

## ARTIKELEN

# Het meten van effecten van de handhaving door de Belastingdienst\*

Sjoerd Goslinga, Maarten Siglé & Lisette van der Hel

De maatschappij verwacht in toenemende mate dat toezichthouders de effecten van hun toezicht inzichtelijk maken. Dat geldt ook voor de Belastingdienst. Dit artikel bespreekt de theorie en de praktijk van effectmeting in het fiscale domein en levert zo een bijdrage aan de ontwikkeling van effectmeting van het (overheids)toezicht. Vastgesteld wordt dat in de praktijk een aantal uitdagingen het hoofd moet worden geboden wil de Belastingdienst daadwerkelijk tot een integrale effectmeting van het toezicht komen. Dit zijn: het expliciteren van de beleidstheorie, het vinden van de juiste determinanten van compliance, het meten van effecten van preventieve activiteiten, het opzetten van methodologisch verantwoord onderzoek en de organisatorische inbedding van effectmeting.

## 1 Inleiding

Effectmeting is in de loop van de jaren steeds belangrijker geworden voor toezichthouders.<sup>1</sup> Het *Tijdschrift voor Toezicht* wijdde er in 2012 een themanummer aan, waarin bijvoorbeeld effectmeting in het toezicht op de financiële sector aan bod kwam.<sup>2</sup> De Wetenschappelijke Raad voor het Regeringsbeleid (WRR) geeft in 2013 aan dat toezichthouders de maatschappelijke opbrengsten van toezicht centraler moeten stellen in hun beleid, wat impliceert dat hij toezichthouders aanspoort om te kijken naar de effecten van het toezicht.<sup>3</sup>

Het belang van effectmeting komt bijvoorbeeld naar voren bij de introductie van nieuwe handhavingsstrategieën. Zo is bij belastingdiensten over de hele wereld een ontwikkeling zichtbaar van het hanteren van een vrijwel uitsluitend repressieve handhavingsfilosofie naar inbedding van een meer 'horizontale' handhavingsfilosofie (*cooperative compliance*).<sup>4</sup> Uit een enquête van de OESO (2013) blijkt echter dat nog maar weinig belastingdiensten de effectiviteit van deze voor-

\* De auteurs hebben het artikel op persoonlijke titel geschreven.

- 1 P. Welp, M. Bokhorst & P. de Goede, Effecten en evaluatie van toezicht, in: F.J.H. Mertens, E.R. Muller & H.B. Winter (red.), *Toezicht: Inspecties en autoriteiten in Nederland*, Deventer: Wolters Kluwer 2015 p. 87-108.
- 2 P.L.C. Hilbers, D.R. Rijsbergen, K. Raaijmakers & F. de Vries, Effectmeting in het toezicht op de financiële sector, *Tijdschrift voor Toezicht* 2012, nr. 3, p. 9-21.
- 3 Wetenschappelijke Raad voor het Regeringsbeleid, *Toeziën op publieke belangen*, Amsterdam: Amsterdam University Press 2013.
- 4 E.C.J.M. van der Hel-van Dijk, *De ontwikkeling van horizontaal toezicht wereldwijd*, Vakblad Tax Assurance 2013, nr. 1, p. 20.

namelijk preventieve inspanningen meten.<sup>5</sup> Dat geldt ook voor Nederland, zo blijkt uit het rapport van de Commissie-Stevens (2012) – bij de evaluatie van de Nederlandse equivalent horizontaal toezicht: ‘Hier wreekt zich ook dat (...) geen nulmetingen zijn gedaan, nauwelijks prestatie-indicatoren en meetmechanismen zijn ontwikkeld. Het is dan ook onduidelijk waaraan de Belastingdienst zelf het al dan niet succesvol zijn van het horizontaal toezicht afmeet.’<sup>6</sup> De WRR onderschrijft dat beeld voor het gehele rijkstoezicht en stelt dat inzicht in de effectiviteit van zowel klassieke (repressieve) als innovatieve (preventieve) vormen van toezicht beperkt is.<sup>7</sup>

De Belastingdienst opereert in een ingewikkeld en dynamisch krachtenveld, waarin belastingplichtigen (burgers en bedrijven), de politiek, de maatschappij, andere toezichthouders en overige stakeholders hun eigen verwachtingen en wensen hebben over handhaving en toezicht. Hoewel effectmeting integraal onderdeel is van de handhavingsaanpak van de Belastingdienst, valt te concluderen dat de Belastingdienst nog nauwelijks inzicht geeft in de effecten van zijn handhaving.<sup>8</sup> Dat is een beeld dat overigens ook te zien is bij andere toezichthouders. Centraal in de handhavingsstrategie van de Belastingdienst staat het beïnvloeden van het nalevingsgedrag van burgers en bedrijven (compliance). Om effect te kunnen hebben zullen de inspanningen of activiteiten van de Belastingdienst in moeten grijpen op de determinanten van (non-)compliance. Veelal verloopt het bereiken van compliance dus via tussendoelen. Om effect te kunnen meten moeten de (tussen)doelen in beeld zijn en moet het duidelijk zijn waar de inzet van handhavinginstrumenten op is gericht.

Dit artikel ontrafelt dit complexe vraagstuk aan de hand van beschikbare beleidsstukken, wetenschappelijke literatuur, vakliteratuur en de praktijkervaring van de auteurs en geeft zo inzicht in hoe effectmeting in het fiscale domein eruit kan zien. Het artikel benoemt de uitdagingen die met effectmeting zijn gemoeid, om zo een bijdrage te leveren aan de ontwikkeling van effectmeting in het fiscale domein en in andere domeinen waar gedragsbeïnvloeding een prominente plek heeft in de handhavingsstrategie. Gezien de toenemende aandacht voor gedragsbeïnvloeding bij toezichthouders is de fiscale context ook relevant voor onderzoekers die onderzoek doen bij en/of naar deze andere toezichthouders.

Centraal in dit (beschrijvende en verkennende) artikel staat de vraag hoe de Belastingdienst de effecten van zijn handhavings- en toezichtactiviteiten meet en wat de uitdagingen hierbij zijn. Om deze vraag te beantwoorden is literatuuronderzoek uitgevoerd. Bijgebracht zijn (onder andere) openbare documenten over missie, visie, beleid en uitvoering van de Belastingdienst, relevante vakliteratuur, wetenschappelijke literatuur over effectmeting bij of door Nederlandse toezichthouders en belastingdiensten in het buitenland en wetenschappelijke litera-

5 OESO, *Co-operative Compliance: A Framework*, Parijs: OECD Publishing 2013.

6 Commissie Horizontaal Toezicht Belastingdienst, *Fiscaal toezicht op maat*. Soepel waar het kan, streng waar het moet, Rijswijk: VijfKeerBlauw 2012, p. 121.

7 Wetenschappelijke Raad voor het Regeringsbeleid, *Toeziën op publieke belangen*, Amsterdam: Amsterdam University Press 2013.

8 Rijksbegroting 2016, Hoofdstuk IX, artikel 1 Belastingen.

Sjoerd Goslinga, Maarten Siglé & Lisette van der Hel

tuur over determinanten van naleving van belastingregels. Het bijeengebrachte materiaal is geordend, samengevat en geanalyseerd.

Het artikel is als volgt opgebouwd: paragraaf 2 gaat in op het handhavingsbeleid van de Belastingdienst en beschrijft verder de relatie tussen de doelstellingen van de Belastingdienst, zijn activiteiten en de beoogde effecten. De relevante beleidsdocumenten vormen daarbij de basis voor een reconstructie van de beleidstheorie en een invulling van de effectketen. Paragraaf 3 gaat dieper in op het daadwerkelijk meten van de effecten van interventies van de Belastingdienst. Hierbij komen de aannames in het fiscale toezicht aan bod, inclusief de literatuur waarop de aannames zijn gebaseerd, en de methoden die de Belastingdienst gebruikt om tot uitvoering van effectmeting te komen. Paragraaf 4 verkent in hoeverre enkele in de literatuur genoemde aandachtspunten bij het implementeren van effectmetingen relevant zijn in het fiscale domein. Op basis van de bevindingen van de eerdere paragrafen sluit paragraaf 5 het artikel af met enkele conclusies en een vijftal uitdagingen gekoppeld aan het meten van effecten van fiscale handhaving.

## 2 Fiscaal beleid en effecten

Deze paragraaf gaat in op het handhavingsbeleid van de Belastingdienst en beschrijft de relatie tussen de doelstellingen van de Belastingdienst, zijn activiteiten en de beoogde effecten. Zonder een beoogd doel is het onmogelijk om effecten van handhavingsactiviteiten te meten. Winter en De Ridder formuleren dat als volgt: 'In algemene zin kan "effectiviteit van toezicht" omschreven worden als de mate waarin de uitoefening van toezicht leidt tot verwezenlijking van de doeleinden van dat toezicht.'<sup>9</sup> Dit roept de vraag op wat de 'doeleinden van toezicht' zijn in het fiscale domein.

### 2.1 Algemene beleidsdoelstelling en handhavingsregie

De Rijksbegroting 2016 verwoordt de algemene beleidsdoelstelling van de Belastingdienst als volgt: '(...) Doeltreffende en doelmatige uitvoering van die wet- en regelgeving zorgt ervoor dat burgers en bedrijven bereid zijn hun wettelijke verplichtingen ten aanzien van de Belastingdienst na te komen (compliance).'<sup>10</sup> De term 'bereid' geeft aan dat de Belastingdienst ernaar streeft dat belastingplichtigen uit zichzelf fiscale regels naleven, zonder (dwingende en kostbare) acties van de kant van de Belastingdienst. Als belastingplichtigen hun wettelijke fiscale verplichtingen nakomen, dan komt belastinggeld de staatskas binnen zoals de wetgever beoogt.<sup>11</sup> De Rijksbegroting benoemt ook de beleidsinstrumenten – dienstverlening, handhaving en massale processen – die de Belastingdienst inzet om de

9 H.B. Winter & J. De Ridder, Toezicht en effectmeting: Het kan!, Groningen: Pro Facto 2010.

10 Rijksbegroting 2016, Hoofdstuk IX.

11 De Rijksbegroting definieert 'compliance' als de bereidheid van burgers en bedrijven om hun wettelijke verplichtingen ten aanzien van de Belastingdienst na te komen. Deze definitie van compliance wijkt af van de definitie die gangbaar is in de wetenschappelijke literatuur over naleving van belastingregels. Daarin staat compliance voor gedrag dat conform de regels of normen is, ongeacht of dat uit eigen beweging of doelbewust zo is gekomen.

compliance doelstelling te realiseren: 'De Minister bevordert (...) compliance door passende dienstverlening te leveren, massale processen juist en tijdig uit te voeren, adequaat toezicht uit te oefenen en waar nodig naleving bestuurs- of strafrechtelijk af te dwingen.'<sup>12</sup>

De Belastingdienst hanteert de term handhaving in een 'brede' betekenis als het gaat om *handhavingsregie*.<sup>13</sup> Handhavingsregie betekent dat de Belastingdienst op een systematische en gestructureerde manier naar (groepen van) belastingplichtigen kijkt en op basis van de oorzaken van het gedrag van belastingplichtigen keuzes maakt voor wat betreft het inzetten van handhavingsinstrumenten om zijn doel (compliance) te bereiken.<sup>14</sup> Met effectmeting vindt de beoordeling plaats of het uitvoeren van de handhavingsactiviteiten daadwerkelijk de determinanten van compliance (zie paragraaf 3.1) heeft beïnvloed en of dit vervolgens effect heeft gehad op de *compliance* (zie voorbeeld 1).

