

Vooringenomenheid voorkomen

Aanbevelingen voor risicoprofilering in
het Controle Uitwonendenbeurs-proces:
een kwantitatieve en kwalitatieve analyse

Inhoudsopgave

1. Samenvatting	3
1.1 Managementsamenvatting	3
1.2 Bevindingen	8
1.3 Aanbevelingen	13
1.4 Leeswijzer	15
2. Achtergrond Controle Uitwonendenbeurs: feitenvaststelling	16
2.1 Werkwijze feitenvaststelling	16
2.2 Chronologie CUB	17
2.3 Overzicht CUB-proces	19
3. Onderzoeksmethodologie	23
3.1 Kwantitatieve analyse	23
3.2 Kwalitatieve analyse	30
3.3 Samenstelling onderzoeksteam	36
4. Resultaten kwantitatieve analyse	37
4.1 Resultaten aselecte steekproef 2014 en 2017	37
4.2 Resultaten biasmeting 2014 en 2019	47
5. Resultaten kwalitatieve analyse	56
5.1 Resultaten kwalitatieve analyse risicoprofiel	56
5.2 Processen om het risico op vooringenomenheid tegen te gaan	60
6. Voorbehouden	62
7. Conclusie: bevindingen en aanbevelingen	63
Appendix A – Verdere verwijzingen	70
Appendix B – Datawarehouse query	66
Appendix C – Aanvullende informatie statistische analyse	71
Appendix D – Documentlijst	72

Over Algorithm Audit

Algorithm Audit is een Europees kennisplatform voor AI bias testing en normatieve AI-standaarden. De doelen van de stichting zijn drieledig:



Normatieve adviescommissies

Adviseren over ethische kwesties in concrete algoritmische toepassingen door het samenbrengen van deliberatieve, diverse adviescommissies, met [algotrudentie](#) als resultaat



Technische hulpmiddelen

Implementeren en testen van technische methoden voor bias-detectie en -mitigatie, zoals onze [bias detection tool](#)



Kennisplatform

Samenbrengen van kennis en experts voor collectief leerproces over verantwoorde inzet van algoritmes, bijvoorbeeld ons [AI Policy Observatory](#) en [position papers](#)

1. Samenvatting

Hieronder volgen de managementsamenvatting van dit rapport (§1.1), de bevindingen (§1.2), de aanbevelingen die aan DUO worden gedaan (§1.3) en een leeswijzer (§1.4).

1.1 Managementsamenvatting

De Dienst Uitvoering Onderwijs (DUO) heeft in de periode 2010-2023 gebruik gemaakt van een effectief en eenvoudig risicoprofiel bij het tegengaan van onrechtmatig gebruik van de uitwonendenbeurs.¹ Aan de hand van drie selectiecriteria prioriteerde DUO de inzet van handhavings- en inspectiemiddelen binnen het proces Controle Uitwonendenbeurs (CUB).² In 2023 is het CUB-proces in de media gekomen. Aanleiding was een onderzoek van Investico, NOS op 3 en het Hoger Onderwijs Persbureau.³ Journalisten rapporteerden dat advocaten hen desgevraagd hadden geïnformeerd dat 97,6% (367 van de 376) van de studenten die zij bijstonden in zaken waarin DUO hen aansprak voor onrechtmatig gebruik van de uitwonendenbeurs een migratieachtergrond heeft. Om mogelijke vooringenomenheid in het CUB-proces te onderzoeken heeft DUO Algorithm Audit opdracht gegeven om 1) kwantitatief onderzoek te doen naar vooringenomenheid in het gehele CUB-proces en 2) kwalitatief onderzoek te doen naar het risicoprofiel dat in het CUB-proces is gebruikt. Andere externe partijen onderzoeken andere aspecten van het CUB-proces. Dit rapport bevat de bevindingen van het kwantitatieve en kwalitatieve onderzoek verricht door Algorithm Audit. Een overzicht van de reikwijdte van dit rapport is hieronder weergegeven.

Wat is dit rapport wel?	Wat is dit rapport niet?
<p>Een beschrijving van de totstandkoming van het risicoprofiel en de manier waarop het profiel in de periode 2010-2023 is gedocumenteerd en geëvalueerd.</p> <p>Een kwalitatieve analyse van de totstandkoming en het gebruik van het risicoprofiel.</p>	<p>Onderzoek naar de vraag of het risicoprofiel rechtmatig is opgesteld en gebruikt. Er wordt dan ook:</p> <ul style="list-style-type: none"> > Niet getoetst aan normen uit de periode 2010-2022, maar vanwege het toekomstgerichte karakter van dit onderzoek aan normen uit 2023;

¹ Het woord effectief wordt hier gebruikt zoals DUO de term intern ook gebruikte: beter in het opsporen van onrechtmatig gebruik dan aselechte steekproeven. Cijfers voor de effectiviteit in de periode 2011-2017 worden gegeven in . Ook voor andere termen gebruikt in het CUB-proces wordt in dit rapport de door DUO intern gehanteerde bewoording gebruikt, zoals 'ouder(s)'.

² De drie selectiecriteria zijn onderwijsvorm, leeftijd en afstand tot ouder(s). De exacte samenstelling van het risicoprofiel wordt toegelicht in [2.3 Overzicht CUB-proces](#).

³ Onder meer gepubliceerd door de [NOS](#), [Investico](#), [De Groene Amsterdammer](#) en [Trouw](#).

Wat is dit rapport wel?	Wat is dit rapport niet?
<p>Een kwantitatieve analyse van:</p> <ul style="list-style-type: none"> > De door het risicoprofiel aan studenten toegekende risicoscores; > De selectie door bureaumedewerkers van studenten voor een huisbezoek. <p>Het rapport is gericht op de toekomst. Het is bedoeld om DUO te helpen duiden hoe het risicoprofiel standhoudt in het licht van de huidige normen. Ook is het bedoeld om DUO te helpen verantwoord om te gaan met algoritmes en risicoprofielen als wordt besloten deze weer in te zetten in het CUB-proces. Er wordt:</p> <ul style="list-style-type: none"> > Getoetst aan normen die op het moment van publiceren van het rapport <i>state-of-the-art</i> zijn; > Geadviseerd om interne normen op te stellen en processen in te richten voor het gebruik van algoritmes. 	<ul style="list-style-type: none"> > Niet ingegaan op mogelijke verantwoordingen voor het gebruik van het risicoprofiel zoals politiek-bestuurlijke druk in de periode 2010-2022. <p>Een kwantitatieve analyse van vooringenomenheid ten aanzien van migratieachtergrond. Dit onderzoek heeft tot op heden niet kunnen plaatsvinden omdat de data die daarvoor nodig zijn tot nog toe niet beschikbaar zijn gesteld door het Centraal Bureau voor de Statistiek (CBS).</p> <p>Onderzoek naar de vraag of het risicoprofiel en/of het CUB-proces discriminatoir waren. Dat is een juridische vraag, waarvan de beantwoording mogelijk gebruik kan maken van de uitkomsten van de kwantitatieve analyse, maar die in dit rapport niet als zodanig wordt beantwoord.</p> <p>Een normatief oordeel over het risicoprofiel en/of het CUB-proces.</p> <p>Een uitputtend onderzoek naar het ontstaan en het gebruik van het CUB-proces. Het onderzoek is gedaan op basis van door DUO beschikbaar gestelde informatie.</p>

Kwantitatieve analyse CUB-proces

Het CUB-proces is op twee manieren kwantitatief geanalyseerd. Voor de eerste analyse zijn de uitkomsten van de aselecte steekproef 2014 en 2017 onderzocht. Het doel van deze analyse is om de selectiecriteria en de daarbij behorende subcategorieën die worden gebruikt in het risicoprofiel te testen op statistische relevantie.⁴ De tweede analyse is een zogeheten biasmeting. Deze biasmeting is gebaseerd op data over het verloop van het CUB-proces in 2014 en 2019 en heeft als doel om te meten of studenten in bepaalde onderwijs-, leeftijds- en afstandscategorieën bovenmatig worden geselecteerd voor een huisbezoek ten opzichte van de toegekende risicoscore door het risicoprofiel. Naast de analyse van de steekproeven en de biasmeting wordt een dataverkenning

⁴ Het risicoprofiel bestaat uit verschillende onderwijs-, leeftijds- en afstandscategorieën. De exacte samenstelling van het risicoprofiel wordt toegelicht in [2.3 Overzicht CUB-proces](#).

gepresenteerd van de onderzochte uitwonende studentenpopulaties uit 2014 en 2019. Ook is de effectiviteit in beeld gebracht van het CUB-proces ten opzichte van aselechte steekproeven om onrechtmatig gebruik van de uitwonendenbeurs te traceren.

Algorithm Audit heeft door onvoldoende beschikbare gegevens nog geen kwantitatief onderzoek kunnen doen naar vooringenomenheid van het CUB-proces met betrekking tot studenten met een migratieachtergrond. Wel is vastgesteld dat er in het gebruikte risicoprofiel geen sprake is van direct onderscheid op basis van migratieachtergrond (noch andere beschermde gronden).

Uit de aselechte steekproeven uit 2014 en 2017 volgt onvoldoende statistische onderbouwing voor de in het risicoprofiel gebruikte selectiecriteria onderwijsvorm en leeftijd. Voor verschillende afstandscategorieën is wel een consistent en statistisch significant verband gevonden met onrechtmatig gebruik van de uitwonendenbeurs, zij het vooral voor specifieke subcategorieën zoals een kleine of juist zeer grote afstand tot ouder(s). Voor de opdeling in zes risicocategorieën is onvoldoende meerwaarde gevonden ten opzichte van een simpele binaire opdeling in een lage en hoge risicocategorie. De resultaten zijn tot stand gekomen door enkelzijdige statistische Z-toetsen uit te voeren op de onrechtmatig gebruik-percentages per onderwijs-, leeftijds- en afstandscategorie. De uitkomsten worden bevestigd door Fishers exacte toets.

Uit de biasmeting van het CUB-proces volgt dat studenten die binnen een afstand van 2km tot hun ouder(s) staan ingeschreven aanzienlijk vaker handmatig zijn geselecteerd voor huisbezoek dan men op basis van de door het risicoprofiel toegekende risicoscores kan verwachten. Dit verschil is gemeten door het percentage geselecteerde studenten voor een huisbezoek uit te splitsen per onderwijs-, leeftijds- en afstandscategorie en af te zetten tegen de toegekende risicoscores. Voor de hand ligt dat specifieke werkinstructies die aansporen tot het handmatig selecteren van studenten die nabij het ouderlijk adres staan ingeschreven hier de oorzaak van zijn.

Deze bevindingen leiden tot de aanbeveling om een meer omvattende onderbouwing van het risicoprofiel op te stellen voordat het mogelijk weer in gebruik wordt genomen. Daarnaast moet nader onderzoek worden gedaan of er een verband bestaat tussen de groepen die bovenmatig vaak handmatig zijn geselecteerd voor huisbezoek en studenten met een migratieachtergrond. Ook moet in de toekomst, mocht het CUB-proces geherintroduceerd worden, gebruik van hetzelfde selectie criterium in verschillende stappen van het CUB-proces worden voorkomen. Daartoe moet een onderzoek en verbetertraject worden ingesteld voor het proces van handmatige selectie, waaronder de werkinstructies.

Kwalitatieve analyse risicoprofiel

Voor de kwalitatieve analyse is het ontstaan en gebruik van het risicoprofiel getoetst aan de normen en richtlijnen die daar in 2023 voor gelden. Dit zijn onder meer de Impact Assessment Mensenrechten en Algoritmes (IAMA) en het Onderzoekskader Algoritmes van de Audit Dienst Rijk (ADR). De feitelijke situatie van het CUB-proces is vastgesteld door

documentanalyse en workshops. Of de onderzoekers van Algorithm Audit alle (destijds) beschikbare stukken hebben ontvangen, is niet na te gaan. Het valt niet uit te sluiten dat er ongeziene documenten zijn die een ander licht werpen op de vaststellingen in dit rapport. DUO wijst erop dat het profiel 13 jaar geleden is ontwikkeld, waardoor niet alle relevante documenten meer beschikbaar zijn.

Uit de kwalitatieve analyse volgen onderstaande conclusies.

Tussen 2010 en 2022 is er binnen DUO onvoldoende aandacht geweest voor mogelijke neveneffecten van het risicoprofiel. Er is niet gebleken dat bij het opstellen en gebruiken van het risicoprofiel is onderzocht of bepaalde demografische groepen over- of ondervertegenwoordigd zijn in de hogere risicocategorieën. Het verdient vermelding dat het hier gaat om onbedoelde vooringenomenheid; van bedoelde vooringenomenheid of opzettelijke discriminatie is in het onderzoek niet gebleken. Nadat de omvang van de kinderopvangtoeslagaffaire aan het licht kwam, is in 2019 door DUO onderzocht of het risicomodel direct discrimineert. Er is geconcludeerd dat daar geen sprake van was. Indirecte discriminatie lijkt destijds niet te zijn onderzocht of besproken.



Toetsen van feiten uit het verleden aan normen uit 2023 lijkt wellicht te streng. Dat zou het voor het beantwoorden van normatieve vragen zoals de schuldvraag ook zijn. Voor politiek-bestuurlijke duiding in 2023 en voor het opstellen van normen en processen waarmee DUO in de toekomst algoritmes en risicomodellen kan gebruiken, zijn dit echter de relevante normen.

De selectiecriteria van het risicoprofiel zijn vanaf 2012 openbaar geweest. Deze criteria zijn vermeld in kamerbrieven [A.19] en in rechtspraak.⁵ Desondanks is er uit het onderzoek niet gebleken dat andere partijen DUO hebben gewezen op het risico van vooringenomenheid bij het gebruik van deze selectiecriteria. Dit gaat in ieder geval over de Tweede Kamer, het Ministerie van Onderwijs, Cultuur en Wetenschap (OCW), de Audit Dienst Rijk (ADR), de journalistiek, de advocatuur en het maatschappelijk middenveld.

Bevindingen en aanbevelingen

Naar aanleiding van deze analyse zijn er verbeterpunten aan te wijzen om in de toekomst vooringenomenheid in het CUB-proces te voorkomen. Dit rapport doet daartoe een aanzet in zeven bevindingen en vier aanbevelingen.

Bevindingen

Bevinding 1 – Er is jarenlang een regel-gebaseerd algoritme gebruikt, dat op basis van onderwijsvorm, leeftijd en de afstand tussen het adres van de student

⁵ Onder meer Rb Limburg 13 december 2013, ECLI:NL:RBLIM:2013:11417, Rb Rotterdam 14 juli 2014, ECLI:NL:RBROT:20145684.

en het adres van zijn/haar ouder(s) een risicoscore heeft toegekend aan studenten die een uitwonendenbeurs ontvingen. Er is geen sprake van een zelflerend algoritme of van de inzet van kunstmatige intelligentie. Gebruik van risicoprofilering is effectief gebleken.

- Bevinding 2** – Uit de aselecte steekproeven uit 2014 en 2017 volgt onvoldoende statistisch verband tussen de selectiecriteria onderwijsvorm en leeftijd en onrechtmatig gebruik van de uitwonendenbeurs. Voor specifieke categorieën binnen het selectie criterium afstand is wel voldoende verband met onrechtmatigheid gevonden. Voor de opdeling in zes risicocategorieën is onvoldoende statistische onderbouwing gevonden ten opzichte van een binaire risicoclassificatie.
- Bevinding 3** – Studenten die binnen een afstand van 2km tot hun ouders(s) staan ingeschreven zijn aanzienlijk vaker handmatig geselecteerd voor een huisbezoek dan men op basis van de door het risicoprofiel toegekende risicoscores kan verwachten. Voor de hand ligt dat specifieke werkinstructies, die aansporen tot het handmatig selecteren van studenten die nabij het ouderlijk adres staan ingeschreven, hier de oorzaak van zijn.
- Bevinding 4** – Er is geen sprake van direct onderscheid op basis van migratieachtergrond (noch op basis van andere beschermde gronden) in het risicoprofiel. Door onvoldoende beschikbare gegevens is het tot op heden niet mogelijk geweest om kwantitatief onderzoek te doen naar indirecte vooringenomenheid ten aanzien van studenten met een migratieachtergrond.
- Bevinding 5** – De selectiecriteria en risicocategorieën die zijn gebruikt in het risicoprofiel zijn naar huidige normen onvoldoende onderbouwd, zeker ten aanzien van mogelijke vooringenomenheid.
- Bevinding 6** – Er is niet gebleken van intern onderzoek naar mogelijke vooringenomenheid van het gebruikte risicoprofiel, noch bij het opstellen van het risicoprofiel noch tijdens gebruik ervan.

Bevinding 7 – Er is niet gebleken van interne normen voor het gebruik van algoritmes.

Deze bevindingen worden nader toegelicht in [1.2 Bevindingen](#).

Aanbevelingen

- Aanbeveling 1** – Zorg voor een meer omvattende onderbouwing van het risicoprofiel voordat het mogelijk weer in gebruik wordt genomen, onder meer met behulp van een normatief kader.

Aanbeveling 2 – Er moet nader onderzoek worden gedaan of er een verband bestaat tussen de groepen die bovenmatig vaak handmatig zijn geselecteerd voor huisbezoek en studenten met een migratieachtergrond.

Aanbeveling 3 – Voorkom het gebruik van hetzelfde selectiecriteria in verschillende stappen van het CUB-proces. Stel daartoe een onderzoek en verbetertraject in voor het proces van handmatige selectie, waaronder de werkinstructies.

Aanbeveling 4 – Stel een algoritmemanagementbeleid op om de risico's te verkleinen die de inzet van algoritmes met zich meebrengt.

Deze aanbevelingen worden nader toegelicht in [1.3 Aanbevelingen](#).

1.2 Bevindingen

In deze sectie worden de genoemde bevindingen toegelicht. De bevindingen komen voort uit de kwantitatieve en kwalitatieve analyse van het risicoprofiel en van hoe dit profiel is opgesteld en beheerd binnen DUO.

Bevinding 1 – Er is jarenlang een regel-gebaseerd algoritme gebruikt, dat op basis van onderwijsvorm, leeftijd en de afstand tussen het adres van de student en het adres van zijn/haar ouder(s) een risicoscore heeft toegekend aan studenten die een uitwonendenbeurs ontvingen. Er is geen sprake van een zelflerend algoritme of van de inzet van kunstmatige intelligentie. Gebruik van risicoprofilering is effectief gebleken.

Tussen 2012 en 2023 is in het CUB-proces een risicoprofiel gebruikt om geautomatiseerd risicoscores toe te kennen aan alle studenten die van DUO een uitwonendenbeurs ontvingen. Het risicoprofiel is een lineair model dat op basis van drie criteria studenten een risicoscore toekent. Deze criteria zijn onderwijsvorm, leeftijd en de afstand tussen het adres waar zij stonden ingeschreven en het adres van hun ouder(s). Ieder criterium is opgedeeld in categorieën, bijvoorbeeld studenten die tussen de 1m-1km van hun ouder(s) wonen, studenten die 50-500km van hun ouder(s) wonen, een mbo 1-2 opleiding volgen etc. Na toepassing van vooraf bepaalde wegingsfactoren per categorie leidt het profiel van een student tot een risicoscore. De risicoscore kan worden verhoogd als de bij DUO bekende leeftijd van de student afwijkt met de leeftijd van de student in de Basis Registratie Personen (BRP). Aan alle studenten zijn risicoscores toegekend. Het toepassen van het CUB-proces, waarvan het risicoprofiel een belangrijk onderdeel uitmaakt, is effectief gebleken. De effectiviteit blijkt uit het feit dat met toepassing van het CUB-proces meer onrechtmatig gebruik van de uitwonendenbeurs is vastgesteld dan bij aselecte steekproeven. Dit betreft respectievelijk 3,6% en 3,8% effectiviteit bij de aselecte steekproeven van 2014 en 2017 en respectievelijk 38,9% en 35,3% effectiviteit bij toepassing van het CUB-proces in 2014 en 2019 waarbij het risicoprofiel is ingezet.

Meer informatie over het risicoprofiel en het CUB-proces wordt gegeven in [2.3 Overzicht CUB-proces](#). Meer informatie over de effectiviteitscijfers kan worden gevonden in [3.1 Kwantitatieve analyse](#).

Bevinding 2 – Uit de aselechte steekproeven uit 2014 en 2017 volgt onvoldoende statistisch verband tussen de selectiecriteria onderwijsvorm en leeftijd en onrechtmatig gebruik van de uitwonendenbeurs. Voor specifieke categorieën binnen het selectie criterium afstand is wel voldoende verband met onrechtmatigheid gevonden. Voor de opdeling in zes risicocategorieën is onvoldoende statistische onderbouwing gevonden ten opzichte van een binaire risicoclassificatie.

Per onderwijs-, leeftijds- en afstandscategorie is geturfd hoe vaak studenten die zijn geselecteerd voor onderzoek middels de aselechte steekproef 2014 (n=387) en 2017 (n=293) onrechtmatig gebruik hebben gemaakt van de uitwonendenbeurs. Aan de hand van deze frequenties is getoetst, middels de Z-toets en Fishers exacte toets, of de verschillen in onrechtmatigheidspercentages voor de gebruikte categorieën in het risicoprofiel statistisch significant zijn. Bijvoorbeeld: is er statistisch significant meer onrechtmatigheid vastgesteld in de afstandscategorie 2-5km dan in de categorie 5-10km? Voor leeftijd zijn onvoldoende aanwijzingen aangetroffen voor statistisch significante verschillen tussen de gebruikte categorieën. Voor onderwijsvorm zijn in de steekproef uit 2014 wel en in de steekproef uit 2017 geen significante verschillen aangetroffen. Dit acht Algorithm Audit een onvoldoende consistent statistisch signaal om risicoprofilering op te baseren. Een nieuwe, grotere steekproef zou hier verandering in kunnen brengen. Voor afstand tot ouder(s) leveren vooral de categorieën 1m-1km, 2-5km en 50-500km een statistisch significant signaal. Dit biedt onderbouwing voor gebruik van afstand als selectiecriteria voor risicoprofilering, mits de specifieke categorisering preciezer wordt vastgesteld. Daarnaast is er onvoldoende onderbouwing om de toegekende risicoscores op te delen in zes risicocategorieën. Voor een binaire risicoclassificatie is wel statistische onderbouwing, en deze simplificering geniet de voorkeur.

