

Bijlage 1. Analyse risicoclassificatiemodel Toeslagen

Hieronder treft u de analyse aan van het risicoclassificatiemodel Toeslagen (hierna: model) zoals die aan uw Kamer is toegezegd in de brief van 8 december 2021¹. In de analyse wordt de volgende indeling gehanteerd:

- A. Terugblik: werking van het model
- B. Aanpak van de analyse
- C. Resultaten
 - I) Kinderopvangtoeslag
 - II) Huurtoeslag
 - III) Het gebruik van nationaliteit in het model
 - IV) Het gebruik van het model voor andere zaken
 - V) Conclusie
- D. Wat leren we hiervan?

A. Terugblik: werking van het model

In de brieven van 26 november 2021 en 8 december 2021 is een uitgebreide toelichting gegeven op het model en de werking ervan.² Onder verwijzing naar die brieven wordt voor nu volstaan met op te merken dat het model elke maand werd ingezet, voorafgaand aan de maandelijkse uitkering van huur- en kinderopvangtoeslagen aan burgers. Voordat het bedrag werd uitbetaald en er daarmee 'formeel' werd beschikt, werd tussen 2013 en eind 2019 alle informatie uit de conceptbeschikkingen ingelezen in het model. Het model bepaalde het risico op onjuistheden in deze beschikkingen aan de hand van vooraf gedefinieerde indicatoren. De conceptbeschikkingen met het hoogste risico op fouten werden maandelijks geselecteerd voor handmatige behandeling. Dit betekende kortgezegd dat het formele beschikmoment en de uitbetaling van de betreffende toeslagen werd uitgesteld, totdat een medewerker de zaak behandeld had. Het model bevatte geen automatische besluitvorming over het recht op toeslag; dit werd altijd gedaan door de medewerkers die de aanvragen beoordeelden.

Het model is steeds in ontwikkeling geweest en de weging en grenswaarden van indicatoren werden door veranderende voorbeelden van een juiste of onjuiste toeslagaanvraag voortdurend aangepast. Indicatoren die door verschuivingen in weging of door veranderende wet- en regelgeving niet meer relevant waren, werden door data-analisten niet meer aan het model aangeboden. Dit had als gevolg dat de waarden van en de samenhang tussen indicatoren per maand kon verschillen.

B. Aanpak van de analyse

Om te komen tot een gedegen analyse van de uitkomsten van het model zijn allereerst de aanvragen van de huurtoeslag en de kinderopvangtoeslag in kaart gebracht die handmatig zijn gecontroleerd door een medewerker, na het gebruik van het model. Dit wordt in het systeem van Toeslagen geregistreerd door middel van 'werkopdrachten' gekoppeld aan een burgerservicenummer (hierna: bsn). Daarbij zijn alle werkopdrachten geïnventariseerd die in het systeem van Toeslagen staan met kenmerken die voortkomen uit het model onder de noemer 'hoog risico'. Vanuit de werkopdrachten is vervolgens bij de geselecteerde aanvragen de bijbehorende informatie gezocht over alle kenmerken die destijds bekend waren en die zijn gebruikt bij de selectie door het model.³⁴

Voor de analyse is een vergelijking gemaakt van de informatie die beschikbaar is over de

¹ Kamerstukken II, 2021-2022, 31 066, nr. 938.

² Kamerstukken II, 2021-2022, 31 066, nr. 923 en 938.

³ Voor 2013 geldt dat eerder is gemeld dat 2013 het jaar van de opstartfase van het model betrof en dat hierdoor de over 2013 beschikbare informatie minder compleet zal zijn. Helaas moet geconstateerd worden dat deze verwachting is uitgekomen.

