggen, als deze kenmerken niet clusteren tot een klokvormige curve, zijn ze niet typisch; maar als ze dat wel doen, vertegenwoordigt de normaal het ideale type van die natuurlijke klasse.

Voor Galton was het geen enkel probleem om biometrische methoden te gebruiken om sociale groepen te bestuderen, omdat hij ervan overtuigd was dat bepaalde sociale groepen natuurlijke klassen waren. Dat wil zeggen Galton nam aan dat de leden van deze sociale groepen biologische of fysiologische kenmerken gemeen hebben. De groepen die hij met dit uitgangspunt bestudeerde waren Joden, criminelen en tuberculosepatiënten. Het statistische werk van Galton was gemotiveerd door eugenetische overwegingen, waarbij sommige typen als superieur werden beschouwd aan andere. Galton definieerde de door hem zelf bedachte term eugenetica als "de wetenschap die zich bezighoudt met alle

⁸ Klein 1997, p. 224.

⁹ Galton 1879, pp. 160-61.

invloeden die de aangeboren eigenschappen van een ras verbeteren.”¹⁰ Eugenetica was gericht op het verbeteren van de genetische kwaliteit van een menselijke populatie door mensen en groepen uit te sluiten die als inferieur werden beoordeeld of door diegenen die als superieur werden beschouwd te promoten.

Het idee dat een klokvormige clustering rond de normaal inzicht geeft over een natuurlijke klasse was gebaseerd op Quetelets idee van *l'homme moyen, de gemiddelde man*. Hij beschouwde de gemiddelde man als een weergave van de ware menselijke natuur.

Voor de Belgische astronoom Adolphe Quetelet waren statistische regelmatigheden tekenen van een diepere sociale realiteit, en de gemiddelde man was daar representatief voor. De gemiddelde man had, volgens hem, meer realiteit dan het diverse scala aan werkelijke individuen die een stad of een land bevolkten. Zo vertegenwoordigt de gemiddelde Brusselaar de typische Brusselaar en de gemiddelde inwoner van België de typische Belg.

Volgens deze eerste statistici heeft een statistisch gemiddelde van een bepaald kenmerk van een sociale groep alleen betekenis als dit kenmerk van de individuele leden van die groep zich clustert rond dat gemiddelde, en als deze clustering dan de vorm van een klok laat zien (zie figuur 1). Het verschijnen van een klokvormige verdeling laat zien dat deze eigenschap een *natuurlijke* eigenschap is en dat het gemiddelde het *type* weergeeft. Dit betekent dat telkens wanneer wordt aangenomen dat de gebruikte data de normale verdeling heeft – iets wat vaak impliciet wordt aangenomen om regressieanalyse te rechtvaardigen – deze aanname ontologische consequenties heeft: het uniformeert niet alleen de sociale groep, maar maakt ook van de betreffende sociale groep een natuurlijke klasse. Dat wil zeggen, de aanname van een normale verdeling *naturaliseert* de sociale groep die statistisch onderzocht wordt.

De gebruikelijke betekenis van naturalisatie is de juridische, namelijk het toekennen van een nationaliteit aan iemand. Hier is het relevant om te weten dat een van de voorwaarden voor het verkrijgen van de Nederlandse nationaliteit het inburgeringsexamen is. Dit examen omvat vragen met betrekking tot de basiskennis van de Nederlandse taal, de Nederlandse cultuur, zoals Sinterklaas, en typisch Nederlandse gewoontes en gebruiken, zoals handen schudden.¹¹ Ik neem aan dat dit de context is, waar de “doe-normaal-of-ga-weg”-brief naar verwees.

De tendens in de sociale wetenschappen, of meer in het bijzonder in de sociale statistiek, om sociale groepen als natuurlijke klassen te zien, was volgens de twintigste-eeuwse filosoof Karl Popper een combinatie van twee trends, *holisme* en *essentialisme*. Volgens het holisme is een sociale groep meer dan alleen de som van haar leden, en is ze ook meer dan alleen de som van de persoonlijke relaties tussen haar leden.¹²

¹⁰ Galton 1904, p. 1.

