Algemene Rekenkamer

Datagedreven selectie van aangiften door de Belastingdienst

2019

Datagedreven selectie van aangiften door de Belastingdienst

De tekst in dit document is vastgesteld op 4 juni 2019. Dit document is op 11 juni 2019 aangeboden aan de Tweede Kamer.

Inhoud

1 Inleiding 5

1.1 Aanleiding van dit onderzoek 5

1.2 Leeswijzer 6

2 Risicomodellen 8

2.1 Inleiding 8

2.2 Het risicomodel van teruggave btw 8

2.1.2 Het risicomodel van de inkomstenbelasting 8

2.1.3 Trefkans 10

2.2 De Investeringsagenda in relatie tot de uitvoerings- en toezichtstrategie 10

2.3 Risicomodellen en de Investeringsagenda 11

2.4 Topicmodellen 12

2.5 Capaciteit handhaving 13

2.5.1 Capaciteit bij de inkomstenbelasting 13

2.5.2 Capaciteit bij btw-teruggave 13

2.5.3 Grenswaarde in het risicomodel van de inkomstenbelasting 14

2.5.4 Afweging capaciteit handhaving versus additionele opbrengsten 15

2.6 Omvang reductie ***tax-gap*** 16

3 Verantwoord omgaan met data 17

3.1 Opzet privacybeleid 17

3.2 Fysieke toegang en informatiebeveiliging 18

3.3 Hoeveelheid verzamelde gegevens 19

3.4 Ongewenste neveneffecten van risicomodellen 19

3.5 Geautomatiseerde besluitvorming bij risicomodellen 20

3.6 Pseudonimiseren van data 21

4 Samenvatting, conclusies en aanbevelingen 22

4.1 Verbeterpunten in de risicomodellen 22

4.2 Uitgaan van risico versus capaciteit 23

4.3 Verantwoord omgaan met data 23

4.4 Reactie van de staatssecretaris 25

4.5 Nawoord van de Algemene Rekenkamer 26

Bijlagen 27

1 Van data naar risicomodel 28

2 Risicomatrix IB 30

3 Bepaling grenswaarde 32

4 Lijst met gebruikte afkortingen 35

5 Literatuur 36

6 Noten 37

1 Inleiding

Ieder jaar zorgt de Belastingdienst voor het heffen, innen en controleren van belastingen en premies. Jaarlijks komen er bij de Belastingdienst alleen al 12 miljoen aangiften voor de inkomstenbelasting binnen. Per werkdag int de Belastingdienst gemiddeld een miljard euro, geld dat vervolgens kan worden besteed aan overheidstaken zoals onderwijs, veiligheid, sociale zekerheid en gezondheidszorg. Daarnaast zorgt de dienst ervoor dat jaarlijks € 11,6 miljard aan toeslagen aan burgers wordt toegekend en uitbetaald.

De Belastingdienst maakt voor zijn toezichthoudende en controlerende taak steeds meer gebruik van risicomodellen die gebaseerd zijn op data. Deze nieuwe methode stelt risico's vast met alle data of kenmerken van de belastingplichtige die op dat moment bij de Belastingdienst bekend zijn. Dit in tegenstelling tot de oorspronkelijke werkwijze van de Belastingdienst die slechts één kenmerk of een klein aantal kenmerken gebruikte om een aangifte te controleren. De Belastingdienst heeft anno 2018 nog maar een beperkt aantal risicomodellen ontwikkeld die gebaseerd zijn op data en ook daadwerkelijk gebruikt worden in de handhaving. De dienst investeert veel in deze nieuwe methoden en verwacht in de toekomst het aantal modellen uit te breiden.

In dit onderzoek naar de datagedreven benadering van de controle van aangiften, onderzoekt de Algemene Rekenkamer deze nieuwe methode en beoordeelt de toegevoegde waarde van de risicomodellen. Wij hebben daartoe twee risicomodellen onderzocht: het risicomodel voor de teruggave van btw aan ondernemers en het risicomodel dat gebruikt wordt bij de inkomstenbelasting van particulieren en kleine en middelgrote ondernemers.

In dit onderzoek komen ook de maatschappelijke en ethische aspecten van het gebruik van data-analyse in de uitvoerings- en toezichtstrategie aan de orde. Het gebruik van data van belastingplichtigen moet immers zorgvuldig gebeuren om te voorkomen dat privacy van belastingplichtigen wordt geschonden.

1.1 Aanleiding van dit onderzoek

In mei 2015 presenteerde de staatssecretaris van Financiën de Investeringsagenda (Financiën, 2015) die de toen gesignaleerde problemen bij de Belastingdienst beoogde op te lossen. Eén van de doelstellingen van de Investeringsagenda was om data, risicomodellen en data-analyse een grotere rol te laten spelen in toezicht en handhaving. De introductie

van deze centrale rol van de datagedreven aanpak heeft onder andere geresulteerd in de oprichting van de afdeling Datafundamenten en analytics (DF&A).

De Belastingdienst verwacht met risicomodellen de aangiften tegen lagere kosten te kunnen controleren en tegelijkertijd de kans op noodzakelijke correcties in bijvoorbeeld de inkomstenbelasting en btw te vergroten. Voor de inkomstenbelasting beoogde de Investeringsagenda een toename van de belastingontvangsten. Voor de btw lag de nadruk op het voorkomen van onjuiste teruggaven bij negatieve btw-aangiften. De staatssecretaris beoogde daarmee het nalevingstekort (de gemiste belastingontvangsten door onvolledige of onjuiste aangifte) te verkleinen, tegen lagere kosten. De inzet van de Investeringsagenda was dus dat de Belastingdienst met risicomodellen tegen geringere kosten de tax-gap kan verkleinen.

De Investeringsagenda is in 2017 opgevolgd door een herijking van de Investeringsagenda en in 2018 is deze agenda door het nieuwe kabinet in afgeslankte vorm voortgezet onder de naam “beheerst vernieuwen”. De oorspronkelijke doelstellingen van de Investeringsagenda voor het gebruik van data in de controle- en toezichthoudende taak zijn niet veranderd en inmiddels belegd bij de afdeling DF&A van de Belastingdienst. Deze nieuwe afdeling verzamelt alle data binnen de Belastingdienst en maakt de data vergelijkbaar zodat ze bruikbaar zijn in een risicomodel en bij de uitvoering van andere taken van de Belastingdienst.

1.2 Leeswijzer

Dit rapport over de werking van twee risicomodellen in het handhavingsbeleid bestaat uit twee onderdelen. We behandelen achtereenvolgens de risicomodellen (hoofdstuk 2) en de waarborgen omtrent het gebruik van data in deze modellen (hoofdstuk 3).

In hoofdstuk 2 beschrijven we de modellen voor de inkomstenbelasting en de btw-teruggave. We onderzoeken de toegevoegde waarde van deze modellen voor de uitvoerings- en toezichtstrategie, het gebruik van deze modellen door de controleurs binnen de Belastingdienst en welke rol de modellen spelen bij de inzet van controlecapaciteit.
De onderzoeksvragen in dit hoofdstuk zijn:

1. Op welke wijze geeft de Belastingdienst invulling aan het gebruik van data-analyse in het handhavingsbeleid in het segment Particulieren en het segment MKB?

2. Kan de Belastingdienst aantonen dat de verwachte opbrengsten in termen van financiële besparingen en een hogere trefkans worden gerealiseerd? Wat is daarmee de meeropbrengst van deze methode vergeleken met klassieke handhavingsmiddelen zoals boekenonderzoeken en controles met selectieregels?

In hoofdstuk 3 beoordelen we de maatschappelijke en ethische aspecten die een rol spelen bij het werken met data-analyse en risicomodellen en hoe daarmee wordt omgegaan binnen de Belastingdienst. De potentie van de toepassing van risicomodellen en data-analyse is groot. Tegelijkertijd bestaat het risico dat onvoldoende rekening wordt gehouden met privacyaspecten van het werken met zo veel persoonlijke gegevens. De onderzoeksvragen in dit hoofdstuk zijn:

1. Zijn maatregelen getroffen waarmee de privacy van de belastingplichtigen binnen de modellen en binnen de processen is gewaarborgd?
2. Welke rode lijnen hanteert de Belastingdienst bij het gebruik van data en wie stelt deze vast?

Hoofdstuk 4 bevat onze conclusies en aanbevelingen.

2 Risicomodellen

2.1 Inleiding

Een volledige controle van alle ontvangen belasting- en premie aangiften is onmogelijk gezien het grote aantal aangiften dat de dienst jaarlijks ontvangt en de omvang van de Belastingdienst . Vanuit het oogpunt van doelmatigheid is een volledige controle ook onwenselijk. De Belastingdienst moet daarom kiezen welke aangiften wel en welke niet gecontroleerd worden. De datagedreven aanpak ondersteunt het selectieproces van de Belastingdienst en heeft dus als oogmerk juist die aangiften te selecteren met de grootste kans op noodzakelijke correcties. De Belastingdienst heeft deze datagedreven aanpak geïmplementeerd door middel van risicomodellen.

Een risicomodel is een berekening op basis van data die de kans op een fout in de aangifte weergeeft met een percentage. De tot standkoming van risicomodellen bij de Belastingdienst is opgenomen in bijlage 1. Na berekening van het risico worden de aangiften gesorteerd van lage naar hoge kansen op een fout in de aangifte. De Belastingdienst kan zo de meest risicovolle aangiften controleren door de aangiften te selecteren met een hoge kans op een fout. Aangiften met een klein risico worden minder vaak geselecteerd.