Meer informatie over de methodologie van de statistische toetsen kan worden gevonden in [3.1 Kwantitatieve analyse](#). De resultaten worden besproken in [4.1 Resultaten aselechte steekproef 2014 en 2017](#).

Bevinding 3 – Studenten die binnen een afstand van 2km tot hun ouders(s) staan ingeschreven zijn aanzienlijk vaker handmatig geselecteerd voor een huisbezoek dan men op basis van de door het risicoprofiel toegekende risicoscores kan verwachten. Voor de hand ligt dat specifieke werkinstructies, die aansporen tot het handmatig selecteren van studenten die nabij het ouderlijk adres staan ingeschreven, hier de oorzaak van zijn.

In de biasmeting is een sterke oververtegenwoordiging waargenomen van studenten die binnen 2km afstand tot hun ouder(s) staan ingeschreven bij handmatige selectie voor huisbezoek. Deze oververtegenwoordiging komt niet voort uit het gehanteerde risicoprofiel.

In 2014 is er voor studenten die 0km van hun ouder(s) staan ingeschreven een dispropor-tionele verhouding tussen de kans op een huisbezoek (8,3x het gemiddelde van alle cate-gorieën) en de toegekende risicoscore (3,7x het gemiddelde). Dat betekent dat studenten uit deze categorie een meer dan 2 keer zo grote kans hebben op een huisbezoek dan vanuit hun risicoscore kan worden verwacht. Voor de 1m-1km categorie is deze verhouding tussen kans op huisbezoek en risicoscore (beide ten opzichte van hun gemiddelden) 6,8x : 3,0x en voor de 1-2km categorie 3,5x : 2,7x. Dit betekent dat de groep met een kleine afstand tot ouder(s) structureel is oververtegenwoordigd in de handmatige selectie voor huisbezoek, ten opzichte van de hun toegekende risicoscore's.

Deze oververtegenwoordiging kan vermoedelijk worden verklaard door de aanwezigheid van specifieke werkinstructies die medewerkers opdragen om studenten met bijzonderheden in de woonsituatie – zoals de combinatie jonge leeftijd, dichtbij ouder(s) wonend en/of wonend bij familieleden – gericht te selecteren voor huisbezoek. De oververtegenwoordiging van studenten (zoals die op een kleine afstand tot ouder(s) staan ingeschreven) in de handmatige selectie voor huisbezoek zou potentieel een oververtegenwoordiging van andere kenmerken, zoals die van migratieachtergrond, kunnen veroorzaken. Er moet daarom nader onderzoek worden gedaan of er een verband bestaat tussen de groepen die bovenmatig vaak handmatig zijn geselecteerd voor huisbezoek en studenten met een migratieachtergrond. Zie [Aanbeveling 2](#). Daarnaast moet een onderzoek en verbetertraject worden ingesteld voor het proces van handmatige selectie, waaronder de werkinstructies. Zie [Aanbeveling 3](#).

De uitkomsten van de biasmeting, inclusief de genoemde cijfers, kunnen worden teruggevonden [Figuur 8](#) en in [4.2. Resultaten biasmeting 2014 en 2019](#).

Bevinding 4 – Er is geen sprake van direct onderscheid op basis van migratieachtergrond (noch op basis van andere beschermde gronden) in het risicoprofiel. Door onvoldoende beschikbare gegevens is het niet mogelijk geweest om kwantitatief onderzoek te doen naar indirecte vooringenomenheid ten aanzien van studenten met een migratieachtergrond.

DUO heeft Algorithm Audit verzocht ook een biasmeting uit te voeren voor mogelijke vooringenomenheid met betrekking tot het bijzondere persoonskenmerk migratieachtergrond. DUO heeft zelf geen data over de migratieachtergrond van studenten. Daarom is aan het CBS verzocht de data van DUO uitsluitend voor het uitvoeren van deze biasmeting te verrijken met data over de migratieachtergrond van studenten. Tot op heden heeft Algorithm Audit deze data niet kunnen verkrijgen. In de toekomst kunnen deze data mogelijk wel beschikbaar worden gesteld voor eventueel vervolgonderzoek. Alternatieve methoden om indirecte vooringenomenheid met betrekking tot migratieachtergrond te meten, zoals met behulp van aggregatiestatistieken per postcodegebied, achtte Algorithm Audit onvoldoende verantwoord in methodologisch opzicht. Algorithm Audit heeft ten tijde van het verschijnen van dit rapport daarom geen kwantitatief onderzoek kunnen uitvoeren naar indirecte vooringenomenheid van het CUB-proces met betrekking tot

studenten met een migratieachtergrond. Mogelijke indirecte vooringenomenheid in het risicoprofiel of in het verdere verloop van het CUB-proces valt niet uit te sluiten.

Hoe de biasmeting voor migratieachtergrond zou zijn uitgevoerd in het geval van voldoende beschikbare gegevens wordt toegelicht in [3.1 Kwantitatieve analyse](#).

Bevinding 5 – De selectiecriteria en risicocategorieën die zijn gebruikt in het risicoprofiel zijn naar huidige normen onvoldoende onderbouwd, zeker ten aanzien van mogelijke vooringenomenheid.

De selectiecriteria van het risicoprofiel zijn grotendeels gebaseerd op ervaringsgegevens en zogenoemd gezond verstand van medewerkers. De onderbouwing voor geschiktheid van deze criteria is niet gedocumenteerd. Het ontstaan van het risicoprofiel is getraceerd naar een serie workshops in 2010. Hoewel ervaring en gezond verstand belangrijk en nuttig zijn, vormen *self fulfilling prophecies*, *confirmation biases* en het onbedoeld gebruiken van proxycriteria (zie [Box 1](#)) hier een reëel risico. De wegingsfactoren zijn gebaseerd op een datastudie uitgevoerd in 2010. Documentatie over de oorsprong en gehanteerde methodologie voor het samenstellen van de controlegroep in deze datastudie ontbreekt. Zonder nader onderzoek naar de representativiteit van deze controlegroep kunnen in een handmatig geselecteerde populatie bepaalde groepen over- of ondervertegenwoordigd zijn. Door op die populatie wegingsfactoren te baseren, bestaat het risico op het ontstaan van negatieve *feedback loops*. Dat betekent dat groepen die in de controlegroep oververtegenwoordigd zijn een hogere wegingsfactor krijgen, waardoor deze groepen later door het risicoprofiel systematisch een overmatige risicoscore wordt toebedeeld. Het onderbouwen van de keuze voor bepaalde selectiecriteria en wegingsfactoren en hun mogelijke invloed op vooringenomenheid wordt voorgeschreven door de nu geldende kaders voor algoritmes zoals de Impact Assessment Mensenrechten en Algoritmes (IAMA) en het Onderzoekskader Algoritmes van de ADR.

Een analyse over de onderbouwing van de gebruikte selectiecriteria en categorieën in het risicoprofiel wordt gegeven in [5.1 Resultaten kwalitatieve analyse risicoprofiel](#).

Bevinding 6 – Er is niet gebleken van intern onderzoek naar mogelijke vooringenomenheid van het gebruikte risicoprofiel, noch bij het opstellen van het risicoprofiel noch tijdens gebruik ervan.

Box 1

Wat zijn proxyvariabelen?

Met proxykenmerken worden ogenschijnlijk neutrale gegevens bedoeld die sterk gecorreleerd zijn aan andere wél gevoelige kenmerken. Door een sterke correlatie kan men door proxykenmerken te gebruiken (mogelijk onbedoeld) ook de gevoelige groep raken. Lidmaatschap van een vrouwelijke studentenvereniging is bijvoorbeeld een proxy voor het persoonskenmerk 'vrouw'; het bezitten van een seizoenkaart voor een voetbalclub is mogelijk een proxy voor het persoonskenmerk 'man'. Bij overheidshandelen moet hiermee rekening worden gehouden. Wellicht is het toegestaan om onderscheid te maken tussen burgers met en zonder seizoenkaart, maar niet tussen mannen en vrouwen.

Risicoprofilering is een krachtig en door de Centrale Raad van Beroep toegestaan middel voor het vergroten van effectiviteit, efficiëntie en kostenbesparing van de bestrijding van onrechtmatig gebruik van sociale voorzieningen.⁶ Aan het gebruik zijn echter voorwaarden verbonden, waaronder dat het gemaakte onderscheid geschikt, noodzakelijk en proportioneel dient te zijn. Dit moet worden gewaarborgd bij het opstellen van het risicoprofiel en de inzet daarvan. Er is niet gebleken van bewustzijn binnen DUO (en/of bij betrokken instanties als het Ministerie van OCW, de ADR, de Tweede Kamer na invoering van CUB in 2012, etc.) dat het gebruik van het risicoprofiel in het CUB-proces een risico op (indirecte) ongelijke behandeling met zich meebrengt. Er is ook niet gebleken dat waarborgen ingesteld zijn om dergelijke ongelijke behandeling te monitoren of tegen te gaan. Er is geen vooringenomenheid gemeten en er is niet onderzocht of overwogen of de gebruikte criteria en wegingsfactoren wellicht proxykenmerken waren voor bepaalde demografieën.

Een toelichting op het ontbreken van intern onderzoek naar vooringenomenheid kan worden gevonden in [5.2 Processen om het risico op vooringenomenheid tegen te gaan](#).

Bevinding 7 – Er is niet gebleken van interne normen voor het gebruik van algoritmes.

In het onderzoek zijn geen richtlijnen gevonden die DUO intern gebruikt voor het tegengaan van vooringenomenheid of het mitigeren van andere risico's bij het gebruik van algoritmes. Er lijkt niet te zijn vastgelegd hoe DUO invulling geeft aan algemene wettelijke kaders en richtlijnen van andere instanties voor de samenstelling van risicoprofielen. Als die richtlijnen er wel zijn, is in ieder geval niet de hele organisatie zich daarvan bewust. Richtlijnen ontbreken op ten minste twee onderdelen die relevant zijn voor het tegengaan van (indirecte) discriminatie. Ten eerste voor de vaststelling bij welke mate van vooringenomenheid sprake is van indirecte discriminatie. Ten tweede om vast te stellen onder welke omstandigheden indirecte discriminatie naar het oordeel van DUO gerechtvaardigd is, omdat wordt voldaan aan de juridische eisen van geschiktheid, noodzakelijkheid en proportionaliteit. In de periode 2012-2022 hebben er wel jaarlijkse effectiviteitsrapportages, privacy-audits en misbruik-en-oneigenlijk-gebruik (M&O) controles plaatsgevonden. Aselecte steekproeven hebben plaatsgevonden in 2010, 2014 en 2017, primair met het doel om de effectiviteit van het CUB-proces en het financieel 'restrisico' in de CUB-populatie vast te stellen. Het gebrek aan richtlijnen voor de omgang met (algoritmische) risicoprofilering heeft eraan bijgedragen dat er bij de controles geen oog is geweest voor mogelijke vooringenomenheid.

Een toelichting op het ontbreken van interne normen voor gebruik van algoritmes wordt gegeven in [5.2 Processen om het risico op vooringenomenheid tegen te gaan](#).

⁶ Onder meer CRvB 8 september 2015, ECLI:NL:CRVB:2015:3249, r.o. 4.5.; zie ook Toetsingskader Discriminatie door Risicoprofielen van het College voor de Rechten van de Mens (2021) Richtlijn 1.

1.3 Aanbevelingen

Uit de genoemde bevindingen volgen de onderstaande aanbevelingen.

Aanbeveling 1 – Zorg voor een meer omvattende onderbouwing van het risicoprofiel voordat het mogelijk weer in gebruik wordt genomen, onder meer met behulp van een normatief kader.

Inherent aan het gebruik van risicoprofilering in het CUB-proces is dat er onderscheid wordt gemaakt tussen groepen uitwonende studenten. Deels is dit de bedoeling: geen *bug*, maar een *feature*. Het maken van onderscheid kan echter ook juridische en maatschappelijke normen overschrijden. Dit is bijvoorbeeld het geval wanneer het onderscheid niet geschikt, noodzakelijk en proportioneel is. Tegelijkertijd kan onderscheid dat juridisch geoorloofd is toch als maatschappelijk onwenselijk worden beschouwd. Hierbij kan gedacht worden aan onderscheid dat zich niet richt op een juridisch beschermd (persoons) kenmerk, maar op een ander kenmerk zoals opleidingsniveau. Het is aan te bevelen om op voorhand kaders op te stellen waaraan getoetst kan worden voor welke *vormen* en in welke *mate* van onderscheid als (on)wenselijk worden beschouwd. Deze kaders hebben specifiek betrekking op het toetsen van criteria die in het risicoprofiel worden opgenomen. Het voorgestelde Fase-2 onderzoek van Algorithm Audit biedt aanknopingspunten om dergelijke normatieve afwegingen te maken. Het toetsen van een risicoprofiel aan een kader moet ook een kwantitatieve onderbouwing bevatten van de gebruikte selectiecriteria, waarvoor de analyses in dit rapport een voorzet doen. Om het selectieprofiel kwantitatief te onderbouwen, en niet alleen de effectiviteit van het CUB-proces te meten, wordt aanbevolen om een grotere aselechte steekproef te trekken dan de aselechte steekproeven uit 2014 en 2017.

Aanbeveling 2 – Er moet nader onderzoek worden gedaan of er een verband bestaat tussen de groepen die bovenmatig vaak handmatig zijn geselecteerd voor huisbezoek en studenten met een migratieachtergrond.

Of groepen die tijdens de handmatige selectie bovenmatig vaak zijn geselecteerd ook bovenmatig vaak een migratieachtergrond hebben, is niet op voorhand duidelijk en moet nader worden onderzocht. De gehanteerde werkinstructies, bijvoorbeeld de richtlijn om bijzonderheden in de woonsituatie nader te onderzoeken (denk aan: combinatie jong en bij familie inwonend), kunnen bijdragen aan mogelijke oververtegenwoordiging van demografische groepen, waaronder migratieachtergrond. Daarnaast is het proces van handmatige selectie vatbaar voor onderscheid op basis van latente kenmerken. Bij de lijsten die medewerkers gepresenteerd krijgen tijdens handmatige selectie zijn, naast de criteria uit het risicoprofiel en de risicoscore, ook andere kenmerken zichtbaar van studenten, zoals naam, adres en geboortedatum. Hier bevindt zich een risico voor (onbewuste) vooringenomenheid ten opzichte van migratieachtergrond of andere kenmerken in handmatige

selectie. De invloed die deze en andere kenmerken spelen tijdens de handmatige selectie voor huisbezoek dient te worden onderzocht. Daarnaast moet een verbetertraject worden ingesteld om werkinstructies voor handmatige selectie te verbeteren. Zie [Aanbeveling 3](#).

Aanbeveling 3 – Voorkom het gebruik van hetzelfde selectie criterium in verschillende stappen van het CUB-proces. Stel daartoe een onderzoek en verbetertraject in voor het proces van handmatige selectie, waaronder de werkinstructies.

Er is vastgesteld dat studenten die binnen een afstand van 2km tot hun ouders(s) staan ingeschreven aanzienlijk vaker handmatig zijn geselecteerd voor een huisbezoek dan men op basis van de door het risicoprofiel toegekende risicoscores kan verwachten. Voor de hand ligt dat specifieke werkinstructies die aansporen tot het handmatig selecteren van studenten die nabij het ouderlijk adres staan ingeschreven hier de oorzaak van zijn. Deze bevindingen geven aanleiding om de werkinstructies voor medewerkers in het selectieproces tegen het licht te houden. Als de door het model toegekende risicoscore al kenmerken zoals kleine afstand tot ouder(s) meeweegt, kan een nadere instructie om diezelfde kenmerken handmatig vaker te selecteren een overreactie veroorzaken. Daardoor worden studenten met bijvoorbeeld een kleine afstand tot ouder(s) onterecht veel vaker geselecteerd voor huisbezoek. Het algemene proces van handmatige selectie moet nader worden onderzocht om verbeterpunten vast te stellen. Het huidige proces is ondoorzichtig, omdat de procedure om handmatig studenten in en uit te sluiten van huisbezoek op basis van hun risicoscore's en relevante kenmerken niet een eenduidig protocol is. Het is lastig te achterhalen hoe medewerkers precies te werk gaan en het proces is daardoor slecht controleerbaar. Daarnaast kunnen latente persoonskenmerken die zichtbaar zijn voor medewerkers (zoals naam, adres, geboortedatum) een risico vormen voor (onbewuste) vooringenomenheid. Het anonimiseren van studenten en het exclusief tonen van alleen de relevante kenmerken en de risicoscore zou een mogelijkheid zijn. Nader onderzoek moet uitwijzen welke verbeteringen kunnen worden aangebracht in het handmatige selectieproces.

Aanbeveling 4 – Stel een algoritmemanagementbeleid op om de risico's te verkleinen die de inzet van algoritmes met zich meebrengt.

Stel beleid op ten behoeve van een consistente en verantwoorde omgang met algoritmes en risicoprofilering. Dergelijk beleid kan opgesteld worden in verschillende dimensies. Hierbij kan gedacht worden aan:

- > **Bestuur:** centralisatie van belangrijke besluitvorming in comités, het afstemmen taken en verantwoordelijkheden binnen de bestaande organisatiestructuur, de implementatie van 3 lines-of-defence (3LoD) risicobeheersingsraamwerk.
- > **Documentatie:** het opstellen van gestandaardiseerde documentatievereisten en werkinstructies.
- > **Algoritme-inventarisatie:** centraal overzicht van alle algoritmes⁷ binnen de organisatie ten behoeve van informatieverzoeken, monitoring en evaluatie.

⁷ Voor een definitie van een algoritme zie: <https://algoritmes.overheid.nl/nl/footer/over-algoritmes>.

- > **Processen:** Opstellen van ontwikkelings- en valideringsprocessen voor ieder algoritme.
- > **Toezicht en rapportage:** Toezien op rapportage over algoritmerisico's bevordert risico-georiënteerde werkwijzen.

Voor een andere opdeling kan ook inspiratie worden opgedaan uit bestaande kaders voor verantwoorde omgang met algoritmes in de publieke sector.⁸

1.4 Leeswijzer

In [Hoofdstuk 2](#) wordt de werkwijze voor de feitenvaststelling toegelicht, wordt de chronologie van het CUB-proces beschreven en worden de verschillende stappen in het CUB-proces behandeld.

In [Hoofdstuk 3](#) worden voor zowel de kwantitatieve als de kwalitatieve analyse de onderzoeksmethodologie en onderzoeksvragen geïntroduceerd.

In [Hoofdstuk 4](#) worden de resultaten van de kwantitatieve analyse besproken.

In [Hoofdstuk 5](#) worden de resultaten van de kwalitatieve analyse besproken.

In [Hoofdstuk 6](#) staan enkele voorbehouden bij dit onderzoeksrapport.

In [Hoofdstuk 7](#) volgt de conclusie in de vorm van bevindingen en aanbevelingen.

In [Appendix A](#) wordt verwezen naar relevante literatuur.

In [Appendix B](#) staat een voorbeeld van een query om data op te halen uit de datawarehouse van DUO.

In [Appendix C](#) wordt aanvullende informatie gegeven over de uitgevoerde statistische toetsen.

In [Appendix D](#) staat de volledige documentenlijst.

⁸ Zie bijvoorbeeld het Implementatiekader 'Verantwoorde inzet van algoritmes' van de Rijksoverheid <https://www.rijksoverheid.nl/documenten/rapporten/2023/06/30/implementatiekader-verantwoorde-inzet-van-algoritmen>, Onderzoekskader algoritme van de Audit Dienst Rijk (ADR) <https://www.rijksoverheid.nl/documenten/rapporten/2023/07/11/onderzoekskader-algoritmes-adr-2023> en het Toetsingskader algoritmes van de Algemene Rekenkamer <https://www.rekenkamer.nl/onderwerpen/algoritmes-digitaal-toetsingskader>.

2. Achtergrond Controle Uitwonendenbeurs: feitenvaststelling

Hieronder volgen de werkwijze van de feitenvaststelling (§2.1), de chronologie van het CUB-proces (§2.2) en een overzicht van het CUB-proces (§2.3).

2.1 Werkwijze feitenvaststelling

De feitenvaststelling heeft plaatsgevonden op basis van een vooraf vastgestelde methode en bestond uit een bureauonderzoek en workshops. DUO heeft documenten beschikbaar gesteld over het CUB-proces. Algorithm Audit heeft die documenten onderzocht en naar aanleiding van dat onderzoek nieuwe documenten opgevraagd. Daarop zijn door DUO nieuwe documenten aangeleverd, die zijn onderzocht. Deze stappen zijn een aantal keer doorlopen. In totaal zijn ruim 125 documenten aangeleverd en onderzocht. Naar deze documenten wordt verwezen volgens de afkortingen zoals bijgehouden in [Appendix D – Documentlijst](#).

Tegelijk met dit bureauonderzoek zijn twee workshops gehouden met medewerkers van DUO die in het verleden bij het CUB-proces betrokken waren. Tijdens deze workshops zijn door Algorithm Audit vragen gesteld en is door DUO achtergrondinformatie gegeven over het CUB-proces. In overleg met DUO is besloten in dit rapport niet te noemen met welke medewerkers er is gesproken en wordt er niet uit de workshops geciteerd. Dit om te waarborgen dat medewerkers vrijuit konden spreken. De voordelen van dergelijke workshops zijn het snel kunnen ontvangen van informatie, het kunnen koppelen van informatie die zich in verschillende delen van de organisatie bevindt en het eenvoudiger kunnen begrijpen van grote lijnen die uit documentanalyse niet altijd naar voren komen. Nadelen zijn het gevaar van bias bij de selectie van de workshopgroep, het risico van verborgen agenda's van de deelnemers en groepsdruk. Hierdoor en door andere invloeden kan er tijdens de workshops onvolledige of onjuiste informatie zijn gedeeld. Doordat de inzichten waarop dit rapport is gebaseerd verkregen zijn uit een combinatie van schriftelijke documenten en de workshop, zijn deze risico's zo veel als mogelijk ondervangen.