¹¹ <https://ind.nl/Nederlanderschap/naturalisatie/Paginas/Naturalisatie.aspx>. Laatst ingezien op 20 december 2021.

¹² Popper 1944, p. 91.

De tweede stroming, sterk gerelateerd aan het holisme, is het essentialisme. Een essentialistische benadering werkt als volgt: we verzamelen eerst een groep individuen en geven deze groep een label, bijvoorbeeld Nederlands. Vervolgens veronderstellen we dat elk afzonderlijk lid van die groep een bepaalde intrinsieke eigenschap heeft die alle leden gemeenschappelijk hebben, aangeduid met de universele term Nederlands-zijn. Deze intrinsieke eigenschap wordt dan beschouwd als iets dat onderzoek verdient. Dit gaat vaak samen met de bewering dat deze universele termen werkelijk bestaande objecten vertegenwoordigen, bijvoorbeeld dat het Nederlands-zijn echt bestaat boven de afzonderlijke individuen. En bovendien wordt zo'n universeel object, zoals Nederlands-zijn, gezien als de *essentie* van die groep.¹³

Popper bekritiseerde voornamelijk een specifieke combinatie van holisme en essentialisme, die hij "historisme" noemde, vanwege de bijzondere functie van de geschiedenis daarin. In het historisme wordt de historische methode gezien als het meest geschikt voor de analyse van de *aard* van sociale groepen. Historisme gaat ervan uit dat een sociale groep, zoals de Nederlanders, een eigen geschiedenis heeft en dat de aard van de Nederlander bepaald is door deze geschiedenis. Een groep wordt namelijk verondersteld zijn oorspronkelijke aard te behouden, zelfs als alle oorspronkelijke leden zijn overleden en zijn vervangen door nieuwe leden. Alle groepen hebben hun eigen tradities, bijvoorbeeld Sinterklaas, hun eigen instellingen, zoals de polder en hun eigen rituelen, zoals handen schudden. Het historisme stelt dat we de geschiedenis van een groep, haar tradities en instellingen moeten bestuderen, als we de essentie ervan willen begrijpen.

Poppers kritiek op het historisme was een belangrijk onderdeel van zijn indrukwekkende verdediging van de liberale democratie, *The Open Society and Its Enemies*. Volgens Popper zijn de "vijanden" van de "open samenleving" de filosofieën van Plato en Hegel. Deze zijn namelijk gebaseerd op het idee dat een staat een "essentie" heeft. Volgens deze essentialistische filosofieën moet iedereen die geacht wordt deze essentie aan te tasten, worden beschouwd als een staatsgevaarlijke vijand die wettelijk kan worden verdreven of zelfs vernietigd.

De reden dat ik historicisme als kernthema van mijn oratie heb gekozen, is dat we momenteel een comeback zien van historicisme in de huidige politieke en maatschappelijke debatten over nationale identiteit door deze identiteit te verankeren in een nationale *canon* of *traditie*. Hoe sterk iemands identiteit verweven is met deze nationale tradities, ook al zijn ze soms van een verrassend recente datum, blijkt uit de soms gewelddadige verdediging van deze tradities. Zoals het historisme zijn vijanden creëert, zo scheidt de roep om normaliteit monsters die men vervolgens wil verdrijven, zoals verwoord in het tweede deel van Ruttes oproep: "ga weg."

¹³ Popper 1944, p. 91.

3 Hier zijn monsters

Poppers bespreking van het historicisme was bedoeld als kritiek op bepaalde trends in de sociale wetenschappen in zijn tijd. Mijn inziens is deze kritiek net zo relevant als het gaat om de huidige toepassingen van de statistiek, vooral waar *machine learning* wordt gebruikt voor patroonherkenning of *profilering*, zoals patroonherkenning wordt genoemd in enkele van zijn toepassingen. In dit deel zal ik laten zien dat deze moderne vormen van historicisme *monsters* scheppen.