2.2 Het risicomodel van teruggave btw

De belasting over de toegevoegde waarde (btw) is een indirecte belasting op de verkoop van producten en diensten. Btw die een bedrijf zelf betaalt bij de aanschaf van bedrijfsmiddelen mag het bedrijf aftrekken van de ontvangen btw van afnemers. Als een bedrijf in een periode meer btw heeft betaald dan het bedrijf heeft ontvangen bij verkoop van producten en diensten ontstaat een zogenaamde ‘negatieve’ btw aangifte. De betaalde btw kan in dit geval worden teruggevraagd. Onjuiste btw-teruggaven die niet door de inspecteur worden opgemerkt, leiden voor de overheid tot een onrechtmatige en onjuiste uitgaande kasstroom. Het risicomodel bepaalt van alle btw-teruggaven het risico op onjuistheden. De inspecteur controleert de aangiften met de hoogste risico’s.

2.1.2 Het risicomodel van de inkomstenbelasting

De inkomstenbelasting wordt geheven over de som van het inkomen uit salaris, pensioen, lijfrente, alimentatie, rente en/of dividend. De inkomstenbelasting bestaat uit drie hoofdbestanddelen, de zogenaamde boxen. Box 1 is het inkomen uit werk en woning, box 2 is het inkomen uit aanmerkelijk belang en box 3 is het inkomen uit vermogen. Deze gegevens

in de aangifte kan de Belastingdienst controleren met gegevens uit externe bronnen. Voor box 1 zijn bijvoorbeeld de loongegevens beschikbaar voor werknemers in loondienst voor een controle van de juistheid van het aangegeven loon en voor box 3 zijn de standen van de spaar- en beleggingsrekeningen beschikbaar voor een controle van de juistheid van het aangegeven vermogen. Met deze gegevens is het mogelijk een risicomodel op te stellen.

De Belastingdienst zet voor de selectie van te controleren aangiften bij de inkomstenbelasting twee instrumenten in: risicomodellen en selectieregels.

Het risicomodel voor de inkomstenbelasting berekent aan de hand van heel veel variabelen de kans op een correctie in de aangifte. Die aangiften met de hoogste kans worden vervolgens gecontroleerd.

Een tweede instrument voor de Belastingdienst om aangiften voor de inkomstenbelasting te selecteren zijn de 'selectieregels'. Selectieregels detecteren afwijkingen van normbedragen of inconsistenties in de aangiften. Deze selectieregels zijn gebaseerd op kennis van fiscale wet- en regelgeving en worden niet afgeleid uit historische data. Deze regels zijn te beschouwen als de voorloper van risicomodellen.

Een voorbeeld van een vergelijking met een normbedrag is de giftenaftrek. Deze aftrek mag niet hoger zijn dan 10% van het (gezamenlijk) belastbaar inkomen. Als de aftrek boven de drempel uitkomt, komt de aangifte in de 'uitworp' terecht. Met uitworp duidt de Belastingdienst die aangiften aan waarvan een selectieregel een signaal heeft gegeven dat ze mogelijk onjuistheden bevatten. Deze aangiften worden apart gelegd voor nadere controle door de controleur of inspecteur.

Een ander voorbeeld is de vergelijking van het opgegeven vermogen met de saldi van spaar- en beleggingsrekeningen met zogeheten contra-informatie. Als het opgegeven vermogen afwijkt van het totaal van de saldi van spaar- en beleggingsrekeningen, komt een aangifte in de uitworp terecht.

De aanvankelijke gedachte van de Investeringsagenda was dat risicomodellen de selectieregels zouden vervangen. Dit idee is inmiddels verlaten. In de praktijk bleek dat de controleurs van aangiften slecht uit de voeten kunnen met de instructie om de gehele aangifte te bekijken. De selectie van de aangiften door de selectieregels wordt nu gecombineerd met het risicomodel in de risicomatrix. De risicomatrix is een visualisatie van de scores op basis van de selectieregels en de scores van het risicomodel. Een voorbeeld van een risicomatrix hebben we opgenomen en nader uitgelegd in bijlage 2.

2.1.3 Trefkans

De vraag is of de nieuwe risicomodellen een grotere trefkans hebben om een aangifte met een fout te detecteren dan de oude manier van werken met selectieregels. Bij de btw-teruggave is deze vraag wel te beantwoorden, maar bij de inkomstenbelasting niet. Bij de btw-teruggave heeft het risicomodel de oude selectieregels vervangen. We kunnen de oude manier van werken daarom vergelijken met het risicomodel.

Wij hebben geconstateerd dat de trefkans bij btw-teruggave met het nieuwe risicomodel groter is dan met het oude model. Dit kan de Belastingdienst aantonen met een steekproef die een integrale controle van de geselecteerde aangiften omvat. Daarbij worden de twee methoden met elkaar vergeleken. Als we de oude manier van handhaven vergelijken met de manier gebaseerd op het risicomodel, blijkt dat meer aangiften worden onderzocht. In 2014 en 2015 zijn jaarlijks 32.000 aangiften onderzocht (Belastingdienst, 2017), met het nieuwe model gebaseerd op risicoanalyse worden jaarlijks 40.000 aangiften onderzocht. De trefkans van het oude model voor btw-teruggave was 17%, met het nieuwe model wordt een trefkans van 34% gehaald. De Belastingdienst schat de meeropbrengst van het nieuwe model op € 60 miljoen per jaar.

Bij de inkomstenbelasting kan de oude manier van werken niet worden vergeleken met het risicomodel. Daar zijn de volgende vier redenen voor:

1. De Belastingdienst gebruikt zowel de risicoscores als de selectieregels bij de selectie van aangiften.
2. Een deel van de controlecapaciteit is gereserveerd voor het uitvoeren van steekproeven waarbij is voorgeschreven op welke punten de aangifte extra gecontroleerd wordt. De steekproef kan hierdoor niet volledig risicogericht worden ingezet.
3. De selectieregels zijn zo geconstrueerd dat ze extra opbrengsten voor de schatkist kunnen detecteren, maar ook kunnen signaleren of de belastingplichtige geld terug-krijgt.
4. Het risicomodel beoogt uitsluitend extra opbrengsten voor de schatkist te genereren; als criterium wordt gebruikt dat mogelijke opbrengsten voor de schatkist worden gedetecteerd.

2.2 De Investeringsagenda in relatie tot de uitvoerings- en toezichtstrategie

De uitvoerings- en toezichtstrategie van de Belastingdienst heeft als doel het vergroten van ‘compliance’. Dat wil zeggen dat burgers en bedrijven fiscale regels naleven, zonder dat daarvoor dwingende (corrigerende) en acties van de Belastingdienst nodig zijn. De

Belastingdienst bevordert naleving bij de belastingplichtigen door onder andere voorlichting te geven en het doen van aangifte te faciliteren. Deze vooral preventieve aanpak maakt het burgers en bedrijven gemakkelijk om de wet- en regelgeving na te leven. Een voorbeeld van deze benadering is de vooringevulde aangifte (VIA), waarbij de Belastingdienst zo veel mogelijk gegevens vooraf invult. Een ander voorbeeld is de voorlichting die de Belastingdienst geeft aan startende ondernemers.

Een tweede element in de uitvoerings- en toezichtstrategie van de Belastingdienst is een doelgroepbenadering. Daarbij wordt onderscheid gemaakt tussen bijvoorbeeld ondernemers en particulieren. Dit onderscheid zien we ook terug bij de ontwikkeling van op deze doelgroepen gerichte data-analyses en risicomodellen.

De risicomodellen bij de Belastingdienst zijn in 2015 geïntroduceerd in de Investeringsagenda (Financiën 2015). Het doel van de introductie en het gebruik van risicomodellen gebaseerd op data-analyse was het verhogen van de efficiëntie van de handhaving en het genereren van extra belastingopbrengsten. De verwachting was dat de maatregelen in de Investeringsagenda de tax-gap met € 750 miljoen op jaarbasis[1] zouden verkleinen.

In de Investeringsagenda werden vijf positieve resultaten van deze nieuwe aanpak genoemd:

- De trefkans ligt aanzienlijk hoger.
- Het systeem geeft de aard van het risico aan.
- De behandeltijd van de medewerker gaat omlaag.
- Elk bekeken geval levert meer op.
- Er kunnen meer gevallen worden bekeken.

De doelstelling voor de introductie van risicomodellen in de Investeringsagenda lijkt dus niet geheel in overeenstemming met de doelstelling van de huidige uitvoerings- en toezichtstrategie, omdat deze strategie meer omvat dan alleen belastingontvangsten te vergroten.

2.3 Risicomodellen en de Investeringsagenda

De risicomodellen, die ontwikkeld zijn op basis van het opbrengstenmaximalisatiedoel uit de Investeringsagenda, signaleren alleen correcties die de opbrengst van de schatkist verhogen en geen correcties waarbij de burger te veel betaalde belasting terugkrijgt. In de uitvoering komen deze risicomodellen samen met de selectieregels die zowel positieve als

negatieve correcties detecteren. Hier bestaat dus een spanningsveld tussen het bevorderen van nalevingsgedrag volgens de uitvoerings- en toezichtstrategie en het verkleinen van het nalevingstekort dat centraal stond in de Investeringsagenda. Aangiften waarbij de burger door onjuist aangifte te doen te veel belasting betaalt, worden door het nu gebruikte risicomodel alleen als bijvangst gecontroleerd.

Doordat het model teruggaven onvoldoende detecteert, hebben deze aangiften een kleinere kans om door een controleur gecorrigeerd te worden. Deze correcties, waarbij de burger geld terugkrijgt, verlagen het belastbaar inkomen. Dit gegeven wordt uiteindelijk opgenomen als geregistreerd inkomensgegeven in de Basisregistratie Inkomen (BRI). Deze registratie wordt gebruikt bij de toekenning van inkomensafhankelijke regelingen zoals toeslagen, studiefinanciering of de eigen bijdrage bij langdurige zorg. Daardoor kan een onjuist geregistreerd inkomen financiële gevolgen hebben.