In aanvulling op de workshops heeft een focusgroep plaatsgevonden waarin met datawarehousing- en CUB-experts de benodigde historische data is opgehaald. Aanvullende data zijn na de focusgroep digitaal gedeeld. Op basis van de initieel gedeelde data is een analyse gemaakt en een eerdere versie van dit rapport gepubliceerd. In februari 2024 bleek dat de gedeelde historische data door het datawarehousing-team incorrect was. In de data die is gedeeld en uitsluitend over 1 januari 2014 en 1 januari 2019 zou gaan, bleek ook data uit respectievelijk 2013 en 2018 aanwezig te zijn. In de huidige versie van dit rapport staat de – voor zover bij Algorithm Audit bekend en voor zover door haar te controleren – correcte data.

Een tussentijdse versie van het feitenonderzoek en de kwalitatieve analyse zijn op 25 oktober en 14 december 2023 aan DUO voorgelegd. Daarop zijn aanvullingen en commentaren ontvangen die in deze versie zijn verwerkt.

Er zijn kwaliteitscontroles gedaan op het eindrapport door medewerkers en bestuurders van Algorithm Audit die verder niet bij het project betrokken zijn geweest.

2.2 Chronologie CUB

Op 31 oktober 2007 ontving de voorganger van DUO signalen van de Sociale Recherche Twente (SRT) over onrechtmatig gebruik met de uitwonendenbeurs. De SRT berichtte aan de Minister dat zij regelmatig dergelijk onrechtmatig gebruik tegenkwamen, maar geen bevoegdheden hadden om hiertegen op te treden [F.7,F.8]. Op dat moment werd onrechtmatig gebruik met de uitwonendenbeurs uitsluitend administratief gecontroleerd door het adres van de als uitwonende geregistreerde student te vergelijken met de adressen van zijn/haar ouder(s) in de Gemeentelijke Basis Administratie (GBA). Huisbezoeken werden niet afgelegd en er lag een zware bewijslast bij DUO [F.8]. In juli 2009 zijn door Tweede Kamerleden vragen aan de Minister van OCW gesteld over onrechtmatig gebruik van de uitwonendenbeurs [A.16]. Daarop is in oktober 2009 een rapport gepubliceerd door de Sociale Inlichtingen- en Opsporingsdienst (SIOD) [o.a. A.17]. Ook heeft in oktober 2009 een workshop plaatsgevonden over de aanpak van onrechtmatig gebruik van de uitwonendenbeurs [o.a. A.17]. Aanwezig waren de voorganger van DUO, de SRT, het Openbaar Ministerie, de Basisadministratie Persoonsgegevens en Reisdocumenten, de SIOD en het Ministerie van OCW. Tijdens en na de workshop zijn verschillende maatregelen bedacht voor het tegengaan van onrechtmatig gebruik. Onder meer is de Pilot Twente gestart [o.a. A.17].

In de Pilot Twente is in een iteratief proces tussen de voorganger van DUO en de SRT het risicoprofiel opgesteld. De voorganger van DUO stelde risicocriteria op. Dat waren in eerste instantie de afstand tussen de uitwonende student en zijn/haar ouder(s), afstand tussen de uitwonende student en de onderwijsinstelling, een groot aantal inschrijvingen op een bepaald adres en bepaalde verhuismomenten. Aan de hand van die risicocriteria heeft de SRT onderzoeken uitgevoerd, waarna de criteria zijn verfijnd [A.17]. Uit gesprekken in 2023 volgt dat de risicocriteria zijn gekozen op basis van ervaring, gezond verstand en door handmatig te zoeken naar patronen in lijsten met uitwonende studenten. Dit ging bijvoorbeeld om situaties waarin kinderen van twee of meerdere gezinnen zich bij het adres van het andere gezin inschreven; kruislings of in een carrousel.

In de eerste helft van 2010 is de pilot Twente uitgebreid met nieuwe deelnemers (te weten de Gemeenten Amsterdam, Rotterdam, Den Haag, Utrecht en de Sociale Recherche Noord-Holland) [F.11]. In maart 2010 is een workshop risicoprofilering geweest waaraan verschillende pilotgemeenten, DUO, OCW en SIOD deelnamen [o.a. E.1-2]. Voor het vaststellen van het risicoprofiel is gekozen voor de methode 'beslisboom' [o.a. E.5]. Daarbij worden inschattingen van onrechtmatig gebruik gemaakt aan de hand van deelpopulaties uitwonende studenten. Ook zijn tussen DUO en de pilotgemeenten op basis van werkafspraken gegevens van studenten en risicoprofielen uitgewisseld [E.5]. Ook is in 2010 door DUO een datawarehouse ingericht met daarin data van alle uitwonende studenten op basis waarvan een analyse van de selectiecriteria is gedaan [F.11].

In 2010 is het risicoprofiel verfijnd door op basis van 934 studenten de wegingsfactoren van drie selectiecriteria en de onderlinge relatie tussen die wegingsfactoren te bepalen [A.47]. De onderbouwing is niet gedocumenteerd.

Ook is in 2010 door ADR een aselecte steekproef gedaan onder alle studenten die een uitwonende beurs ontvingen [A.47].

In 2012 is gebruik van het risicoprofiel gestart [A.19].

In 2012 is een wetswijziging in werking getreden waardoor DUO extra middelen kreeg om onrechtmatig gebruik met de uitwonendenbeurs tegen te gaan [o.a. A.19].

In 2013 zijn 366 studerenden gecontroleerd die in Groningen in een studentenhuis woonden. Gezien de praktische en juridische moeilijkheden bij het controleren van studentenhuisen en de lage hoeveelheid onrechtmatig gebruik die bij de huisbezoeken in 2013 zijn vastgesteld, zijn na 2013 studentenhuisen van controle uitgezonderd [A.11].

In 2014 is opnieuw door de ADR een aselecte steekproef gedaan [A.40]. De gehele populatie studenten die een uitwonende beurs ontvingen is onderzocht. Een aanvullende steekproef is gedaan binnen de vier laagste risicocategorieën. Na deskresearch en huisbezoeken is van 3,52 procent van de onderzochte studenten vastgesteld dat zij onrechtmatig een uitwonende studiebeurs ontvangen. Dit is met 95 procent zekerheid statistisch relevant. Een en ander leidt tot de inschatting dat er in 2010 een restrisico was van EUR 12 á 22 miljoen per jaar. DUO concludeert dat het onrechtmatig gebruik tussen 2011 en 2014 is gedaald met EUR 30 miljoen.

In 2015 is het leenstelsel ingevoerd. De basisbeurs – en daarmee de uitwonendenbeurs – wordt afgeschaft in het hoger onderwijs. Er wordt overwogen het risicoprofiel aan te passen, maar dit wordt niet gedaan [A.33].

In 2017 wordt opnieuw een aselecte steekproef uitgevoerd [A.15]. Gezien de afschaffing van het leenstelsel ziet deze steekproef uitsluitend op de groepen mbo 1-2 en mbo 3-4. Bij 3,65 procent van de studenten is vastgesteld dat zij onrechtmatig een uitwonende studiebeurs ontvangen. Dat komt neer op een restrisico EUR 3,75 miljoen per jaar voor mbo-studenten. Inclusief het hoger onderwijs wordt het restrisico in 2017 op EUR 8 miljoen geschat.

In 2018 is het risicoprofiel naar aanleiding van de kinderopvangtoeslagaffaire geherevalueerd. DUO zag geen aanleiding om het CUB-proces aan te passen.

Van 2021 tot en met 2023 zijn door het Hoger Onderwijs Persbureau vragen gesteld aan DUO over het CUB-proces. Er is ook een Woo-verzoek gedaan door Investico. Deze vragen zijn door DUO beantwoord en aan het Woo-verzoek is voldaan.

In 2023 zijn Kamervragen gesteld over het CUB-proces naar aanleiding van nieuwsartikelen gebaseerd op onderzoek van NOS op 3, Investico en het Hoger Onderwijs Persbureau [o.a. A.42]. Daarop is selectie door middel van het risicoprofiel stilgelegd.

2.3 Overzicht CUB-proces

Het CUB-proces is op te delen in zeven stappen:

- > **Stap 0 – Brondata:** Vaststellen van alle ontvangers van een uitwonendenbeurs voor een gegeven peildatum. Naar deze groep wordt gerefereerd als de *uitwonendenbeurspopulatie*.
- > **Stap 1 – Risicoprofiel:** Toekennen van een risicoscore tussen de 0-180 aan alle studenten in de uitwonendenbeurspopulatie aan de hand van het risicoprofiel zoals weergegeven in Tabel 1. Het risicoprofiel werkt als volgt:
 - > Voor de onderwijsvorm van een student wordt een wegingsfactor R_1 toegekend (mbo 1-2: 1,2, mbo 3-4: 1,1, hbo: 1, wo: 0,8);
 - > Voor de combinatie van de leeftijd⁹ van een student en de afstand¹⁰ tussen de postcode van een student en de postcode van een van de ouder(s) wordt een risico R_2 toegekend. De risicofactor voor alle combinaties van leeftijds- en afstands categorieën wordt vermeld in Tabel 1. Merk op: 0km tot ouder(s) houdt in dat een student in hetzelfde postcodegebied (4 cijfers + 2 letters) als de ouder(s) staat ingeschreven maar niet op hetzelfde adres. Een onbekend adres houdt in dat óf het adres van een student niet bekend is óf het adres van de ouder(s) niet bekend is. Dit kan het geval zijn wanneer de ouder(s) niet bekend zijn, overleden zijn of er een hardheidsclausule van toepassing is¹¹;
 - > Aan de hand van een mogelijke afwijking van de leeftijd van de student zoals bekend bij DUO en de leeftijd bekend bij de Gemeente Basis Administratie (GBA), gecombineerd met informatie hoe lang de student al uitwonend is, wordt de risicofactor R_3 bepaald. De volgende leeftijden worden gebruikt om een de risicofactor R_3 te bepalen:
 - > Actuele_leeftijd: leeftijd van student zoals bekend bij DUO op het moment van berekenen risicoscore;
 - > Leeftijd_uitwonend: leeftijd op het moment dat de persoon uitwonend werd;
 - > Leeftijd_huidig_GBA: leeftijd op het moment van de ingangsdatum van het (op dat moment) actuele woonadres uit de Gemeente Basis Administratie (GBA).
 De risicofactor R_3 voor alle combinaties van deze soorten leeftijd wordt gegeven in Tabel 1 [A.53].
 - > De risicoscore wordt als volgt bepaald:

$$\text{Risicoscore} = R_1 * (R_2 + R_3).$$

⁹ De volgende leeftijdscategorieën worden gehanteerd: 15-18, 19-20, 21-22, 23-24, 25-50.

¹⁰ De volgende afstandscategorieën worden gehanteerd: 0km, 1m-1km, 1-2km, 2-5km, 5-10km, 10-20km, 20-50km, 50-500km, onbekend. Merk op dat de exacte afbakening in Tabel 1 gespecificeerd is.

¹¹ Er zijn ook technische redenen waarom het adres van een student op een peildatum (nog) niet beschikbaar is. Zo kan het zijn dat op het moment van inschrijven het adres van de student niet bekend is, maar enkele dagen, weken of maanden later wel.

Risicofactor (R1)		Risicocategorie				
Onderwijsvorm	Factor	Codering	Beschrijving	Ondergrens	Bovengrens	
mbo 1-2	1,2	1	Zeer hoog risico	80	180	
mbo 3-4	1,1	2	Hoog risico	60	79	
hbo	1,0	3	Gemiddeld risico	40	59	
wo	0,8	4	Laag risico	20	39	
		5	Zeer laag risico	1	19	
R1*(R2+R3)		6	Onbekend risico	0	0	
Afstand tot ouder(s)	Leeftijd	Risicoscore (R2)	mbo 1-2	mbo 3-4	hbo	wo
0	15-18	120	144	132	120	96
0	19-20	110	132	121	110	88
0	21-22	105	126	116	105	84
0	23-24	100	120	110	100	80
0	25-50	80	96	88	80	64
1-1000	15-18	100	120	110	100	80
1-1000	19-20	95	114	105	95	76
1-1000	21-22	85	102	94	85	68
1-1000	23-24	75	90	83	75	60
1-1000	25-50	65	78	72	65	52
1001-2000	15-18	95	114	105	95	76
1001-2000	19-20	85	102	94	85	68
1001-2000	21-22	75	90	83	75	60
1001-2000	23-24	65	78	72	65	52
1001-2000	25-50	60	72	66	60	48
2001-5000	15-18	85	102	94	85	68
2001-5000	19-20	75	90	83	75	60
2001-5000	21-22	65	78	72	65	52
2001-5000	23-24	55	66	61	55	44
2001-5000	25-50	45	54	50	45	36
5001-10000	15-18	75	90	83	75	60
5001-10000	19-20	65	78	72	65	52
5001-10000	21-22	55	66	61	55	44
5001-10000	23-24	45	54	50	45	36
5001-10000	25-50	35	42	39	35	28
10001-20000	15-18	50	60	55	50	40
10001-20000	19-20	40	48	44	40	32
10001-20000	21-22	30	36	33	30	24
10001-20000	23-24	25	30	28	25	20
10001-20000	25-50	20	24	22	20	16
20001-50000	15-18	35	42	39	35	28
20001-50000	19-20	25	30	28	25	20
20001-50000	21-22	20	24	22	20	16
20001-50000	23-24	15	18	17	15	12
20001-50000	25-50	10	12	11	10	8
50001-500000	15-18	20	24	22	20	16
50001-500000	19-20	20	24	22	20	16
50001-500000	21-22	15	18	17	15	12
50001-500000	23-24	10	12	11	10	8
50001-500000	25-50	5	6	6	5	4
Onbekend	15-18	0	0	0	0	0
Onbekend	19-20	0	0	0	0	0
Onbekend	21-22	0	0	0	0	0
Onbekend	23-24	0	0	0	0	0
Onbekend	25-50	0	0	0	0	0

ACTUELE_LEEFTIJD_VAN	ACTUELE_LEEFTIJD_TM	LEEFTIJD_UTWONEND_VAN	LEEFTIJD_UTWONEND_TM	LEEFTIJD_HUIDIG_GBA_VAN	LEEFTIJD_HUIDIG_GBA_TM	R3
21	22	17	18		17	5
21	22	17	18		19	0
21	22	19	20		19	0
23	24	17	18		17	15
23	24	17	18		19	10
23	24	17	18		21	0
25	65	17	18		17	30
25	65	17	18		19	25
25	65	17	18		21	15
25	65	17	18		23	0
25	65	17	18		25	0
25	65	19	20		19	25
25	65	19	20		21	0
25	65	19	20		23	0
25	65	19	20		25	0
25	65	21	22		21	15
25	65	21	22		23	0
25	65	23	24		23	0

Tabel 1 – Scoringstabellen voor R_1 , R_2 en R_3 . R_1 is de risicofactor die wordt toegekend aan verschillende onderwijsvormen. R_2 is de risicofactor die wordt bepaald door de combinatie van de afstandscategorie en de leeftijdscategorie van de student. In combinatie met R_3 – of er een afwijking is tussen de leeftijd van de student die bekend is bij DUO en de leeftijd die bekend is bij de Gemeente Basis Administratie (GBA), gecombineerd met informatie hoe lang de student al uitwonend is – wordt de uiteindelijke risicoscore bepaald.

- > **Stap 2 – Opdeling per regio:** Opdelen van de uitwonendenbeurspopulatie, inclusief risicoscore, naar regio. Naar deze groep wordt gerefereerd als de *risicoscore-populatie per regio*.
- > **Stap 3 – Handmatige check door medewerker:** Een medewerker van DUO doorloopt een risicoscore-populatie per regio. De populatie wordt met de DUO-medewerker gedeeld in een document waarin voor iedere student in die regio een lijst met persoonskenmerken staat. Dit gaat onder meer om de risicoscore, de woonsituatie, het adres van de ouder(s), de onderwijsinstelling, de onderwijsvorm van de student.¹² Een medewerker stelt aan de hand van uitsluitingsgronden vast bij welke studenten geen extern onderzoek gerechtvaardigd is.¹³ Uitsluitingsgronden zijn onder meer stelletjes (2 personen die rond dezelfde leeftijd gelijktijdig zijn ingeschreven in de Basis Registratie Personen (BRP)), het wonen in een studentenhuus, wonen bij het Leger des Heils, of het zijn van een alleenstaande ouder. Vervolgens wordt de lijst van studenten in principe op volgorde van hoge naar lage risicoscores doorlopen en geselecteerd voor een huisbezoek. Van deze volgorde (hoog naar laag) kan worden afgeweken. Zo kan een student met een wat lagere risicoscore eerder worden behandeld dan een student met een hogere risicoscore omdat er sprake is van een bijzondere woonsituatie (denk aan: combinatie jong en bij familie inwonend) [H.1 p.10]. Bij de beoordeling van de woonsituatie wordt de BRP geraadpleegd voor (familie-)relaties en de Basisregistratie Adressen en Gegevens (BAG) voor de woonfunctie en woonoppervlakte van een gebouw. Voor een huisbezoek worden volgens de instructies aan de medewerkers specifiek geselecteerd:
 - > Studenten met een risicofactor 1 t/m 3 met een inschrijving nabij hun ouderlijk adres, een inschrijving bij een gezin met kleine kinderen, een inschrijving bij (hoog) bejaarden zonder familierelatie, een inschrijving op een adres zonder woonfunctie, een lange reistijd (>1,5 uur per enkele reis) tussen het woonadres en de onderwijsinstelling en geen stage, inschrijving op een adres met meer bewoners dan 12m² per bewoner, of een onlogische samenstelling van bewoners zoals een jong persoon op hetzelfde adres als meerdere alleenstaande oudere personen in combinatie met een klein woonoppervlak.
 - > Om een nadere beoordeling vragen studenten waarin er geen koppeling met hun ouders is in het Studie Financiering Stelsel, maar wel in de BRP. Hetzelfde geldt voor situaties met een lange reistijd tussen het woonadres en de onderwijsinstelling [H.1 p.10].

De uiteindelijke selectie van studenten voor huisbezoek kan op basis van bovenstaan-

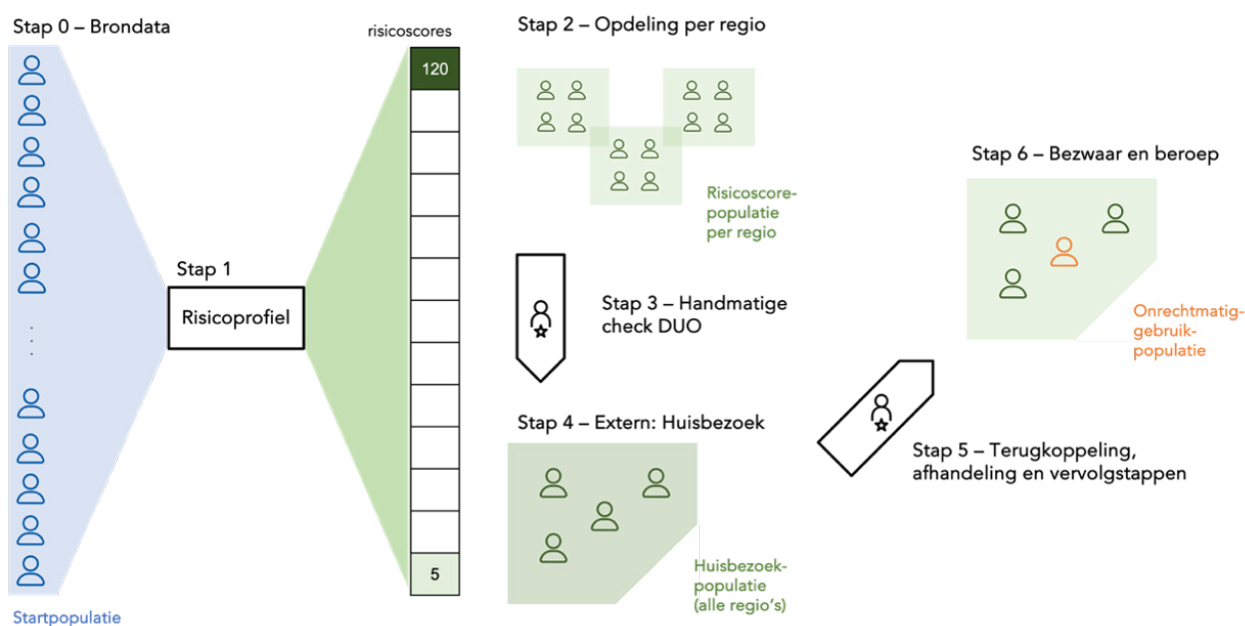
¹² In [A.53] staan alle getoonde kenmerken. Dit zijn, het correspondentienummer, het BSN-nummer, de voor naam, de voorletters, het tussenvoegsel, de achternaam, het geslacht, de geboortedatum, de straatnaam van het GBA-adres, het huisnummer, de toevoeging aan het huisnummer, de aanduiding van het woonadres, de postcode, de woonplaats, de ingangsdatum van het GBA-adres, de vormen van studiefinanciering die de student in de afgelopen maanden ontving, de eerste maand waarin de student als uitwonend geregistreerd stond, of de student inwonend of uitwonend is in de voorafgaande maanden, het studievorm, de studierichting, de instelling waar de student studeert, van beide ouders: alle adresgegevens en of de ouder weigerachtig of overleden is, de risicoscores en risicofactoren voor de huidige en aankomende maand.

¹³ Er is een uitputtende lijst welke studenten uitgesloten worden uit het CUB-proces [H.1].

de criteria dus afwijken van de gesorteerde lijst op risicoscore van hoog naar laag. Hoeveel studenten er worden geselecteerd is afhankelijk van de capaciteitsafspraken die gemaakt zijn met externe onderzoeksbureaus per regio. Naar de groep die is geselecteerd voor een huisbezoek (samengenomen voor alle regio's) wordt gerefereerd als de *huisbezoek-populatie*.

- > **Stap 4 – Huisbezoek:** Externe partij verricht een huisbezoek, verzamelt indien mogelijk gegevens over de rechtmatigheid van de aanspraak op een uitwonende beurs en koppelt resultaten terug aan DUO.¹⁴
- > **Stap 5 – Terugkoppeling, afhandeling en vervolgstappen:** Verwerken van uitkomsten. Vaststellen van rechtmatig en onrechtmatig gebruik. Naar de populatie die onrechtmatig gebruik heeft gemaakt van de uitwonende beurs wordt gerefereerd als *onrechtmatig-gebruik-populatie*.
- > **Stap 6 – Bezwaar- en beroepsprocedures:** De onrechtmatig-gebruik-populatie kan in bezwaar gaan tegen de beschikking van DUO. Naar de studenten die bezwaar indienen wordt gerefereerd als de *bezwaarpopulatie*.