• *Voorbeeld 1*<sup>15</sup>

*De Belastingdienst voert jaarlijks in maart een aangiftecampagne in verband met het tijdig (voor 1 april) doen van de aangifte inkomstenbelasting. Het doel van deze campagne is om de compliance (in de vorm van tijdig aangifte doen) van burgers te verhogen door middel van voorlichting.*

*Aan een dergelijke aangiftecampagne ligt de aanname ten grondslag dat voorlichting de kennis bij belastingplichtigen verhoogt. De tweede aanname is dat deze verhoging van kennis leidt tot een verbetering van de compliance. Door effectmeting kan worden getoetst of deze aannames kloppen. In de praktijk vindt het toetsen van het effect van dergelijke communicatie-uitingen niet structureel plaats.*

## 2.2 Beleidstheorie

Het geheel aan beleidsdoelen en instrumenten en de veronderstellingen en aannames over hoe deze samenhangen wordt aangeduid als een *beleidstheorie*. Een beleidstheorie geeft aan *hoe* activiteiten (de inzet van mensen en middelen) bijdragen aan het realiseren van de doelstellingen. In de praktijk moet een beleidstheorie gereconstrueerd worden op basis van verschillende beleidsdocumenten. Hiervoor kan een zogenoemde *doelenboom* worden gebruikt.<sup>16</sup> Voor de Belasting-

12 Rijksbegroting 2016, Hoofdstuk IX.

13 Handhaving omvat in deze betekenis zowel preventieve maatregelen, zoals informatievoorziening of ondersteuning, als repressieve acties, zoals een controle of een sanctie nadat een overtreding is geconstateerd.

14 Handhavingsregie valt te zien als een *compliancestrategie* die in belangrijke mate is gebaseerd op gedragswetenschappelijke inzichten, vooral economische en (sociaal)psychologische.

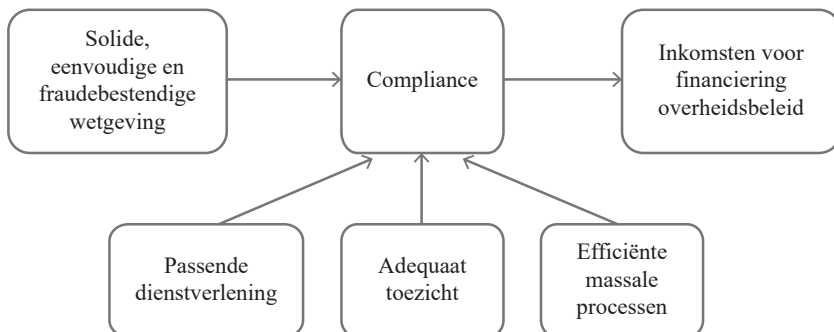
15 E.C.J.M. van der Hel-van Dijk & M.A. Siglé, Effectmeting van horizontaal toezicht bij grote ondernemingen, Vakblad Tax Assurance 2013, nr. 1, p. 63.

16 In de Engelstalige literatuur wordt gesproken van 'logic models' en van 'theory of change', zie bijvoorbeeld J.A. McLaughlin & G.B. Jordan, Logic models: a tool for telling your program's performance story, Evaluation and Program Planning 1999, 22, p. 65-72; D.H. Taplin & H. Clark, Theory of change basics, a primer on theory of change, New York: Actknowldge 2012; C.L.S. Coryn, L.A. Noakes, C.D. Westine & D.C. Schröter, A Systematic review of theory-driven evaluation practice from 1990 to 2009, American Journal of Evaluation 2011, 32, p. 199-226.

Sjoerd Goslinga, Maarten Siglé & Lisette van der Hel

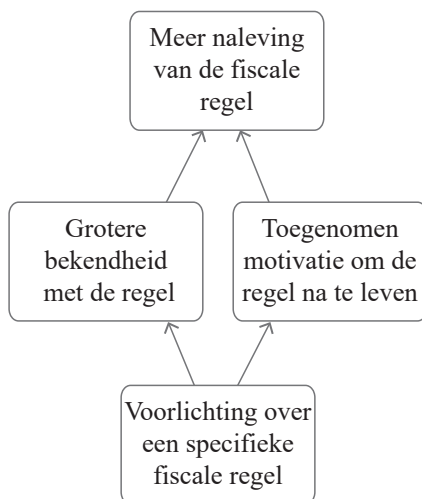
dienst als geheel (strategisch niveau) zou de doelenboom versimpeld, dat wil zeggen zonder tussendoelen, als volgt kunnen worden weergegeven.<sup>17</sup>

**Figuur 1** Voorbeeld doelenboom op algemeen niveau



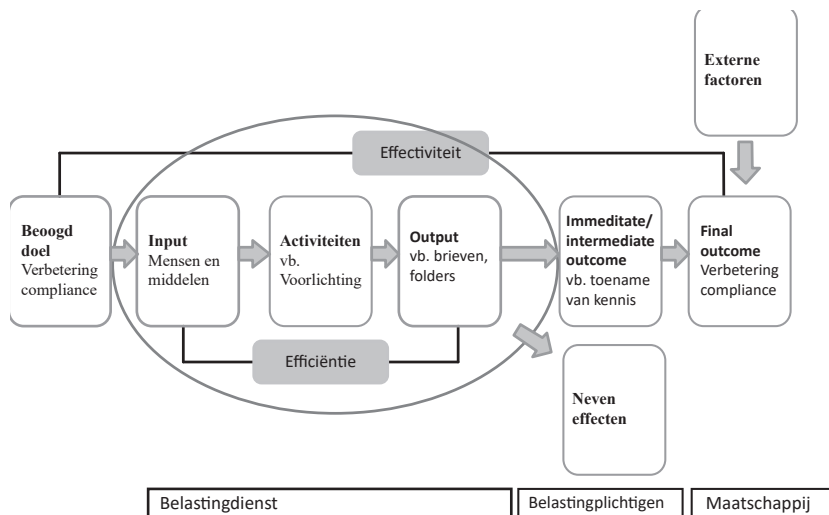
Het is van belang om vast te stellen of de Belastingdienst zijn beoogde doelen – zowel op strategisch, tactisch als operationeel niveau – ook bereikt. Die vaststelling gebeurt in principe door het meten van effecten. Ook op *operationeel niveau* kan voor het beoogde effect van een individueel handhavingsinstrument een doelenboom worden opgesteld. Voor bijvoorbeeld voorlichting aan een bepaalde groep belastingplichtigen over een specifieke verplichting of regel zou een doelenboom er zo uit kunnen zien:

**Figuur 2** Voorbeeld doelenboom op operationeel niveau



Als een doelenboom is opgesteld, is ook bekend welke tussendoelen eerst moeten worden bereikt en kan worden gemeten of de tussendoelen zijn gerealiseerd. Na

<sup>17</sup> Rijksbegroting 2016, Hoofdstuk IX.

**Figuur 3** Effectketen

een voorlichtingssessie kan de Belastingdienst bijvoorbeeld vaststellen of de bekendheid met en kennis over een bepaalde fiscale regel inderdaad is toegenomen, bijvoorbeeld door deelnemers te enquêteren. Als de voorlichting er niet in is geslaagd om de kennis bij de deelnemers te vergroten, lijkt het ook niet realistisch om te verwachten dat belastingplichtigen de aangeboden kennis toepassen en daardoor de compliance van de voorgelichte groep belastingplichtigen toeneemt.

### 2.3 Van beleidstheorie naar effectketen

Om een uitspraak te kunnen doen over effecten is het niet alleen nodig om het nalevingsniveau te weten, maar ook om te bepalen in hoeverre inspanningen van de Belastingdienst bepalend zijn (geweest) voor dat nalevingsniveau (causaal verband). Voor een effectmeting in het fiscale domein moeten derhalve beoogd doel, inzet van mensen en middelen, uitgevoerde activiteiten, concrete resultaten en de veranderingen in gedrag van belastingplichtigen met elkaar in verband worden gebracht. Dat kan worden geïllustreerd aan de hand van een zogenaamde *effectketen* (Figuur 3).<sup>18</sup>

‘Outcome’ representeert in deze effectketen de uiteindelijke impact van de activiteiten van de Belastingdienst op zijn strategische doel: compliance.<sup>19</sup> ‘Outcome’ heeft daarmee betrekking op (gedrag van) belastingplichtigen en kan worden

18 Afhankelijk van het niveau (strategisch, tactisch of operationeel) waarop de effectketen wordt uitgewerkt, zullen de verschillende elementen anders zijn ingevuld.

19 De term ‘final outcome’ of impact wordt gebruikt om op een hoger, strategisch niveau het effect op de *gehele samenleving* (i.c. borging van de continuïteit van belastingopbrengsten) aan te duiden.

Sjoerd Goslinga, Maarten Siglé & Lisette van der Hel

onderscheiden in ‘immediate outcome’, ‘intermediate outcome’ en ‘final outcome’. Dit onderscheid geeft aan dat het bereiken van het beoogde effect veelal in etappes of via een aantal tussendoelen verloopt. ‘Output’ daarentegen is datgene wat door de inspanningen van de Belastingdienst is geproduceerd (zoals het aantal verstuurd brieven of uitgevoerde controles (Figuur 3 – rode cirkel). Hierop – en op de inzet van mensen en middelen en de uit te voeren activiteiten – kan de Belastingdienst relatief gemakkelijk sturen. De veronderstelling is daarbij dat de (resultaten van) inspanningen – in fasen – een verandering teweeg brengen in houding en/of gedrag van belastingplichtigen. Inspanningen kunnen ook leiden tot niet-beoogde effecten, ook wel neveneffecten genoemd.

Om vast te stellen in hoeverre inspanningen van de Belastingdienst bepalend zijn (geweest) voor dat nalevingsniveau is het noodzakelijk om een vergelijking te maken met het nalevingsniveau in een situatie waarin de inspanningen niet zouden zijn geleverd (*counterfactual*).<sup>20</sup> Er kunnen ook andere (externe) factoren debet zijn aan een verbeterd nalevingsniveau, bijvoorbeeld verbeterde economische omstandigheden voor belastingplichtigen. Op het hoogste niveau, waar het gaat om de effecten van het algehele beleid en het totaal aan inspanningen van de Belastingdienst, is het vrijwel onmogelijk om een *counterfactual* te vinden. In wezen zou een vergelijking gemaakt moeten worden met de fiscale compliance in Nederland zonder inspanningen van de Belastingdienst. Op dat niveau kan alleen aannemelijk worden gemaakt dat het beleid zijn vruchten afwerpt, maar bewezen worden kan dat niet. Op tactisch niveau en uitvoeringsniveau is dat wel te realiseren.

#### 2.4 Ten slotte

Uit deze paragraaf blijkt dat heldere doelen en inzicht in de samenhang tussen beleidsdoelen en handhavingsinstrumenten noodzakelijk zijn voor het uitvoeren van een effectmeting. In het fiscale domein is de bedoelde samenhang vaak logisch en beredeneerd, maar veelal niet in een document geëxpliciteerd en beschreven. Voor het uitvoeren van effectmetingen zal de Belastingdienst derhalve de beleidstheorie moeten reconstrueren. Het strategisch doel van de Belastingdienst is ‘compliance’; het bereiken van dat doel verloopt in de praktijk veelal via allerlei tussendoelen. Hier gaat de volgende paragraaf dieper op in.