2.4 Topicmodellen

Wij hebben geconstateerd dat het gebruik van het risicomodel bij de inkomstenbelasting voor ondernemers kan worden verbeterd. In de praktijk blijken de controleurs van de dossiers moeilijk met het risicomodel te kunnen werken, omdat de reden van de selectie door het risicomodel niet bij hen bekend is. Zij moeten dus de gehele aangifte door om te achterhalen of en, zo ja, welke correctie nodig is. Bij het werken volgens de oorspronkelijke selectiemethode met selectieregels is de reden van selectie altijd bekend bij de controleurs en daardoor voor hen makkelijker in de praktijk te gebruiken.

De introductie en het gebruik van de risicomodellen verloopt bij de inkomstenbelasting dus moeizaam, omdat de uitkomsten van het risicomodel geen specifiek onderdeel in de aangifte aanwijzen. De Belastingdienst heeft dit probleem gesignaleerd en probeert het te ondervangen door zogeheten topicmodellen te bouwen. Een topicmodel is een risicomodel voor een gedeelte van de aangifte. Een topicmodel voorspelt een fout in een specifiek gedeelte van de aangifte, bijvoorbeeld in de rubriek eigen woning. Er zijn meerdere topicmodellen, die het risico voor elk topic berekenen. Het is mogelijk dat een aangifte op een of meerdere topics een hoog risico heeft. Deze topicmodellen zijn er voor de aangifte inkomstenbelasting voor particulieren, maar nog niet voor de aangifte inkomstenbelasting voor ondernemers en aangiften van directeuren-grootaandeelhouders. Bij deze aangiften gebruiken de controleurs de risicoscores van het risicomodel slechts zeer beperkt en soms zelfs helemaal niet. Deze aangiften zijn vaak ingewikkeld zodat een integrale behandeling arbeidsintensief is. De controleurs vallen daardoor vaak terug op de oude selectieregels.

De introductie van topicmodellen kan het gebruik van het risicomodel voor ondernemers en directeuren-grootaandeelhouders verbeteren.

2.5 Capaciteit handhaving

In deze paragraaf bespreken we achtereenvolgens de controlecapaciteit bij de inkomstenbelasting, de capaciteit bij de btw-teruggave en de relatie tussen de grenswaarde van het risico in de aangifte en de capaciteit.

2.5.1 Capaciteit bij de inkomstenbelasting

Jaarlijks detecteert de Belastingdienst bij ruim een miljoen aangiften in de inkomstenbelasting, één of meer mogelijke fouten in de aangifte. Deze aangiften waarbij mogelijk iets mis is, worden aangeduid met de term 'uitworp'. Op peildatum 25 oktober 2016 bedroeg het aantal uitgeworpen aangiften bij de inkomstenbelasting meer dan een miljoen (Belastingdienst, 2016a). Dit betrof de uitworp voor het belastingjaar 2015. Eind 2016 bleek dat deze werkvoorraad te hoog was en is besloten het aantal te controleren aangiften te reduceren. De voorraad van 816.000 aangiften is daardoor gereduceerd tot 350.000. Bij deze reductie is gebruikgemaakt van de risicomatrix, een combinatie van het risicomodel met de selectieregels (zie § 2.1.2 en bijlage 2).

Bij de reductie van de werkvoorraad aangiften is de beschikbare handhavingscapaciteit leidend. Het bedrag aan gemiste belastinginkomsten speelt slechts een kleine rol bij de 'doorselectie'. De Belastingdienst selecteert met de selectieregels en het risicomodel veel meer aangiften dan er daadwerkelijk (kunnen) worden gecontroleerd. De Belastingdienst heeft te weinig capaciteit om alle aangiften waar mogelijk iets mis mee is te controleren. Dat is ook niet per definitie doelmatig. Om de optimale balans tussen handhavingscapaciteit en doelmatigheid te kunnen bepalen is het begrip grenswaarde van belang. In § 2.5.3 gaan wij daar nader op in.

2.5.2 Capaciteit bij btw-teruggave

In 2018 vonden 40.000 onderzoeken op de uitworp van btw-teruggave plaats, voor 2019 staan er vooralsnog 50.000 gepland. Deze uitworp wordt bepaald met het risicomodel. De uitworp wordt onderzocht met een 'kantoortoets'. Dit zijn onderzoeken waarbij gebruikgemaakt wordt van informatie die al bij de Belastingdienst bekend is, zoals aangiften van e onderneming uit het verleden. Als de kantoortoets daar aanleiding toe geeft, volgt een boekenonderzoek. De capaciteit van het aantal kantoortoetsen is een gegeven vooraf, er wordt geen afweging gemaakt tussen extra capaciteit en extra opbrengst. De Belastingdienst verdeelt zijn toezicht en handhavingscapaciteit over belastingplichtigen, belastingmiddelen

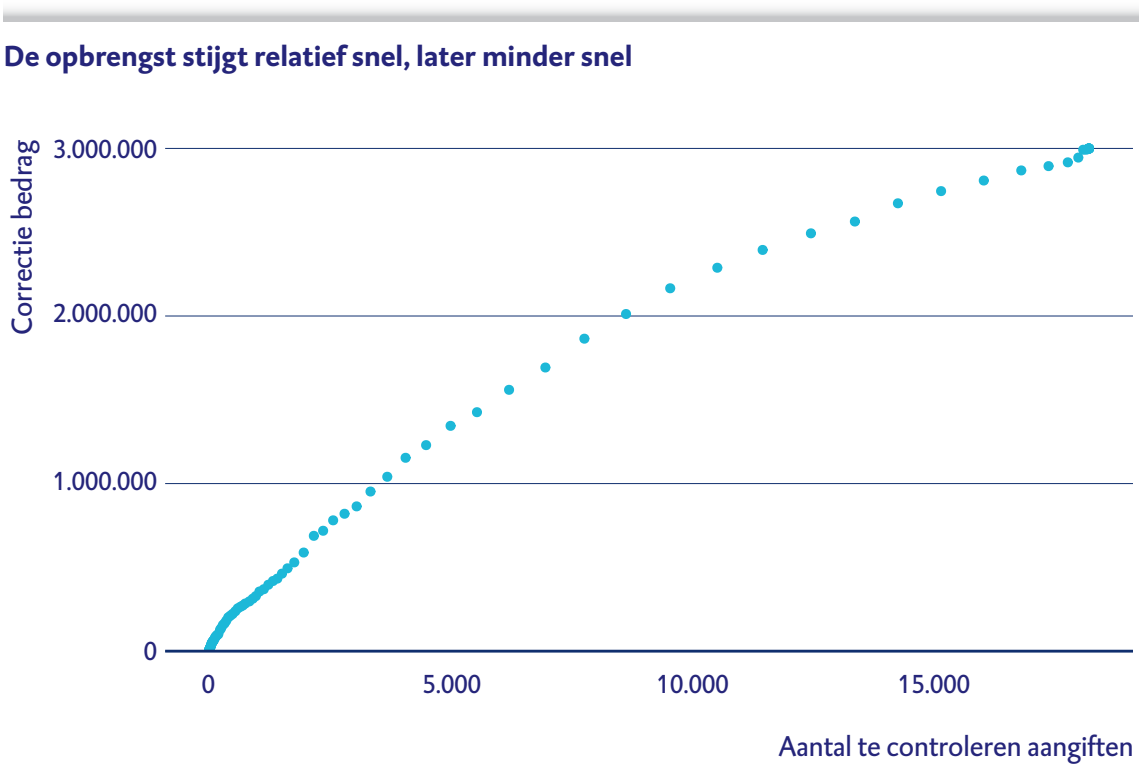
en handhavingsmiddelen. Meer capaciteit voor de controle van de btw-teruggave gaat ten koste van de andere werkzaamheden, zoals startende ondernemers begeleiden met de aangifte of controlewerkzaamheden op de btw-aangiften met een wellicht te laag positief bedrag.

2.5.3 Grenswaarde in het risicomodel van de inkomstenbelasting

Om te bepalen of de aangifte gecontroleerd moet worden, is een 'grenswaarde' voor het risico nodig. In deze paragraaf bespreken we de grenswaarde bij de inkomstenbelasting, voor de grenswaarde bij de btw verwijzen we naar bijlage 3. Aangiften met een risico lager dan de grenswaarde worden niet gecontroleerd en aangiften met een risico groter dan de grenswaarde worden wel gecontroleerd. De grenswaarde bepaalt welk risico nog acceptabel is maar tegelijkertijd dus ook hoeveel aangiften gecontroleerd moeten worden.

De grenswaarde kan zo gekozen worden dat de hoeveelheid te controleren aangiften precies aansluit op de capaciteit. In dat geval worden de hoogste risico's geselecteerd en gecontroleerd. Een tweede manier om de grenswaarde te kiezen is door deze als norm te hanteren. Aangiften met een risico hoger dan de grenswaarde worden altijd gecontroleerd om ervoor te zorgen dat de belastingplichtige juist en volledig aangifte doet. Dan wordt de capaciteit afgestemd op het risico en niet andersom.

Deze analyse hebben we toegepast op het model met de gegevens van de Belastingdienst, door de grenswaarde van de risicoscore in stapjes te verhogen. Bij iedere grenswaarde is de verwachte correctieopbrengst en het aantal te beoordelen aangiften bepaald. De grenswaarden zijn toegepast op het door de Belastingdienst gebruikte risicomodel. De resultaten staan in figuur 1.