Deze stappen zijn schematisch weergegeven in [Figuur 1](#). Het kwalitatieve deel van dit onderzoek richt zich uitsluitend op stap 1 – risicoprofiel. Het kwantitatieve deel van dit onderzoek richt zich op stap 0-6.



Figuur 1 – Schematische weergave van het CUB-proces.

¹⁴ Indien een huisbezoek niet heeft kunnen plaatsvinden vindt een alternatief onderzoek plaats om te bepalen of een student uit de huisbezoek-populatie wel of niet tot de onrechtmatig-gebruik-populatie behoort. Zie ook [\[A.3\]](#).

3. Onderzoeksmethodologie

Hieronder volgen de werkwijze van de kwantitatieve analyse (§3.1) en de werkwijze van de kwalitatieve analyse (§3.2). Het doel, de opzet en de onderzoeksvragen van beide analyses worden besproken.

3.1 Kwantitatieve analyse

Eerst wordt het doel van de kwantitatieve analyse beschreven, gevolgd door de methodologie van deze analyse die wordt toegelicht aan de hand van de aselechte steekproeven van 2014 en 2017 en de biasmetingen van 2014 en 2019 en bijbehorende onderzoeksvragen.

Doel van de analyse

Het doel van de kwantitatieve analyse is het traceren van mogelijke vooringenomenheid met betrekking tot 1. onderwijsvorm, 2. leeftijd, 3. afstand tot ouder(s) of 4. migratieachtergrond in verschillende stappen van het CUB-proces. Een overzicht van algoritmische handelingen (stap 1-2) en menselijke handelingen (stap 3-5) in het CUB-proces wordt gegeven in [2.3 Overzicht CUB-proces](#). Door per stap een meting uit te voeren kan gericht vervolgonderzoek worden gedaan naar waar in het CUB-proces mogelijk (indirecte) vooringenomenheid zich voordoet. Deze kwantitatieve analyse wordt ook wel een *biasmeting* genoemd.

Opzet van de analyse

Om mogelijke vooringenomenheid in het CUB-proces te onderzoeken is de biasmeting in drieën opgedeeld:

A. Toetsen geschiktheid risicoprofiel voor detectie onrechtmatig gebruik – Een eerste vraag over mogelijke vooringenomenheid in het CUB-proces is of er een verband bestaat tussen onrechtmatig gebruik van de uitwonendenbeurs en de kenmerken waar door het risicoprofiel onderscheid op wordt gemaakt: 1. onderwijsvorm, 2. leeftijd en 3. afstand tot ouder(s). Deze vraag kan worden beantwoord door het analyseren van een *aselechte steekproef*. De werkwijze van deze analyse wordt hieronder toegelicht in [A. Aselechte steekproef 2014 en 2017](#).

B. Toetsen proportionaliteit CUB-risicoprofiel – Mocht de aselechte steekproef wijzen op een verband tussen onrechtmatig gebruik van de uitwonendenbeurs en kenmerken 1-3, dan is de vraag of er proportioneel onderscheid wordt gemaakt op basis van deze kenmerken. Bijvoorbeeld: uit de aselechte steekproef kan blijken dat voor studenten binnen een bepaalde onderwijsvorm de kans op onrechtmatig gebruik 10% groter is, maar dat het risicomodel studenten van deze onderwijsvorm gemiddeld genomen een risicoscore toekent die veel meer dan 10% afwijkt ten opzichte van het gemiddelde van

studenten met andere onderwijsvormen. Deze analyse wordt voor stap 1 (risicoprofiel) en stap 3 (handmatige selectie) van het CUB-proces uitgevoerd. Een toelichting op deze proportionaliteitstoets wordt hieronder gegeven in B. [Biasmeting 2014 en 2019](#).

C. Indirect onderscheid in CUB-proces met betrekking tot migratieachtergrond –

Na analyse van de aselechte steekproef en de proportionaliteitstoets rijst de vraag of onderscheid op basis van 1. onderwijsvorm, 2. leeftijd en 3. afstand tot ouder(s) indirect studenten met een migratieachtergrond benadeelt en, zo ja, in welke fase van het CUB-proces deze vooringenomenheid ontstaat. Om indirecte vormen van vooringenomenheid goed te onderzoeken is toegang nodig tot bijzondere persoonsgegevens, namelijk data of een individuele student wel of geen migratieachtergrond heeft. Hiertoe is bij het CBS een verzoek ingediend om deze microdata te verkrijgen. Tot op heden heeft het CBS deze data niet verstrekt. In de toekomst kunnen deze data mogelijk wel beschikbaar worden gesteld voor eventueel vervolgonderzoek. Alternatieve methoden om data over migratieachtergrond op individueel niveau te verwerken zijn niet voorhanden. Hierdoor is het tot op heden niet mogelijk om deel C van de kwantitatieve analyse uit te voeren.

Overwogen is om publiek toegankelijke data over migratieachtergrond geaggregeerd per postcodegebied van het CBS¹⁵ als alternatief te gebruiken. Algorithm Audit heeft hier om de volgende reden van afgezien: de publiek toegankelijk data zijn gebaseerd op de gehele Nederlandse bevolking. Het is niet na te gaan of de CBS-data representatief is voor de onderzochte populatie uitwonende studenten. Algorithm Audit is van mening dat er sterke bezwaren aan kleven om deze representativiteit wel aan te nemen. Het woongedrag van studenten, zoals woonvorm en verspreiding over het land, maar ook de sociaaleconomische status van studenten is dusdanig specifiek dat deze kenmerken hoogstwaarschijnlijk niet representatief zijn voor de Nederlandse bevolking. Dit betekent dat wanneer er N studenten zijn geselecteerd voor een huisbezoek in een postcodegebied met $x\%$ inwoners met een migratieachtergrond, het niet duidelijk is dat onder deze N studenten dezelfde $x\%$ een migratieachtergrond heeft. Door deze beperking kunnen aan de hand van deze data geen betekenisvolle uitspraken worden gedaan hoe studenten met een migratieachtergrond in het CUB-proces zijn behandeld.

A. Aselechte steekproef 2014 en 2017

De geschiktheid van het gebruikte risicoprofiel in het CUB-proces wordt onderzocht aan de hand van de volgende drie onderzoeksvragen:

Onderzoeksvraag 1

Zijn er op basis van de aselechte steekproef 2014 en 2017 aanwijzingen voor een verband tussen de kenmerken van het risicoprofiel en onrechtmatig gebruik van de uitwonendenbeurs?

¹⁵ Zie <https://www.cbs.nl/nl-nl/cijfers/detail/83503NED>.

Onderzoeksvraag 2

Zijn er op basis van de aselecte steekproef 2014 en 2017 aanwijzingen voor een verband tussen de gehanteerde risicocategorieën 1-6 en onrechtmatig gebruik met de uitwonendenbeurs?

Onderzoeksvraag 3

Hoeveel onrechtmatig gebruik is er geconstateerd in de aselecte steekproef populatie 2014 en 2017?

De werkwijze om deze onderzoeksvragen te beantwoorden worden hieronder toegelicht.

Bij een aselecte steekproef worden willekeurige individuen gekozen uit de uitwonendenbeurspopulatie van studenten. Willekeur is belangrijk omdat het hierdoor mogelijk is conclusies te trekken over de gehele populatie. Om het verband te meten tussen bepaalde kenmerken (bijvoorbeeld onderwijsvorm, leeftijd en afstand tot ouder(s)) en de kans op onrechtmatig gebruik met de uitwonendenbeurs, wordt om deze reden gebruik gemaakt van een aselecte steekproef. Dergelijke verbanden kunnen niet worden gebaseerd op uitkomsten van het CUB-proces, aangezien hierbij niet gegevens van willekeurige individuen zijn verzameld, maar de gegevens van studenten die door DUO zijn geselecteerd voor een huisbezoek. Deze groep studenten die in het CUB-proces zijn geselecteerd is niet met zekerheid representatief voor de gehele (start)populatie, en dus is er voor deze data een kans dat er verkeerde conclusies worden getrokken rondom het verband tussen de kenmerken van het risicoprofiel en de kans op onrechtmatig gebruik.

In 2014 en in 2017 heeft DUO in samenwerking met de ADR steekproeven genomen met als doel de omvang van onrechtmatig gebruik van de uitwonendenbeurs vast te stellen. Details over de getrokken steekproeven kunnen worden gevonden in [A.40]. Deze steekproeven kunnen ook worden gebruikt om onderzoeksvraag 1-3 te beantwoorden. De resultaten van de steekproeven uit 2014 en 2017 worden apart gerapporteerd, omdat de populaties tussen de twee steekproeven veranderd zijn door de introductie van het leenstelsel.

Om onderzoeksvragen 1 en 2 te onderzoeken wordt het verschil in de onrechtmatig gebruik-percentages getest van verschillende onderwijs-, leeftijd- en afstandscategorieën. Stel dat voor groep A het onrechtmatig gebruik-percentage p_A bedraagt, en voor groep B p_B . Vervolgens wordt getest of het verschil tussen p_A en p_B statistisch significant is. Hiermee wordt vastgesteld of studenten met een bepaald profiel vaker onrechtmatig gebruik maken van de uitwonendenbeurs. Formeel gezien worden de volgende nulhypothese H_0 en alternatieve hypothese H_1 getoetst voor een groep A en B met onrechtmatig gebruik-percentages p_A en p_B :

$$H_0: p_A = p_B$$

$$H_1: p_A > p_B$$

Box 2

Z-toets en Fishers exacte toets

Een Z-toets is een statistische toets die wordt gebruikt om de proporties (bijvoorbeeld onrechtmatig gebruik-percentage) tussen twee onafhankelijke groepen te vergelijken. De toets wordt vaak toegepast voor binaire uitkomsten (bijvoorbeeld rechtmatig of onrechtmatig gebruik). De toets wordt gebruikt om te beoordelen of de waargenomen verschillen in proporties tussen de groepen statistisch significant zijn.

Fishers exacte toets is een andere statistische test die ook gebruikt kan worden om de proporties (bijvoorbeeld onrechtmatig gebruik-percentage) tussen twee groepen te vergelijken. Fishers exacte toets is vooral geschikt wanneer de steekproefgrootte klein is.

p_A kan als volgt worden vastgesteld. Zij $N_{A, \text{onrechtmatig gebruik}}$ het aantal individuen voor groep A die onrechtmatig gebruik maken van de uitwonendenbeurs. Zij N_A het totaal aantal individuen behorende tot groep A in de aselecte steekproef. Dan: $p_A = N_{A, \text{onrechtmatig gebruik}} / N_A$. Op dezelfde manier kan p_B worden bepaald. Vervolgens wordt H_0 onderworpen aan een eenzijdige Z-toets (zie Box 2). Gegeven de mogelijk kleine groepsgrootte wordt ter controle Fisher's eenzijdige toets uitgevoerd (zie ook Box 2). Een verdere toelichting op deze statistische toetsen kan worden gevonden in [Appendix C – Aanvullende informatie statistische analyse](#).

Deze statistische toets wordt toegepast om alle onderwijs-, leeftijds- en afstandscategorieën ten opzichte van elkaar te testen. Dus, bijvoorbeeld voor onderwijsvorm worden de volgende scenario's onderworpen aan de Z-toets:

- I. A=mbo 1-2, B=mbo 3-4
- II. A=mbo 1-2, B=hbo
- III. A=mbo 1-2, B=wo
- IV. A=mbo 3-4, B=hbo
- V. A=mbo 3-4, B=wo
- VI. A=hbo, B=wo.

Er is gekozen voor deze stapsgewijze vergelijking omdat het de logica van het gebruikte CUB-risicoprofiel volgt: verschillende onderwijsvormen krijgen stapsgewijs een andere risicoscore. Aan de hand van [Tabel 2](#) kunnen p_A en p_B voor scenario I-VI worden bepaald.

	Onderwijsvorm			
	mbo 1-2	mbo 3-4	hbo	wo
Onrechtmatig gebruik	$N_{\text{mbo 1-2, onrechtmatig}} / N_{\text{mbo 1-2}}$	$N_{\text{mbo 3-4, onrechtmatig}} / N_{\text{mbo 3-4}}$	$N_{\text{hbo, onrechtmatig}} / N_{\text{hbo}}$	$N_{\text{wo, onrechtmatig}} / N_{\text{wo}}$

Tabel 2 – Benodigde informatie om Z-toets uit te voeren om de statistische significantie van verschil in onrechtmatig gebruik-percentage per onderwijsvorm te toetsen.

Dezelfde werkwijze wordt gehanteerd voor het gemaakte onderscheid op basis van de gehanteerde leeftijds-, afstands- en risicocategorieën. Deze analyse geeft antwoord op onderzoeksvragen 1-3. De resultaten worden gepresenteerd in [4.1 Resultaten aselechte steekproef 2014 en 2017](#).

Merk op dat in 2010 een poging is ondernomen om [Tabel 2](#) op te stellen op basis van n=934 studenten [\[A.47\]](#). De uitkomsten van deze analyse zijn echter onvoldoende zorgvuldig gedocumenteerd om nu te kunnen hergebruiken of analyseren.¹⁶ Zie ook [5.1 Resultaten kwalitatieve analyse risicoprofiel](#).

B. Biasmeting 2014 en 2019

De over- of onderreactie van het risicoprofiel (stap 1) en handmatige selectie door medewerkers (stap 3) met betrekking tot kenmerken 1. onderwijsvorm, 2. leeftijd, 3. afstand tot ouder(s) wordt onderzocht aan de hand van een biasmeting. Voor deze biasmeting zijn twee uitwonendenbeurspopulaties relevant.

De eerste uitwonendenbeurspopulatie behelst studenten die voor de introductie van het leenstelsel een uitwonendenbeurs hebben ontvangen. Naar deze populatie, bestaande uit mbo-, hbo- en wo-studenten, wordt gerefereerd als *uitwonendenbeurspopulatie-2014*. Als peildatum voor het vaststellen van deze populatie is 01-01-2014 gekozen. Deze peildatum is gekozen omdat het merendeel van de CUB-onderzoeken per kalenderjaar wordt afgerond. Voor alle studenten die op 01-02-2014 een uitwonendenbeurs voor januari 2014 zouden ontvangen, zijn uitkomsten van het CUB-proces in kalenderjaar 2014 verzameld. De query die in de SQL-database is uitgevoerd om uitwonendenbeurspopulatie-2014 op te halen wordt weergegeven in [Appendix B – Datawarehouse query](#).

De tweede uitwonendenbeurspopulatie betreft studenten die na de introductie van het leenstelsel een uitwonendenbeurs hebben ontvangen. Naar deze populatie, die uit mbo-studenten en uitfaserende hbo- en wo-studenten bestaat, wordt gerefereerd als *uitwonendenbeurspopulatie-2019*. Als peildatum voor het vaststellen van deze populatie is 01-02-2019 gekozen. Kalenderjaar 2019 is gekozen om twee redenen. Allereerst omdat in 2019 zoveel mogelijk hbo- en wo-studenten die nog een uitwonendenbeurs zouden ontvangen – omdat zij zijn begonnen met studeren in het de periode voor het afschaffen van de basisbeurs – zijn afgestudeerd. Ten tweede omdat 2019 het laatste jaar is waarop CUB-onderzoeken regulier hebben plaatsgevonden voordat de onderzoeken zijn aangepast door de coronamaatregelen in 2020 en 2021.

Zoals in [2.3 Overzicht CUB-proces](#) is toegelicht, worden uitwonendenbeurspopulatie-2014 en uitwonendenbeurspopulatie-2019 door de verschillende stappen in het CUB-proces gevolgd. De onderzoeksvragen voor de proportionaliteitstoets worden geformuleerd met

¹⁶ Zo is onbekend hoe de 934 studenten zijn geselecteerd. Het kan niet worden nagegaan of dit aselekt is gedaan. Ook zijn in het beschikbare Excelbestand gebruikte formules weggevalen waardoor het niet te traceren is hoe tot bepaalde cijfers is gekomen, wat het reconstrueren van de analyse compliceert.

behulp van twee kernbegrippen uit de literatuur rondom eerlijke algoritmes die allereerst worden geïntroduceerd.

1. Demografische pariteit – Wanneer de actie tussen twee stappen in het CUB-proces onafhankelijk is van 1. onderwijsvorm, 2. leeftijd of 3. afstand tot ouder(s). Bijvoorbeeld: de gemiddelde score toegekend door het risicoprofiel aan studenten met onderwijsvorm A is hetzelfde als de gemiddelde score toegekend aan studenten met onderwijsvorm B. In het geval van stap 1 – risicoprofiel per onderwijsvorm impliceert dat:

$$\text{gemiddelde}(\text{risicoscore} \mid \text{onderwijsvorm}) = \text{gemiddelde}(\text{risicoscore}).$$

Voor stap 3 – kans op handmatige selectie door medewerker per onderwijsvorm impliceert dat:

$$p(\text{selectie voor huisbezoek} \mid \text{onderwijsvorm}) = p(\text{selectie voor huisbezoek}).$$

2. Conditionele demografische pariteit – Ook wel *equalized odds* genoemd. Wanneer de actie tussen twee stappen in het CUB-proces onafhankelijk is van 1. onderwijsvorm, 2. leeftijd of 3. afstand tot ouder(s), gegeven dat er onrechtmatig gebruik is gemaakt van de uitwonendenbeurs. Bijvoorbeeld: de gemiddelde score toegekend door het risicoprofiel aan studenten met onderwijsvorm A afgezet tegen de mate van onrechtmatig gebruik in de subpopulatie studenten met onderwijsvorm A is hetzelfde als de gemiddelde score toegekend aan studenten met onderwijsvorm B afgezet tegen de mate van onrechtmatig gebruik in de subpopulatie studenten met onderwijsvorm B. In het geval van stap 1 – risicoprofiel per onderwijsvorm impliceert dat:

$$\text{gemiddelde}(\text{risicoscore} \mid \text{onderwijsvorm, onrechtmatig gebruik}) = \text{gemiddelde}(\text{risicoscore} \mid \text{onrechtmatig gebruik}).$$

Merk op het meten van conditionele pariteit voor stap 3 – handmatige selectie door medewerker niet geschikt is, aangezien de kans om geselecteerd te worden door een medewerker gegeven dat er onrechtmatig gebruik is gemaakt van de uitwonendenbeurs altijd 1 is.

Demografische pariteit en conditionele demografische pariteit worden beide gemeten om mogelijke vooringenomenheid in het CUB-proces op een zo compleet mogelijke manier te onderzoeken. Aan de hand van deze begrippen worden de volgende onderzoeksvragen geformuleerd.

Onderzoeksvraag 4

Is er sprake geweest van demografische pariteit voor de risicoscore-populatie (na stap 1 – risicoprofiel) en de huisbezoekpopulatie (na stap 3 – handmatige selectie door medewerker)?

Merk op: aan de hand van onderzoeksvraag 4 wordt getracht te meten of er afwijkende patronen zich voordoen bij het toekennen van risicoscores door het risicoprofiel of bij de handmatige selectie van studenten voor een huisbezoek door medewerkers.

Onderzoeksvraag 5

Is er sprake geweest van conditionele demografische pariteit voor de risicoscore-populatie (na stap 1 – risicoprofiel)?

Merk op: aan de hand van onderzoeksvraag 5 wordt getracht te meten welke risicoscores zijn toegekend aan studenten waarvan bleek dat zij onrechtmatig gebruik hebben gemaakt van de uitwonendenbeurs. Aan de hand hiervan kunnen lacunes in het risicomodel worden gevonden.

Onderzoeksvraag 6

Hoeveel onrechtmatig gebruik wordt er geconstateerd in de populatie geselecteerd door het CUB-proces in 2014 en 2019?

Voortbouwend op de eerder geïntroduceerde notatie in [A. Aselecte steekproef 2014 en 2017](#) kunnen deelvragen 4-6 met betrekking tot onderwijsvorm worden beantwoord door Tabel 3 vast te stellen voor uitwonendenbeurspopulatie-2014 en uitwonendenbeurspopulatie-2019. Merk op dat $N_{startpopulatie-X} = N_{risicoscore-populatie-X}$ voor $X = 2014$ en 2019 aangezien iedere student die een uitwonendenbeurs ontvangt een risicoscore krijgt.

#	Selectie voor huisbezoek	Onderwijsvorm	Onrechtmatig gebruik
1	Ja	A	Ja
2	Nee	B	Nee
...			
N	Ja	C	Nee

Tabel 3 – Benodigde informatie om de (conditionele) demografische pariteit van stap 3 – handmatige selectie van studenten te bepalen. Om de (conditionele) demografische pariteit voor stap 1 – risicoprofiel te bepalen is ook de toegekende risicoscore per student (#) nodig.

Op basis van de data uit [Tabel 3](#) kan $p(\text{selectie voor huisbezoek})$, $p(\text{selectie voor huisbezoek} \mid \text{onderwijsvorm})$ en $p(\text{selectie voor huisbezoek} \mid \text{onderwijsvorm, onrechtmatig gebruik})$ als volgt worden vastgesteld:

$$p(\text{selectie voor huisbezoek}) = \frac{N_{\text{selectie voor huisbezoek} = \text{Ja}}}{N}$$

$$p(\text{selectie voor huisbezoek} \mid \text{onderwijsvorm}) = \frac{N_{\text{selectie voor huisbezoek} = \text{Ja, onderwijsvorm} = A}}{N_{\text{onderwijsvorm} = A}}$$

$$p(\text{selectie voor huisbezoek} \mid \text{onderwijsvorm, onrechtmatig gebruik}) = \frac{N_{\text{selectie voor huisbezoek} = \text{Ja, onderwijsvorm} = A, \text{onrechtmatig gebruik} = \text{Ja}}}{N_{\text{onderwijsvorm} = A, \text{onrechtmatig gebruik} = \text{Ja}}}$$

De (conditionele) demografische pariteit voor stap 1 – risicoprofiel kan worden bepaald door de gemiddelde risicoscore te bepalen voor bovengenoemde groepen.