⁴ Ten aanzien van 2014 moet daarbij opgemerkt worden dat de werkopdrachten wel zijn gevonden, maar dat de bijbehorende scores van de indicatoren niet uit het archief van Toeslagen gehaald kunnen worden. Als het gaat om de kenmerken van de behandelde populatie wordt dan ook gebruik gemaakt van gegevens van 2015 en later.

aanvragen van burgers die handmatig gecontroleerd zijn door een behandelaar (hierna: geselecteerde groep) en van de volledige populatie aan burgers die op dat moment toeslag ontvingen (hierna: totale populatie). Omdat in sommige gevallen extreme (hoge of lage) waarden een vertekend beeld kunnen geven van het gemiddelde is ook gekeken naar: de mediaan (de middelste), de minimale en de maximale waarde van een score op een bepaalde indicator. Uit elk jaar tussen 2015 en 2019 is één run halverwege het jaar geanalyseerd. Het doel is om een referentie te geven van de 'gemiddelde' toeslagbetrokkene. De ervaring leert dat de meeste persoonskenmerken over de tijd niet snel veranderen. Zo wordt er dan ook vanuit gegaan dat de toeslagpopulatie halverwege het jaar representatief is voor de toeslagpopulatie in het hele jaar. Zoals toegelicht is van 2013 niets en van 2014 enkel de totaal aantallen van de aanvragen meegenomen in deze analyse.

C. Resultaten

Hoeveel aanvragen zijn geselecteerd?

In totaal zijn van 2014 tot en met 2019 ongeveer 79.000 unieke individuen handmatig behandeld na de selectie door gebruik van het model. Sommigen hiervan zijn vaker dan één keer handmatig behandeld. In totaal waren er ongeveer 92.000 behandelingen: zo'n 66.000 aanvragen van huurtoeslag en zo'n 26.000 aanvragen van kinderopvangtoeslag.

Welke indicator(en) speelde(n) een rol, en had die veel of weinig invloed?

Bij brief van 8 december 2021 is naar aanleiding van het verzoek van uw Kamer in de bijlage van deze brief een overzicht opgenomen van indicatoren die van 2014 tot en met 2019 in het kader van de kinderopvangtoeslag zijn gebruikt in het model.⁵ Dit overzicht is opnieuw aan de bijlagen van de huidige brief toegevoegd (zie bijlage 2). Tevens is het overzicht van de gebruikte indicatoren van 2015 tot en met 2019 in het kader van de huurtoeslag toegevoegd (zie bijlage 3).

Het is niet mogelijk om de invloed van elke losstaande indicator zelfstandig te duiden. De indicatoren zelf zijn, net als de scores die aan de indicatoren worden toegewezen, door de jaren heen vaak veranderd. Dit maakt een volledige kwantitatieve vergelijking van de invloed van enkele indicatoren tussen verschillende periodes onmogelijk. Indicatoren werden bovendien door het model in onderlinge samenhang in plaats van individueel gewogen; de verschillende indicatoren konden elkaar versterken of juist verzwakken. Het is daarom niet mogelijk om aan te geven wat het effect van een bepaalde individuele indicator is op de hoogte van de risicoscore. Daar komt nog bij dat de risicoscore van een aanvraag ten slotte ook afhankelijk is van de score van de rest van de populatie. Een individuele risicoscore zegt op zichzelf dus niet zoveel over de kans op selectie. De uitkomst van het model voor de individuele burger wordt pas inzichtelijk als er in samenhang wordt gekeken naar de groep met de hoogste risicoscore die zijn geselecteerd voor handmatige beoordeling.

In bijlage 4 treft u het overzicht van de gemiddelde waardes op alle indicatoren van de behandelde aanvragen en de gehele populatie uit het jaar 2015 tot en met 2019. In dit overzicht is zichtbaar dat bij veel indicatoren een duidelijk verschil zit tussen de gemiddelde waardes van de geselecteerde groep en van de totale populatie. Dit betekent dat als één van de kenmerken van een geselecteerde burger anders was geweest, hij of zij waarschijnlijk alsnog tot de hoogste risicogroep behoorde omdat ook de andere kenmerken tot een verhoogde risicoscore leidden. Burgers kwamen dus niet in de hoogste risicocategorie als ze enkel op één indicator hoog scoorden. Hieronder wordt verder ingegaan op een aantal bijzonderheden van de analyse.

Wat zijn de specifieke kenmerken van de groep die geraakt is door het toezicht?