Figuur 1 *Aantallen te controleren aangiften en correctiebedragen bij de inkomstenbelasting particulieren*

Als de Belastingdienst meer aangiften controleert, levert dat altijd meer opbrengsten op. Linksonder in de grafiek staan de meest risicovolle aangiften afgebeeld en stijgt de opbrengst snel. Daarna stijgt de opbrengst wel, maar minder snel. Uit de grafiek blijkt, dat een controle van 1.600 aangiften nodig is om een opbrengst van € 500.000 te realiseren. Om de opbrengst te verdubbelen, moet de Belastingdienst 3.700 aangiften met het hoogste risico onderzoeken. De benodigde capaciteit om de opbrengst twee keer zo groot te maken moet meer dan verdubbelen. Er is dus sprake van afnemende meeropbrengsten, maar dit is geen bijzonder sterk effect. De conclusie uit ons onderzoek is dat er vanuit het oogpunt van doelmatigheid van tevoren geen bepaald aantal aangiften aan te wijzen is waarbij de extra opbrengsten niet meer opwegen tegen de extra controle inspanningen. Voor een vergelijkbare analyse van de btw-teruggave verwijzen we naar bijlage 3.

2.5.4 Afweging capaciteit handhaving versus additionele opbrengsten

Selectie van aangiften met een grote kans op een fout levert een kleine werkvoorraad op, waarbij relatief weinig aangiften gecontroleerd worden. Als de Belastingdienst de werkvoorraad vergroot, betekent dit dat aangiften geselecteerd worden waarbij een fout in de aangifte minder zeker is.

Wij zijn geen analyses tegengekomen waarin de extra kosten tegen de extra opbrengsten worden afgewogen. Wel weegt de Belastingdienst de opbrengsten van verschillende modellen tegen elkaar af, waarbij ook het aantal controles een rol speelt. DF&A analyseert hoeveel een selectie oplevert en hoeveel belastingplichtigen gecontroleerd worden, maar de beschikbare capaciteit is in deze analyses altijd leidend. De vraag hoeveel controles nodig zijn om de uitvoerings- en toezichtstrategie succesvol te laten zijn, wordt in deze analyses niet gesteld.

2.6 Omvang reductie *tax-gap*

De Belastingdienst meet de meeropbrengsten van de nieuwe risicomodellen voor de inkomstenbelasting en btw-teruggave. De bepaling van de meeropbrengsten van het model voor de btw-teruggave is verder gevorderd dan voor het model voor de inkomstenbelasting. De Belastingdienst heeft een methode ontwikkeld om de extra opbrengsten bij het gebruik van deze modellen in de praktijk vast te stellen. Hiervan zijn op dit moment nog geen resultaten bekend.

Het aantal risicomodellen dat de Belastingdienst tot zijn beschikking heeft, is nog beperkt en de modellen dekken slechts een klein deel van de belastinginkomsten af. Bij de btw-aangiften wordt het model voor btw teruggave gebruikt in de handhaving. Het risicomodel voor alle btw-aangiften is nog in ontwikkeling. Het gaat om een stroom van ongeveer € 50 miljard belastinginkomsten op jaarbasis. Ook voor de vennootschapsbelasting is nog geen risicomodel ontwikkeld om de meest risicovolle aangiften te selecteren.

De correcties die met deze modellen bij de inkomstenbelasting gevonden worden, zijn relatief klein. Meestal gaat het om onjuistheden in de aftrekposten. De vraag of de tax-gap met € 750 miljoen verkleind kan worden, zoals in 2015 werd verwacht, kan op dit moment niet worden beantwoord.

3 Verantwoord omgaan met data

De gegevens die de Belastingdienst beheert, zijn privacygevoelig en specifiek bij bedrijven concurrentiegevoelig. Deze gegevens kunnen alleen gebruikt worden als de juiste technische en organisatorische beheersmaatregelen zijn getroffen. Op dit moment is geen normenkader beschikbaar voor data-analyse en het verwerken van gegevens in risicomodellen. We hebben daarom in kaart gebracht welke normen in het maatschappelijk verkeer worden gehanteerd met betrekking tot een zorgvuldige omgang met (persoons)gegevens in risicomodellen of – breder – algoritmes. Eén van de meest relevante normen voor een zorgvuldige omgang met persoonsgegevens is de Algemene Verordening Gegevensbescherming (AVG), die per 25 mei 2018 in werking is getreden.

In dit hoofdstuk bespreken we of de afdeling DF&A voldoende technische en organisatorische beheersmaatregelen heeft getroffen om verantwoord om te gaan met data. Ook gaan we in op een aantal dilemma's dat het gebruik van risicomodellen oproept.

3.1 Opzet privacybeleid

Een belangrijk instrument in het naleven van de AVG is de Data Protection Impact Assessment of Privacy Impact Assessment (DPIA), in dit rapport wordt de meer gangbare afkorting PIA (Privacy Impact Assesment) gebruikt. Hiervoor bestaat een rijksbreed sjabloon (Rijksoverheid, 2017). Met een PIA wordt op een gestructureerde en gestandaardiseerde wijze de gevolgen van verwerking van persoonsgegevens voor de betrokkenen beoordeeld. Een PIA moet voorafgaand aan het gebruik van data worden uitgevoerd. De uitkomsten ervan bepalen welke maatregelen moeten worden genomen om te voldoen aan de privacyregelgeving.

Een PIA bestaat uit de volgende vier elementen (Rijksoverheid, 2017):

1. Een beschrijving van de voorgenomen verwerkingen en de verwerkingsdoeleinden.
2. Een beoordeling van de rechtsgrond, de noodzaak, evenredigheid en verenigbaarheid van de voorgenomen verwerkingen in relatie tot de verwerkingsdoeleinden.
3. Een beoordeling van de gevolgen en risico's van de voorgenomen verwerkingen voor de rechten en vrijheden van de betrokkenen.
4. De voorgenomen maatregelen om deze gevolgen en risico's van de voorgenomen verwerkingen aan te pakken.

Wij constateren dat de Belastingdienst de rijksbreed vastgestelde procedure voor het uitvoeren van PIA's volgt en daarmee in opzet voldoet aan de gestelde eisen. Tegelijkertijd zien wij dat de PIA-procedure verbeterd kan worden: het management moet actief en zichtbaar betrokken zijn bij het vaststellen en handhaven van de resultaten van de PIA.

Conform de binnen de Belastingdienst vastgestelde procedure is de lijndirecteur verantwoordelijk voor de naleving van de privacy en bescherming van de persoonsgegevens binnen de afdeling. De lijndirecteur hoort de PIA's vast te stellen en de aanbevelingen op te volgen. De PIA-rapportages worden in de meeste gevallen aan het managementteam van DF&A gericht. Wij hebben voor de onderzochte periode echter niet kunnen vaststellen dat het managementteam van DF&A de PIA-rapportages heeft besproken en de opvolging van de aanbevelingen heeft bewaakt. Voor zover wij konden vaststellen, heeft het gebrek aan betrokkenheid van het managementteam ook niet geleid tot problematische situaties met betrekking tot de bescherming van de privacy van Nederlandse burgers.

3.2 Fysieke toegang en informatiebeveiliging

We kregen een aantal maanden toegang tot de afdeling om het onderzoek uit te voeren. Uit observaties ter plaatse blijkt dat de fysieke toegang tot informatie goed is beveiligd. Zo zijn de kantoren van de afdeling waar met data wordt gewerkt alleen te betreden met een specifiek kenmerk op de toegangspas.

Een medewerker heeft alleen toegang tot die data die hij op dat moment nodig heeft. Data worden opgeslagen in datagebieden die per rol van de medewerker geautoriseerd worden. Wanneer medewerkers op grond van hun rol niet geautoriseerd zijn, dan kunnen zij deze data niet inzien of bewerken. Medewerkers die zich bezighouden met het datagebied omzetbelasting hebben bijvoorbeeld geen toegang tot het datagebied inkomensheffing. Daarnaast is er toezicht op het gebruik van data door de medewerkers. Zo wordt bijvoorbeeld gemonitord of er op specifieke burgerservicenummers (BSN's) gezocht wordt, zodat later onderzocht kan worden of dit voor het werk noodzakelijk was.
Het mailverbod zoals dat op de afdeling gold was een organisatorische beheersmaatregel ('je mag niet mailen') en geen technische beheersmaatregel ('je kan niet mailen'). Ten tijde van het onderzoek bleek dat het nog mogelijk was bestanden naar email adressen buiten de Belastingdienst te mailen. Naar aanleiding van ons onderzoek is dit volgens de Belastingdienst technisch onmogelijk gemaakt.

3.3 Hoeveelheid verzamelde gegevens

Het aanleggen van grote gegevensverzamelingen door bedrijven zonder dat de noodzaak daartoe is aangetoond, wordt ook wel aangeduid met de term ‘datahonger’. Die datahonger wordt bij de Belastingdienst ingeperkt door het vereiste dat voor het verzamelen, gebruiken en bewerken van gegevens altijd een wettelijke basis nodig is. Dit geldt dus ook voor de data die DF&A verwerkt in de risicomodellen.

De Belastingdienst mag wettelijk veel gegevens verzamelen en gebruiken, zoals bankgegevens, gegevens van het kadaster, loongegevens van de werkgever et cetera. Voor zover wij hebben kunnen vaststellen, gebruikt DF&A geen gegevens waarvoor de wettelijke basis ontbreekt.

3.4 Ongewenste neveneffecten van risicomodellen

Bij het gebruik van een risicomodel spelen naast technische kwaliteitscriteria ook maatschappelijke criteria een rol. Een maatschappelijk vereiste is dat de Belastingdienst risicomodellen gebruikt die alleen risico’s in de aangifte detecteren en verder neutraal zijn voor andere kenmerken.