Beperkingen van gekozen onderzoeksmethodologie

Voor de gekozen onderzoeksmethodologie om vooringenomenheid in het CUB-proces kwantitatief te onderzoeken gelden enkele beperkingen.

Ten eerste, door enkel de gekozen data uit de datawarehouse van DUO te analyseren (1. onderwijsvorm, 2. leeftijd en 3. afstand tot ouder(s)) kan vooringenomenheid van het CUB-proces met betrekking tot migratieachtergrond slechts zeer beperkt worden onderzocht. Afwijkingen in de data over het CUB-proces kunnen alleen duiden op een mogelijk verband met vooringenomenheid met betrekking tot migratieachtergrond. Dit verband kan niet hard worden gemaakt.

De biasmeting meet alleen afwijkingen met betrekking tot de kenmerken die zijn gebruikt in het risicoprofiel (onderwijsvorm, leeftijd en afstand tot ouder(s)). Andere variabelen, die niet in het risicoprofiel zijn opgenomen, maar mogelijk wel voorspellende waarde hebben om onrechtmatig gebruik met de uitwonendenbeurs te detecteren zijn niet bij dit onderzoek betrokken.

3.2 Kwalitatieve analyse

Doel van de analyse

In de kwalitatieve analyse wordt de geschiktheid van het risicoprofiel in het CUB-proces getoetst aan de normen die daar op het moment van uitkomen van dit rapport voor gelden. Het doel is om te reflecteren op het risicoprofiel en DUO handvatten te geven voor het effectief tegengaan van onrechtmatig gebruik met de uitwonendenbeurs in de toekomst.

Opzet van de analyse

De onderzoeksvragen voor de kwalitatieve analyse luiden:

Onderzoeksvraag 7

Voldeed het risicoprofiel in het CUB-proces aan de normen die tegenwoordig worden gesteld aan door de overheid gebruikte algoritmen?

Onderzoeksvraag 8

Als het antwoord op onderzoeksvraag 7 negatief is, wat is er nodig om in de toekomst verantwoord het risicoprofiel te gebruiken?

Om deze onderzoeksvragen te beantwoorden wordt de IRAC-methode gebruikt – een acroniem voor *issue, rule, analysis, conclusion*.

Issue. Eerst is vastgesteld wat de onderzoeksvragen en de kaders zijn van het onderzoek.

Vervolgens is feitelijk onderzoek naar het risicoprofiel gedaan. Zie daartoe [2.2 Chronologie CUB](#).

Rule. Dit rapport beoogt niet een normatief of juridisch antwoord te geven op de vraag of DUO goed of fout heeft gehandeld. Toch acht Algorithm Audit het relevant om het risicoprofiel langs normatieve standaarden te leggen. Deze standaarden worden hieronder toegelicht.

Analysis. De analyse is primair uitgevoerd door de vastgestelde feiten langs de gevonden normen te leggen. Waar de vastgestelde feiten niet aan de normen voldoen, wordt dat genoteerd en waar mogelijk verklaard.

Conclusion. Uit de analyse volgen bevindingen. Deze worden in de conclusie opgesomd in combinatie met aanbevelingen voor vervolgstappen.

Verantwoording werkwijze

Het CUB-proces wordt op het moment van publiceren van dit rapport niet langer gebruikt. Het doel van dit rapport is DUO te assisteren bij haar heroverweging van het gebruik van risicoprofilering; in het bijzonder bij het tegengaan van onrechtmatig gebruik van de uitwonendenbeurs. Daarvoor is het van belang om vast te stellen hoe het CUB-proces tot stand is gekomen en gebruikt en die bevindingen te analyseren (zie ook [2. Achtergrond Controle Uitwonendenbeurs: feitenvaststelling](#)).

De kwalitatieve analyse wordt gedaan aan de hand van de normen die op het moment van publiceren van dit rapport gelden. Dat zijn immers de normen die relevant zijn voor nieuw beleid; voor de in de toekomst gebruikte risicoprofielen of juist de afschaffing van dit gebruik.

De keerzijde van deze werkwijze is dat er uit de in dit rapport gemaakte kwalitatieve analyse geen normatieve conclusies kunnen worden getrokken over het gebruik van het CUB-proces in het verleden. Daarvoor zou het CUB-proces gelegd moeten worden langs de normen die golden op het moment dat het CUB-proces werd gebruikt. Dat is de relevante toets voor de vraag of DUO zich in het verleden ergens schuldig aan heeft gemaakt. Het beantwoorden van de schuldvraag is uitdrukkelijk niet het doel van dit rapport en zo moet het rapport niet worden gelezen of gebruikt. Het mogelijke Fase-2 onderzoek is wel geschikt voor een normatief oordeel. Met DUO is de optie overeengekomen om na publicatie verder te gaan in Fase-2 waarin op basis van de bevindingen van dit rapport een normatief oordeel over het CUB-proces (en mogelijke toekomstige processen) kan worden geveld.

Normen en richtlijnen

Op de inzet van risicoprofilering door uitvoeringsorganisaties zijn verschillende normen en richtlijnen van toepassing. Hieronder volgt een overzicht.

De normen en richtlijnen zijn ingedeeld in juridische normen waar DUO zich aan moet houden bij het opsporen van onrechtmatig gebruik, *soft law* en interne normen.

Juridische normen

In het kader van dit rapport zijn drie categorieën juridische normen relevant: de algemene normen voor gelijke behandeling, de algemene beginselen van behoorlijk bestuur zoals opgenomen in de Algemene wet bestuursrecht (Awb) en normen uit de Algemene Verordening Gegevensbescherming (AVG). Hieronder worden deze drie categorieën kort besproken.

Ten eerste gelden de algemene normen voor gelijke behandeling uit de Nederlandse wetgeving en internationale instrumenten.¹⁷ In artikel 1 van de Grondwet staat het discriminatieverbod. De Algemene Wet Gelijke Behandeling (AWGB) geeft daar invulling aan. Zo bepaalt artikel 7a lid 1 AWGB dat bij sociale bescherming (waaronder de toekenning van de uitwonendenbeurs) onderscheid op grond van ras verboden is. Internationaal gelden onder meer artikel 14 van het Europees Verdrag tot bescherming van de Rechten van de Mens en de fundamentele vrijheden (EVRM) en artikel 1 van Protocol 12 bij het EVRM.

¹⁷ Allereerst zijn er de discriminatieverboden uit artikel 14 EVRM en art. 1 twaalfde protocol EVRM. Daarnaast gelden er vanuit het Unierecht enkele gelijkebehandelingsrichtlijnen, die in Nederlandse wetgeving geïmplementeerd zijn, te weten de Anti-rassendiscriminatie richtlijn (2000/43/EG), de Kaderrichtlijn gelijke behandeling in arbeid en beroep (2000/78/EG), de Richtlijn gelijke behandeling van mannen en vrouwen bij de toegang tot en het aanbod van goederen en diensten (2004/113/EG), en de herschikte Richtlijn gelijke behandeling van mannen en vrouwen in arbeid en beroep (2006/54/EG).

Box 3 Normen, juridisch kader en algoritmedefinitie

Normen

In dit rapport wordt naar onderstaande regels, richtlijnen, handreikingen, etc. verwezen als 'normen'. Voor de duidelijkheid: dit gaat niet uitsluitend over afdwingbare normen. Het gaat om normen in de zin van gedragsregels die door de leden van een gemeenschap als min of meer vanzelfsprekend worden beschouwd en waar gedrag aan kan worden getoetst. Er wordt bij de vaststelling van de normen aangenomen dat de wet, door de overheid en de EU opgestelde toetsingskaders en handreikingen vanuit de wetenschap en het College voor de rechten van de mens de in de gemeenschap als vanzelfsprekend beschouwde normen bevatten.

Juridisch kader

Een student moet uitwonend zijn om in aanmerking te komen voor de uitwonendenbeurs (art. 1.5 Wet studiefinanciering 2000 (Wsf)). Het toezicht op deze verplichting is opgedragen aan de afdeling Handhaving en Inspectie van DUO (art. 9.1a Wsf en Besluit aanwijzing personen toezicht naleving ex art. 1.5 Wsf). Een student die niet uitwonend is, maar wel een uitwonendenbeurs heeft aangevraagd kan een boete worden opgelegd. Te veel ontvangen studiefinanciering kan worden teruggevorderd (art. 9.9 Wsf).

Algoritmedefinitie

In dit onderzoek wordt het risicoprofiel aangemerkt als algoritme. Het risicoprofiel voldoet aan de definitie van een algoritme dat onder meer de Algemene Rekenkamer en het Algoritmeregister hanteren: "Een set van regels en instructies die een computer geautomatiseerd volgt bij het maken van berekeningen om een probleem op te lossen of een vraag te beantwoorden". Zie bijvoorbeeld: <https://algoritmes.overheid.nl/nl/footer/over-algoritmes>.

De Centrale Raad van Beroep (CRvB) heeft zich uitgelaten over risicoprofilering in het kader van de bovengenoemde instrumenten.¹⁸ De CRvB staat risicoprofilering onder voorwaarden toe voor het vergroten van effectiviteit, efficiëntie en kostenbesparing van overheidshandelen en vanwege het belang van bestrijding van onjuist gebruik van sociale voorzieningen. Deze rechtspraak van de CRvB ging over het buitenlandse vermogen van bijstandsgerechtigden. De CRvB overwoog:

Zoals de Raad in die uitspraken heeft overwogen, kunnen ervaringen met groepen bijstandsgerechtigden en criteria die objectief een risicoprofiel voor niet gemeld vermogen in het buitenland kunnen vormen, rechtvaardigen dat de algemene onderzoeksbevoegdheid wel wordt ingezet ten aanzien van bijstandsgerechtigden met een niet-Nederlandse herkomst, een bepaalde leeftijd en een bepaald vakantiegedrag, en niet ten aanzien van de overige bijstandsgerechtigden. Indien een dergelijk risicoprofiel, zoals thans in geding, echter niet gericht is op alle bijstandsgerechtigden met niet-Nederlandse herkomst, maar slechts gericht is op bijstandsgerechtigden, afkomstig uit een bepaald geboorteland, dan kan het gaan om een onderscheid, dat volgens de Raad in zijn onder 4.5 genoemde rechtspraak als “verdacht” wordt aangemerkt. Een zodanig onderscheid moet worden gerechtvaardigd door zeer gewichtige redenen.

Ten tweede zijn op de inzet van risicoprofilering de algemene beginselen van behoorlijk bestuur uit de Awb van toepassing.¹⁹ Voor risicoprofilering zijn in het bijzonder het gelijkheidsbeginsel en het zorgvuldigheidsbeginsel van belang.

- > **Gelijkheidsbeginsel** – Gelijke gevallen moeten gelijk worden behandeld. Ongelijke gevallen moeten ongelijk worden behandeld naar de mate waarin zij verschillen. Het gelijkheidsbeginsel kan bij risicoprofilering onder andere in het gedrang komen wanneer de profilering leidt tot (indirecte) discriminatie.
- > **Zorgvuldigheidsbeginsel** – Dit beginsel strekt ertoe de omstandigheden te scheppen waarin een bestuursorgaan een juist besluit kan nemen. Een bestuursorgaan moet zich op de hoogte stellen van de relevante feiten en af te wegen belangen (art. 3:2 Awb). Er moet een geschikte methode worden gebruikt om de belangen af te wegen en er dient een volledige belangenafweging plaats te vinden.²⁰ Het zorgvuldigheidsbeginsel

¹⁸ ECLI:NL:CRVB:2015:3249, ECLI:NL:CRVB:2018:1541, en ECLI:NL:CRVB:2020:1664 zoals toegepast in lagere rechtspraak, zie onder meer ECLI:NL:RBLIM:2013:11417, ECLI:NL:RBOBR:2013:BZ6037, ECLI:NL:RBROT:2014:5684, ECLI:NL:RBROT:2013:2359, ECLI:NL:RBAMS:2012:BV6364, ECLI:NL:RBAMS:2018:9659 en ECLI:NL:RBLIM:2023:1325.

¹⁹ Deze beginselen zijn het verbod van vooringenomenheid (artikel 2:4 Awb), het zorgvuldigheidsbeginsel (artikel 3:2 Awb), het verbod van willekeur (artikel 3:4 lid 1 Awb), het evenredigheidsbeginsel (artikel 3:4 lid 2 Awb), het verbod van *détournement de pouvoir* (artikel 3:3 Awb) en het motiveringsbeginsel (artikelen 3:46 en 3:47 Awb). Toepasselijke normen uit het ongeschreven bestuursrecht zijn het vertrouwensbeginsel, het gelijkheidsbeginsel, het rechtszekerheidsbeginsel. Zie bijv. Barkhuysen e.a., 'Bestuursrecht in het Awb-tijdperk', p. 110 e.v.

²⁰ Chris Adriaansz, De rechtmatigheid van algoritmische besluitvorming in het licht van het zorgvuldigheidsbe-

kan bij risicoprofilering onder andere in het gedrang komen als de gebruikte gegevens incompleet of incorrect zijn en wanneer het risicoprofiel niet alle relevante feiten of belangen meeneemt.

In 2018 heeft de Afdeling advisering van de Raad van State geadviseerd om bij geautomatiseerde besluitvorming door de overheid de beginselen van behoorlijk bestuur, en in het bijzonder het motiveringsbeginsel en het zorgvuldigheidsbeginsel, verscherpt te interpreteren.²¹ Dit kan volgens de Afdeling onder andere betekenen dat aan burgers wordt uitgelegd welke beslisregels zijn gebruikt. De Afdeling noemt bij het gebruik van profilering ook uitdrukkelijk dat de kans bestaat dat de burger geconfronteerd wordt met een "*statistische werkelijkheid die afwijkt van de concrete feiten*".

Voor de uitwisseling van gegevens tussen overheidsinstanties bij het tegengaan van onrechtmatig gebruik is verder de SyRI-uitspraak van de Rechtbank Den Haag relevant.²²

Ten derde wordt voor de volledigheid genoteerd dat ook de AVG voor het risicoprofiel relevante normen bevat. De betrokkene wiens gegevens worden verwerkt heeft het recht om geïnformeerd te worden over het bestaan van (semi-)geautomatiseerde besluitvorming, zoals risicoprofilering, en "*nuttige informatie over de onderliggende logica*", en het belang en de verwachte gevolgen van deze verwerking voor de betrokkene (artikel 15 lid 1(h) AVG). Toetsing van het CUB-profiel aan de AVG valt echter buiten de reikwijdte van dit rapport. Artikel 22 van de AVG bevat een verbod op volledig geautomatiseerde besluitvorming. Hiervan is geen sprake bij de CUB.

Instrumenten van soft law

Er zijn kaders voor de verantwoorde inzet van algoritmes door publieke sector organisaties. Deze kaders bieden handvatten bij vraagstukken over gelijke behandeling bij de inzet van algoritmes. In Nederland gezaghebbende kaders zijn:

- > Het Rijksbrede Implementatiekader *Verantwoorde inzet van algoritmes*²³ (2023);
- > Het *Onderzoekskader algoritmes*²⁴ (2023) van de ADR ('Onderzoekskader ADR');
- > Het *Toetsingskader algoritmes* (2021) van de Algemene Rekenkamer;
- > De *Impact Assessment Mensenrechten en Algoritmes*²⁵ (IAMA) van het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (2021);
- > De *Handreiking non-discriminatie by design*²⁶ (2021) van het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties.

ginsel en het motiveringsbeginsel, NTB 2020/100.

²¹ Kamerstukken II 2017/2018, 26 643, 557.

²² Rb Den Haag 20 februari 2020, ECLI:NL:RBDHA:2020:865.

²³ <https://www.rijksoverheid.nl/documenten/rapporten/2023/06/30/implementatiekader-verantwoorde-inzet-van-algoritmen>

²⁴ <https://www.rijksoverheid.nl/documenten/rapporten/2023/07/11/onderzoekskader-algoritmes-adr-2023>

²⁵ <https://www.rijksoverheid.nl/documenten/rapporten/2021/02/25/impact-assessment-mensenrechten-en-algoritmes>

²⁶ <https://www.rijksoverheid.nl/documenten/rapporten/2021/06/10/handreiking-non-discriminatie-by-design>

Speciale aandacht verdient het Implementatiekader 'Verantwoorde inzet van algoritmen'. Bij de analyse wordt dit document niet gebruikt, omdat er nog geen definitieve versie van beschikbaar is. Zodra het Implementatiekader in definitieve vorm is gepubliceerd, zal het van nut zijn bij de inzet van algoritmes door uitvoeringsorganisaties. Voor nu volstaat het volgende citaat dat ook van toepassing is op de andere *soft law* normen:

"Het implementatiekader is geen 'checklist'. Het nemen van de maatregelen in dit kader leidt veelal niet rechtstreeks tot het voldoen aan een verplichting en niet alle maatregelen in dit kader zijn verplicht. Gebruikers van dit implementatiekader zullen dus tot een eigen afweging komen. Waar nodig zullen zij ook aanvullend advies moeten vragen, bijvoorbeeld bij waardenspanningen en ethische dilemma's".

Verder staan relevante uitgangspunten in buiten de Nederlandse overheid opgestelde kaders. Relevant zijn onder meer:

- > Het *Toetsingskader Discriminatie door Risicoprofielen* van het College voor de Rechten van de Mens (2021);
- > Het rapport *De mens in de Machine* van WAAG (februari 2020);
- > M.J. Vetz, J.H. Gerards en R. Nehmelman, *Algoritmes en Grondrechten* (2018);
- > De Europese Commissie (2019), *High-level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI*;
- > Europese Commissie (2020), *White Paper on Artificial Intelligence: a European approach to excellence and trust*.

Op de twee laatstgenoemde stukken hebben de Algemene Rekenkamer en de ADR hun aanbevelingen grotendeels gebaseerd.

Interne normen

Er is onderzocht of DUO interne normen heeft voor het gebruik van risicoprofielen, voor het onderzoeken van mogelijke vooringenomenheid en/of voor de nadere invulling van de bestuursrechtelijke normen. Er is niet gebleken dat die normen er zijn.

Het strekt dan ook tot aanbeveling concretere normen op te stellen voor het gebruik van risicoprofilering binnen DUO. Ook het Onderzoekskader ADR schrijft dit voor: *"het algoritme voldoet aan de intern vastgestelde Beleidskaders"* (SV.8) en *"de rollen en verantwoordelijkheden bij de ontwikkeling en inzet van het algoritme zijn belegd"* (SV.9).

Er zijn methodes voor het opstellen een dergelijk normatief kader. Deze methodes vallen buiten de kaders van dit onderzoek, maar hier kan zo nodig nader over worden geadviseerd.

Selectie van normen waar het risicoprofiel in dit onderzoek aan wordt getoetst

Niet alle hierboven genoemde normen zijn geschikt voor beantwoording van de onderzoeksvraag. De juridische normen geven een relevante buitengrens, maar zijn onvoldoende

de specifiek om relevant te zijn voor een analyse van het risicoprofiel. Aangezien interne normen ontbreken, wordt ook hier geen houvast gevonden.

Algorithm Audit acht het IAMA en het Onderzoekskader Algoritmes het meest geschikt als normen voor dit onderzoek. Deze kaders zijn specifiek opgesteld voor verantwoord gebruik van algoritmes door overheidsinstanties. Deze kaders zullen worden gebruikt voor de kwalitatieve analyse. Specifiek zal worden gerefereerd naar Hoofdstuk 1 Sturing & Verantwoording (SV) en Hoofdstuk 3 Data & Model (DM) uit het Onderzoekskader Algoritmes.

3.3 Samenstelling onderzoeksteam

Het onderzoek is uitgevoerd door onderzoekers verbonden aan Algorithm Audit. Er is getracht rekening te houden met representativiteit. Onder meer door onderzoekers met een verschillende gender, culturele achtergrond en professionele expertises te betrekken. Het onderzoeksteam was divers en multidisciplinair.

4. Resultaten kwantitatieve analyse

Hieronder volgen de resultaten van de aselecte steekproeven uit 2014 en 2017 (§4.1) en de resultaten van de biasmetingen van 2014 en 2019 (§4.2). Aan de hand van deze resultaten worden onderzoeksvragen 1-6 beantwoord.

4.1 Resultaten aselecte steekproef 2014 en 2017

In 2014 en 2017 zijn respectievelijk 387 en 293 studenten met een uitwonendenbeurs willekeurig geselecteerd [A.15,A.40]. Van deze studenten is onderzocht of zij rechtmatig gebruik maakten van de uitwonendenbeurs. DUO heeft in samenspraak met de ADR de omvang van de deelpopulaties samengesteld om onrechtmatig gebruik van de uitwonendenbeurs met 95% zekerheid in de gehele studentenpopulatie vast te stellen. De omvang van de gehele studentenpopulatie, bestaande uit mbo-, hbo- en wo-studenten, op het moment dat de aselecte steekproef plaatsvond in maart 2014 bedroeg 268.030 [A.40]. De omvang van de gehele studentenpopulatie, bestaande enkel uit mbo-studenten, op het moment dat de aselecte steekproef plaatsvond in oktober 2017 bedroeg 45.025 [A.15].

Het bestand van de aselecte steekproef waar de onderzoekers toegang toe hebben gekregen bestaat uit respectievelijk 387 en 293 studenten voor de jaren 2014 en 2017²⁷ [G.1,G.2]. De resultaten van deze steekproeven zijn gebruikt om onderzoeksvragen 1-3 te beantwoorden. Voor al deze studenten zijn de relevante kenmerken voor het risicomodel (onderwijsvorm, leeftijd en afstand tot ouder(s)) gecodeerd.

Resultaten onderzoeksvraag 1

Onderzoeksvraag 1:

Zijn er op basis van de aselecte steekproef 2014 en 2017 aanwijzingen voor een verband tussen de kenmerken van het risicoprofiel en onrechtmatig gebruik van de uitwonendenbeurs?