### 3 Uitvoeren effectmeting fiscaal toezicht

In de praktijk geeft de Belastingdienst vorm en inhoud aan de *handhavingsregie* door zijn interventies zo goed mogelijk af te stemmen op de oorzaken van (non-)compliance. Hij maakt hierbij gebruik van een grote hoeveelheid (wetenschappelijke) literatuur waarin onderzoek is gedaan naar deze oorzaken en de effecten van handhaving/toezicht op het gedrag van individuele belastingplichtigen en van bedrijven.

20 S.L. Morgan & C. Winship, *Counterfactuals and causal inference: methods and principles for social research*, Cambridge: Cambridge University Press 2007.



Om effect te kunnen hebben zullen de inspanningen of activiteiten van de Belastingdienst in moeten grijpen op de achterliggende oorzaken (determinanten) van (non-)compliance. In termen van de effectketen gaat het hier om de immediate en intermediate outcome-variabelen. Die variabelen zijn gericht op de tussendoelen van de Belastingdienst. In deze paragraaf gaan we eerst in op de determinanten van (non-)compliance. Vervolgens behandelt deze paragraaf de methoden die de Belastingdienst gebruikt om tot uitvoering van effectmeting te komen en komt het daadwerkelijk ontwikkelen van meetindicatoren aan bod.

### 3.1 Determinanten van compliance

Deze paragraaf bespreekt de determinanten van compliance van individuele belastingplichtigen en van ondernemingen. De determinanten van compliance van *individuele belastingplichtigen* zijn in drie categorieën in te delen: economische, sociaalpsychologische en economisch-psychologische determinanten. Aangezien *ondernemingen* een significant deel van de Nederlandse belastinginkomsten genereren, is een belangrijke vraag of de besproken determinanten van compliance van individuele belastingplichtigen ook bepalend zijn voor de compliance van ondernemingen of dat bij ondernemingen (ook) andere factoren een rol spelen. Hier blijkt echter relatief weinig onderzoek naar gedaan. De economische determinanten hebben daarbij de meeste aandacht gekregen. De economische determinanten in de context van (grote) ondernemingen lijken niet of weinig af te wijken van die in de context van individuen.<sup>21</sup> Onderzoek naar de sociaal- en economisch-psychologische determinanten bij ondernemingen is zeer beperkt. De weinige onderzoeken die zijn uitgevoerd lijken te bevestigen dat ook deze determinanten bij ondernemingen niet of weinig of afwijken van die bij individuen.<sup>22</sup> Een mogelijke verklaring kan wellicht gevonden worden in het feit dat ondernemingen – in niet-juridische zin – zelf geen besluiten nemen; dit doen de managers van die ondernemingen.<sup>23</sup> Deze managers zijn individuen met hun eigen perceptie en houding ten aanzien van het betalen van belastingen. Voor ‘kleine’ ondernemingen, waar bijvoorbeeld een directeur-groootaandeelhouder een bepalende invloed heeft, is de verwachting gerechtvaardigd dat de determinanten van compliance van deze ondernemingen niet of weinig afwijken van de determinanten van compliance van individuen.<sup>24</sup> Ten aanzien van ‘grote’ ondernemingen zou dat anders kunnen zijn, vanwege relevante verschillen met ‘kleine’ ondernemingen,

- 21 Zie o.a. J. Alm, C. Blackwell & M. McKee, Audit selection and firm compliance with a broad based sales tax, *National Tax Journal* 2004, 57, p. 209-27; J. DeBacker, B.T. Heim, A. Tran & A. Yuskavage, Legal enforcement and corporate behavior: An analysis of corporate tax aggressiveness after an audit, Working paper 2013. <<http://ssrn.com/abstract=2262586>>.
- 22 Zie o.a. B. Ariel, Deterrence and moral persuasion effect on corporate tax compliance: findings from a randomized controlled trial, *Criminology* 2011, 50, p. 27-69; S.D. Dyreng, W.J. Mayew & C.D. Williams, Religious social norms and corporate financial reporting, *Journal of Business Finance and Accounting* 2012, 39, p. 845-875.
- 23 D. Joulfaian, Corporate income tax evasion and managerial preferences, *Review of Economics and Statistics* 2000, 82, p. 698-701.
- 24 Alhoewel wel sprake kan zijn van verschillende determinanten voor verschillende belastingmiddelen. Naar de determinanten van compliance voor niet-winstbelastingen is weinig onderzoek gedaan.



zoals de *scheiding van eigendom en leiding* en specifieke belangen, zoals *aandelenkoersen* en *imago*.<sup>25</sup> We bespreken derhalve achtereenvolgens de economische, de sociaalpsychologische en de economisch-psychologische determinanten van compliance, aangevuld met determinanten die specifiek voor (grote) ondernemingen gelden.

• *Economische determinanten*

De basis voor veel wetenschappelijk onderzoek naar de determinanten van compliance is het economische model van criminaliteit van Becker en de toepassing hiervan op de fiscaliteit door Allingham en Sandmo.<sup>26,27</sup> Dergelijke modellen zien belastingbetalers als rationele beslissers die de keuze om compliant te zijn baseren op economische berekeningen.<sup>28</sup> De potentiële opbrengsten van non-compliance worden afgezet tegen de potentiële kosten. Indien de opbrengsten de kosten overtreffen, zo veronderstellen deze modellen, besluiten belastingplichtigen om niet compliant te zijn. Hierdoor spelen *de hoogte van het belastingtarief* (van belang bij het bepalen van de potentiële opbrengsten), *de pakkans* en *de hoogte van eventuele sancties* (beide van belang bij het bepalen van de potentiële kosten) een centrale rol.

Dat de aannames dat belastingbetalers zich enkel laten leiden door economische afwegingen of in staat worden geacht om de potentiële opbrengsten en kosten van hun gedrag te bepalen niet houdbaar zijn, blijkt uit het feit dat het werkelijk niveau van compliance – in veel landen – ver boven de voorspellingen van de (neoklassieke) economische modellen ligt.<sup>29</sup> Verschillende factoren zijn daarom aan de economische modellen toegevoegd om het gat tussen de voorspellingen en de praktijk te verklaren. Voorbeelden hiervan zijn het gebruik van *een belastingadviseur*, de *daadwerkelijke ervaring van een belastingcontrole*, de *complexiteit van het belastingsysteem* (dan wel *kennis van het belastingsysteem*) en de *mogelijkheden om non-compliant te zijn*.<sup>30</sup> Alhoewel deze uitgebreide economische modellen het werkelijke niveau van compliance beter voorspellen, blijft de verklarende waarde hiervan beperkt.<sup>31</sup>

- 25 B. Ariel, Deterrence and moral persuasion effect on corporate tax compliance: findings from a randomized controlled trial, *Criminology* 2011, 50, p. 27-69.
- 26 G.S. Becker, Crime and punishment: An economic approach, *Journal of Political Economy* 1968, 76, p. 169-217.
- 27 M.G. Allingham & A. Sandmo, Income tax evasion: A theoretical analysis, *Journal of Public Economics* 1972, 1, p. 323-338.
- 28 J. Alm, A Perspective on the experimental analysis of taxpayer reporting, *Accounting Review* 1991, 66, p. 577-593.
- 29 L.P. Feld & B.S. Frey, Tax compliance as the result of a psychological tax contract: The role of incentives and responsive regulation, *Law & Policy* 2007, 29, p. 102.
- 30 Overigens hangen de twee laatstgenoemde sterk samen. Het dichten van mogelijkheden voor non-compliant gedrag kan leiden tot een hogere complexiteit en een hogere mate van complexiteit kan leiden tot meer mogelijkheden om non-compliant te zijn.
- 31 C.M. Fischer, M. Wartick & M. Mark, Detection probability and taxpayer compliance: A review of the literature, *Journal of Accounting Literature* 1992, 11, p. 1-46.

- *Sociaalpsychologische determinanten*

Vanuit de sociale psychologie zijn verschillende andere determinanten van fiscale compliance aangedragen die voortkomen uit de sociale dynamiek rondom de compliancebeslissingen van belastingplichtigen.

Iedere belastingplichtige heeft op dagelijkse basis op verschillende manieren te maken met de overheid (interactie overheid – belastingplichtigen). Op basis van deze ervaringen ontstaat een bepaalde *mate van vertrouwen in de overheid*. De mate waarin dit vertrouwen aanwezig is bij een belastingplichtige kan de mate van compliance beïnvloeden.<sup>32</sup> De overheid bepaalt tevens de hoogte van de te betalen belastingen en daarmee de belastingdruk. De ervaren rechtvaardigheid van deze belastingdruk (de zogenoemde *distributieve rechtvaardigheid*) is onder andere afhankelijk van de belastingdruk van andere belastingplichtigen en beïnvloedt de mate van compliance.<sup>33</sup> Bijvoorbeeld, wanneer een belastingplichtige het idee heeft dat andere belastingplichtigen, die zich in een vergelijkbare situatie bevinden, anders (gunstiger) worden behandeld, verlaagt dit de compliance.<sup>34</sup> Naast een gelijke behandeling van *gelijke* gevallen is ook de behandeling van *ongelijke* gevallen relevant. Veel belastingplichtigen zijn van mening dat sterkere schouders zwaardere lasten moeten dragen. De facto betekent dit dat de perceptie van belastingplichtigen van de mate waarin ‘de rijken’ hun *fair share* betalen de compliance kan beïnvloeden.<sup>35</sup>

De meeste belastingplichtigen hebben met een zekere regelmaat rechtstreeks contact met de Belastingdienst (interactie Belastingdienst – belastingplichtigen). De ervaren rechtvaardigheid van de werkwijze van de Belastingdienst (de zogenoemde *procedurele rechtvaardigheid*) heeft tevens impact op compliance. Belangrijke aspecten hiervan zijn betrouwbaarheid, objectiviteit en respect.<sup>36</sup> Praktisch gezien betekent dit, bijvoorbeeld, dat de wijze waarop een belastingplichtige de contacten met de Belastingdienst ervaart van invloed is op de mate van compliance.<sup>37</sup>

Mensen laten zich in hun gedrag (sterk) beïnvloeden door wat sociaal geaccepteerd of wenselijk is (interactie tussen belastingplichtigen); dit zijn *sociale nor-*

32 C. Kogler e.a., Trust and power as determinants of tax compliance: Testing the assumptions of the slippery slope framework in Austria, Hungary, Romania and Russia, *Journal of Economic Psychology* 2013, 34, p. 169-180. Als voorbeeld hiervan: de Amerikaanse revolutie kwam deels voort uit het verlies in vertrouwen in de Britse overheid en de daarop volgende bezwaren tegen belastingheffing (‘no taxation without representation’).

33 De wijze waarop de betaalde belastingen worden toegedeeld zijn ook van invloed op de ervaren distributieve rechtvaardigheid. Een voorbeeld hiervan is de recente weigering van auteur Leon de Winter om een deel van zijn belastingschuld, gelijk aan het deel van het overheidsbudget dat werd besteed aan hulp aan de voormalige Egyptische president Morsi, te betalen.