Om de cijferanalyse van bijlage 3 beter te kunnen duiden wordt hieronder aan de hand van enkele voorbeelden ingegaan op de uitkomsten van de analyse. Hierbij is gekozen voor de indicatoren die met name zien op de persoonskenmerken van de aanvragers van kinderopvangtoeslag en huurtoeslag.

⁵ Kamerstukken II, 2021-2022, 31 066, nr. 938.

I) Kinderopvangtoeslag

In reactie op vragen van de Adviescommissie Uitvoering Toeslagen (AUT) is in november 2019 gekeken naar de 1.000 hoogste risicoscores ten aanzien van de kinderopvangtoeslag over een willekeurige maand van 2019 bij gebruik van het model. Dit overzicht (zie figuur 1) maakte destijds inzichtelijk dat in deze 1.000 hoogste scores een oververtegenwoordiging zichtbaar is van burgers waarop (één of meer) specifieke kenmerken van toepassing zijn. Figuur 1 is eerder als bijlage van de brief op 26 november 2021 met uw Kamer gedeeld⁶.

Figuur 1. Reactie op vragen AUT

Groep	4 grote steden	Bezit Nederlandse nationaliteit	Gezinsinkomen onder €20.000	Alleenstaand	Minstens 3 kinderen in opvang	Minimaal 200 uur dagopvang	Woont meer dan 10 km van opvang
1000 hoogste risico's per run	30,8%	78,8%	82,3%	86,9%	12,0%	34,7%	3,8%
Overig	12,5%	95,5%	7,3%	14,0%	6,2%	1,2%	0,9%

Hieronder laat figuur 2 zien dat de analyse op de uitkomsten van het model door de jaren heen een vergelijkbare uitkomst heeft van geselecteerde groep burgers bij de kinderopvangtoeslag als in 2019 voor de AUT is gedaan (en zoals zichtbaar in figuur 1). Daarmee wordt bevestigd dat er sprake is van een oververtegenwoordiging van een groep burgers waarop (één of meer) specifieke kenmerken van toepassing zijn. De in figuur 1 en 2 gebruikte titels verschillen van elkaar omdat bij de AUT specifiek is gekeken naar persoonskenmerken en bij de analyse is enkel gekeken naar de gebruikte indicatoren waar de persoonskenmerken zoals weergegeven in figuur 1 van invloed op zijn. Het is daarbij van belang om te melden dat over de geografische spreiding van de geselecteerde groep op basis van het model niets gezegd kan worden, aangezien het model geen indicator bevatte die kon duiden op iemands woonplaats.

Figuur 2. Gemiddelde uitkomsten analyse model door de jaren heen voor de kinderopvangtoeslag

	Heeft Nederlandse Nationaliteit	Jaarinkomen van de aanvrager	Heeft partner	Heeft aantal kinderen	Max. aantal uren*	Afstand woning tot opvang
Geselecteerde groep	79%	€ 14.948	23%	1.8	147	1.4 km
Totale populatie	96%	€ 38.485	86%	1.8	57	0.8 km

* Aantal uur dat de minst werkende partner zou moeten werken per maand om het aantal aangevraagde uren kinderopvang te kunnen rechtvaardigen. De achterliggende gedachte daarbij was dat kinderopvangtoeslag bedoeld is voor ouders die werken of een andere dagtaak hebben: als een hoog aantal uur opvang wordt aangevraagd, dan wordt geacht dat de ouders dat aantal uur daadwerkelijk werken.

II) Huurtoeslag

Voor de huurtoeslag zijn niet altijd dezelfde indicatoren gebruikt als bij de kinderopvangtoeslag, zoals te zien in bijlage 1 en 2. Er kan daarom niet exact dezelfde tabel worden opgesteld met vergelijkbare kenmerken als destijds voor de AUT is gedaan, dit ook aangezien voor de AUT enkel gekeken is naar de kinderopvangtoeslag. Veel indicatoren van het model die zien op de huurtoeslag zijn meer technisch van aard en wijzen minder op persoonskenmerken dan de meeste indicatoren bij de kinderopvangtoeslag.