Omdat het statistische gemiddelde werd beschouwd als het *ware type* van een categorie, is de volgende vraag hoe we de afwijkingen van het gemiddelde moeten beschouwen. Quetelet had altijd het gemiddelde geïdealiseerd als het punt van deugd tussen ondeugdelijke uitersten.¹⁴ De gemiddelde man, bijvoorbeeld *de Belg*, was het ideale type in vergelijking met wie alle inwoners van België in gebreke zijn. Vergelijkbaar met de foutentheorie in de astronomie, was de gemiddelde man de ware positie van een ster (waar de fout nul is) en de afwijkingen van dat gemiddelde veroorzaakt door fouten. Volgens Quetelet zijn echte individuen onvolmaakte kopieën van de deugdzame gulden middenweg. Hij liet zien hoe we deze afwijkingen kunnen categoriseren met behulp van vaste afstanden van het gemiddelde. Deze afstanden noemde hij “limieten.”¹⁵ Voor de “lengte van de man” resulteerde dit in de volgende categorieën:

monsters | dwergen | kleine mannen | de gewone maat | grote mannen | reuzen | monsters

Monsters speelden geen rol in Poppers kritiek op het historicisme, maar de term speelde wel een belangrijke rol in een uiteenzetting over de ontwikkeling van de wiskunde door de wetenschapsfilosoof Imre Lakatos.¹⁶ Popper en Lakatos hadden een vergelijkbare filosofische en historische kijk op wetenschap en deelden dus een afkeer van essentialisme. In de wiskunde neemt het essentialisme de vorm aan van het platonisme. In het platonisme bestaan alle wiskundige concepten al in een wiskundige wereld die onafhankelijk is van onze kennis.¹⁷ Daarom zien platonisten de ontwikkeling van de wiskunde als de ontdekking van die reeds bestaande objectieve wiskundige wereld.

In de platonische wereld is er geen plaats voor monsters, ze worden gezien als bedreigingen van *harmonie en orde*. Ze worden beschouwd als *pathologische gevallen* die *walging* oproepen omdat ze *anarchie en chaos* propageren.¹⁸ Volgens de dominante platonische opvatting “bevorderen monsters nooit groei, noch in de wereld van de natuur, noch in de wereld van het denken,” en volgt evolutie altijd een harmonieus en geordend patroon.¹⁹

¹⁴ Gigerenzer e.a. 1989, p. 53.

¹⁵ Quetelet 1849, p. 102.

¹⁶ Lakatos 1976.

¹⁷ Brown 1999, p. 11.

¹⁸ Lakatos 1976, p. 19.

¹⁹ Lakatos 1976, p. 21.

De Franse wiskundige Henri Poincaré was waarschijnlijk de eerste die het concept van monsters gebruikte om de grondslagen crisis in de wiskunde aan het einde van de negentiende eeuw te bespreken. Deze monsters waren wiskundige objecten die botsten met de bestaande wiskundige theorie. Volgens Poincaré waren ze “vals” en “bizar.”²⁰ Hij vergeleek ze met voorwerpen van een teratologisch museum. Sinds het einde van de achttiende eeuw begonnen anatomen verzamelingen van misvormde menselijke en dierlijke foetussen te verzamelen. De vader van Poincaré was hoogleraar in de geneeskunde, dus we kunnen aannemen dat hij goed bekend was met dergelijke collecties. Maar het is ook relevant om te weten dat de term “teratologie” was bedacht door de schoonvader van Poincaré, de natuuronderzoeker Isidore Geoffroy Saint-Hilaire, om de studie van aangeboren misvormingen te benoemen.²¹

In tegenstelling tot deze platonische visie zag Lakatos de ontwikkeling van de wiskunde als een proces van het oplossen van problemen, *concept-stretching* en het adopteren van monsters. Lakatos liet zien dat in de geschiedenis van de wiskunde, naast de platonische strategieën om monsters te verbannen, ook nominalistische strategieën zoals *concept-stretching* kunnen worden onderscheiden. In tegenstelling tot essentialisme pleitten Popper en Lakatos voor nominalisme.