34 C. Bazart & A. Bonein, Reciprocal relationships in tax compliance decisions, *Journal of Economic Psychology* 2014, 40, p. 83-102.

35 J. Braithwaite, A good century for tax? Globalisation, redistribution and tax avoidance, *Public Policy Research* 2005, 12, p. 85-92.

36 M. Wenzel, The impact of outcome orientation and justice concerns on tax compliance: The role of taxpayers’ identity, *Journal of Applied Psychology* 2002, 87, p. 629-645.

37 K. Gangl, S. Muehlbacher, M. de Groot, S. Goslinga, E. Hofmann, C. Kogler, G. Antonides & E. Kirchler, “How can I help you?” Perceived service orientation of tax authorities and tax compliance, *Finanz Archiv/Public Finance Analysis* 2013, 69, p. 487-510.

men.<sup>38</sup> Belastingplichtigen worden beïnvloed door de sociale normen van de groepen waarvan zij deel uitmaken.<sup>39</sup> Dit kunnen meerdere groepen tegelijk zijn, zoals familie, collega's, geloofsgenoten, mensen met dezelfde etnische achtergrond, enzovoort. De mate waarin een belastingplichtige zich identificeert met een bepaalde groep, bepaalt de mate waarin de compliance wordt beïnvloed door de sociale normen van die groep.<sup>40</sup> Ook iemands *persoonlijke normen* zijn een belangrijke sociaalpsychologische determinant van compliance. Persoonlijke normen zijn ieders eigen standaarden ten aanzien van wat passend gedrag is.<sup>41</sup> Het effect van persoonlijke normen houdt verband met het effect van sociale normen. Sociale normen kunnen namelijk geleidelijk worden geïnternaliseerd en daarmee worden het persoonlijke normen.<sup>42</sup> Anderzijds kunnen belastingplichtigen non-compliant gedrag rationaliseren door selectieve interpretatie van sociale normen.<sup>43</sup> De interactie tussen sociale en persoonlijke normen is zodoende complex. Wat wel geconcludeerd kan worden is dat persoonlijke normen de mate van compliance beïnvloeden.<sup>44</sup>

• *Economisch-psychologische determinanten*

De aanname uit de klassieke economische modellen dat mensen volledig geïnformeerde wezens zijn die in staat zijn om te 'berekenen' wat voor hen het meest profijtelijke gedrag is, is niet houdbaar gebleken. Mensen zijn slechts beperkt in staat informatie te verwerken en maken gebruik van allerlei simpele redeneringen en vuistregels om een beslissing te nemen. Dit worden *heuristieken en biases* genoemd.<sup>45</sup> Deze kunnen van invloed zijn op de compliancebeslissingen van belastingplichtigen.<sup>46</sup> Het gaat echter te ver om deze hier alle te behandelen. Belangrijk in dit licht is ook dat in de praktijk zelden alle informatie beschikbaar is. Zo is informatie omtrent de pakkans, sancties en het effectieve belastingtarief bij veel mensen onbekend. Deze *onzekerheid* beïnvloedt de compliance, al is nog

38 M. Wenzel, An analysis of norm processes in tax compliance, *Journal of economic psychology* 2004, 25, p. 213-228.

39 Schaamte en schuld spelen hier een belangrijke rol in.

40 J.S. Ashby, P. Webley & A.S. Haslam, The role of occupational taxpaying cultures in taxpaying behaviour and attitudes, *Journal of Economic Psychology* 2009, 30, p. 216-2277.

41 D.D. Bobek, R.W. Roberts & J.T. Sweeney, The social norms of tax compliance: Evidence from Australia, Singapore, and the United States, *Journal of Business Ethics* 2007, 74, p. 49-64.

42 M. Wenzel, An analysis of norm processes in tax compliance, *Journal of Economic Psychology* 2004, 25, p. 213-228.

43 M. Wenzel, Motivation or rationalisation: Causal relations between ethics, norms and tax compliance, *Journal of Economic Psychology* 2005, 26, p. 491-508.

44 M. Wenzel, The multiplicity of taxpayer identities and their implications for tax ethics, *Law & Policy* 2007, 29, p. 31-50.

45 Heuristieken zijn *decision aids or cognitive short-cuts that humans use to minimize cognitive effort to make decisions that may not merit the full cognitive effort assumed in normative decision theory* (J.T. Scholz & N. Pinney, Duty, fear, and tax compliance: The heuristic basis of citizenship behavior, *American Journal of Political Science* 1995, 39, p. 490-512). Biases zien wij, in deze context, als een neiging naar of een voorkeur voor een standpunt dat afwijkt van het – in economische zin – meest rationele standpunt.

46 E.J. McCaffery & J. Baron, Heuristics and biases in thinking about tax, *Proceedings of the Annual Conference on Taxation* 2004, p. 434-443.

niet geheel duidelijk op welke wijze.<sup>47</sup> Ook iemands *risicobereidheid* is een (economisch-psychologische) determinant van compliance. Indien een belastingplichtige de intentie en mogelijkheid heeft om non-compliant te zijn, moet hij ook bereid zijn de risico's te lopen die met dit gedrag samenhangen. Uit empirisch onderzoek blijkt dan ook dat belastingplichtigen met een hogere risicobereidheid gemiddeld minder compliant zijn.<sup>48</sup>

• *Aanvullende determinanten van compliance van (grote) ondernemingen*

Bij (grote) ondernemingen zijn eigendom en leiding vaak gescheiden. Een uitvloeisel hiervan is bijvoorbeeld het gebruik van een bonussysteem. Bonussen dienen ervoor de belangen van aandeelhouders en bestuur zo veel mogelijk op één lijn te plaatsen. Als bijvoorbeeld de bonus wordt bepaald aan de hand van de nettowinst, is de bonus deels afhankelijk van de betaalde belasting.<sup>49</sup> Een mogelijk bijeffect kan zijn dat het bestuur er belang bij heeft de te betalen belasting te minimaliseren. De aanwezigheid van *bonussystemen* – voor het bestuur, middenmanagement of de tax director – kan daardoor de compliance negatief beïnvloeden.<sup>50</sup> Een ander belangrijk kenmerk van (grote) ondernemingen is de aanwezigheid van 'governance'-structuren. Ten eerste kan de *samenstelling van het bestuur* (raad van bestuur en raad van commissarissen) sterk verschillen. Deze verschillen kunnen de compliance beïnvloeden.<sup>51</sup> Ten tweede wordt vaak een *externe accountant* aangesteld om de jaarrekening te controleren. Ook dit kan invloed hebben op de compliance van de onderneming.<sup>52</sup> Ten slotte kan de *kwaliteit van de AO/IB* (waaronder beheersing van 'tax') sterk verschillen, wat ook de compliance kan beïnvloeden.<sup>53</sup> Ondernemingen kennen daarnaast verschillende soorten eigendomsstructuren. Zo kan sprake zijn van een beursgenoteerde onderneming, een

47 A.L. Christensen & P.A. Hite, A study of the effect of taxpayer risk perceptions on ambiguous compliance decisions, *Journal of the American Taxation Association* 1997, 19, p. 1-18.

48 K. Gërkhani & A. Schram, Tax evasion and income source: A comparative experimental study, *Journal of Economic Psychology* 2006, 27, p. 402-422.

49 K.J. Crocker & J. Slemrod, Corporate tax evasion with agency costs, *Journal of Public Economics* 2005, 89, p. 1593-1610.

50 Zie o.a. J.D. Phillips, Corporate tax-planning effectiveness: The role of compensation-based incentives, *The Accounting Review* 2003, 78, p. 847-874; C.S. Armstrong, J.L. Blouin & A.D. Larcker, The incentives for tax planning, *Journal of Accounting and Economics* 2012, 53, p. 391-411; S. Mahenthiran & J. Kasipillai, Influence of ownership structure and corporate governance on effective tax rates and tax planning: Malaysian evidence, *Australian Tax Forum* 2012, 27, p. 941-969; S.O. Rego & R. Wilson, Equity risk incentives and corporate tax aggressiveness, *Journal of Accounting Research* 2012, 50, p. 775-810.

51 Zie o.a. G. Taylor & G. Richardson, The determinants of thinly capitalized tax avoidance structures: Evidence from Australian firms, *Journal of International Accounting, Auditing and Taxation* 2013, 22, p. 12-25; G. Richardson, G. Taylor & R. Lanis, The impact of board of director oversight characteristics on corporate tax aggressiveness: An empirical analysis, *Journal of Accounting and Public Policy* 2013, 32, p. 68-88.

52 G. Taylor & G. Richardson, The determinants of thinly capitalized tax avoidance structures: Evidence from Australian firms, *Journal of International Accounting, Auditing and Taxation* 2013, 22, p. 12-25.

53 G. Richardson, G. Taylor & R. Lanis, The impact of board of director oversight characteristics on corporate tax aggressiveness: An empirical analysis, *Journal of Accounting and Public Policy* 2013, 32, p. 68-88.

Sjoerd Goslinga, Maarten Siglé &amp; Lisette van der Hel

**Tabel 1** De determinanten van compliance van individuele belastingplichtigen en ondernemingen

Economisch	Sociaalpsychologisch	Economisch-psychologisch	Bedrijfsspecifiek
<b>Belastingtarief</b>	Vertrouwen in de overheid	Onzekerheid	Bonussystemen
<b>Pakkans</b>	Persoonlijke normen	Risicobereidheid	Samenstelling bestuur
<b>Boetes</b>	Sociale normen	Heuristieken en biases	Externe accountant
<b>Belastingadviseur</b>	Procedurele rechtvaardigheid		AO/IB (incl. fiscale beheersing)
<b>Ervaringen eerdere controles</b>	Distributieve rechtvaardigheid		Eigendomskenmerken
<b>Complexiteit</b>			
<b>Mogelijkheden</b>			

familiebedrijf, een not-for-profit-organisatie, een hedge fund, enzovoort. Uit verschillend onderzoek blijkt dat zulke *eigendomskenmerken* van invloed kunnen zijn op de mate van compliance.<sup>54</sup>

- *Gebruik determinanten in handhavingsregie*

In Tabel 1 is een overzicht opgenomen van de besproken determinanten van compliance van *individuele belastingplichtigen en ondernemingen*.

De genoemde determinanten zijn in deze bijdrage grotendeels afzonderlijk besproken. De verschillende determinanten kunnen echter met elkaar interacteren en dat maakt toepassing van de bevindingen uit de beschreven onderzoeken in de praktijk complex. Een voorbeeld is dat het afschrikwekkende effect van boetes sterker is indien een belastingplichtige een hogere mate van procedurele rechtvaardigheid ervaart.<sup>55</sup> Ook kan de inzet van indringende onderzoeken en straffen de persoonlijke norm verdringen, waardoor wederzijds wantrouwen tussen de belastingdienst en de belastingplichtigen toeneemt.<sup>56</sup>

### 3.2 Operationaliseren en meten

In de vorige paragraaf passeerden de belangrijkste determinanten (oorzaken of omstandigheden die bepalend zijn voor) van *compliance* de revue. In termen van effectmeting gaat het hier om *variabelen* die gericht zijn op de *tussendoelen* van de Belastingdienst en die vooral de (subjectieve) beleving van belastingplichtigen

54 Zie o.a. D. Joulfaian, Corporate Income Tax Evasion and Managerial Preferences, *Review of Economics and Statistics* 2000, 82, p. 698; C. Cheng e.a., The effect of hedge fund activism on corporate tax avoidance, *Accounting Review* 2012, 87, p. 1493-1526.