Volgens de Wetenschappelijke Raad voor het regeringsbeleid (WRR) is profilering, de ene bevolkingsgroep intensiever controleren dan de andere, een onbedoelde en ongewenste uitkomst van het gebruik van risicomodellen (Wetenschappelijke Raad voor het Regeringsbeleid, 2016). Wanneer dat ongewenste onderscheid niet direct in het risicomodel wordt gemaakt, maar via een indirect kenmerk in het model wordt opgenomen, spreken we van een proxy. Proxies zijn dus kenmerken die neutraal lijken, bijdragen aan het voorspellen van het risico, maar indirect leiden tot een hogere selectiekans van leden van een bepaalde bevolkingsgroep. Een voorbeeld van een proxy bij de Belastingdienst is “personen met een recent afgegeven BSN”. Het jaar van afgifte is informatief want een BSN ontvang je na inschrijving in de Basisregistratie Personen. Een recent verstrekt BSN behoort dus toe aan pasgeborenen of aan migranten. Bij een belastingaangifte is een jong BSN daarom een duidelijke indicatie dat de belastingplichtige een migrant kan zijn. Het gebruik van dit gegeven in een risicomodel leidt tot een hogere selectiekans als migranten vaker fouten maken of een lagere selectiekans als migranten minder fouten maken dan zij die in Nederland geboren zijn. Van tevoren kunnen we niet vaststellen of de selectiekans hoger of lager is, omdat daarvoor eerst het risicomodel moet worden berekend op basis van historische gegevens.

DF&A erkent dat het gebruik van dergelijke proxies ongewenst kan zijn; binnen het privacybeleid van DF&A wordt daar op gelet. In dit specifieke geval van een 'jong BSN' wordt deze proxy in het risicomodel vervangen door een variabele "is dit een eerste aangifte". Dit gegeven geldt niet alleen voor migranten, maar voor iedereen die voor het eerst aangifte doet. De vraag of op deze manier het risico van profilering is afgewend, is niet eenduidig te beantwoorden. Het gebruik van proxies moet worden afgewogen tegen (het belang van) de doelstelling van het model en de vergroting van de selectiekans van het risicomodel. Bekende proxies, zoals de postcode of de voor- of achternaam van de belastingplichtige, zijn we in de onderzochte modellen niet tegengekomen.

In de huidige situatie wegen de projectleider en de privacyfunctionaris op basis van een ingevulde PIA af in hoeverre sprake is van profilering of het gebruik van een proxy en of DF&A daarbij handelt binnen maatschappelijk aanvaarde normen. Gezien het maatschappelijk belang van dergelijke beslissingen, verdient het naar onze mening de voorkeur dat het managementteam van de afdeling DF&A deze afweging maakt, waarbij ook de Functionaris Gegevensbescherming van het kerndepartement wordt betrokken. Juist vanwege de potentiële gevolgen van dit soort afwegingen en het lopende maatschappelijk debat over het gebruik van algoritmes door de overheid, zouden deze afwegingen afgeleid moeten worden van een rijksbreed kader. Dit kader ontbreekt op dit moment.

3.5 Geautomatiseerde besluitvorming bij risicomodellen

Van geautomatiseerde besluitvorming is sprake wanneer een besluit wordt genomen op basis van de uitkomst van een risicomodel, zonder menselijke tussenkomst.

Bij DF&A hebben wij geen geautomatiseerde besluitvorming op grond van de risicomodellen aangetroffen. De risicomodellen leveren wel een risicoscore op, op basis waarvan menselijke controleurs (bijvoorbeeld inspecteurs) besluiten om de aangifte al dan niet verder te onderzoeken. Na controle van de aangifte bepaalt de inspecteur of hij deze moet corrigeren en hoe hoog de correctie is.

Op dit moment zijn de risicomodellen nog niet in staat tot geautomatiseerde besluitvorming. De trefkans is namelijk niet hoog genoeg of betrouwbaar genoeg om blindelings te kunnen vertrouwen op het systeem. Gezien de snelle ontwikkelingen van risicomodellen is het de vraag of dat in de toekomst zo blijft.

We zien overigens wel automatische besluitvorming bij andere onderdelen van de Belastingdienst. Besluitvorming bij Massaal Automatisch Corrigeren (MAC) leidt tot het sturen van een brief waarna de belastingplichtige bezwaar kan maken tegen de correctie. Correcties die op deze manier worden opgelegd zijn vaak eenvoudig van aard, bijvoorbeeld als een aantal deelbedragen niet optellen tot het totaalbedrag. Als de burger het niet met deze correctie eens is, kan hij bezwaar maken.

3.6 Pseudonimiseren van data

Bij het bewerken van data om te komen tot de risicomodellen is het vaak niet nodig om weten op wie de gegevens betrekking hebben. Het gaat immers om het herkennen van patronen en verbanden. Het verwijderen van het BSN en de naam, adres, woonplaats (NAW) gegevens verhoogt de privacy van de personen van wie de gegevens worden verwerkt. Het doel van deze verwijdering is dat medewerkers die toegang hebben tot deze bestanden gegevens niet meer kunnen herleiden tot personen. Zij kunnen dan ook geen gegevens van hen bekende personen opvragen. Het BSN kan in een dergelijk geval vervangen worden door een pseudoniem, een unieke sleutel die in een later stadium koppeling met adresgegevens mogelijk maakt. Door het BSN los te koppelen van data is niet meer na te gaan om wie het gaat. Pas als actie moet worden ondernomen, zoals het sturen van een brief, wordt de versleuteling ongedaan gemaakt zodat duidelijk is om wie het gaat en waar de brief naar toe moet.

Op dit moment zijn de data bij de Belastingdienst nog niet gepseudonimiseerd. Er zijn al wel projecten gestart om dit op korte termijn te realiseren, en zo de privacy van de belasting-betaler te waarborgen.

4 Samenvatting, conclusies en aanbevelingen

De Belastingdienst is begonnen een werkwijze te ontwikkelen voor het gebruik van data en risicomodellen. De dienst is in staat moderne technieken te ontwikkelen en die toe te passen. Vooral de invoeging van dit nieuwe model voor btw-teruggave in het oude legacy-systeem van de btw toont het succes. Daarnaast constateren we dat de verantwoordelijkheid voor privacy aspecten van het werken met data serieus wordt opgepakt met technische maatregelen en men begonnen is met organisatorische maatregelen.

Het gebruik van risicomodellen zoals opgenomen in de Investeringsagenda sluit echter niet aan op de huidige uitvoerings- en toezichtstrategie van de Belastingdienst: het handhavingsbeleid is gericht op het bevorderen van compliant gedrag van de belastingplichtige, de inzet van de Investeringsagenda was gericht op het verkleinen van de tax-gap. Deze tegenstelling is terug te zien in de uitvoering van het handhavingsbeleid. De modellen voor de inkomstenbelasting worden getraind op positieve correcties, wat in lijn is met de doelstelling van de Investeringsagenda. Compliance veronderstelt een correcte vaststelling van het belastbaar inkomen, waarvoor zowel positieve als negatieve correcties nodig zijn. De modellen zijn daarvoor niet toegerust.

4.1 Verbeterpunten in de risicomodellen

De selectie van aangiften met behulp van risicomodellen kan een nuttige toevoeging in de uitvoerings- en toezichtstrategie zijn. Zo is de trefkans van een incorrecte btw-aangifte bij het datagedreven btw-model groter dan bij de oude selectieregels. Wij zien twee verbeterpunten bij het gebruik van deze modellen in de inkomstenbelasting. Ten eerste zijn de risicomodellen bij de inkomstenbelasting voor ondernemers te weinig specifiek en hebben controleurs meer inzicht nodig welke onderdelen in de aangifte risicovol zijn. Ten tweede zijn deze modellen in de inkomstenbelasting gericht op het verhogen van belastingopbrengsten en komen hogere belastingteruggaven hooguit als bijvangst in de te controleren aangiften terecht. Hierdoor sluiten deze modellen onvoldoende aan op de doelstelling van de uitvoerings- en toezichtstrategie.

Aanbevelingen bij de risicomodellen

We bevelen de staatssecretaris aan de risicomodellen niet alleen te gebruiken voor positieve correcties, maar ook geschikt te maken voor belastingteruggave (negatieve correcties). Hierdoor sluiten de modellen beter aan op het bevorderen van compliance, het doel van de uitvoerings- en toezichtstrategie van de Belastingdienst.

Verder bevelen we de staatssecretaris de ontwikkeling aan van meer topicmodellen bij de inkomstenbelasting voor ondernemers. Die helpen de controleurs beter op weg en maken de controle voor hen minder tijdrovend. Vooral bij de aangiften voor ondernemers kan dit het risicogericht controleren een stimulans geven.

4.2 Uitgaan van risico versus capaciteit

Verhoging van de capaciteit bij de controle van aangiften leidt in de praktijk bij de Belastingdienst tot een reductie van capaciteit bij andere processen. De totale controle-capaciteit is vooraf gegeven en wordt verdeeld over de verschillende toezichtsactiviteiten. De gegeven controlecapaciteit heeft gevolgen voor het aantal aangiften dat wordt gecontroleerd. De capaciteit bepaalt de omvang van het aantal gecontroleerde aangiften, niet de grenswaarde van het risico.
Wij hebben geconstateerd dat de risicomodellen meer aangiften met grote kans op correcties kunnen opleveren dan de controlecapaciteit aankan. Als de capaciteit krapper wordt, controleert de Belastingdienst minder aangiften en wordt de controle dus soepeler dan vooraf noodzakelijk wordt geacht. Er is geen systematiek die het aantal benodigde controleurs of inspecteurs afleidt uit de geanalyseerde risico's in de ontvangen aangiften. Het verdient volgens ons een zelfstandige afweging of een verhoging van het aantal controles automatisch ten koste zou moeten gaan van andere handhavingsinstrumenten zoals voorlichting, een controle van de eerste aangifte of de inzet van menskracht bij bezwaar, en daarom achterwege kan blijven.

Aanbevelingen bij uitgaan van risico versus capaciteit
We bevelen de staatssecretaris aan de afweging tussen de behandelcapaciteit en het risico in de aangifte transparant te maken. Deze afweging moet voor de Tweede Kamer helder zijn zodat op een verantwoorde wijze een afweging kan worden gemaakt over de gewenste toezicht- en handhavingscapaciteit.