Een essentialist stelt vragen als: “wat is?,” bijvoorbeeld “wat is de aard van?” Een nominalist daarentegen stelt “hoe”-vragen. Volgens de nominalisten Popper en Lakatos is de taak van de wetenschap om het gedrag van verschijnselen te beschrijven en niet te vragen wat de aard ervan is; en als dat nodig is mogen voor die beschrijvingen nieuwe termen geïntroduceerd of oude termen opnieuw gedefinieerd worden. Geen enkele definitie is heilig. Daarom zag Lakatos *concept-stretching* als een meer vruchtbare strategie dan *monster-barring*.

Net zo als elk wiskundig systeem zijn misvormingen voortbrengt, dat wil zeggen vormen die niet in het systeem passen, zal elke andere wetenschappelijke ordening of categorisering buitenbeentjes creëren. De relevante vraag is: hoe gaan we met ze om? Zien we een monster als iets walgelijks of zelfs gevaarlijks, of als iets hoopgevends, iets dat nieuwe wegen opent en tot nieuwe ontdekkingen leidt, zoals Lakatos suggereerde? Het antwoord op deze vraag kan niet los worden gezien van de vraag of de bestaande orde als essentieel of zelfs goddelijk en heilig wordt beschouwd, of als een ordening die cultureel bepaald is. Een prominente pleitbezorger van deze laatste opvatting is de antropoloog Mary Douglas. Volgens haar opvatting “zijn anomalieën niet in de natuur geïnstalleerd, maar komen ze voort uit bepaalde kenmerken van classificatieschema’s.”²² In haar boek met de veelzeggende titel *Purity and Danger* bespreekt Douglas gevallen waarin mensen als “vuil” worden behandeld.²³

²⁰ Poincaré 1913, p. 435.

²¹ Aberdein 2019, p. 394.

²² Douglas 1996, p. 126.

²³ Douglas 2002.

Douglas laat zien dat vuil in wezen wanorde is en dat reiniging en zuivering onze omgeving opnieuw ordent. Maar volgens Douglas bestaat er niet zoiets als absoluut vuil, “het bestaat alleen in de ogen van de kijker.”²⁴ “Vuil is het bijproduct van een systematische ordening en classificatie van materie, voor zover ordening inhoudt dat ongepaste elementen worden afgewezen.”²⁵ Reiniging is de veroordeling van elk object of idee dat gekoesterde classificaties kan verwarren of tegenspreken; reiniging zou deze dissonantie moeten verminderen.²⁶

Door wanorde te vergelijken met vuil, kan Douglas verklaren waarom en wanneer wanorde als gevaarlijk wordt beschouwd. Vuil wordt namelijk als gevaarlijk gezien als het bestaat uit stukjes en beetjes die nog een zekere identiteit hebben; wanneer sommige van deze stukjes onthullen waar ze vandaan kwamen, bijvoorbeeld als ze restanten tonen van haar, voedsel of verpakkingsmateriaal. Als vuil deze halve identiteit heeft, wordt het als het gevaarlijkst gezien. Pas als na een proces van verpulveren, oplossen en rotten alle identiteit verdwenen is, wordt vuil niet meer als gevaarlijk beschouwd – denk maar aan de mest die je in je tuin gebruikt. Maar het is onaangenaam om in het afval rond te snuffelen om te proberen iets terug te vinden, want dit doet de identiteit herleven. Zolang identiteit ontbreekt, is afval niet gevaarlijk.²⁷