55 M. Verboon & P. van Dijke, When do severe sanctions enhance compliance? The role of procedural fairness, *Journal of Economic Psychology* 2011, 32, p. 120-130.

56 S. Leviner, An overview: A new era of tax enforcement – From 'big stick' to responsive regulation, *Regulation & Governance* 2008, 2, p. 360-380.

meten, zoals ervaren rechtvaardigheid, inschatting van de pakkans, enzovoort. Het operationaliseren begint echter met het operationaliseren van het strategisch doel van de Belastingdienst: ‘compliance’.

• *Indicatoren voor compliance*

De Belastingdienst omschrijft zijn strategische doelstelling, *compliance*, als de bereidheid van burgers en bedrijven om hun fiscale verplichtingen na te komen. De OESO (2008) concretiseert compliance als het nakomen van de verplichting:

- 1 te registreren voor belastingplicht;
- 2 tijdig belastingaangifte te doen;
- 3 juiste en volledige belastingaangifte te doen; en
- 4 tijdig en volledig te betalen van de verschuldigde belasting.

De OESO-definitie gaat over concrete gedragingen, wat het makkelijker maakt om het nalevingsniveau op die vier aspecten te meten. De Rijksbegroting rept van de *bereidheid* tot het nakomen van belastingverplichtingen. Dat betekent dat naast gedrag ook de *houding* van belastingplichtigen onderdeel is van de doelstelling. Dat is van belang omdat verschil kan bestaan tussen de houding en het gedrag van belastingplichtigen. Onbedoelde vergissingen van mensen bijvoorbeeld gelden op gedragsniveau als non-compliant, terwijl ze op houdingsniveau als compliant kunnen worden aangemerkt. Het is dan van belang dat compliance niet alleen wordt gemeten op (1) gedragsniveau, maar ook op (2) houdingsniveau. En dat werkt een andere manier van meten die hierna wordt uitgewerkt.<sup>64</sup>

(1) De Belastingdienst heeft verschillende indicatoren voor de aspecten van compliance ontwikkeld voor uiteenlopende projecten en processen. Toch – of misschien daardoor – wordt niet gewerkt met standaardindicatoren. In hoeverre belastingplichtigen *juiste en volledige aangiftes* doen, stelt de Belastingdienst vast door middel van steekproefsgewijze (integrale) controles op basis waarvan uitspraken kunnen worden gedaan over het nalevingsniveau van een groep van belastingplichtigen. Dat gebeurt onder andere aan de hand van het percentage aangiftes dat na controle is gecorrigeerd en het bedrag aan belastingen dat daarmee is gemoeid. De kwaliteit van deze indicatoren is moeilijk te bepalen: bij controle ontdekt de Belastingdienst niet alle fouten en onbekend is wat de omvang is van wat hij niet ziet.

Los van het probleem dat niet alle belastingplichtigen bekend zijn bij de Belastingdienst (onvolledige registratie), is *tijdigheid van aangiftes* in principe een gesloten systeem. Van belastingplichtigen die de Belastingdienst uitnodigt om aangifte te doen of een lopende verplichting daartoe hebben, is duidelijk of ze de datum hebben overschreden waarop de aangifte binnen moet zijn. Met die informatie is een indicator te ontwikkelen. Daarbij is de vraag aan de orde wanneer sprake is van ‘te laat’. Belastingplichtigen kunnen bijvoorbeeld uitstel aanvragen of hun aangifte door een fiscaal dienstverlener laten doen. De operationalisatie zal daarom per project en per belastingmiddel kunnen verschillen. Ook *tijdigheid van betalingen* is in principe een gesloten systeem. Bij een uitnodiging tot betalen hoort immers een vastgestelde termijn. De Belastingdienst kan echter een beta-



Sjoerd Goslinga, Maarten Siglé & Lisette van der Hel

lingsregeling treffen of er kan een coulancetermijn gelden, enzovoort. Dergelijke aspecten moeten betrokken worden bij het bepalen van wat precies compliant en niet-compliant is en welke indicator daar dan het meest informatief voor is.

(2) Om de houdingsaspecten van compliance, zoals de bereidwilligheid, in kaart te brengen voert de Belastingdienst zoals gezegd survey-onderzoek onder belastingplichtigen uit.<sup>57</sup> Het operationaliseren van variabelen geschiedt door vragen of stellingen waarop belastingplichtigen antwoorden. Zo is bijvoorbeeld een belangrijke vraag die de bereidheid of houding van belastingplichtigen inventariseert de vraag: ‘Wat is uw gevoel bij belasting betalen?’, waarbij de antwoordmogelijkheden zijn: 1. Ik draag iets bij; 2. Ik sta iets af; 3. Mij wordt iets afgenomen.

Voorbeeld 2 laat zien hoe de Australische belastingdienst (ATO) het begrip compliance – gericht op het opgeven van contante ontvangsten in de belastingaangiften van ondernemers – in het kader van een effectmeting heeft geoperationaliseerd.

• *Voorbeeld 2*<sup>58</sup>

*De ATO constateerde dat het niet opgeven van contante ontvangsten in belastingaangiften van ondernemers een belangrijke vorm van non-compliance was. De ATO startte daarom een project genaamd Cash Economy, om de compliance op dit gebied te bevorderen. In dit kader is een mix aan instrumenten ingezet, waaronder het communiceren over de sancties die staan op deze vorm van non-compliance, het geven van voorlichting aan belastingplichtigen uit sectoren met veel contante betalingen, samenwerken met stakeholders zoals brancheorganisaties en adviseurs en tot slot het repressief optreden tegen belastingplichtigen die contante betalingen buiten de boeken hielden. Gesteld kan worden dat de ATO de volgende twee aannames heeft getoetst:*

- 1 *Het communiceren over de gevolgen van non-compliance, het geven van voorlichting, samenwerken met andere stakeholders zoals brancheorganisaties en adviseurs en tot slot het repressief optreden tegen belastingplichtigen verhoogt de gepercipieerde pakkans van het niet aangeven van contante betalingen en verlaagt de maatschappelijke acceptatie van contante betalingen.*
- 2 *Deze verhoogde gepercipieerde pakkans en verlaagde maatschappelijke acceptatie verhogen de (voluntary) compliance.*

*Om de effecten van dit project te kunnen meten moest de ATO indicatoren benoemen voor de gepercipieerde pakkans, maatschappelijke acceptatie en de mate van compliance. Hiervoor benoemde de ATO verschillende indicatoren:*

- *Maatschappelijke acceptatie*
  - a *Trend in het aantal meldingen bij de ‘tax evasion hotline’.*
  - b *Trend in het aantal en de aard van bezoeken aan webpagina’s die gerelateerd zijn aan de ‘cash economy’.*

57 Zie <www.fiscalemonitor.nl>.

58 OESO, Evaluating the effectiveness of compliance risk treatment strategies, Parijs: OECD Publishing 2010.



- c *Responses op surveys, inclusief onderzoek naar de perceptie van de maatschappij, het bedrijfsleven en de fiscale adviseurs.*
- d *Frequentie en toon van mediacommentaar gerelateerd aan de 'cash economy'.*

- **Gepercipieerde pakkans**

- a *Responses op surveys, inclusief onderzoek naar de perceptie van de maatschappij, het bedrijfsleven en de fiscale adviseurs.*

- **Mate van compliance**

- a *De mate van betrokkenheid van het bedrijfsleven bij beheersing van de 'cash economy'.*
- b *Wijzigingen in prestatie-indicatoren in relatie tot een individuele belastingplichtige en het bedrijfsleven.*
- c *Wijzigingen bij het indienen van zogenoemde 'activity statements' van 'cash economy' deelnemers.*
- d *Betaling van uitstaande verplichtingen door 'cash economy' deelnemers.*

- **Uitkomsten:**

*De belangrijkste uitkomst van dit onderzoek was dat de compliance, in de vorm van het indienen van een juiste en volledige aangifte en het tijdig betalen van belastingen, in de controlegroep significant lager was dan in de experimentele groep. Op basis hiervan concludeert de ATO dat het communiceren over de gevolgen van non-compliance met betrekking tot contante ontvangsten de compliance op dit terrein verhoogt.*

- **Determinanten van compliance**

Het meten van de determinanten van compliance gebeurt door middel van het stellen van vragen aan belastingplichtigen. Achtergrondkenmerken van individuen en ondernemingen kunnen voor sommige van deze variabelen een indicator zijn. Als dat bekend is, kan dat een voordeel zijn omdat in registraties en bestanden van de Belastingdienst verschillende achtergrondkenmerken zijn vastgelegd, bijvoorbeeld leeftijd, geslacht, inkomen, omzet, aantal werknemers, enzovoort. Kanttekening daarbij is wel dat het aantal achtergrondkenmerken in de bestanden beperkt is en dat niet altijd duidelijk is hoe deze samenhangen met bijvoorbeeld motieven en mogelijkheden. Vragenlijst- of survey-onderzoek kan inzicht bieden in de relatie of samenhang tussen achtergrondkenmerken en determinanten van compliance. Dat inzicht kan vervolgens weer input zijn voor een nieuwe operationalisatie.

Voor het meten van determinanten van compliance kan vaak gebruik worden gemaakt van operationalisaties die eerder zijn gebruikt in de wetenschappelijke onderzoeksliteratuur (zie paragraaf 3.1). Zo zijn voor veel variabelen meetinstrumenten ontwikkeld die soms direct kunnen worden gebruikt en soms aanpassing of vertaling behoeven (zie voorbeeld 3).

- **Voorbeeld 3**

*Een voorbeeld is de – in paragraaf 3.1 besproken – determinant **procedurele rechtvaardigheid**. Deze determinant is reeds in veel wetenschappelijk onderzoek naar het belas-*

Sjoerd Goslinga, Maarten Siglé & Lisette van der Hel

*tinggedrag van individuen onderzocht. Als gevolg daarvan kan de Belastingdienst gebruikmaken van een reeds gevalideerde vragenlijst. Ter illustratie, de operationalisatie van procedurele rechtvaardigheid zoals gebruikt door (onder andere) Verboon en Van Dijke (2011).<sup>59</sup>*

- *De Belastingdienst zorgt ervoor dat hij alle benodigde informatie heeft alvorens een beslissing te nemen.*
  - *De Belastingdienst behandelt iedereen gelijk.*
  - *De Belastingdienst neemt de specifieke omstandigheden van belastingplichtigen in aanmerking bij het nemen van beslissingen.*
  - *De Belastingdienst werkt nauwkeurig.*
  - *Belastingplichtigen die het oneens zijn met de Belastingdienst krijgen de gelegenheid hun standpunt toe te lichten.*
- *De interventie*

Het uitvoeren van handhavingsactiviteiten wordt ook wel omschreven als een interventie. In termen van de effectketen gaat het hier om *input*, *proces* en *output*. Handhavingsactiviteiten beïnvloeden de compliance indien ze minstens een van de determinanten van compliance beïnvloeden. Het doel van een effectmeting is te onderzoeken of en in welke mate dat ook het geval is. Het is dan ook zaak dat de Belastingdienst de activiteit of interventie scherp en gedetailleerd omschrijft. Zonder dat is niet goed te achterhalen waarom of door welk ingrediënt de interventie heeft gewerkt.<sup>60</sup>

Scherpe definities van de verschillende aspecten van compliance zijn noodzakelijk om compliance te operationaliseren en indicatoren te ontwikkelen.<sup>61</sup> Een vraag is bijvoorbeeld of compliance enkel het sec naleven van wet- en regelgeving betreft of dat ook de geest van de wet van belang is. En, ten aanzien van het compliance-aspect 'juistheid en volledigheid', bij welke geconstateerde afwijking is een aangifte 'aanvaardbaar'?<sup>62</sup> Heldere definities van de aspecten van compliance helpen ook om compliance-indicatoren te ontwikkelen die zo veel mogelijk gelijk of vergelijkbaar zijn voor de verschillende groepen van belastingplichtigen (segmenten). Daarmee zou de gehele Belastingdienst op verschillende niveaus kunnen werken met dezelfde indicatoren.