Een voorbeeld hiervan is 'huur deelbaar door 10'. De achterliggende gedachte bij deze indicator was destijds waarschijnlijk dat een huurprijs vaak op de cent nauwkeurig is. De aanname was dan dat als de aanvrager een huurprijs heeft ingevuld die deelbaar is door 10, dat er bijvoorbeeld op zou kunnen wijzen dat de huurprijs niet direct is overgenomen van het huurcontract en dus een fout zou kunnen bevatten bevat. Bij de geselecteerde groep was dit in 42% van de aanvragen het geval, bij de totale populatie was dit 6%.

⁶ Kamerstukken II, 2021-2022, 31 066, nr. 923.

Om de specifieke kenmerken van de geselecteerde groep van de huurtoeslag goed inzichtelijk te maken is figuur 3 opgesteld. Ook hier wordt duidelijk dat sprake is van een oververtegenwoordiging van een groep burgers waarop (één of meer) specifieke kenmerken van toepassing zijn.

Figuur 3. Gemiddelde uitkomsten analyse model door de jaren heen voor de huurtoeslag

	Heeft Nederlandse Nationaliteit	Gezinsinkomen	Aantal ingeschreven volwassenen	Verhouding huurprijs t.o.v. gezinsinkomen	Openstaande vorderingen*	Bedrag aan openstaande vorderingen
Geselecteerde groep	69%	€ 8.971	5.2	56%	2.9	€ 4.450
Totale populatie	93%	€ 16.036	1.7	13%	0.8	€ 947

* Openstaande vorderingen kunnen erop wijzen dat in het verleden al terugvorderingen zijn ontstaan als gevolg van correcties. Dit maakt vroegtijdige behandeling belangrijk om meer (hoge) terugvorderingen te voorkomen.

III) Het gebruik van nationaliteit in het model

Over het gebruik van de indicator 'Nederlandschap Ja/Nee' is uw Kamer bij brieven van 26 november en 8 december⁷ jl. geïnformeerd. Tijdens de analyse is in kaart gebracht hoe groot de populatie is die door het gebruik van deze indicator een grotere kans op selectie hebben gehad. De aantallen zijn opgenomen in figuur 4.

Als wordt gekeken naar het effect van het gebruik van deze indicator in beide toeslagen blijkt dat de geselecteerde groep aanzienlijk vaker niet de Nederlandse nationaliteit bezat:

- Bij de huurtoeslag blijkt dat 31% van de door het model geselecteerde groep geen Nederlandse nationaliteit had, tegenover 7% van de hele populatie.
- Ten aanzien van de kinderopvangtoeslag blijkt dat 21% van de geselecteerde groep geen Nederlandse nationaliteit had, tegenover 4% van de totale populatie.

Figuur 4. Overzicht gebruik 'Nederlandschap Ja/Nee' door het model

Totaal aantal behandelde posten	+/- 92.000
Aantal unieke personen	+/- 79.000
Aantal behandelde posten zonder bezit Nederlandse nationaliteit	+/- 24.000
- Waarvan aanvragers KOT	+/- 5.000
- Waarvan aanvragers HT	+/- 19.000
Aantal unieke personen zonder bezit Nederlandse nationaliteit	+/- 20.000*
- Waarvan aanvragers KOT	+/- 4.000
- Waarvan aanvragers HT	+/- 17.000

* Dat het aantal unieke personen KOT (+/- 4.000) en unieke personen HT (+/- 17.000) niet optelt tot +/- 21.000 komt omdat burgers voor beide toeslagen handmatig behandeld kunnen zijn en dus in beide aantallen als unieke personen voorkomen.