Douglas gebruikt zelf niet de term monster in haar antropologie van vuil. Maar als men kijkt naar de definities van monster, dan zijn die vergelijkbaar met haar beschrijvingen van vuil vooral wanneer vuil als het gevaarlijkst wordt beschouwd. Het lemma voor monster in de *Van Dale* bevat trefwoorden zoals gedrocht, wanstaltig, bovenmatig en buitensporig.²⁸ De vergelijking tussen monsters en vuil laat zien waarom het niet-normaal-zijn maakt dat mensen willen dat je weggaat. Monsters verzetten zich – net als vuil – tegen een bestaande orde, en worden daarom beschouwd als een gevaarlijke bedreiging. En als ze uit verschillende componenten zijn samengesteld, als ze twee of meer identiteiten hebben, worden ze als het gevaarlijkst beschouwd. Om dit te zien hoef ik alleen maar te verwijzen naar de *toeslagenaffaire*, waarin de mensen met twee nationale identiteiten het meest gewantrouwd werden.

Het meest iconische, maar fictieve monster dat de wetenschap heeft voortgebracht, is de creatie van Frankenstein. In haar beroemde boek *Frankenstein* laat Mary Shelley zien dat het niet het *wezen* van dit schepsel is dat het tot een monster maakt, maar zijn “ellendige misvorming.”²⁹ Ondanks het feit dat zijn karakter alle kenmerken van een Victoriaanse gentleman heeft, het is zachtaardig, nieuwsgierig en verfijnd, wordt het vanwege zijn halve identiteit als mens als een monster beschouwd. Alleen vanwege zijn uiterlijk wordt het als

²⁴ Douglas 2002, p. 2.

²⁵ Douglas 2002, p. 44.

²⁶ Douglas 2002, p. 340.

²⁷ Douglas 2002, pp. 197-8.

²⁸ Van Dale 1999.

²⁹ Shelley 2018.

wreed en gevaarlijk gezien. Pas nadat het jarenlang als een gevaarlijk monster behandeld was, begon het zich als zodanig te gedragen.

Moderne technieken van statistische inductie, zoals *machine learning*, kunnen de nieuwe Frankensteins zijn wanneer de resulterende profileringen iets menen te zeggen over wat *typisch* of *natuurlijk* is voor een sociale groep. Elke orde impliceert het bestaan van misfits, maar zodra deze ordeningen als natuurlijk worden gezien, worden deze vreemden als gevaarlijk gezien.

Elke vaststelling van wat normaal is, roept de vraag op hoe we de afwijkingen daarvan moeten beoordelen. Belangstelling voor deze afwijkingen van het normale – ook wel variatie genoemd – was cruciaal voor de ontwikkeling van de moderne statistiek. Deze geschiedenis van de statistiek kan niet los worden gezien van de opkomst van de eugenetica. De grondleggers van de moderne statistiek, Francis Galton, Karl Pearson en Ronald Fisher, hielden zich intensief bezig met de eugenetische controle van de menselijke evolutie. Deze gedeelde wortel van de moderne statistiek en eugenetica is welbekend in de wetenschapsgeschiedenis en daarover zijn uitstekende werken verschenen. Sinds de dagen van Galton tot onze tijd van *machine learning* wordt klontering van data in een klokvorm nog steeds zo opgevat dat we een aantal echte natuurlijke kenmerken van een sociale groep hebben gevonden. Dit is essentialisme in de statistiek. We gaan ervan uit dat de statistisch gevonden kenmerken *natuurlijk*, *typisch* of *generiek* zijn en zo een echte *essentie* onthullen. Hoe hoger de mate van clustering, des sterker dit geloof.

Monsters zijn de *punten* die op enige afstand van de centrale cluster verschijnen. Als het er maar een paar zijn, zijn ze ongevaarlijk. Ze worden dan gezien als *uitbijters* die genegeerd kunnen worden. Maar als het aantal groot is, worden ze wel gezien als een bedreiging voor de orde die bepaald wordt door het centrale cluster, de normaal. Elke overeenkomst met het normale, ook al is het maar gedeeltelijk, maakt ze gevaarlijk.