### 3.3 Methodieken

Een verbetering van de mate van naleving van een bepaalde fiscale verplichting, bijvoorbeeld een toename van het percentage tijdige betalingen, zegt er op zich nog niets over of dat komt door (intensivering van) de inspanningen van de

59 P. Verboon & M. van Dijke, When do severe sanctions enhance compliance? The role of procedural fairness, *Journal of Economic Psychology* 2011, 32, p. 120-130. De opgenomen vragen zijn door de auteurs van deze bijdrage vertaald vanuit het Engels.

60 Ten behoeve van onderzoek naar beïnvloeding van gezondheidsgerelateerd gedrag zijn hiervoor standaarden ontwikkeld. Zie bijvoorbeeld K.W. Davidson e.a., Evidence-based behavioral medicine: what is it and how do we achieve it?, *Annals of Behavioral Medicine* 2003, 26, p. 161-171.

61 OESO, *Measures of Tax Compliance Outcomes: A Practical Guide*, Parijs: OECD Publishing 2014.

62 Een aanvaardbare aangifte is een aangifte die voldoet aan wet- en regelgeving en vrij is van materiële fouten.

Belastingdienst. Er zijn verschillende onderzoeksmethoden en -designs waarmee 'hard' is te maken dat een verbetering van het nalevingsniveau daadwerkelijk kan worden toegeschreven aan de inspanningen van de Belastingdienst. De essentie daarvan is dat kan worden uitgesloten dat andere factoren dan de inspanningen of interventies die zijn gedaan, verantwoordelijk zijn voor het effect.

Onderzoeksdiseins kunnen met de zogenoemde *Maryland Scientific Methods Scale* (MSMS) worden geordend naar de mate waarin ze het oorzakelijke of causale verband hard maken.<sup>63</sup> De MSMS onderscheidt vijf niveaus van oplopende onderzoeksvaliditeit:

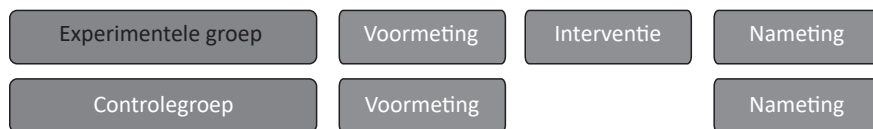
- 1 samenhang tussen een interventie en een uitkomstvariabele nadat de interventie is uitgevoerd;
- 2 score op de uitkomstvariabele voor en nadat de interventie is uitgevoerd, zonder (vergelijkbare) controlegroep;
- 3 score op de uitkomstvariabele voor en nadat de interventie is uitgevoerd, gemeten in een experimentele en vergelijkbare controlegroep (quasi-experimenteel design);
- 4 score op de uitkomstvariabele voor en nadat de interventie is uitgevoerd, gemeten in een experimentele en vergelijkbare controlegroep, waarbij gecontroleerd voor andere invloeden op de uitkomstvariabele (quasi-experimenteel design);
- 5 score op de uitkomst voor en nadat de interventie is uitgevoerd, waarbij de interventie alleen is uitgevoerd bij een experimentele groep die 'ad random' is onderscheiden van een controlegroep (experimenteel design).

Onderzoek op de laagste twee niveaus van de MSMS kan geen uitspraken opleveren over de *causaliteit*. Anders gezegd: het uitvoeren van alleen een voor- en nameting geeft geen uitsluitel of de uitkomsten zijn veroorzaakt door de interventie. Bij de hogere niveaus van onderzoeksvaliditeit is dat in oplopende mate wel het geval. Dit betekent echter niet dat de eerste twee niveaus in de context van effectmeting nutteloos zijn. Integendeel, er zijn veel situaties denkbaar waarin het niet mogelijk is een hoger niveau van onderzoeksvaliditeit te bereiken; de eerste twee niveaus kunnen dan zeker van toegevoegde waarde zijn. In die situaties is het echter van belang dat de te onderzoeken verbanden een stevig theoretisch fundament hebben (zie paragraaf 2).

Om uit te sluiten dat er een andere factor in het spel is die een eventuele verbetering van het nalevingsniveau heeft veroorzaakt, is een controlegroep nodig; dat wil zeggen, een groep die de interventie niet heeft ondergaan. Met een experimenteel design (niveaus 3, 4 en 5), waarbij een groep die de interventie ondergaat (de experimentele groep) kan worden vergeleken met een (vergelijkbare) groep die de interventie niet ondergaat (de controlegroep), kunnen uitspraken worden gedaan over het effect van een interventie. Experimentele onderzoeksdiseins zijn er in allerhande vormen.<sup>64</sup>

63 Zie bijvoorbeeld D.P. Farrington, *Methodological quality standards for evaluation research*, The Annals of the American Academy of Political and Social Science 2003, 587, p. 49-68.

64 W.R. Shadish, T.D. Cook & D.T. Campbell, *Experimental and quasi-experimental designs for generalized causal inference*, Boston: Houghton Mifflin 2002.

**Figuur 4** MSMS-niveau 5

Het ideale onderzoeksdesign is de zogenoemde *randomized controlled trial of gecontroleerd veldexperiment* (zie Figuur 4). Het gaat dan om *niveau 5* van de MSMS-ordening. De kracht van een dergelijk design zit in de vergelijkbaarheid van de experimentele groep en de controlegroep. Om dat te realiseren worden uit de populatie waarop de interventie is gericht entiteiten willekeurig (ad random) toegewezen aan ofwel de experimentele groep, ofwel de controlegroep. Deze willekeurige toewijzing zorgt ervoor dat beide groepen op alle kenmerken behalve de interventie vergelijkbaar zijn, waardoor een verschil in compliance na de interventie (nameting) kan worden toegeschreven aan de interventie. Met een dergelijk onderzoeksdesign wordt dus uitgesloten dat andere factoren dan de interventie een eventueel verschil tussen de twee groepen hebben veroorzaakt.

Voorbeeld 4 is een voorbeeld van een experimenteel onderzoek dat is uitgevoerd door de Belastingdienst naar het effect van zogenoemde startersbezoeken.

- *Voorbeeld 4*

*Dit onderzoek had als doel te achterhalen wat het effect is van een dienstverlenend bezoek van een medewerker van de Belastingdienst aan een recentelijk gestarte ondernemer (een zogenoemd 'startersbezoek') op zijn of haar belastingcompliance. Het onderzoek is opgezet als een gecontroleerd veldexperiment, waarbij één groep willekeurig geselecteerde starters door een medewerker werd bezocht en een willekeurig geselecteerde controlegroep starters geen bezoek kreeg. Van de ondernemers werd gedurende de gehele onderzoeksperiode de tijdigheid van de aangiftes en de tijdigheid van betalen vastgelegd. Daarnaast werden de ervaringen, opvattingen en houdingen ten aanzien van belasting betalen en de Belastingdienst verzameld met een enquête onder de ondernemers in zowel de experimentele groep als de controlegroep. Voorts werd na drie jaar ondernemerschap bij beide groepen een volledig boekenonderzoek uitgevoerd door de Belastingdienst, waarmee de juistheid en volledigheid van de fiscale aangiftes kon worden beoordeeld.*

- *Uitkomsten:*

*Met betrekking tot de effecten van startersbezoeken werd verwacht dat ondernemers die een bezoek hadden gekregen aan het begin van het ondernemerschap (de experimentele groep), hun belastingzaken sneller en beter op de rails zouden hebben dan de groep die geen bezoek van de Belastingdienst had gehad (de controlegroep) en dat dus de experimentele groep meer compliant zou zijn dan de controlegroep. Dat bleek echter niet of nauwelijks het geval, met name met betrekking tot het tijdig indienen van de aangifte*

en het tijdig betalen. Uit de enquête bleek wel dat de startersbezoeken door een ruime meerderheid van de ondernemers werden gewaardeerd en nuttig bevonden. Ondernemers met een startersbezoek gaven aan meer rekening te houden met (onverwachte) toekomstige (belasting)uitgaven, dan ondernemers in de controlegroep. Uit de boekenonderzoeken bleek dat het bezoek bij één specifieke subgroep resulteert in een lager aantal correcties.<sup>65</sup> Bij de andere ondernemers was geen effect waarneembaar.

In de praktijk kan het moeilijk of zelfs onmogelijk zijn om een controlegroep te selecteren die volledig vergelijkbaar is met de experimentele groep. Ook kan het zo zijn dat een interventie niet precies gericht kan worden, waardoor ook de controlegroep met de interventie in aanraking kan komen, bijvoorbeeld als informatie of voorlichting wordt verstrekt via de massamedia. Als dat het geval is, spreekt men van quasi-experimentele designs. Achteraf kan, wanneer het onderzoeksdesign niet perfect is, in de statistische analyse van de verschillen tussen de experimentele groep en de controlegroep op de uitkomstvariabele nog een en ander worden gecorrigeerd.

- *Innovaties*

Een van de uitdagingen waarvoor belastingdiensten zich gesteld zien, is hoe zij de effecten van innovatieve (vaak preventieve) maatregelen zichtbaar kunnen maken.<sup>66</sup> Het gaat dan met name om maatregelen die fouten moeten voorkomen (*compliance by design*)<sup>67</sup> of meer horizontale strategieën, waarbij het toezicht achteraf van een belastingdienst is gebaseerd op de mate waarin een bedrijf zelf 'in control' is over zijn fiscale risico's (*co-operative compliance*).<sup>68</sup>

Traditioneel gebruiken belastingdiensten met name outputindicatoren (zoals aantallen controles, correctieopbrengsten, boetes, enzovoort) om zich te verantwoorden over hun (voornamelijk repressieve) handhavingsactiviteiten (gebaseerd op de economische inzichten). Voor belastingdiensten geldt echter dat de handhavingsstrategieën zich de afgelopen jaren hebben ontwikkeld (door de introductie van sociaalpsychologische en economisch-psychologische inzichten). Het meten van effecten heeft echter geen gelijke tred gehouden met deze ontwikkelingen: belastingdiensten gebruiken nog steeds voornamelijk outputindicatoren gerelateerd aan repressieve activiteiten om zich te verantwoorden, terwijl een groot deel van hun activiteiten proactief en preventief is. Uit een enquête van de OESO (2013) bleek bijvoorbeeld dat nog maar weinig belastingdiensten de effectiviteit van de *cooperative compliance* inspanningen meten.<sup>69</sup> Dat geldt ook voor

65 Het betrof de ondernemers die in het eerste jaar van de onderneming geen loon ontvingen naast de inkomsten uit de onderneming.