4.3 Verantwoord omgaan met data

De Belastingdienst verzamelt veel data over burgers en bedrijven bij de controle van aangiften. Dit brengt de verantwoordelijkheid met zich mee om op een zorgvuldige, correcte en maatschappelijk aanvaarde manier met deze data om te gaan. We constateren dat men zich bewust is van die verantwoordelijkheid en deze serieus oppakt door het gebruik van data te borgen met technische maatregelen en begonnen is met de implementatie van organisatorische maatregelen.

De organisatorische maatregelen die privacy moeten waarborgen zijn in opzet aanwezig. De rol van het management behoeft nog versterking bij het bewaken van de rode lijnen: de betrokkenheid van het management moet zichtbaar zijn in de besluitvorming over het gebruik van data en privacy.

Sommige van de geschetste dilemma's over het gebruik van proxies, profilering en geautomatiseerde besluitvorming moeten centraal of zelfs op rijksniveau worden vastgelegd. Wij achten het van belang dat er Belastingdienstbreed of zelfs rijksbreed toegewerkt wordt naar een set normen voor de toepassing en het gebruik van modellen gebaseerd op data-analyse. De staatssecretaris heeft de mogelijkheid hier met de binnen de Belastingdienst beschikbare kennis en ervaring een belangrijke rol in te vervullen.

De fysieke informatiebeveiliging is op orde. Het is niet mogelijk zonder autorisatie de afdeling te betreden en medewerkers hebben uitsluitend toegang tot die gegevens die zij voor hun werk nodig hebben. De toegang tot andere gegevens is technisch geblokkeerd. De gegevens die input zijn voor de risicomodellen verkrijgt de Belastingdienst op grond van zijn wettelijke taak. Medewerkers van DF&A verrijken de beschikbare data van de dienst niet op eigen initiatief met externe gegevens, zoals socialmedia-data.

We hebben geen gevallen aangetroffen van geautomatiseerde besluitvorming met behulp van deze modellen. Het is altijd de inspecteur die bepaalt hoe hoog de correctie wordt en dit motiveert aan de belastingplichtige.

Aanbevelingen bij privacybeleid

Wij bevelen de staatssecretaris aan al bij de opzet en ontwikkeling van een risicomodel, dus vooraf, de gevolgen voor de privacy en de bescherming van persoonsgegevens op een gestructureerde manier in kaart te brengen en daarmee niet te wachten tot het model in de praktijk uitgeprobeerd wordt.

Daarnaast bevelen wij de staatssecretaris aan de actieve betrokkenheid van het management te versterken bij het vaststellen en handhaven van de rode lijnen die in de Privacy Impact Assesments (PIA) door de afdeling Data Fundamenten & Analytics (DF&A) worden vastgelegd rondom het gebruik en het inzetten van data-analyse.

4.4 Reactie van de staatssecretaris

De staatssecretaris begint zijn reactie met een toelichting op de Uitvoerings- en Toezichtstrategie van de Belastingdienst. Deze strategie vertaalt de Belastingdienst naar activiteiten die gericht zijn op de naleving bij burgers en bedrijven. Op basis van het inzicht in het niveau van de naleving maakt de Belastingdienst keuzes in de inzet van instrumenten en het uitvoeren van activiteiten. Dit keuzeproces volgt de lijnen uit de strategie, namelijk het voorkomen van fouten, het vergemakkelijken van het doen van aangifte en het aanvragen van toeslagen, het afstemmen van de interventie van de Belastingdienst op het gedrag van burgers en bedrijven en, bij bewust niet naleven, het afdwingen van naleving. Data gedreven selectie is onderdeel van dit keuzeproces en de staatssecretaris wil deze modellen niet alleen gebruiken voor de selectie van aangiften maar ook om de keuzes per doelgroep te onderbouwen.
De staatssecretaris neemt de door de Algemene Rekenkamer gesignaleerde verbeterpunten van de risicomodellen over. Datafundamenten en Analytics (DF&A) is gestart met de ontwikkeling van een model dat niet alleen positieve correcties (positief voor de schatkist) maar ook negatieve correcties (waarbij de burger geld terugkrijgt) detecteert. Daarnaast onderzoekt deze afdeling hoe de controleurs ondersteund kunnen worden. Topicmodellen zijn een mogelijke manier om die ondersteuning te geven bij het gebruik van risicomodellen. Bij de aanbeveling de afweging tussen het risico en de behandelcapaciteit transparant te maken verwijst de staatssecretaris naar het Jaarplan Belastingdienst. De voortgangsrapportages over het jaarplan verschijnen halfjaarlijks, waarbij de eerste rapportage in juni 2019 zal verschijnen.

De Belastingdienst volgt de maatschappelijke ontwikkelingen die relevant zijn voor datagedreven toepassingen en participeert in een rijksbreed overleg dat het gebruik van algoritmes bespreekt. De staatsecretaris wijst op de Privacy Impact Assessments en het starten van het programma “Privacy by Design” door DF&A. Dit programma beoogt de bescherming van gegevens verder te versterken. De aanbeveling privacy aspecten vooraf te beoordelen en niet te wachten tot het model al in de praktijk wordt uitgeprobeerd is al geïmplementeerd en de aanbeveling de actieve betrokkenheid van het management te versterken wordt onderschreven. Ook ten aanzien van deze aanbeveling zijn reeds maatregelen genomen.
Ten slotte bevestigt de staatssecretaris dat het inmiddels technisch onmogelijk is voor medewerkers van DF&A om bestanden naar externe e-mailadressen te mailen.

De volledige reactie van de staatssecretaris is te vinden op de website van de Algemene Rekenkamer.

4.5 Nawoord van de Algemene Rekenkamer

De reactie van de staatssecretaris bevestigt ons beeld dat de Belastingdienst serieus inzet op het gebruik van risicomodellen bij het aangifte proces en zich bewust is van de maatschappelijke gevoeligheden bij het gebruik van algoritmes en data. De Algemene Rekenkamer wijst er volledigheidshalve nogmaals op dat juist vanwege de potentiele gevolgen van afwegingen over het gebruik van data en het lopende maatschappelijke debat over het gebruik van algoritmes door de overheid deze afwegingen afgeleid zouden moeten worden uit een rijksbreedkader.

Wij volgen de staatssecretaris niet in zijn redenering dat het jaarplan voldoende inzicht biedt bij de afweging tussen de behandelcapaciteit en het risico van de aangifte. Het Jaarplan Belastingdienst 2019 biedt inderdaad inzicht in de keuzes die de Belastingdienst maakt in de uitvoering en het toezicht. Maar deze keuzes zien vooral toe op de toedeling van capaciteit aan de verschillende activiteiten die vastliggen op grond van een wettelijke taak of prestatieafspraken binnen een segment. Capaciteit die na invulling van deze vaste activiteiten overblijft, kan worden ingezet voor de behandeling van risicovolle aangiften. Onze aanbeveling heeft betrekking op de gewenste toezichts- en handhavingscapaciteit in relatie tot het risico van fouten in de aangiften. Op dit moment is de beschikbare behandelcapaciteit leidend voor de keuze van het aantal te controleren aangiften, waarbij de meest risicovolle aangiften worden geselecteerd. Onze aanbeveling is dit proces om te draaien door eerst een acceptabel risiconiveau vast te stellen en op basis daarvan de benodigde toezichts- en handhavingscapaciteit af te leiden. Daarmee komt de Belastingdienst tot een betere inzet van zijn handhavingsinstrumenten vanuit een risicogerichte benadering.

Bijlagen

1 Van data naar risocomodel
2 Risicomatrix IB
3 Bepaling grenswaarde
4 Lijst met gebruikte afkortingen
5 Literatuur
6 Noten

Bijlage 1 Van data naar risicomodel

De Belastingdienst gebruikt data en data-analyse bij het bepalen welke aangiften een controleur bekijkt. De Belastingdienst beschikt op grond van zijn wettelijke taak over veel gegevens uit verschillende processen. Door deze gegevens te combineren en in verband te brengen met historische gegevens kan de Belastingdienst een risicomodel opstellen. Een risicomodel voorspelt de kans op een fout in een aangifte, op basis van kenmerken van de aangifte en van de belastingplichtige. Een belangrijk onderdeel in het maken van schattingen in risicomodellen is de selectie van de kenmerken die in het model worden opgenomen.

Het berekenen van een risicomodel kent de volgende stappen:

- Stel een bestand van aangiften samen waarvan de controle-uitkomst met een grote mate van zekerheid bekend is. Dit kan worden bepaald met een steekproef of met recente gegevens uit het controleproces.
- Registreer welke aangiften volgens deze controle een fout bevatten.
- Splits de steekproef op in twee bestanden, een trainingsbestand en een testbestand. Het trainingsbestand wordt gebruikt voor het berekenen van het model en het testbestand voor de controle van dit model.
- Bereken het risicomodel met het trainingsbestand.
- Test het risicomodel met het testbestand.

Deze vijf stappen zijn terug te vinden in de risicomodellen van de Belastingdienst die we onderzochten. Een steekproef is een willekeurige representatieve selectie van aangiften die integraal gecontroleerd worden door een ervaren controleur. Daardoor is van de steekproef bekend welke aangiften een fout bevatten en ook welk bedrag daarmee gemoeid is. Deze steekproef kan ook vervangen of aangevuld worden met gesimuleerde controles, zoals een berekening op basis van contra-informatie of controles uit eerdere jaren.