Naast 'Nederlandschap Ja/Nee' is ook de indicator 'Leeftijd Burgerservicenummer (bsn)' gebruikt door het model voor zowel de huurtoeslag als de kinderopvangtoeslag: 'Hoe lang geleden is het Burgerservicenummer geactiveerd (gemeten van het moment van de aanvraag)?'. In de Gegevensbeschermingseffectbeoordeling van het model, die als bijlage is gedeeld met de brief van 26 november 2021, staat bij de omschrijving van deze indicator opgenomen: "Om te bepalen hoe lang iemand al in Nederland is."⁸ Het effect van het gebruik van deze indicator is als volgt:

- Ten aanzien van het effect van het gebruik van deze indicator in de huurtoeslag wordt duidelijk dat de geselecteerde groep gemiddeld een jonger bsn heeft (10 jaar) dan de hele populatie (16 jaar).
 - De mediaan van de geselecteerde groep is 2 jaar tegenover 22 jaar bij de hele populatie. De mediaan betreft de middelste waarde wanneer alle waarden van laag naar hoog op een rij worden gezet. Het geeft in sommige gevallen een beter beeld

⁷ Kamerstukken II, 2021-2022, 31 066, nr. 923 en 938.

⁸ Kamerstukken II, 2021-2022, 31 066, nr. 923.

van de uitkomsten dan het gemiddelde, omdat het gemiddelde door extreme waarden kan worden beïnvloed.

- Ten aanzien van het effect van het gebruik van deze indicator in de kinderopvangtoeslag wordt duidelijk dat de geselecteerde groep gemiddeld een jonger bsn heeft (10 jaar) dan de hele populatie (18 jaar).
 - De mediaan van de geselecteerde groep is 2 jaar tegenover 28 jaar bij de hele populatie.

IV) Het gebruik van het model voor andere zaken

Naar aanleiding van recente vragen uit de media is gebleken dat de risicoscores van het model breder zijn gebruikt dan enkel voor het selecteren van aanvragen die voor een handmatige behandeling in aanmerking kwamen. Eerder is al geconstateerd dat de datapreparatie van het model breder is gebruikt, uw Kamer is hierover ook geïnformeerd in de brieven van 8 december 2021 en 24 februari 2022⁹.

a. Breder gebruik risicoscores Screeningsteam

De risicoscore van het model werd onder andere toegevoegd aan 'beoordelingen van het zogenaamde "screeningsteam" van Toeslagen. Dit team beoordeelde signalen van misbruik en oneigenlijk gebruik en was betrokken bij het instellen van een CAF-onderzoek binnen Toeslagen. De informatie kon vervolgens door dit screeningsteam worden gedeeld met de betrokkenen bij CAF-onderzoeken binnen Toeslagen, de handmatige beoordeling binnen Intensief subjectgericht toezicht (IST) en Fraudeteams.

Landelijke Stuurgroep Interventieteams

De risicoscore van het model werd door Toeslagen gebruikt voor de selectie van adressen in het kader van projecten van de Landelijke Stuurgroep Interventieteams (LSI). De risicoscore van het model was daarbij één van de elementen op basis waarvan een adres in aanmerking kon komen voor selectie binnen een project van de LSI. Vanaf eind 2019 tot medio juli 2020 werd gebruik gemaakt van de risicoscores van de laatste run van het model eind 2019. Per medio juli 2020 is de risicoscore niet meer gebruikt om adressen te selecteren.

Financieel Expertise Centrum

De risicoscore van het model is door Toeslagen ook gebruikt naar aanleiding van een door het Financieel Expertise Centrum (FEC) aangedragen subject ten aanzien van de aanpak van terrorismefinanciering. De risicoscore werd hierbij mede gebruikt om te bepalen of er ook aanleiding was voor onderzoek door Toeslagen. Het betreft uitsluitend eigen onderzoek van Toeslagen. Toeslagen (en dus ook de risicoscores) was niet betrokken bij de selectie van subjecten door het FEC en deelde de resultaten van het eigen Toeslagen onderzoek dan ook niet met het FEC. Ook hiervoor geldt dat vanaf eind 2019 tot medio juli 2020 gebruik werd gemaakt van de risicoscores van eind 2019. Per medio juli 2020 is de risicoscore niet meer gebruikt.