66 OESO, *Measures of Tax Compliance Outcomes: A Practical Guide*, Parijs: OECD Publishing 2014.

67 OESO, *Tax Compliance by Design: Achieving Improved SME Tax Compliance by Adopting a System Perspective*, Parijs: OECD Publishing 2014.

68 E.C.J.M. van der Hel-van Dijk, *De ontwikkeling van horizontaal toezicht wereldwijd*, Vakblad Tax Assurance 2013, nr. 1, p. 20.

69 OESO, *Co-operative Compliance: A Framework*, Parijs: OECD Publishing 2013.

Sjoerd Goslinga, Maarten Siglé & Lisette van der Hel

Nederland, zo blijkt uit het rapport van de Commissie-Stevens (2012) bij de evaluatie van de Nederlandse equivalent horizontaal toezicht.<sup>70</sup>

Naar onze mening is de kern van het probleem dat belastingdiensten überhaupt niet gewend zijn om effecten te meten, maar zich vooral beperken tot output omdat dat veelal gemakkelijker vast te stellen is dan effecten. Bovendien kan output gemakkelijk achteraf worden vastgesteld, terwijl voor een goede effectmeting al voorafgaand aan de activiteit nagedacht moet zijn over de uitvoering ervan. De in deze paragraaf beschreven methoden zijn naar onze mening echter evenzeer van toepassing op preventieve maatregelen. Dat blijkt bijvoorbeeld ook uit het hierboven beschreven voorbeeld 4.

### 3.4 *Ten slotte*

In de praktijk stemt de Belastingdienst in de handhavingsregie zijn interventies zo goed mogelijk af op de oorzaken van (non-)compliance. De Belastingdienst kan hier gebruikmaken van een grote hoeveelheid (internationale) (wetenschappelijke) literatuur, maar zal echter steeds in de context van de Nederlandse praktijk vast moeten stellen of de inspanningen of activiteiten van de Belastingdienst daadwerkelijk tot de verwachte effecten leiden.

Het onderzoek naar de determinanten van compliance is tot op heden hoofdzakelijk gericht op de determinanten die betrekking hebben op het (on)juist of (on)volledig doen van aangifte. De overige aspecten van (non-)compliance, zoals het (niet) registreren, het tijdig (niet of niet tijdig) doen van aangifte en het tijdig (niet of niet tijdig) betalen zijn slechts beperkt onderzocht. Of de besproken determinanten zich laten generaliseren naar deze andere aspecten van fiscale compliance is een belangrijke vraag voor nader onderzoek.

Een van de uitdagingen van belastingdiensten wereldwijd is het vaststellen van het effect van innovatieve (veelal preventieve) maatregelen om compliance te bevorderen en non-compliance tegen te gaan. Hier speelt een rol dat belastingdiensten van oudsher zich met name hebben verantwoord in termen van 'output' gerelateerd aan repressieve maatregelen. Ook het opzetten van een methodologisch verantwoord onderzoek is in de praktijk een uitdaging. Het is bijvoorbeeld niet altijd mogelijk tot een onderzoeksdesign te komen waarmee causaliteit is vast te stellen tussen activiteiten van de Belastingdienst en het realiseren van zijn doelen. In dit kader is in deze bijdrage benadrukt dat ook onderzoeken op de lagere schalen van de MSMS – gegeven een sterk theoretisch fundament – van meerwaarde kunnen zijn.

## 4 Implementatie van effectmeting

Deze paragraaf verkent de vraagstukken die gemoeid zijn met de daadwerkelijke implementatie van effectmeting in de organisatie van de Belastingdienst. Bespro-

70 Commissie Horizontaal Toezicht Belastingdienst, Fiscaal toezicht op maat. Soepel waar het kan, streng waar het moet, Rijswijk: VijfKeerBlauw 2012.

ken worden: beleidsvorming (over effectmeting), de benodigde competenties en verantwoordelijkheden, de stakeholders en continuïteit en timing.

### *Beleidsvorming*

Een uitdaging voor veel toezichthouders is om effectmeting op een structurele manier te borgen in de organisatie en onderdeel te maken van de manier van werken. Winter (2012) noemt een drietal verklaringen voor de achterblijvende ontwikkeling van effectmeting bij toezichthouders:<sup>71</sup>

- 1 de complexiteit van effectmetingen en de investering in tijd en geld die het uitvoeren van effectmetingen met zich mee brengt;
- 2 toezichthouders die 'risicogericht' werken, zijn geneigd direct te willen handelen en zitten minder te wachten op 'langdurige' analyse en onderzoek;
- 3 de cultuur binnen de organisaties van toezichthouders is onvoldoende stimulerend voor het uitvoeren van effectmetingen omdat de kans bestaat dat 'slecht nieuws' over prestaties van toezichthouders naar buiten kan komen.

De WRR stelt in zijn rapport *Toeziën op publieke belangen* (2013) dat 'goed toezicht' een consistente en samenhangende beleidsvisie vraagt. Dat geldt voor de overheid als geheel maar naar onze mening ook voor individuele toezichthouders. 'Effectmeting' is een onmisbaar element van 'goed toezicht' en dient als zodanig derhalve geborgd te zijn in het toezichtbeleid van een individuele toezichthouder. Dat de Belastingdienst volgens de uitgangpunten van handhavingsregie (zie paragraaf 2.1) werkt, kan een belangrijk voordeel zijn om effectmeting te implementeren en te borgen. Kort samengevat betekent handhavingsregie dat de Belastingdienst op verschillende niveaus in de organisatie volgens een 'vast stramen' zijn handhavings- en toezichttaken uitvoert en als onderdeel daarvan de effecten van zijn handelen meet. Het doel hierbij is om te leren over de effectiviteit van inspanningen en om de inzet van mensen en middelen te verantwoorden. De investeringen die dat vergt in tijd en geld zouden al ingecalculeerd moeten zijn. Omarming van handhavingsregie impliceert derhalve omarming van effectmeting als onderdeel van de dagelijkse manier van werken. Dat vraagt van de Belastingdienst wel nog meer om deze manier van werken en denken op alle niveaus in de organisatie consequent toe te passen.

### *Competenties en verantwoordelijkheden*

Handleidingen en handreikingen voor effectmeting van handhaving besteden in de regel weinig aandacht aan de benodigde competenties en aan de organisatie van verantwoordelijkheden om tot effectmeting te komen.<sup>72</sup> Impliciet gaat men er doorgaans van uit dat bijvoorbeeld projectverantwoordelijken ook de effectmeting voor hun rekening nemen, waarbij dan wel wordt aangegeven dat voor het ontwerp van het onderzoek en de analyse van verzamelde gegevens specialisati-

71 H.B. Winter, *Metten van de effecten van toezicht: 'Yes we can!?', TvT 2012, nr. 2, p. 63-80.*

72 Zie bijvoorbeeld: Inspectieraad, *Leidraad effectmeting bij inspecties*, Den Haag: Inspectieraad 2005; Centrum voor Criminaliteitspreventie en Veiligheid (CCV), *Effecten van toezicht en handhaving meten, een handreiking*, Utrecht: CCV 2011.



sche kennis nodig is. Wie die kennis zou moeten leveren en hoe de organisatorische inrichting eruit zou moeten of kunnen zien, blijft echter onbesproken.

Ook effectmeting in het fiscale domein is geen simpel kunstje dat medewerkers naast hun reguliere werkzaamheden kunnen doen. Voor een effectmeting zijn specifieke vaardigheden, kennis van wetenschappelijke methodieken en theorieën en van de werkwijze en werkzaamheden van de Belastingdienst nodig. Dat betekent dat de verantwoordelijkheid voor de opzet en uitvoering van effectmeting naar onze mening belegd dient te zijn bij een beperkte groep van mensen. Dat laat onverlet dat in de praktijk een grote groep van medewerkers in aanraking met effectmeting komt aangezien effectmeting een vast onderdeel is van het werken volgens de uitgangspunten van handhavingsregie. Toezichtmedewerkers op uitvoeringsniveau zouden daarom wel basiskennis moeten hebben van het werken volgens handhavingsregie en van effectmeting.

Centrale coördinatie van effectmeting kan voorkomen dat verschillende initiatieven tegelijk worden gestart waardoor effecten elkaar kunnen beïnvloeden en het lastig is om de effecten van specifieke interventies te meten. Het voorkomt ook dat meerdere projecten worden opgestart om hetzelfde te meten (zo zijn er bijvoorbeeld meerdere initiatieven om de effecten van horizontaal toezicht te meten) en kan zorgen voor een prioritering van effectmetingsonderzoeken en de daarmee gemoeide tijd en (schaarse) capaciteit en competenties. Ten slotte zorgt centrale coördinatie voor een beter overzicht van de effecten van verschillende interventies. Dit stelt de onderzoekers in staat in een vroeg stadium mogelijke neveneffecten te identificeren. De praktijk zal verder moeten leren hoe de verantwoordelijkheden voor de opzet en uitvoering van effectmetingen het best kan worden vormgegeven binnen de Belastingdienst.

### *Stakeholders*

Om effectmeting te stimuleren is het van belang om bij *interne stakeholders*, zoals management of medewerkers draagvlak te creëren voor het uitvoeren van effectmetingen.<sup>73</sup> Een consistente boodschap dat effectmeting noodzakelijk is voor de verdere ontwikkeling van de handhaving- en toezichtstrategie is daarbij behulpzaam. Voor het interne draagvlak is het van belang dat op uitvoeringsniveau kleinschalige projecten in samenhang worden uitgevoerd, waarvan de resultaten worden gebruikt om interventies te verbeteren. Communicatie hierover helpt voor vergroting van het draagvlak. Ook kunnen resultaten van effectmeting ertoe bijdragen dat individuele medewerkers beter inzicht krijgen in hun eigen bijdrage aan het fiscaal toezicht.

*Externe stakeholders* hebben verschillende belangen. Of, en zo ja in welke mate, deze stakeholders door de Belastingdienst geïnformeerd dienen te worden over de 'effecten' van de handhaving, hangt af van hun positie ten opzichte van de Belastingdienst. Zo zal bijvoorbeeld de politiek op een hoger aggregatieniveau geïnformeerd willen worden dan de Rijksaccountantsdienst of de Algemene Rekenkamer. Om in kaart te brengen hoe met verschillende stakeholders kan worden gecom-

73 Zie ook: M.C.J.L. Klerks & P.B.M. Robben, De inspecteur over de effecten van het toezicht: een onderzoek onder de inspecteurs van twaalf rijksinspecties, TvT 2012, nr. 3, p. 19-30.

municieerd, kan de Belastingdienst een stakeholderanalyse uitvoeren. De informatie uit effectmetingen kan overigens gelijktijdig meerdere doelen dienen: de Belastingdienst verantwoordt zich extern conform gemaakte afspraken door het openbaar maken van bepaalde resultaten maar gebruikt dezelfde informatie ook als input voor zijn *compliance-strategie*.