Met een risicomodel kan nu bepaald worden welke kenmerken van de aangiften en de belastingplichtigen samenhangen met de kans op het aantreffen van een fout. Hoe hoger de kans, hoe waarschijnlijker de aangifte een te corrigeren post bevat. De data die nodig zijn om het model te bepalen zijn ingevulde aangiften en de uitkomsten van het controleproces. Bij een nieuwe aangifte is niet bekend of de aangifte een fout bevat. Met het model kan wel de kans op een fout in een nieuwe aangifte bepaald worden. Deze vertaling van risico's die bepaald worden met een steekproef naar risico's van aangiften die niet in de steekproef zijn opgenomen, heet generalisatie. Met deze methode van generaliseren is het

dus ook mogelijk om de kans op een fout in een nieuwe, niet-gecontroleerde aangifte te bepalen. Door alleen aangiften met een hoge kans op een fout te selecteren, is het mogelijk de meest kansrijke aangiften voor een correctie uit alle ontvangen aangiften te selecteren. Het bepalen van de kans op een fout wordt het trainen van een model genoemd.

Door de selectie van aangiften te baseren op deze kans, is het mogelijk alleen die aangiften te selecteren die met een redelijke mate van zekerheid een correctie bevatten en de aangiften waarbij dit minder aannemelijk is terzijde te leggen. Dit wordt gedaan door een grenswaarde vast te stellen: alleen die aangiften met een correctie-kans groter dan een bepaalde grenswaarde worden gecontroleerd. Door deze grenswaarde te variëren, kan vooraf ingeschat worden hoe groot de werkvoorraad wordt.

De kwaliteit van het model kan gecontroleerd worden door vooraf een aantal aangiften apart te zetten, dit wordt het testbestand genoemd. Op basis van het testbestand wordt bepaald in hoeveel gevallen de selectieregel het bij het juiste eind had; dit is de nauwkeurigheid (‘accuracy’ in vaktermen). De nauwkeurigheid bestaat uit twee delen:

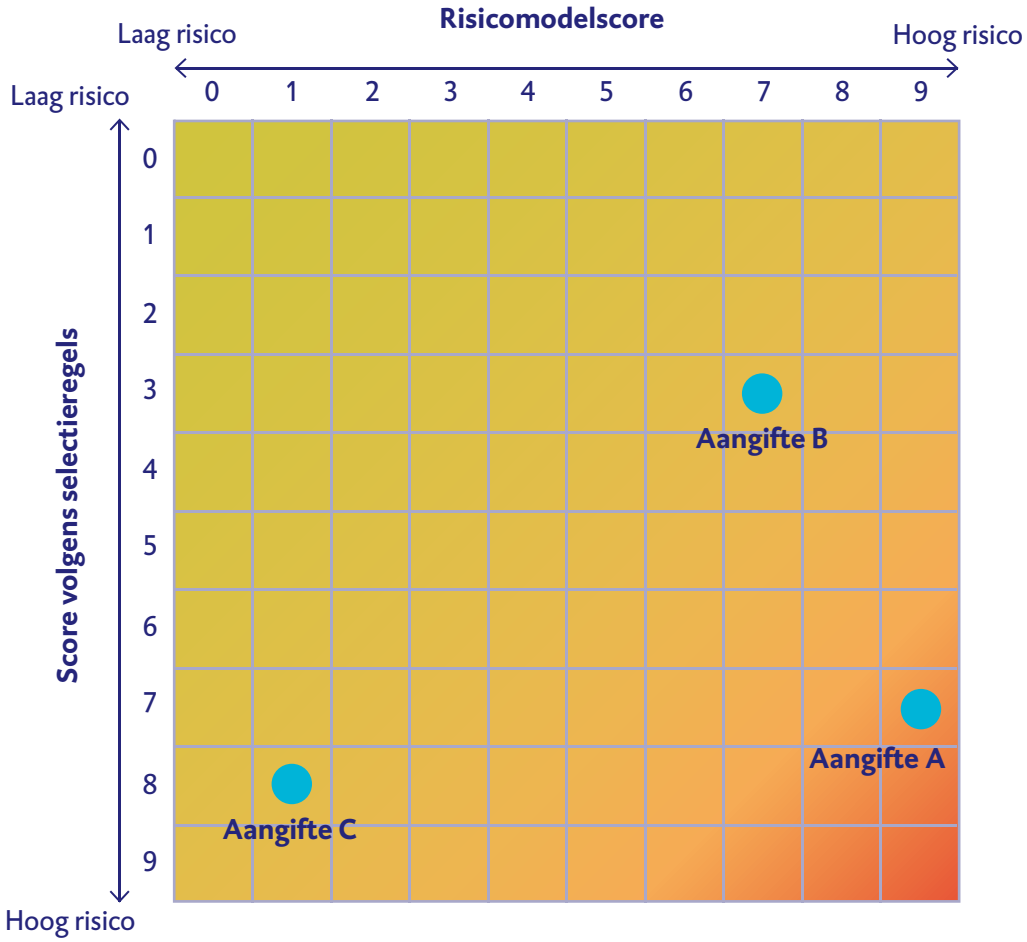
- het risicomodel voorspelde een correctie en dat kwam overeen met de uitkomst van de controle;
- en het risicomodel voorspelde dat een aangifte geen correcties bevatte en dat was eveneens de uitkomst van de controle.

Het is belangrijk dat deze laatste berekening op een apart bestand (het testbestand) uitgevoerd wordt, omdat dit bestand niet gebruikt is om het model te bepalen. In de praktijk wordt het model gebruikt om van nieuwe aangiften de kans op een correctie te bepalen en de veronderstelling is dat de met het testbestand bepaalde nauwkeurigheid beter overeenkomt met resultaten in de praktijk, dan wanneer het trainingsbestand zou worden gebruikt. De splitsing in een trainings- en testbestand is een vereiste om een bruikbaar risicomodel te maken. De Belastingdienst volgt deze procedure.

Bijlage 2 Risicomatrix IB

In deze bijlage leggen we de risicomatrix voor de inkomstenbelasting uit (Belastingdienst, 2016b). In de rijen van deze matrix staan de scores van de selectieregels. De rijen hebben de scores van 0 tot en met 9, waarbij een hoger nummer een grotere kans op een correctie betekent. In de kolommen staan de scores van 0 tot en met 9 van het risicomodel. Ook hier geldt weer hoe hoger het nummer van de kolom, hoe hoger de kans op een correctie. De risicoscores en de scores van de selectieregels worden zo gekozen dat de vakjes in de risicomatrix allemaal evenveel aangiften bevatten. De matrix bestaat uit honderd vakjes (10 maal 10). Van de aangiften in de risicomatrix zijn en de uitkomsten van de selectie-regels en de risicoscores bekend.

De risicomatrix is een visuele weergave van scores volgens selectieregels en scores van het risicomodel



Figuur 2 *De risicomatrix van de Belastingdienst (Bron: Belastingdienst, 2016b)*

In bovenstaande figuur wordt een voorbeeld gegeven van de risicomatrix met drie aangiften. Aangifte C heeft een hoog risico volgens de selectieregels, maar een laag risico volgens het risicomodel. Dit is een dus een aangifte die wel wordt aangewezen door de selectieregels, maar niet door het risicomodel. Aangifte B heeft een hoog risico volgens het risicomodel, maar een laag risico volgens de selectieregels. Deze aangifte heeft volgens het risicomodel wel een grote correctiekans, maar komt niet terecht in de uitworp op basis van de selectieregels. Aangifte A wordt zowel door de selectieregels als het risicomodel aangewezen als een aangifte die na behandeling een correctie zal bevatten.

Bijage 3 Bepaling grenswaarde

De modellen van DF&A berekenen een risicoscore. Om met deze risicoscore een selectie te maken moet een grenswaarde worden bepaald. Aangiften met een score hoger dan de grenswaarde worden geselecteerd en doorgestuurd naar een controleur, die met een score lager dan de grenswaarde worden automatisch afgedaan. Bij elke grenswaarde hoort een aantal aangiften. Door de grenswaarde hoger of lager te kiezen, wordt het aantal te behandelen aangiften kleiner of groter en de kans dat een controleur een aangifte terecht behandelt groter of kleiner.

Bij het instellen van de grenswaarde worden altijd twee verschillende soorten fouten gemaakt. De eerste fout is een 'false positive'. Dit is een aangifte die ten onrechte wordt behandeld, maar geen correctie bevat. De tweede fout is een 'false negative'. Dit is een aangifte die niet geselecteerd is voor behandeling, maar wel een fout bevat. Geselecteerde aangiften die daadwerkelijk een correctie bevatten worden 'true positives' genoemd en aangiften die niet geselecteerd zijn en ook geen correctie bevatten worden 'true negatives' genoemd. Deze vier mogelijkheden kunnen als volgt worden weergegeven.

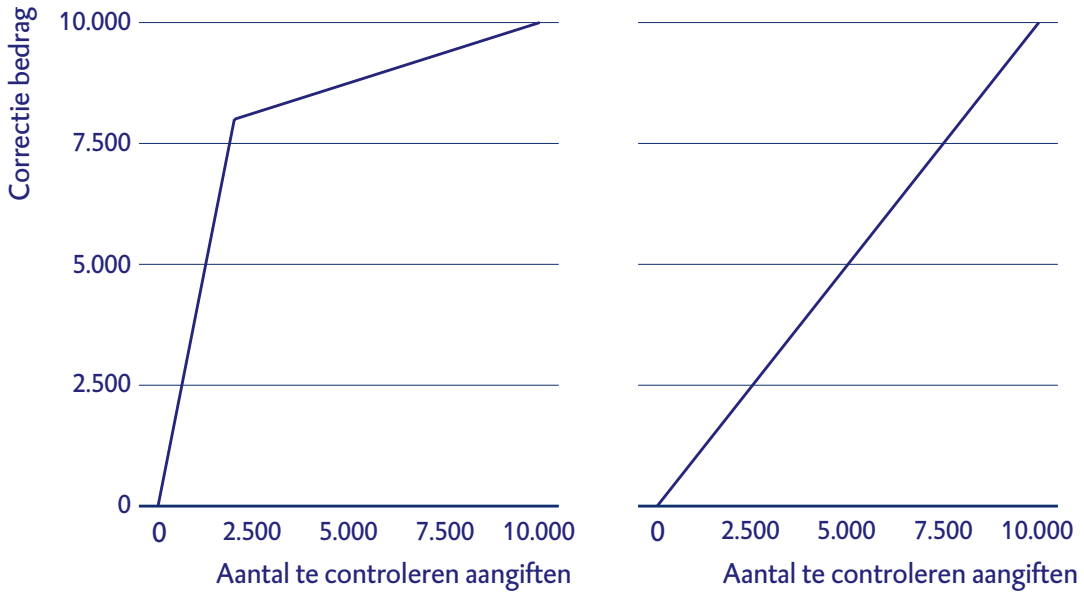
		Correctie in de aangifte?	
		Ja	Nee
Risico volgens model?	Ja	true positive	false positive
	Nee	false negative	true negative

Aantallen aangiften en correctiebedragen bij IB

Met een eigen berekening op de testdata, waarbij we gebruikgemaakt hebben van de risicoscores[2] die gegenereerd worden door de modellen van DF&A, hebben we het verband tussen aantallen behandelingen en correctiebedragen onderzocht.