Antwoord onderzoeksvraag1:

Uit de aselecte steekproeven uit 2014 en 2017 blijkt geen verband tussen de meeste kenmerken die gebruikt worden voor het risicoprofiel en onrechtmatig gebruik van de uitwonendenbeurs. Er is duidelijke aanwijzing dat in de aselecte steekproef uit 2014 studenten die 1m-1km en 2-5km tot hun ouder(s) vaker en studenten die 50-500 km tot hun ouder(s) minder vaak onrechtmatig gebruik maken van de uitwonendenbeurs. Daarnaast volgt uit de aselecte steekproef uit 2014 dat mbo 3-4 studenten statistisch significant vaker onrechtmatig gebruik hebben gemaakt van de uitwonendenbeurs dan wo-studenten. Uit de aselecte steekproef uit 2017 blijkt dat 15-18-jarige studenten statistisch significant vaker misbruik maken dat 21-22-jarigen. Deze resultaten ondersteunen echter niet de granulaire verschillen die zijn gehanteerd in het risicoprofiel.

²⁷ Het verschil met de genoemde aantallen uit [A.15] en [A.40] is niet verklaard kunnen worden.

Toelichting onderzoeksvraag 1:

Onderzoeksvraag 1 wordt per kenmerk onderwijsvorm, leeftijd en afstand tot ouder(s) beantwoord.

Zoals toegelicht in [3.1 Kwantitatieve analyse](#) worden twee aselecte steekproeven gebruikt: de steekproeven uit 2014 (n=387) en uit 2017 (n=293), waarvan de laatste uitsluitend studenten bevat die aan het mbo studeren. Dit omdat ten tijde van de steekproef in 2017 de uitwonendenbeurs exclusief werd toegekend aan mbo-studenten.

Onderwijsvorm

Tabellen 4-5 geven het onrechtmatig gebruik-percentages weer per onderwijsvorm voor studenten die zijn getrokken voor de aselecte steekproef uit 2014 en 2017.

Steekproef 2014	Grootte van de groep	Aantal onrechtmatig	Percentage
mbo 1-2	15	1	6,7%
mbo 3-4	53	4	7,5%
hbo	150	5	3,3%
wo	169	4	2,4%
Totaal	387	14	3,6%

Tabel 4 – Overzicht van groeps-grootte en onrechtmatig gebruik-percentages per onderwijsvorm in de aselecte steekproef 2014 (n=387).

Steekproef 2017	Grootte van de groep	Aantal onrechtmatig	Percentage
mbo 1-2	53	4	7,5%
mbo 3-4	240	7	2,9%
Totaal	293	11	3,8%

Tabel 5 – Overzicht van groeps-grootte en onrechtmatig gebruik-percentages per onderwijsvorm in de aselecte steekproef 2017 (n=293).

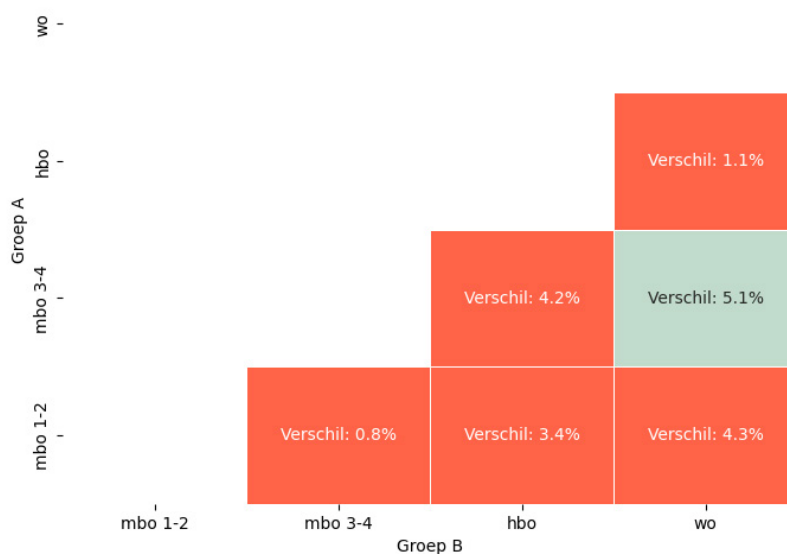
De resultaten van de eenzijdige statistische Z-toets voor de verschillende onrechtmatig gebruik-percentages per onderwijsvorm zijn weergegeven in [Figuur 2](#). Een beschrijving van uitgevoerde statistische toets kan worden gevonden in [3.1 Kwantitatieve analyse](#).

Uit [Figuur 2](#) volgt dat alleen het verschil tussen onrechtmatig gebruik-percentages van mbo 3-4- en wo-studenten statistisch significant is. Dit acht Algorithm Audit een onvoldoende consistent statistisch signaal om risicoprofilering op te baseren. Dit betekent dat er naar aanleiding van de steekproeven 2014 en 2017 geen statistische onderbouwing is om onderscheid te maken op basis van de onderwijsvormen mbo, hbo en wo. Uit de steekproef van 2014 is het patroon waarneembaar dat het % onrechtmatig gebruik afneemt naarmate het opleidingsniveau 'stijgt'. Echter, deze verschillen zijn niet statistisch significant. Merk op dat de relatief beperkte omvang van deelpopulaties in de steekproef van 2014, bijvoorbeeld 15 geselecteerde mbo 1-2 studenten, de resultaten kunnen

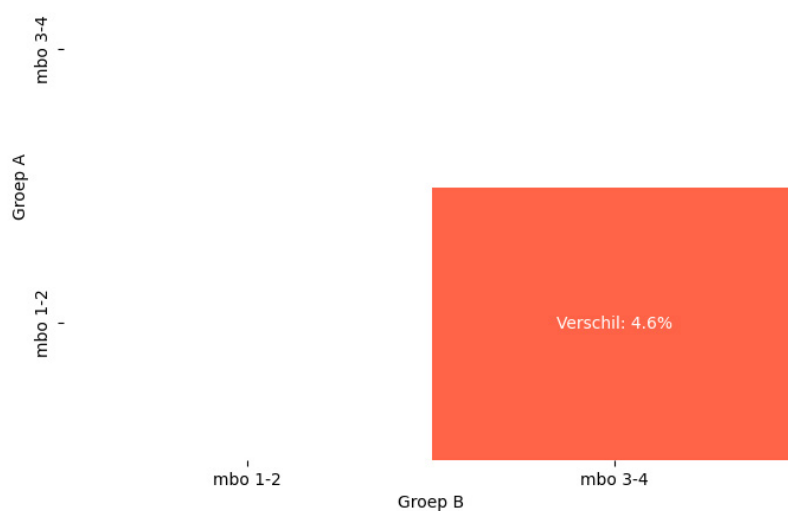
vertroebelen. Een grotere steekproef is nodig om te bevestigen of er aanwijzingen zijn voor verschillen in onrechtmatig gebruik percentage per onderwijsvorm.

Fishers exacte toets – een methode geschikt voor kleine steekproefgrootte – bevestigt bovenstaande uitkomsten. Aanvullende informatie over Fishers exacte toets wordt gegeven in [Appendix C – Aanvullende informatie statistische analyse](#).

Aselecte steekproef 2014: Statistisch significant onderscheid in groen (n=387)
(resultaten gebaseerd op eenzijdige Z-toets)



Aselecte steekproef 2017: Statistisch significant onderscheid in groen (n=293)
(resultaten gebaseerd op eenzijdige Z-toets)



Figuur 2 – Uitkomsten van eenzijdige Z-toetsen, op basis van een betrouwbaarheidsniveau van 5%, voor verschil in onrechtmatig gebruik-percentage per onderwijsvorm (Groep A vs Groep B) gebaseerd op de aselecte steekproef uit 2014 (n=397) en de aselecte steekproef uit 2017 (n=293). Statistisch significant onderscheid tussen twee onrechtmatig gebruik-percentages per onderwijscategorie in groen. Statistisch insignificant onderscheid tussen twee onrechtmatig gebruik-percentages per onderwijscategorie in rood. Een positief verschil betekent dat Groep A een hoger onrechtmatig gebruik-percentage heeft dan Groep B. Een negatief verschil betekent dat Groep B een hoger onrechtmatig gebruik-percentage heeft dan Groep A.

Leeftijd

Tabellen 6-7 geven het onrechtmatig gebruik-percentage weer per leeftijdscategorie voor studenten die zijn getrokken voor de aselecte steekproef uit 2014 en 2017.

Steekproef 2014	Grootte van de groep	Aantal onrechtmatig	Percentage
15-18	24	0	0%
19-20	115	1	0,9%
21-22	149	8	5,4%
23-24	62	3	4,8%
25-50	37	2	5,4%
Totaal	387	14	3,6%

Tabel 6 – Overzicht van groepsgrootte en onrechtmatig gebruik-percentage per leeftijdscategorie in de aselecte steekproef 2014 (n=387).

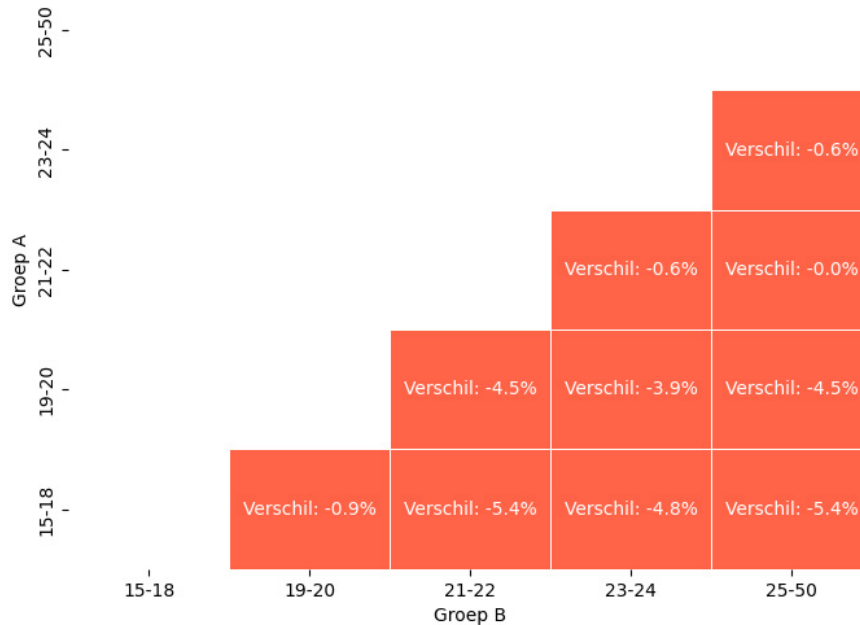
Steekproef 2017	Grootte van de groep	Aantal onrechtmatig	Percentage
15-18	24	2	8,3%
19-20	105	5	4,8%
21-22	76	1	1,3%
23-24	43	2	4,7%
25-50	45	1	2,2%
Totaal	293	11	3,8%

Tabel 7 – Overzicht van groepsgrootte en onrechtmatig gebruik-percentage per leeftijdscategorie in de aselecte steekproef 2017 (n=293).

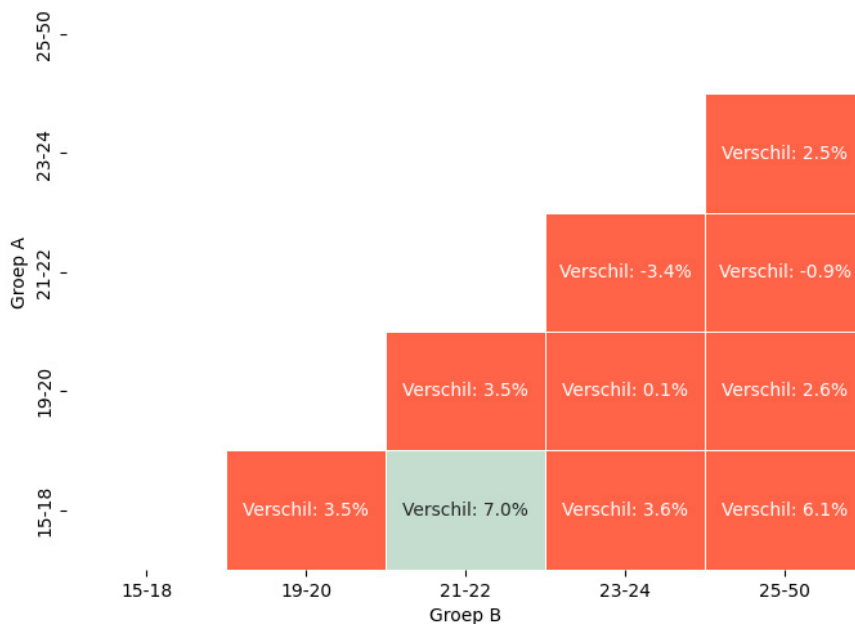
De resultaten van de eenzijdige statistische Z-toets voor de verschillende onrechtmatig gebruik-percentages per leeftijdscategorie zijn weergegeven in [Figuur 3](#). Een beschrijving van uitgevoerde statistische toetsen kan worden gevonden in [3.1 Kwantitatieve analyse](#).

Uit [Figuur 3](#) volgt dat alleen het verschil tussen onrechtmatig gebruik-percentages van 15-18-jarigen en 21-22-jarigen in 2017 statistisch significant is. Dit acht Algorithm Audit een onvoldoende consistent statistisch signaal om risicoprofilering op te baseren. De percentages uit [Tabel 6](#) wijzen zelfs op een patroon dat tegen de logica van het CUB-risicoprofiel ingaat. Voor oudere studenten wordt een hoger onrechtmatig gebruik-percentage waargenomen. Terwijl in het CUB-profiel wordt aangenomen dat jonge studenten een hoger risico hebben om onrechtmatig gebruik te maken van de uitwonendenbeurs. De logica van het risicoprofiel is wel terug te zien in de resultaten van de aselecte steekproef uit 2017 ([Tabel 8](#)). Merk op dat de steekproef van 2017 alleen mbo-studenten betreft, en het dus niet duidelijk is of deze bevinding generaliseert naar hbo- en wo-studenten.

Aselecte steekproef 2014: Statistisch significant onderscheid in groen (n=387)
(resultaten gebaseerd op eenzijdige Z-toets)



Aselecte steekproef 2017: Statistisch significant onderscheid in groen (n=293)
(resultaten gebaseerd op eenzijdige Z-toets)



Figuur 3 – Uitkomsten van eenzijdige Z-toetsen, op basis van een betrouwbaarheidsniveau van 5%, voor verschil in onrechtmatig gebruik-percentages per leeftijdscategorie (Groep A vs Groep B) gebaseerd op de aselecte steekproef uit 2014 (n=397) en de aselecte steekproef uit 2017 (n=293). Statistisch significant onderscheid tussen twee onrechtmatig gebruik-percentages per leeftijdscategorie in groen. Statistisch insignificant onderscheid tussen twee onrechtmatig gebruik-percentages per onderwijscategorie in rood. Een positief verschil betekent dat Groep A een hoger onrechtmatig gebruik-percentages heeft dan Groep B. Een negatief verschil betekent dat Groep B een hoger onrechtmatig gebruik-percentages heeft dan Groep A.

Tabellen 8-9 geven het onrechtmatig gebruik-percentage weer per afstandscategorie voor studenten die zijn getrokken voor de aselecte steekproef uit 2014 en 2017.

Steekproef 2014	Grootte van de groep	Aantal onrechtmatig	Percentage
0km	8	0	0%
1m-1km	21	5	23,8%
1-2km	11	0	0%
2-5km	31	5	16,1%
5-10km	24	1	4,2%
10-20km	31	1	3,2%
20-50km	58	1	1,7%
50-500km	137	0	0%
onbekend	66	1	1,5%
Totaal	387	14	3,6%

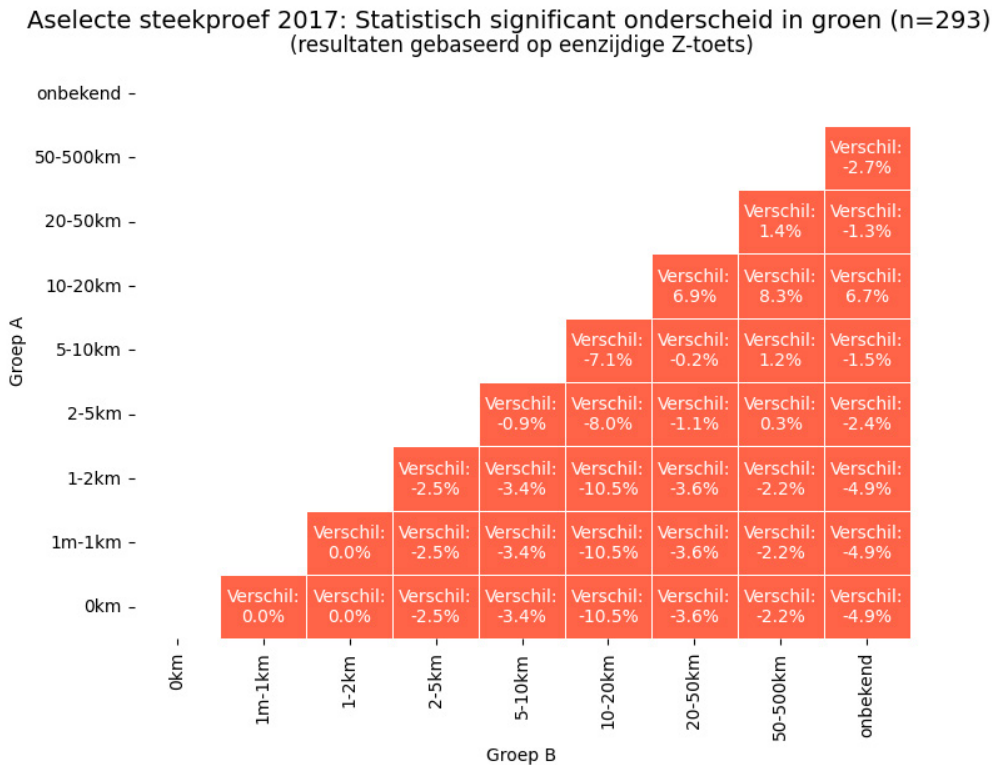
Tabel 8 – Overzicht van groepsgrootte en onrechtmatig gebruik-percentage per afstandscategorie in de aselecte steekproef 2014 (n=387).

Steekproef 2017	Grootte van de groep	Aantal onrechtmatig	Percentage
0km	4	0	0%
1m-1km	24	0	0%
1-2km	24	0	0%
2-5km	40	1	2,5%
5-10km	29	1	3,4%
10-20km	38	4	10,5%
20-50km	28	1	3,6%
50-500km	45	1	2,2%
onbekend	61	3	4,9%
Totaal	293	11	3,8%

Tabel 9 – Overzicht van groepsgrootte en onrechtmatig gebruik-percentage per afstandscategorie in de aselecte steekproef 2017 (n=293).

De resultaten van de eenzijdige statistische Z-toets voor de verschillende onrechtmatig gebruik-percentages per afstandscategorie zijn weergegeven in [Figuur 4](#). Een beschrijving van uitgevoerde statistische toets kan worden gevonden in [3.1 Kwantitatieve analyse](#).

Uit [Figuur 4](#) volgt dat er veel statistisch significante verschillen zijn tussen de gehanteerde afstandscategorieën gebaseerd op de steekproef uit 2014. Voor onderscheid op basis van 1m-1km, 2-5km en 50-500km bestaat voldoende statistische onderbouwing. De duidelijke verschillen die volgen uit de steekproef uit 2014 zijn echter niet terug te zien in de steekproef uit 2017. Voor de aselecte steekproef uit 2017 zijn geen van de verschillen in



Figuur 4 – Uitkomsten van eenzijdige Z-toetsen, op basis van een betrouwbaarheidsniveau van 5%, voor verschil in onrechtmatig gebruik-percentage per afstandscategorie (Groep A vs Groep B) gebaseerd op de aselechte steekproef 2014 (n=397) en de aselechte steekproef 2017 (n=293). Statistisch significant onderscheid tussen twee onrechtmatig gebruik-percentages per afstandscategorie in groen. Statistisch insignificant onderscheid tussen twee onrechtmatig gebruik-percentages per onderwijscategorie in rood. Een positief verschil betekent dat Groep A een hoger onrechtmatig gebruik-percentage heeft dan Groep B. Een negatief verschil betekent dat Groep B een hoger onrechtmatig gebruik-percentage heeft dan Groep A.

Resultaten onderzoeksvraag 2

Onderzoeksvraag 2:

Zijn er op basis van de aselechte steekproef 2014 en 2017 aanwijzingen voor een verband tussen de gehanteerde risicocategorieën 1-6 en onrechtmatig gebruik met de uitwonendenbeurs?

Antwoord op onderzoeksvraag 2:

Er is onvoldoende onderbouwing om de toegekende risicoscores op te delen in zes risicocategorieën. Voor een binaire risicoclassificatie is wel statistische onderbouwing, mits toegepast op uitwonende hbo- en wo-studenten. Een simplificering van de gehanteerde risicoklassen geniet de voorkeur.

Toelichting onderzoeksvraag 2:

Het risico op onrechtmatig gebruik van de uitwonendenbeurs is ingedeeld in verschillende 'risicogroepen', variërend van 'Zeer hoog (1)' tot en met 'Zeer laag (5)', of 'Onbekend (6)'. In 2014 is een steekproef gedaan om een 'uitspraak te doen over het misbruik in de lagere risico-categorieën 3 t/m 6 als geheel' [A.40]. Op basis van deze analyse werd destijds geconcludeerd: 'Op grond van de resultaten in de categorieën 3 t/m 6 lijkt het erop dat DUO terecht minder aandacht besteed aan deze groepen'.

Gebaseerd op de aselechte steekproef van 2014 wordt deze conclusie deels ondersteund.

Zoals in Tabel 10 weergeven is het onrechtmatig gebruik-percentages hoger voor de groep 'Zeer hoog (1) en 'Hoog (2)' dan voor de groep 'Midden (3)', 'Laag (4)' en 'Zeer laag (5)' eventueel aangevuld met 'Onbekend (6)'. De eenzijdige Z-toets, waarvan de toepassing is toegelicht in 3.1 Kwantitatieve analyse, acht het verschil tussen de groep 1+2 en de groep 3+4+5 plus eventueel 6 als statistisch significant (zie de groene percentages in Tabel 10). Het verschil in onrechtmatig gebruik-percentages tussen de individuele groepen 1-6 is echter niet statistisch significant (zie rode percentages in Tabel 10). Samenvattend, de aselechte steekproeven uit 2014 en 2017 bieden een statistische onderbouwing voor een binaire hoge- en lage risicoclassificering. Er zijn echter geen duidelijke redenen om voor de gehanteerde risicocategorieën 1-6, zoals gehanteerd in het CUB-proces, te kiezen.