Gezien het voorgaande gaat het belang van stakeholders verder dan alleen het voldoen aan verwachtingen en het afleggen van verantwoording door de Belastingdienst. De verschillende stakeholders kunnen ook een rol spelen bij het opzetten van effectmetingen, onder andere door het benoemen van relevante variabelen, neveneffecten en randvoorwaarden. Een andere rol die stakeholders kunnen spelen bij een effectmeting is het duiden van bevindingen. Verschillende belastingdiensten zetten al (voorzichtige) stappen om op deze manier met externe stakeholders om te gaan.<sup>74</sup>

#### *Continuïteit en timing*

De Belastingdienst is sterk gericht op en georganiseerd rond de aangiftecycli, zoals de jaarlijkse aangiftes voor de vennootschapsbelasting en inkomstenbelasting. Werkzaamheden gericht op de compliance van belastingplichtigen volgen deze cycli niet. Horizontaal toezicht is bijvoorbeeld niet specifiek gericht op een aangifte of een set van aangiftes. Het meten van de effecten van deze werkzaamheden dient zodoende de vaste cycli te overstijgen en met een langere tijdshorizon te werken.<sup>75</sup> Indien dit niet gebeurt, loopt de Belastingdienst het risico dat potentieel succesvolle strategieën te vroeg worden beëindigd. Anderzijds loopt de Belastingdienst het risico dat potentieel minder succesvolle strategieën als een succes worden aangemerkt, bijvoorbeeld doordat op korte termijn sprake is van verhoogde belastingopbrengsten.

Effectmeting is een *continu proces* omdat het bij het realiseren van de (algemene) beleidsdoelstelling gaat om een 'duurzame' verandering in het gedrag van belastingplichtigen en de borging van de continuïteit van belastingopbrengsten. De inspanningen van de Belastingdienst zullen moeten bijdragen aan een maatschappij waarin sprake is van een hoge belastingmoraal en een hoog percentage burgers en bedrijven die zonder noemenswaardige inzet vanuit het fiscale toezicht aan hun fiscale verplichtingen voldoen.<sup>76</sup> Het is daarom van belang om effecten op verschillende niveaus en verschillende termijnen (kort, middellang en lang) te meten. Bijkomend aandachtspunt is het optreden van positieve of negatieve neveneffecten, die zich vaker pas op de middellange en lange termijn manifesteren. Als onderdeel van het continue proces van effectmeting dient men zodoende alert te blijven op mogelijke neveneffecten zodat deze mogelijk bij latere effectmetingen in kaart kunnen worden gebracht.

74 OESO, Together for better outcomes: engaging and involving SME taxpayers and stakeholders, Parijs: OECD Publishing 2013.

75 OESO, Measures of Tax Compliance Outcomes: A Practical Guide, Parijs: OECD Publishing 2014.

76 Zie voor een verdere onderbouwing bijvoorbeeld: S. James & C. Alley, Tax compliance, self-assessment and tax administration, Journal of Finance and Management in Public Services 2004, 2, p. 27-42.

Ook de *timing* van effectmeting is belangrijk. Het is essentieel om bij de implementatie van een nieuwe interventie vooraf na te denken op welke wijze de effecten gemeten kunnen worden en een beslissing te nemen over de aanpak. Achteraf zijn de mogelijkheden om uitspraken te kunnen doen over effecten immers beperkt.<sup>77</sup> Het is bovendien van belang om een voor- of nulmeting uit te voeren, alvorens te beginnen met werkzaamheden waarvan de effecten gemeten moeten worden. Een dergelijke meting geeft een weergave van de situatie voorafgaand aan de implementatie en kan zodoende worden gebruikt als vergelijkingspunt voor de latere metingen. Voor het aantonen van causale verbanden, bijvoorbeeld door middel van een veldexperiment, moet vooraf een onderzoeksopzet worden opgesteld en zal de interventie volgens die opzet moeten worden uitgevoerd.

## 5 Tot slot

Op verschillende momenten en vanuit verschillende gremia is in het afgelopen decennium nadrukkelijk de eis of verwachting geventileerd dat toezichthouders de effecten van hun inspanningen meer inzichtelijk moeten maken. Die verwachting bestaat ook als het gaat om de Belastingdienst. Ook de Belastingdienst zelf onderschrijft het belang van effectmeting van zijn activiteiten. Tot op heden lijkt effectmeting in het fiscale domein, zowel in binnen- als buitenland, echter nog niet erg uit de verf te komen.<sup>78</sup> De verkenning in deze bijdrage wijst op een aantal *uitdagingen* dat samenhangt met het meten van effecten in het fiscale domein:

- 1 Een van die uitdagingen betreft het feit dat de beleidstheorie – en daarmee de relaties tussen de doelen en de instrumenten die de Belastingdienst kan inzetten – onvoldoende expliciet is beschreven. Explicitering van de beleidstheorie is een van de voorwaarden voor het kunnen opzetten van een succesvolle effectmeting.
- 2 Om effect te kunnen hebben zullen de inspanningen of activiteiten van de Belastingdienst in moeten grijpen op de achterliggende oorzaken (determinanten) van (non-)compliance. Voor het opzetten van een goede effectmeting moet de Belastingdienst bij het inzetten van een handhavingsinstrument helder hebben op welke determinanten dit instrument is gericht. Hoewel in het fiscale domein al betrekkelijk veel wetenschappelijk onderzoek is gedaan naar de determinanten van (non-)compliance, is ook sprake van lacunes in deze literatuur. Wij noemden in dit kader reeds het gebrek aan onderzoek naar de determinanten van de compliance van grote ondernemingen en de determinanten van het (niet) registreren, het tijdig (niet of niet tijdig) doen van aangifte en het tijdig (niet of niet tijdig) betalen. Het is op deze terreinen een uitdaging voor de Belastingdienst om handhavingsinstrumenten op een gefundeerde manier in te zetten en daarmee ook om de effecten van de inspanningen te meten.

77 OESO, *Measures of Tax Compliance Outcomes: A Practical Guide*, Parijs: OECD Publishing 2014.

78 OESO, *Measures of Tax Compliance Outcomes: A Practical Guide*, Parijs: OECD Publishing 2014.

- 3 Een andere uitdaging die het meten van de effecten in de weg lijkt te staan – maar die tegelijkertijd ook juist het belang van goede effectmetingen benadrukt – is de verschuiving van traditionele, meer repressieve activiteiten naar meer preventieve activiteiten. Dit vergt van belastingdiensten dat zij nadenken over nieuwe soorten indicatoren, welke in toenemende mate verder weg lijken te staan van de outputindicatoren waar belastingdiensten voorheen sterk op stuurden.
- 4 Ook het opzetten van een methodologisch verantwoord onderzoek is in de praktijk een uitdaging. Vanuit verschillende inherente beperkingen aan het meten van de effecten van fiscale handhaving, is het niet altijd mogelijk te komen tot een onderzoeksdesign waarmee causaliteit is vast te stellen tussen activiteiten van de Belastingdienst en het realiseren van zijn doelen. Dit lijkt belastingdiensten (en andere toezichthouders) er soms – ten onrechte – van te weerhouden om een effectmeting op te zetten. In dit kader is in deze bijdrage benadrukt dat ook onderzoeken op de lagere schalen van de MSMS – gegeven een sterk theoretisch fundament – van meerwaarde kunnen zijn. Verder is het zaak om de mogelijkheden te benutten wanneer zich een situatie voordoet van een natuurlijk experiment.<sup>79</sup> Bijvoorbeeld als een activiteit of interventie zich richt op een deel van een populatie en er (achteraf) een controlegroep is te creëren.
- 5 De laatste uitdaging die uit deze bijdrage naar voren komt, is de organisatorische inbedding van effectmetingen binnen belastingdiensten. Deze inbedding kan om reden van financiën, capaciteit en cultuur problematisch zijn. In deze bijdrage hebben wij argumenten aangedragen die pleiten voor centrale coördinatie als het gaat om effectmeting.

Uit deze verkenning komt ook naar voren dat effectmeting verschillende *doelen* kan hebben. Afhankelijk van de na te streven doelen kan de opzet en uitvoering van effectmeting verschillen. De ontwikkelingen in de handhaving(ssstrategie) van de Belastingdienst, waarbij het accent verschuift naar het voorkomen van fouten, vergroot de noodzaak van het kunnen laten zien van effecten. Hierbij staan de doelen *leren* en *verantwoorden* voorop. De Belastingdienst dient aannames te toetsen en kan niet volstaan met het rapporteren van de output (bijvoorbeeld het aantal uitgevoerde controles) of de immediate outcome (bijvoorbeeld het gecorrigeerde bedrag aan belastingen) als (nieuw) beleid zich richt op het voorkomen van fouten. Daarnaast kan effectmeting een rol vervullen ten behoeve van de *sturing*. De Belastingdienst maakt in de uitvoering van de handhavingsregie bij voortduring keuzes over de uitvoering van activiteiten (korte en lange termijn) om compliance te bevorderen en non-compliance tegen te gaan en over de inzet van capaciteit. Kennis van en inzicht in de determinanten van (niet-)naleving zijn

79 Bij een natuurlijk experiment worden de experimentele groep en de controlegroep gevormd door een 'toevallige' (in de zin dat deze niet door de onderzoekers is bepaald) gebeurtenis in de werkelijkheid. Hoewel de onderzoekers geen controle hebben gehad over de condities waarin het experiment is uitgevoerd, zoals wel het geval is in een gecontroleerd experiment, kan doorgaans wel worden voldaan aan voorwaarden zoals het ad random toewijzen van de deelnemers aan de verschillende condities binnen het experiment.

Sjoerd Goslinga, Maarten Siglé & Lisette van der Hel

daarbij van groot belang. Die bepalen immers hoe en in hoeverre het gedrag van burgers en bedrijven (het best) kan worden beïnvloed. Tot slot kan effectmeting een rol vervullen bij de *reflecterende* rol van de Belastingdienst, waarin hij de effecten van zijn toezicht gebruikt als graadmeter voor de uitvoerbaarheid en handhaafbaarheid van wetgeving en om aan te geven wie in actie moet komen om de onderliggende problematiek aan te pakken: de wetgever, beleidsmakers, de Belastingdienst zelf, andere toezichthouders of belastingplichtigen.

De gesignaleerde uitdagingen en doelen van effectmeting zijn in de kern niet uniek voor of voorbehouden aan de Belastingdienst, maar gelden voor toezichthouders in algemene zin. Wel bepaalt de context van de betreffende toezichthouder in hoeverre deze uitdagingen en doelen relevant zijn bij het uitvoeren van effectmetingen in het specifieke domein van die toezichthouder. Tussen verschillende domeinen zijn verschillen te verwachten ten aanzien van de beleidsdoelen en het geformuleerde beleid, de stand van zaken in de wetenschappelijke literatuur, het (voor)werk dat al is gedaan op het gebied van operationalisaties en indicatoren, de beschikbaarheid van gegevens en van de expertise van de toezichthouder. Het fiscale domein kenmerkt zich onder andere door veel beschikbare kennis van determinanten van nalevingsgedrag van individuen en (in mindere mate van) bedrijven en door de beschikbaarheid van operationalisaties voor compliance die de OECD heeft ontwikkeld. Deze kennis, maar ook het benoemen en in kaart brengen van de randvoorwaarden om tot effectmeting te komen, kunnen behulpzaam zijn voor andere toezichthouders. Door ook bij andere toezichthouders een exercitie als in dit artikel uit te voeren, kunnen mogelijk kennis en aanvullende inzichten aan het licht komen die de ontwikkeling van effectmeting in het fiscale domein verder kunnen vergroten.