b. Breder gebruik datapreparaties

Eerder is al geconstateerd dat de datapreparatie van het model is gebruikt voor het proces 'hoge toeslag, hoog risico' (Hothor) en het voorspellend model problematische terugvorderingen. Dat betekent dat de brongegevens die werden verzameld en bewerkt voor het model ook zijn gebruikt voor de Hothor en het voorspellend model problematische terugvorderingen. Hierbij waren zij dus weliswaar afhankelijk van de code van het model voor datapreparatie, maar niet van de scores of uitkomsten van het model. Vanaf januari 2022 wordt de datapreparatie van het model niet meer gebruikt voor de Hothor en het voorspellend model problematische terugvorderingen zal bij de eerstvolgende training eind 2022 gebruik maken van een onafhankelijke, op zichzelf staande code. Over de Hothor is uw Kamer in de brief van 8 december 2021 geïnformeerd¹⁰ en in de brief van 24

⁹ Kamerstukken II, 2021-2022, 31 066, nr. 938 en 976.

¹⁰ Kamerstukken II, 2021-2022, 31 066, nr. 938.

februari 2022 is uw Kamer geïnformeerd over het voorspellend model problematische terugvorderingen.¹¹

V) Conclusie

De uitkomsten van deze analyse laten zien dat door het model een specifieke groep het meest voor handmatige beoordeling is geselecteerd. Daarbij moet wel worden opgemerkt dat de output van het model door de jaren heen gedeeltelijk te verklaren is vanuit demografisch perspectief onder meer omdat burgers met lage inkomens de hoogste toeslagen krijgen uitgekeerd en dus het grootste risico kunnen lopen op hoge terugvorderingen.

Aangezien de uitkomsten van het model voor andere zaken zijn gebruikt dan enkel de selectie voor handmatige beoordeling, kan niet worden uitgesloten dat ook deze groep burgers onevenredige consequenties ondervonden kon hebben na gebruik van het model: een verhoogde kans op toezicht kon leiden tot een verhoogde kans op een onevenredig nadeel, doordat de constatering van een kleine fout tijdens de handmatige beoordeling grote gevolgen had.

D. Wat leren we hiervan?

Het model is ontwikkeld in een tijd dat binnen Toeslagen nauwelijks ervaring was opgedaan met dergelijke modellen. Toeslagen had bijvoorbeeld nog geen eigen data-afdeling of een adequate vaktechnische structuur. Inmiddels is dat wel het geval.

De analyse van het model leert dat het bij de ontwikkeling van een dergelijk model belangrijk is om de informatie rondom een ontwikkeling - in alle fasen van ontwikkeling tot en met gebruik - goed te documenteren en te archiveren. Naast de regelmatige evaluatie op de uitkomsten van het model is het ook van belang dat er een structurele feedback-loop bestaat tussen de dataspecialisten die het model ontwikkelen en de medewerkers die in de dagelijkse praktijk de zaken behandelen die met behulp van het model zijn geselecteerd. Dit moet ook bijdragen aan het voorkomen van ongewenste onbedoelde effecten, zoals het focussen op bepaalde specifieke groepen. De inzet van aselecte steekproeven en/of het werken met controlegroepen kan de methodiek rondom het werken met een model optimaliseren. Ten slotte moeten ook de signalen vanuit de burgers een vaste plek krijgen in de evaluaties van het model en de effecten ervan in de praktijk.

Wanneer een dergelijk model wordt ontwikkeld, is het van cruciaal belang om dit te doen met inachtneming van de gestelde waarborgen rondom onder andere privacy, mensenrechten en vaktechniek. Het ontwikkelen van een model vraagt om een bestendige organisatie-inrichting met bijbehorende vaktechnische structuur. Dit jaar wordt de handhavingsstrategie van Toeslagen verder uitgewerkt, resulterend in een toekomstbestendige strategie voor 2023 en verder. Voor de zomer wordt uw Kamer geïnformeerd over de stand van zaken hieromtrent. Deze strategie zal samenhangen met het integraal waarborgenkader dat in ontwikkeling is voor Belastingdienst, Toeslagen en Douane en waarover uw Kamer separaat nog nader geïnformeerd zal worden. Ook bestaande kaders en initiatieven zullen bij Toeslagen gebruikt worden, zoals het Impact Assessment Mensenrechten en Algoritmes.

¹¹ Kamerstukken II, 2021-2022, 31 066, nr. 976.