Groep (aselechte steekproef 2014)	% onrechtmatig gebruik	Aantal onrechtmatig gebruik	Totaal aantal
1 - Zeer hoog	10,0%	2	20
2 - Hoog	17,2%	5	29
3 - Midden	13,2%	5	38
4 - Laag	0,0%	0	106
5 - Zeer laag	0,8%	1	125
6 - Onbekend	1,4%	1	69
Zeer hoog (1) + Hoog (2)	14,3%	7	49
Midden (3) + Laag (4) + Zeer laag (5)	2,2%	6	269
Midden (3) + Laag (4) + Zeer laag (5) + Onbekend (6)	2,1%	7	338
Vershil Groep 1 + 2 & Groep 3 + 4 + 5	12,1%	-	-
Vershil Groep 1 + 2 & Groep 3 + 4 + 5 + 6	12,2%	-	-
Vershil Zeer hoog (1) & hoog (2)	-7,2%	-	-
Vershil Hoog (2) & Midden (3)	4,1%	-	-
Vershil Midden (3) & Laag (4)	13,2%	-	-
Vershil Laag (4) & Zeer laag (5)	-0,8%	-	-

Tabel 10 – Analyse van gehanteerde risicocategorieën in het CUB-proces naar aanleiding van uitkomsten van de aselechte steekproef uit 2014 (n=387). Alle verschillen in het groen zijn statistisch significant volgens een eenzijdige Z-toets, op basis van een betrouwbaarheidsniveau van 5%. Alle verschillen in het rood zijn statistisch insignificant op basis van een betrouwbaarheidsniveau van 5%.

De aselechte steekproef uit 2017 geeft een ander beeld over de gehanteerde risicocategorieën dan de hierboven gegeven analyse op basis van de aselechte steekproef uit 2014. Zoals in Tabel 11 weergeven zijn er geen statistisch significante verschillen tussen de groepen. Sterker nog, het percentage van onrechtmatig gebruik gaat tegen de logica in van het CUB-profiel – enkele keren is het onrechtmatig gebruik-percentages hoger voor lagere risicocategorieën. Het lijkt dat het risicoprofiel niet geschikt is voor de onderzochte populatie naar aanleiding van de aselechte steekproef uit 2017, die uitsluitend uit mbo-stu-

denten bestaat. Een nieuwe, grotere aselechte steekproef moet uitsluitsel geven of, en zo ja hoe, mbo-studenten wel op een statistisch onderbouwde manier onderworpen kunnen worden aan risicoprofilering.

Groep (aselechte steekproef 2017)	% onrechtmatig gebruik	Aantal onrechtmatig gebruik	Totaal aantal
1 - Zeer hoog	3,4%	2	59
2 - Hoog	0,0%	0	52
3 - Midden	6,9%	2	29
4 - Laag	3,3%	2	60
5 - Zeer laag	4,1%	2	49
6 - Onbekend	6,8%	3	44
Zeer hoog (1) + Hoog (2)	1,8%	2	111
Midden (3) + Laag (4) + Zeer laag (5)	4,3%	6	138
Midden (3) + Laag (4) + Zeer laag (5) + Onbekend (6)	4,9%	9	182
Vershil Groep 1 + 2 & Groep 3 + 4 + 5	-2,5%	-	-
Vershil Groep 1 + 2 & Groep 3 + 4 + 5 + 6	-3,1%	-	-
Vershil Zeer hoog (1) & hoog (2)	3,4%	-	-
Vershil Hoog (2) & Midden (3)	-6,9%	-	-
Vershil Midden (3) & Laag (4)	3,6%	-	-
Vershil Laag (4) & Zeer laag (5)	-0,7%	-	-

Tabel 11 – Analyse van gehanteerde risicocategorieën in het CUB-proces naar aanleiding van uitkomsten van de aselechte steekproef uit 2017 (n=293). Alle verschillen in het groen zijn statistisch significant volgens een eenzijdige Z-toets, op basis van een betrouwbaarheidsniveau van 5%. Alle verschillen in het rood zijn statistisch insignificant op basis van een betrouwbaarheidsniveau van 5%.

Resultaten onderzoeksvraag 3

Onderzoeksvraag 3:

Hoeveel onrechtmatig gebruik is er geconstateerd in de aselechte steekproef populatie 2014 en 2017?

Antwoord op onderzoeksvraag 3:

Uit de [Tabellen 4-5](#) kan worden afgeleid dat gebaseerd op de aselechte steekproef uit 2014 er sprake is van een onrechtmatig gebruik-percentages van 3,6%. Voor de aselechte steekproef uit 2017 volgt een onrechtmatig gebruik-percentages van 3,8%.

4.2 Resultaten biasmeting 2014 en 2019

Zoals in 3.1 *Kwantitatieve analyse* is toegelicht, wordt de biasmeting uitgevoerd voor uitwonendenbeurspopulatie-2014 en uitwonendenbeurspopulatie-2019. De eigenschappen van deze populaties worden eerst verkend. Daarna wordt besproken hoe de data van de uitwonendenbeurspopulaties uit de datawarehouse van DUO zijn verkregen. De sectie wordt afgesloten met antwoorden op onderzoeksvragen 4-6.

Dataverkenning startpopulatie-2014 en startpopulatie-2019

De omvang van uitwonendenbeurspopulatie-2014, vastgesteld aan de hand van peildatum 01-02-2014, is 248.649. De omvang van uitwonendenbeurspopulatie-2019, vastgesteld aan de hand van peildatum 01-02-2019, is 50.233. Deze populaties zijn getraceerd in de datawarehouse van DUO [G.3,G.4]. Een voorbeeld van een query om deze populaties te traceren wordt gegeven in *Appendix B – Datawarehouse query*. In deze sectie volgt een verkenning van de eigenschappen van deze populaties.

De verdeling van het aantal studenten met een uitwonendenbeurs per onderwijsvorm, leeftijdscategorie en afstandscategorie in uitwonendenbeurspopulatie-2014 en uitwonendenbeurspopulatie-2019 is weergegeven in *Figuur 5*. Merk op dat in 2014 zowel uitwonende wo-, hbo- en mbo-studenten recht hebben op een uitwonendenbeurs. In 2019 is dat anders door de komst van het leenstelsel, in dit jaar hebben alleen mbo-studenten recht op een uitwonendenbeurs. Hbo- en wo-studenten die nog een uitwonendenbeurs ontvangen in 2019 zijn uitfaserende studenten die voor de introductie van het leenstelsel in 2015 al een uitwonendenbeurs hebben aangevraagd. In 2019 zijn in het CUB-proces vrijwel uitsluitend mbo-studenten geselecteerd voor huisbezoek.

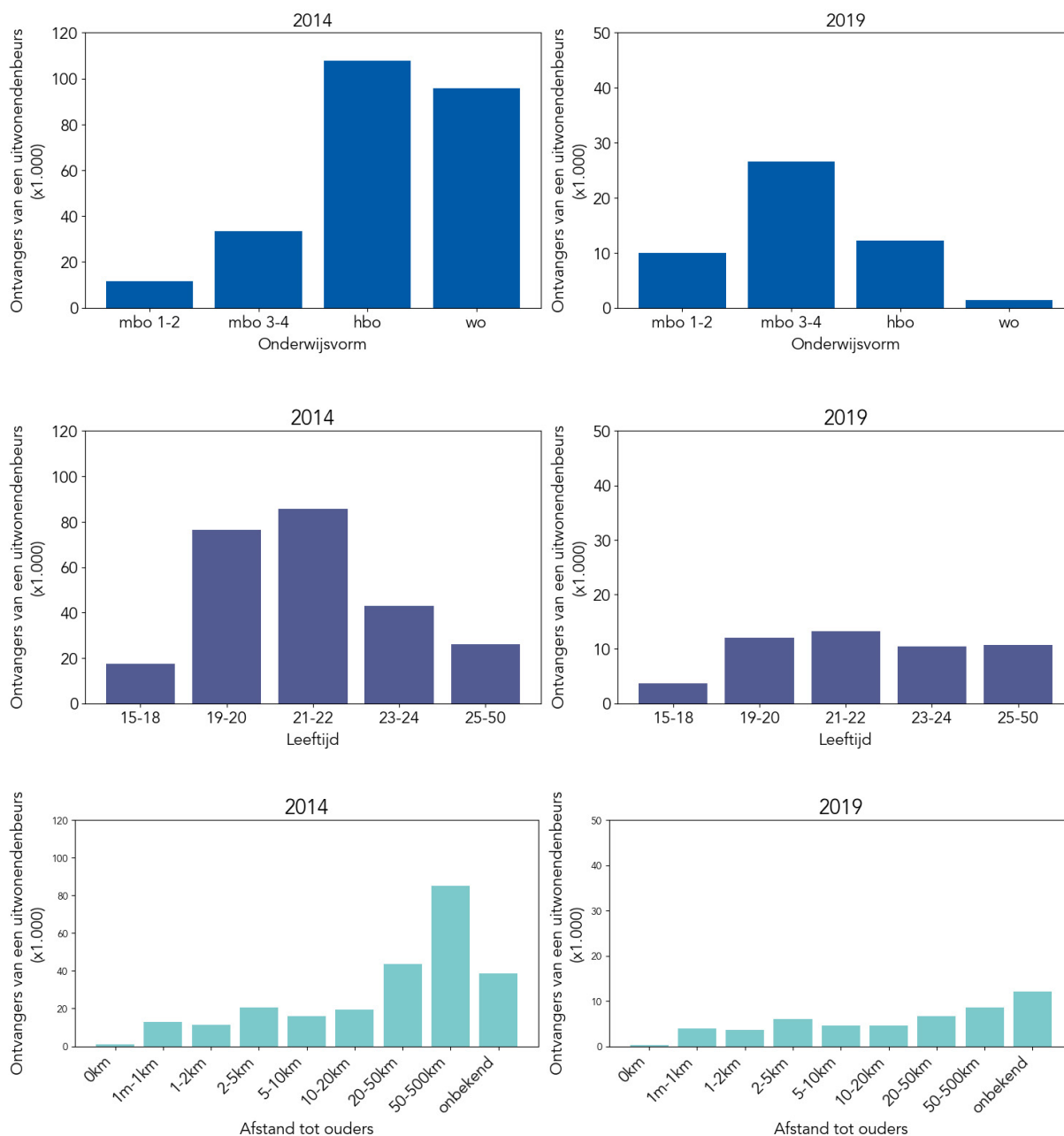
In oktober 2014 stonden er respectievelijk 479.800, 446.434 en 255.661²⁸ studenten ingeschreven bij mbo-, hbo- en wo-onderwijsinstellingen. In 2014 ontving respectievelijk 15,5%, 35,2% en 41,0% van deze studenten een uitwonendenbeurs. Deze cijfers komen op hoofdlijnen overeen met de Landelijke Monitor Studentenhuisvesting 2014.²⁹

Merk op dat de afstandscategorie 'onbekend' relevant is voor studenten waarvan het adres van de ouder(s) niet bekend is, de ouder(s) niet bekend zijn, overleden zijn of waar bij een hardheidsclausule van toepassing is.¹¹

²⁸ Nederlands Jeugdinstituut – Cijfers over hoger onderwijs (ho) <https://www.nji.nl/cijfers/hoger-onderwijs-ho>

²⁹ Landelijke Monitor Studentenhuisvesting (2014) <https://www.kences.nl/publicaties/landelijke-monitor-studentenhuisvesting-2014/>

Verdeling van uitwonendenbeurspopulatie-2014 (n=248.649) en uitwonendenbeurspopulatie-2019 (n=50.233) per onderwijsvorm, leeftijd en afstand tot ouder(s)



Figuur 5 – Verdeling van onderwijsvorm, leeftijd en afstand tot ouder(s) in uitwonendenbeurspopulatie-2014 (n=248.649) en uitwonendenbeurspopulatie-2019 (n=50.233).

In uitwonendenbeurspopulatie-2014 en uitwonendenbeurspopulatie-2019 bevinden zich respectievelijk 3.179 en 934 studenten die zijn geselecteerd voor een controle. Het resultaat van een controle kent drie uitkomsten: de student maakt rechtmatig gebruik van de uitwonendenbeurs (rechtmatig), student maakt onrechtmatig gebruik van uitwonendenbeurs (onrechtmatig) en de uitkomst van het onderzoek is onbekend (onbekend). De uitkomst van een controle kan onbekend zijn, bijvoorbeeld omdat een huisbezoek wordt geweigerd of dat een student niet thuis is. De resultaten van uitgevoerde controles in 2014 en 2019 worden weergegeven in [Tabel 12-13](#).

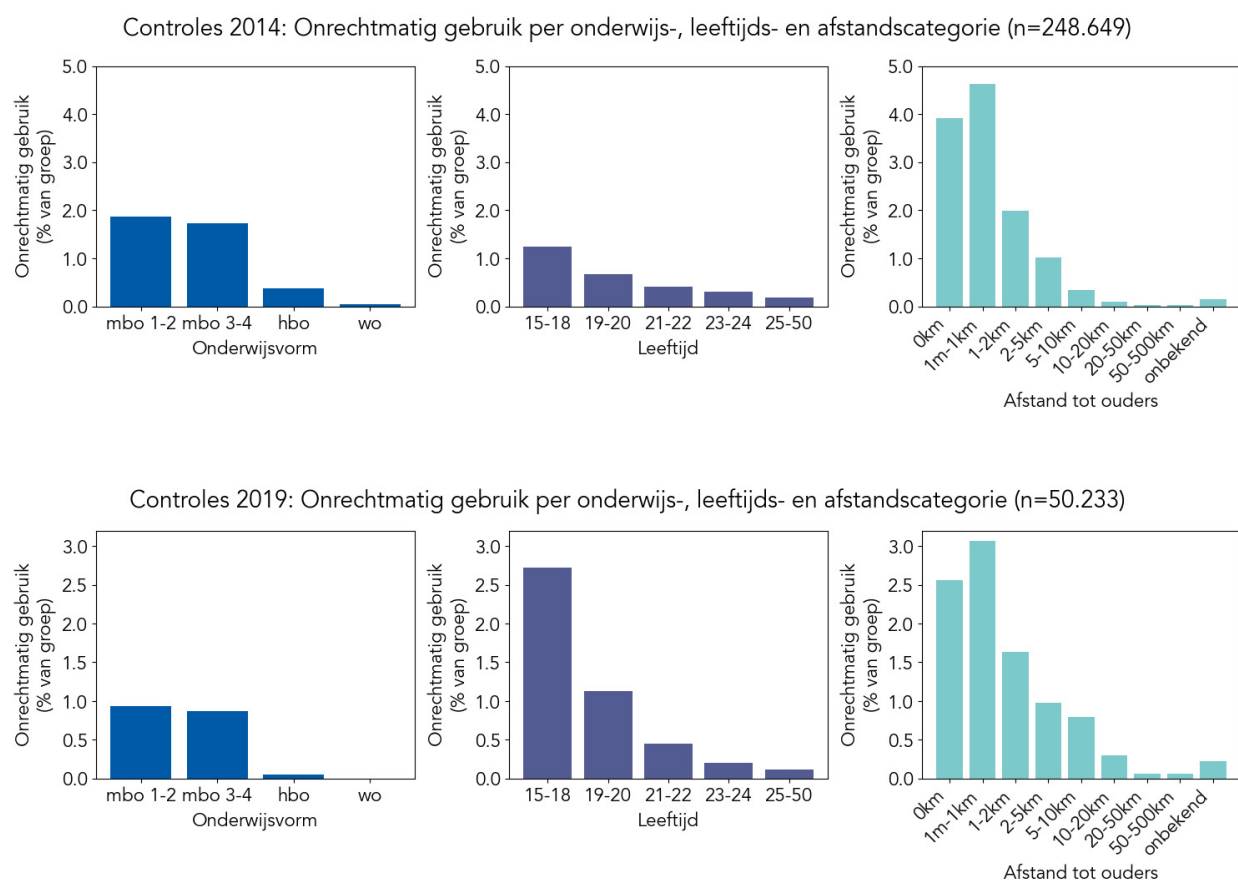
Controles 2014	Aantal	Percentage
Rechtmatig	1.566	49,2%
Onrechtmatig	1.238	38,9%
Onbekend	375	11,9%
Totaal	3.179	100%

Tabel 12 – Uitkomsten van CUB-controles in 2014.

Controles 2019	Aantal	Percentage
Rechtmatig	406	43,5%
Onrechtmatig	330	35,3%
Onbekend	198	21,2%
Totaal	934	100%

Tabel 13 – Uitkomsten van CUB- controles in 2019.

De uitkomsten van de CUB-controles uit 2014 en 2019 afgezet tegen onderwijsvorm, leeftijd en afstand tot ouder(s) worden weergegeven in [Figuur 6](#).



Figuur 6 – Onrechtmatig gebruik van de uitwonendenbeurs uitgesplitst per onderwijs-, leeftijd- en afstandscategorie gebaseerd op uitwonendenbeurspopulatie-2014 (n=248.649) en uitwonendenbeurspopulatie-2019 (n=50.233).

De resultaten die uit [Figuur 6](#) kunnen worden afgeleid, zijn in lijn met de logica van het risicoprofiel: studenten maken vaker onrechtmatig gebruik van de uitwonendenbeurs als zij een praktische onderwijsvorm genieten, jong zijn en dicht bij hun ouder(s) staan ingeschreven. Deze verschillen worden echter geobserveerd aan het einde van het CUB-proces; de toegekende risicoscores in stap 1 van het CUB-proces en de handmatige selectie van studenten in stap 3 van het CUB-proces hebben hier invloed op gehad. Dit werpt de vraag op welke invloed het algoritme (stap 1) en de menselijke selectie (stap 3) precies hebben op de uitkomst van het CUB-proces. Deze vraag staat centraal in de volgende secties van dit onderzoek.

Resultaten onderzoeksvraag 4

Onderzoeksvraag 4 wordt geformuleerd aan de hand van twee kernbegrippen uit de wetenschappelijke literatuur rondom eerlijke algoritmes (conditionele demografische pariteit en demografische pariteit). Deze begrippen worden toegelicht in geïntroduceerd in [3.1 Kwantitatieve analyse](#).

Onderzoeksvraag 4:

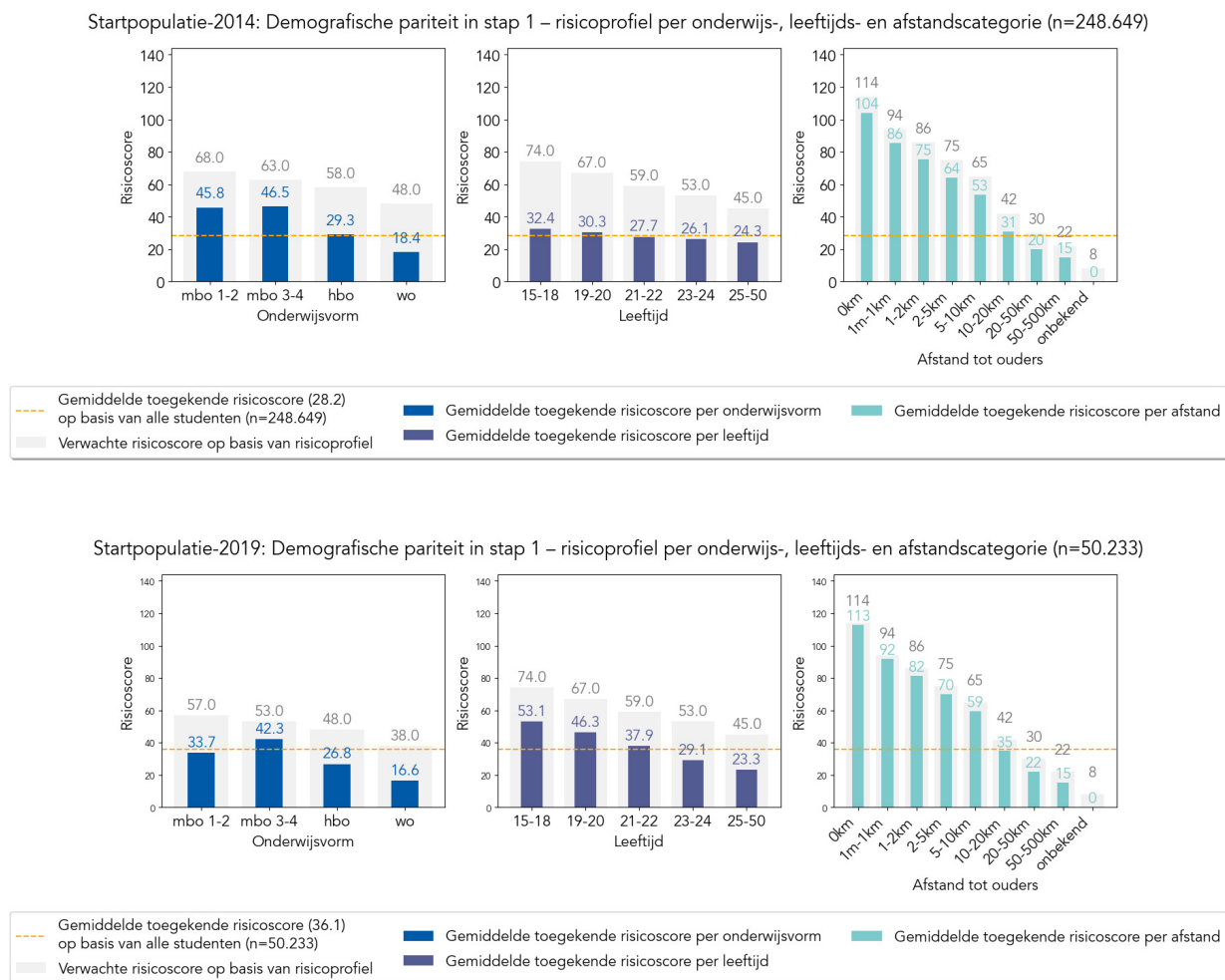
Is er sprake geweest van demografische pariteit voor de risicoscore-populatie (na stap 1 – risicoprofiel) en de huisbezoekpopulatie (na stap 3 – handmatige selectie door medewerker)?