In plaats van het abnormale te verbieden en monsters als gevaarlijk te beschouwen, doe ik een oproep om open te staan voor monsters en ze te verwelkomen, zij zijn de mogelijkheden voor groei en ontwikkeling. Dit verwelkomen van monsters is mijn interpretatie van de kernwaarden van de Universiteit Utrecht: open blik, open houding en open wetenschap.

Auteur

Marcel Boumans (e-mail: m.j.boumans@uu.nl) is hoogleraar Geschiedenis van de Economische Wetenschap aan de Universiteit Utrecht.

Literatuur

- Aberdein, Andrew, 2019, Mathematical Monsters, in: Diego Compagna en Stefanie Steinhart (eds) *Monsters, Monstrosities, and the Monstrous in Culture and Society* (pp. 391-413), Wilmington, DE: Vernon Press.
- Brown, James Robert, 1999, *Philosophy of Mathematics: An Introduction to the World of Proofs and Pictures*, New York: Routledge.
- Daston, Lorraine, 2020, *Tegen de Natuur in*, Amsterdam: Octavo.
- Douglas, Mary, 1996, *Thought Styles. Critical Essays on Good Taste*, London: Sage.
- Douglas, Mary, 2002, *Purity and Danger. An Analysis of the Concept of Pollution and Taboo*. London and New York: Routledge.
- Galton, Francis, 1877, Typical laws of heredity, *Nature*, vol. 15(April 5): 492-5, 512-4, 532-3.
- Galton, Francis, 1879, Generic images, *Nineteenth Century*, vol. 6(July): 157-69.
- Galton, Francis, 1904, Eugenics: Its Definition, Scope, and Aims, *American Journal of Sociology*, vol. 10(1): 1-6.
- Gigerenzer, Gerd, Zeno Swijtink, Theodore Porter, Lorraine Daston, John Beatty en Lorenz Krüger, 1989, *The Empire of Chance. How Probability Changed Science and Everyday Life*, Cambridge: Cambridge University Press.
- Keynes, John Maynard, 1939, Professor Tinbergen's method, *The Economic Journal*, vol. 49(195): 558-68.
- Keynes, John Maynard, 1973a, A Treatise on Probability, in: D. Moggridge (ed.), *The Collected Writings of John Maynard Keynes, VIII*. London: MacMillan.
- Keynes, John Maynard, 1973b, The General Theory and After. Part II Defence and Development, in: D. Moggridge (ed.), *The Collected Writings of John Maynard Keynes, XIV*. London: Macmillan.
- Klein, Judy L., 1997, *Statistical Visions in Time. A History of Time Series Analysis 1662-1938*, Cambridge: Cambridge University Press.
- Lakatos, Imre, 1976, *Proofs and Refutations: The Logic of Mathematical Discovery*, Edited by John Worral and Elie Zahar. Cambridge, UK: Cambridge University Press.
- Poincaré, Henri, 1913, *The Foundations of Science*, Translated by G.H. Halstead. New York: The Science Press.
- Popper, Karl, 1944, The Poverty of Historicism, I., *Economica*, New Series 11 (42): 86-103.
- Popper, Karl, 1994, *The Open Society and Its Enemies*, London and New York: Routledge.
- Quetelet, L.A., 1849, *Letters Addressed to H.R.H. the Grand Duke of Saxe Coburg and Gotha, on the Theory of Probabilities, as Applied to the Moral and Political Sciences*, London: Charles and Edwin Layton.
- Shelley, Mary, 2018, *Frankenstein*, London, UK: VIVI Books.
- Van Dale, 1999, *Groot Woordenboek der Nederlandse Taal*, dertiende herziene uitgave. Utrecht en Antwerpen: Van Dale Lexicografie.
- Whewell, William, 1858, *Novum Organon Renovatum. Second Part of Philosophy of Inductive Sciences*, 3rd edition, London: Parker.