We hebben onderzocht of het mogelijk is met deze datagedreven benadering een groter deel van de correcties op te leggen met een kleiner aantal controleurs. Uit onze analyse blijkt dat het onderscheidend vermogen van het gebruikte risicomodel niet groot is. Er is geen grenswaarde aan te wijzen die de meest riskante aangiften selecteert en aangiften met weinig risico buiten beschouwing laat. In figuur 3 zetten wij dat schematisch uiteen.

Links een voorbeeld van de 80/20-regel, rechts een voorbeeld waarbij de extra opbrengst evenredig stijgt met de extra controles



Figuur 3 *Aantallen te corrigeren aangiften aangiften en correctiebedragen in een fictief voorbeeld*

De figuur toont het verband tussen het aantal aangiften en het aantal correcties in twee hypothetische situaties, een waarin de opbrengsten evenredig zijn met het aantal aangiften en een waarin de opbrengsten na een bepaald aantal aangiften afnemen. In de linkergrafiek zien we de opbrengsten snel stijgen en daarna afvlakken, de lijn vertoont een knik. Deze grafiek illustreert wat wel de 80/20-regel wordt genoemd, om 80% van de correcties te realiseren moet 20% van de aangiften worden gecontroleerd. In deze situatie zou de Belastingdienst door 2.000 aangiften te controleren het grootste deel van de correcties kunnen realiseren. In de rechtergrafiek zien we een situatie waarin het aantal correcties evenredig is met het aantal aangiften. Hier zou de Belastingdienst 80% van de aangiften moeten controleren om 80% van de correcties te realiseren. Deze lijn vertoont geen knik zoals bij de 80/20-regel en elke extra controle geeft evenveel verwachte opbrengst.

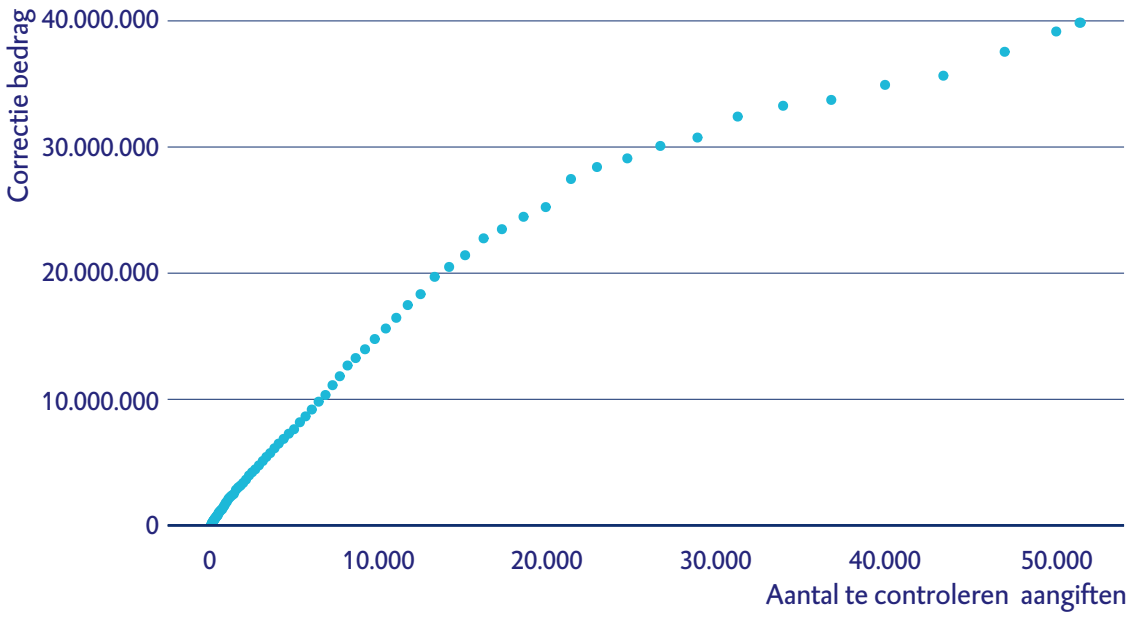
Bij de selectie van aangiften moet rekening worden gehouden met het aantal, zodat een beheersbare werkvoorraad ontstaat, en met de kans op een correctie. Selectie van aangiften met een hoge kans op correctie levert een kleine werkvoorraad op. Als we de werkvoorraad vergroten, betekent dit dat aangiften geselecteerd worden waarbij een correctie minder zeker is.

Wij zijn geen analyses tegengekomen waarin de extra kosten tegen de extra opbrengsten worden afgewogen. Wel weegt de Belastingdienst de opbrengsten van verschillende modellen tegen elkaar af, waarbij ook het aantal behandelingen een rol speelt. Deze analyses optimaliseren zowel de opbrengst als het aantal behandelingen van aangiften. DF&A voert wel analyses uit en selecteert dus een werkvoorraad met gewenste opbrengst of trefkans, waarbij de capaciteit een gegeven is. Er vindt dus geen optimalisatie van de capaciteit gegeven een gewenst risico plaats maar optimalisatie van de belastingopbrengst gegeven de capaciteit.

Aantallen aangiften en correctiebedragen bij btw-teruggave

Bij de btw-teruggave kunnen we een vergelijkbare analyse maken en de bijbehorende grafiek en tabel opstellen. Dit bestand bestaat uit 51.459 aangiften en een opbrengst van bijna € 40 miljoen. We hebben hier het model voor aantallen dat de Belastingdienst in productie gebruikt gereconstrueerd. Uit de grafiek blijkt dat een controle van 13.000 aangiften een correctie opbrengst van € 20 miljoen oplevert. Als de Belastingdienst de controle inspanning zou verdubbelen naar 26000 aangiften haalt de dienst in totaal € 30 miljoen op, oftewel € 10 miljoen extra . Dit bedrag is hoger dan een controle van 13.000 aangiften, maar geen verdubbeling van het verwachtte bedrag.

Als de Belastingdienst meer btw-teruggaven controleert, levert dat meer opbrengsten op



Figuur 4 *Aantallen te corrigeren aangiften en correctiebedragen bij de btw-teruggave*

Bijlage 4 Lijst met gebruikte afkortingen

AVG	Algemene Verordening Gegevensbescherming
BSN	burgerservicenummer
BRI	Basisregistratie Inkomen
DF&A	Datafundamenten en analytics
DPIA	Data Protection Impact Assessment
PIA	Privacy Impact Assessment
IB	inkomstenbelasting
MAC	Massaal Automatisch Corrigeren
mkb	midden- en kleinbedrijf
VIA	vooringevulde aangifte

Bijlage 5 Literatuur

Belastingdienst (2016a). *Reductie voorraad IH NW 2015*. Den Haag: eigen beheer.

Belastingdienst (2016b). *IH risicomatrix 2016 gebruikershandleiding, versie 1.0*. Den Haag: eigen beheer.

Belastingdienst (2017). *Businesscase OB negatief*. Den Haag: eigen beheer.

BZK (2017). Model gegevensbeschermingseffectbeoordeling rijksdienst (PIA). Den Haag: eigen beheer.

Financiën (2015). *Investeringsagenda Belastingdienst*. Tweede Kamer, vergaderjaar 2014–2015, 31 066, nr. 236. Den Haag: Sdu.

Financiën (2017). *Handhavingsbrief Belastingdienst 2018*. Tweede Kamer, vergaderjaar 2019–2018, 31 066, nr. 390. Den Haag: Sdu.

Tweede Kamer (2015). *Wijziging van enkele belastingwetten en enige andere wetten (Belastingplan 2016), Nota naar aanleiding van het verslag*. Tweede Kamer, vergaderjaar 2015–2016, 34 302, nr. 11. Den Haag: Sdu.

Wetenschappelijke Raad voor het Regeringsbeleid (2016). B*ig data in een vrije en veilige samenleving*. Amsterdam: Amsterdam University Press.

Bijage 6 Noten

1. De afname wordt gemeten ten opzichte van het niveau in mei 2015. Om aan te tonen dat deze doelstelling is gerealiseerd, veronderstelt de staatssecretaris dat zich geen externe ontwikkelingen voordoen, zoals groei of krimp van de economie. Dit wordt ook wel de ceteris paribus conditie genoemd.
2. Voor de inkomstenbelasting is gebruikgemaakt van programmatuur van de Belastingdienst.

Voorlichting

Afdeling Communicatie
Postbus 20015
2500 EA Den Haag
telefoon (070) 342 44 00
voorlichting@rekenkamer.nl
www.rekenkamer.nl

Omslag

Ontwerp: Corps Ontwerpers
Foto: Rights Managed/HH

Den Haag, juni 2019