Antwoord op onderzoeksvraag 4:

Voor zowel stap 1 (risicoprofiel) als stap 3 (handmatige selectie door medewerker) is er in 2014 en 2019 geen sprake geweest van demografische pariteit. Van zuivere demografische pariteit is in de realiteit zelden sprake. Demografische pariteit geldt vooral als een referentiepunt om gedrag van een risicoprofiel of van menselijke handelen tegen af te zetten. Aan de hand van onderstaande analyse wordt waargenomen dat studenten die dicht bij hun ouders ingeschreven staan, met name in de 0km, 1m-1km en 1-2km afstandscategorie, uitzonderlijk vaak handmatig worden geselecteerd voor een huisbezoek in stap 3 van het CUB-proces. Voor de hand ligt dat specifieke werkinstructies, die aansporen tot het handmatig selecteren van studenten die nabij het ouderlijk adres staan ingeschreven, hier de oorzaak van zijn.

Toelichting antwoord onderzoeksvraag 4:

[Figuur 7](#) toont de gemiddelde risicoscore die door het risicoprofiel is toegekend per onderwijs-, leeftijds- en afstandscategorie. De grijze balken tonen de verwachte score zoals kan worden bepaald met behulp van het risicoprofiel zoals gegeven in [Tabel 1](#). Dit is de score die het risicoprofiel geeft aan een bepaalde groep, aangenomen dat alle andere kenmerken willekeurig worden gekozen.



Figuur 7 – Demografische pariteit stap 1 – risicoprofiel per onderwijs-, leeftijds- en afstands categorie voor uitwonendenbeurspopulatie-2014 (n=248.649) en uitwonendenbeurspopulatie-2019 (n=50.233).

Observaties met betrekking tot demografische pariteit na stap 1 (risicoprofiel) worden per selectie criterium besproken.

- > **Onderwijsvorm** – Het risicomodel maakt een sterk onderscheid op basis van de onderwijsvorm. In 2014 is de risicoscore voor mbo 1-2 en mbo 3-4 studenten maar liefst 1,6x hoger dan de gemiddelde toegekende risicoscore (45,8/28,2 en 46,5/28,2 respectievelijk). Opvallend is dat dit patroon niet zichtbaar is 2019, waar er minder onderscheid gemaakt wordt op basis van onderwijsvorm – voor mbo 1-2 is de score 0,9x het gemiddelde (33,7/36,1) en voor mbo 3-4 slechts 1,2x (42,3/36,1).
- > **Leeftijd** – Het risicomodel maakt een onderscheid op basis van leeftijd, maar met name in 2014 lijkt dit onderscheid minder groot dan bij de kenmerken onderwijsvorm en afstand tot ouder(s). Bijvoorbeeld, in 2014 is de risicoscore voor iemand van 15-18 maar liefst 1,1x de gemiddelde risicoscore (32,4/28,2). Echter, dit patroon lijkt anders voor 2019, waarin een hogere factor waarneembaar is op basis van leeftijd. In 2019 is de risicoscore voor iemand van 15-18 maar liefst 1,5x de gemiddelde risicoscore (53,1/36,1).
- > **Afstand tot ouder(s)** – Het risicomodel maakt een zeer sterk onderscheid met betrekking tot de afstand tot ouders. In 2014 is de toegekende risicoscore aan iemand die 0km van zijn/haar ouders woont 3,7x hoger dan de gemiddelde toegekende risicoscore

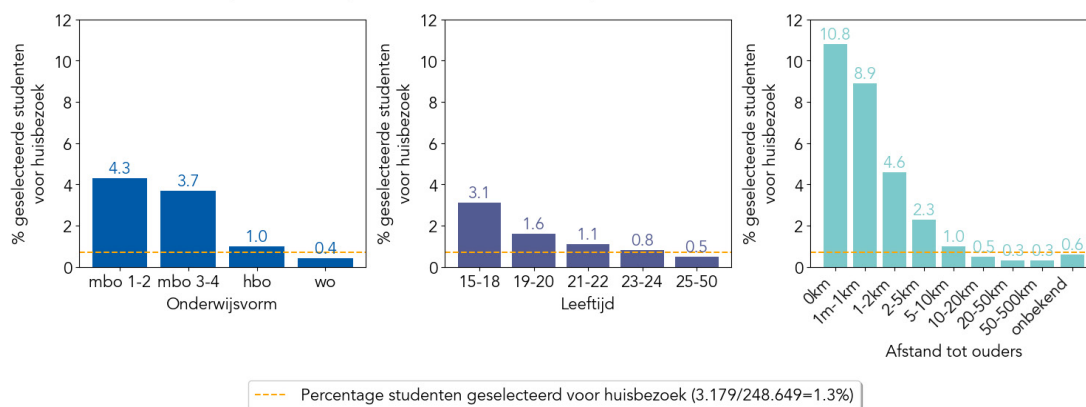
voor alle afstandscategorieën (104/28,2). Voor de 1m-1km categorie is dit 3,0x hoger (86/28,2) en voor de 1-2km categorie is dit 2,7x hoger (75/28,2). We observeren een vergelijkbaar patroon voor 2019.

Daarnaast wordt de verschillen in toegekende en verwachte risicoscores per selectie criterium besproken:

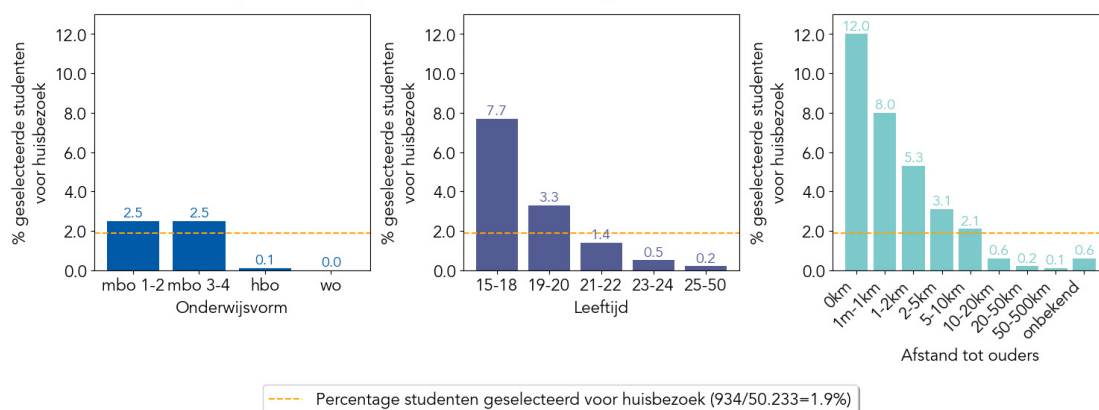
- > **Onderwijsvorm** – De gemiddelde risicoscores per onderwijsvorm (blauwe balken) vallen ruim onder de verwachte risicoscores (grijze balken). Dit kan worden verklaard door de groep studenten met onbekende afstand tot ouder(s) die risicoscore 0 krijgen toegekend, wat de gemiddelde scores aanzienlijk drukt. In 2014 waren er 38.657 ontvangers van de uitwonendenbeurs met een onbekend adres. In 2019 waren dit er 12.059.
- > **Leeftijd** – De gemiddelde risicoscores per leeftijdscategorie (paarse balken) vallen onder de verwachte risicoscores (grijze balken). Dit kan ook worden verklaard door de groep studenten met onbekende afstand tot ouder(s) die risicoscore 0 krijgen toegekend, wat de gemiddelde scores aanzienlijk drukt.
- > **Afstand tot ouder(s)** – De gemiddelde risicoscores per afstandscategorie (groene balken) vallen voor 2014 en 2019 grotendeels in lijn met de verwachte risicoscores (grijze balken).

Figuur 8 toont de kans om geselecteerd te worden voor een huisbezoek per onderwijs-, leeftijds- en afstandscategorie in 2014 en 2019.

CUB 2014: Demografische pariteit in stap 3 – handmatige selectie per onderwijs-, leeftijds- en afstandscategorie (n=248.649)



CUB 2019: Demografische pariteit in stap 3 – handmatige selectie per onderwijs-, leeftijds- en afstandscategorie (n=50.233)



Figuur 8 – Demografische pariteit stap 3 – handmatige selectie door DUO medewerker in CUB-proces 2014 (n=248.649) en CUB-proces 2019 (n=50.233).

Observaties met betrekking tot demografische pariteit na stap 3 (handmatige selectie medewerker) worden per selectie criterium besproken.

- > **Onderwijsvorm** – Voor 2014 valt op dat er bovenmatig veel mbo 1-2- en mbo 3-4-studenten worden geselecteerd. Voor mbo 1-2-studenten is dit 3,3x zo vaak als het gemiddelde (4,3%/1,3%). Voor mbo 3-4-studenten is dit 2,8x zo vaak als het gemiddelde (3,7% /1,3%). Dit is opmerkelijk aangezien het risicoprofiel voor deze groepen respectievelijk een risicoscore toekent die 0,9x en 1,3x hoger is dan het gemiddelde.
- > **Leeftijd** – Voor 2019 valt verder op dat er bovenmatig veel 15-18-jarigen worden geselecteerd: 4,1x meer dan het gemiddelde (7,7%/1,9%). Het risicoprofiel kent voor deze groep een 1,5x hogere risicoscore toe dan het gemiddelde.
- > **Afstand tot ouder(s)** – Voor zowel 2014 als 2019 is de kans om geselecteerd te worden voor een huisbezoek als studenten dicht bij hun ouder(s) ingeschreven staan (0km, 1m-1km, 1-2km) uitzonderlijk hoog. Voor zij die in 2014 op 0km afstand wonen is de kans op een huisbezoek (10,8%) 8,3x groter dan het gemiddelde voor alle categorieën (1,3%). Voor de 1m-1km-categorie is dit 6,8x hoger (8,9%/1,3%) en voor de 1-2km-categorie is dit 3,5x hoger (4,6%/1,3%). Een soortgelijk patroon wordt waargenomen in de data over het CUB-proces uit 2019. Dit is een sterk contrast met de toegekende scores door het risicoprofiel – bijvoorbeeld, iemand die in 2014 op 0km afstand woont heeft slechts een 3,7x hogere score dan het gemiddelde, wat relatief minder groot verschil is dan de kans op huisbezoek (8,3x hoger dan het gemiddelde). Dit verschil wijst op een sterke overreactie van medewerkers op basis van een afstand nabij het ouderlijk adres tijdens de handmatige selectie van studenten voor een huisbezoek.

Resultaten onderzoeksvraag 5

Merk op dat de twee kernbegrippen uit de wetenschappelijke literatuur rondom eerlijke algoritmes (conditionele demografische pariteit en demografische pariteit), aan de hand waarvan onderzoeksvraag 4 is geformuleerd, worden geïntroduceerd in [3.1 Kwantitatieve analyse](#).

Onderzoeksvraag 5:

Is er sprake geweest van conditionele demografische pariteit voor de risicoscore-populatie (na stap 1 – risicoprofiel)?

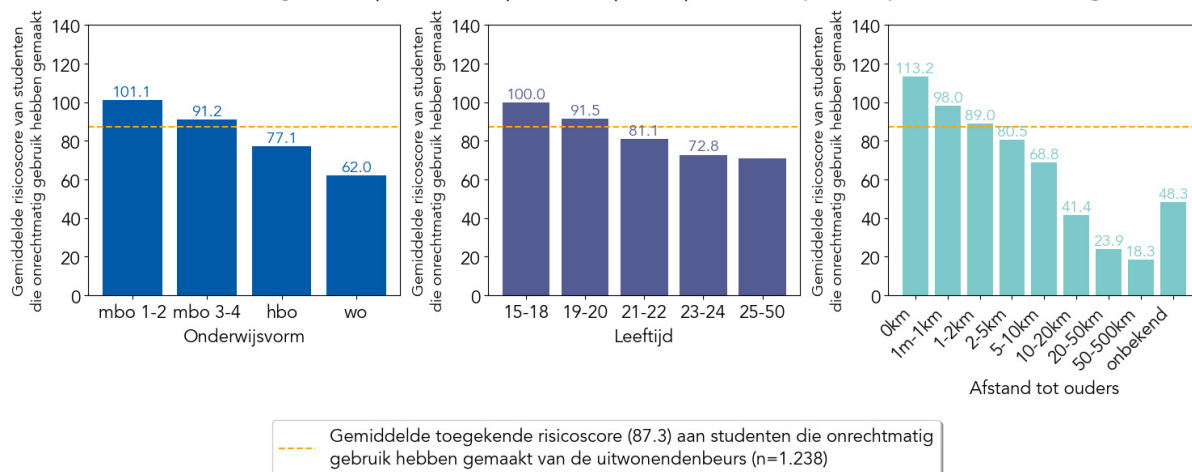
Antwoord op onderzoeksvraag 5:

In zowel 2014 en 2019 is er geen sprake geweest van conditionele demografische pariteit voor stap 1 (risicoprofiel) van het CUB-proces. Van zuivere conditionele demografische pariteit (zoals gedefinieerd in [3.1 Kwantitatieve analyse](#)) is zelden sprake. Het meten van conditionele demografische pariteit geldt vooral als een referentiepunt om het gedrag van het risicoprofiel tegen af te zetten. Aan de hand van deze analyse is waargenomen dat sommige groepen die ver van hun ouder(s) ingeschreven staan een lagere risicoscore krijgen dan verwacht kan worden op basis van de uitkomsten van onrechtmatigheidsonderzoek. Deze observatie bevestigt het beeld dat het risicoprofiel niet optimaal functioneert.

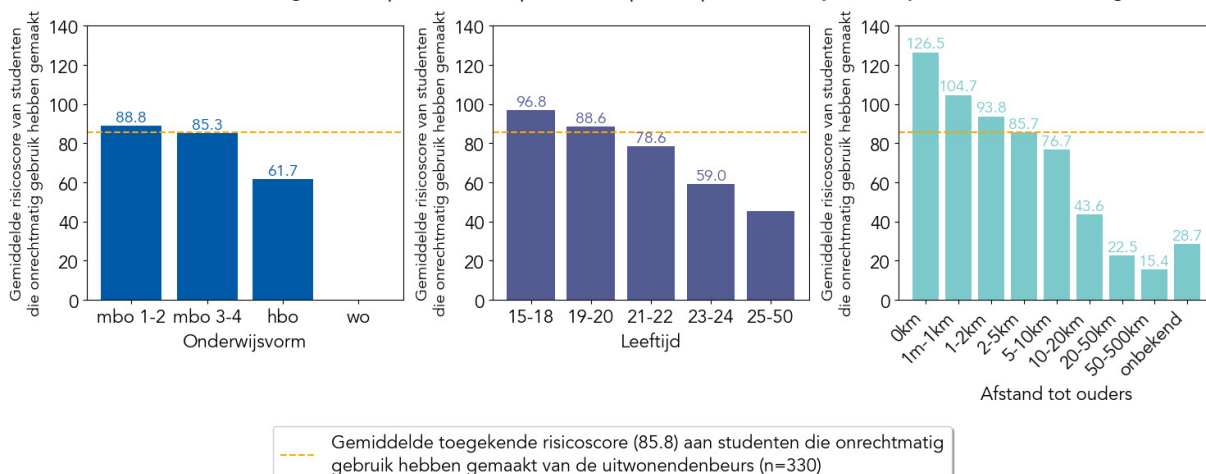
Toelichting onderzoeksvraag 5:

Verschillen opgemerkt in de vorige sectie in termen van demografische pariteit zijn eventueel te rechtvaardigen omdat studenten binnen bepaalde groepen vaker onrechtmatig gebruik maken van de uitwonendenbeurs. Conditionele demografische pariteit compenseert hiervoor door relatieve verschillen weer te geven in de groep met studenten die onrechtmatig gebruik hebben gemaakt van de uitwonendenbeurs. Uit **Figuur 9** volgt dat aan de hand van de uitkomsten van het CUB-proces in 2014 er in zekere zin sprake is van conditionele demografische pariteit voor onderwijs- en leeftijdscategorieën. Alleen wo-studenten in 2014 hebben een lagere score toegekend gekregen dan te verwachten is voor studenten die onrechtmatig gebruik hebben gemaakt van de uitwonendenbeurs. Daarnaast springt de discrepantie in het oog voor hogere afstandscategorieën. Dit betekent dat er relatief veel studenten zijn met een lage risicoscore (onder andere omdat ze ver van hun ouder(s) ingeschreven staan) die toch onrechtmatig gebruik hebben gemaakt van de uitwonendenbeurs. Uit **Figuur 9** volgt voor de uitkomsten van het CUB-proces in 2019 hetzelfde patroon. Behalve dat leeftijdscategorieën waarin oudere studenten vallen verder onder het gemiddelde liggen dan in 2014. Deze uitkomsten bevestigen het beeld dat het risicomodel en bijbehorende risicoclassificatie niet optimaal presteert.

CUB 2014: Conditionele demografische pariteit in stap 1 – risicoprofiel per onderwijs-, leeftijds- en afstandscategorie (n=1.238)



CUB 2019: Conditionele demografische pariteit in stap 1 – risicoprofiel per onderwijs-, leeftijds- en afstandscategorie (n=330)



Figuur 9 – Gemiddelde risicoscores door risicoprofiel (stap 1) toegekend aan studenten die onrechtmatig gebruik hebben gemaakt van de uitwonendenbeurs opgesplitst per onderwijs-, leeftijds- en afstandscategorieën in 2014 (n=613) en 2019 (n=61).

Resultaten onderzoeksvraag 6

Onderzoeksvraag 6:

Hoeveel onrechtmatig gebruik wordt er geconstateerd in de populatie geselecteerd door het CUB-proces in 2014 en 2019?

Antwoord op onderzoeksvraag 6:

Uit [Tabellen 12-13](#) kan worden opgemaakt dat de effectiviteit van het CUB-proces in 2014 en 2019 respectievelijk 38,9% en 35,3% is gebleken (n=3.179 en n=934).

5. Resultaten kwalitatieve analyse

Hieronder worden de resultaten van het kwalitatieve onderzoek besproken. In §5.1 wordt de ontwikkeling van het risicoprofiel besproken. In §5.2 worden processen om vooringenomenheid tegen te gaan behandeld. Aan de hand van deze resultaten worden onderzoeksvragen 7 en 8 beantwoord.

5.1 Resultaten kwalitatieve analyse risicoprofiel

Onderzoeksvraag 7:

Voldeed het risicoprofiel in het CUB-proces aan de normen die tegenwoordig worden gesteld aan door de overheid gebruikte algoritmen?

De onderzoeksvraag wordt beantwoord aan de hand van twee onderdelen. Onderdeel 1: Tegengaan van vooringenomenheid bij de ontwikkeling van het risicoprofiel. En onderdeel 2: Processen om het risico op vooringenomenheid tegen te gaan.

Onderdeel 1: Tegengaan van vooringenomenheid bij de ontwikkeling van het risicoprofiel

Het risicoprofiel bestaat uit twee onderdelen: de selectiecriteria en de wegingsfactoren.

De selectiecriteria zijn opgesteld op basis van een handmatige analyse van lijsten met uitwonende studenten in de pilotregio's in combinatie met ervaringsgegevens en wat in de workshops werd genoemd 'gezond verstand' [A.46,A.47].

Het College voor de Rechten van de Mens waarschuwt voor het gebruik van ervaringsgegevens:

Hoewel ervaringsgegevens kunnen bijdragen aan een efficiënte uitoefening van toezicht- en handhavingsbevoegdheden, kunnen zij stigmatisering en discriminatie van bepaalde bevolkingsgroepen ook verder in de hand werken. Ervaringsgegevens zijn namelijk niet altijd zo neutraal of objectief als ze lijken: een correlatie tussen factor X en gedraging Y is niet hetzelfde als een causaal verband daartussen en dat wordt nogal eens verward. Daardoor klinken in ervaringsgegevens menigmaal vooroordelen en vooronderstellingen door. Bovendien: wanneer een bepaalde bevolkingsgroep vaker gecontroleerd wordt, zullen mensen uit deze groep ook vaker aan het licht komen als daders of overtreders, wat vervolgens weer leidt tot intensievere controle van deze bevolkingsgroep, en tot bevestiging van de noodzaak hiervan. Op die manier werkt een te zware nadruk op ervaringsgegevens een zichzelf vervullende voorspelling ('selffulfilling prophecy') in de hand. Mede om deze reden moet uiterst voorzichtig worden omgegaan met ervaringsgegevens en misdaadstatistieken: de totstandkoming hiervan is in veel gevallen allesbehalve objectief.³⁰

³⁰ College voor de Rechten van de Mens, *Discriminatie door risicoprofielen. Een mensenrechtelijk toetsingskader*, 2021, p.14.

Risicoprofilering kan juist ook een methode zijn om te selecteren op basis van meer objectieve criteria dan ervaringsgegevens. Een risicoprofiel gebaseerd op objectieve gegevens is tot op zekere hoogte neutraler dan een handmatige selectie. Deze objectiviteit kan echter juist ook uitsluitend schijnbaar zijn. Ook risicoprofielen die niet op ervaringsgegevens zijn gebaseerd, zijn niet vanzelfsprekend neutraal. Daar is het onderhavige risicoprofiel een voorbeeld van. Hoewel de selectiecriteria op zichzelf objectief zijn, laat de keuze voor deze selectiecriteria zien hoe de opstellers van het risicoprofiel over fraude denken. De selectiecriteria zijn niet 'gevonden' door een analyse van een groot aantal neutrale selectiecriteria; zij zijn door de opstellers van het risicoprofiel op basis van gezond verstand geselecteerd en daarna getoetst. Het is om deze redenen dat normatieve afwegingen onderliggend aan de samenstelling van een risicoprofiel zorgvuldig gedocumenteerd dienen te worden. Zie ook *Algoritmes en Grondrechten* (Universiteit Utrecht) p.49 met verdere verwijzingen. Waarborgen voor het opstellen en gebruiken van risicoprofileringsalgoritmes worden gegeven in het Onderzoekskader ADR en in het IAMA.

De wegingsfactoren zouden tot stand zijn gekomen door een 'data-analyse' van de verhoudingen tussen verschillende groepen studenten. Deze data-analyse is niet geheel meer te reproduceren. In [4. Resultaten kwantitatieve analyse](#) wordt nader op de totstandkoming van de wegingsfactoren ingegaan.

Het Onderzoekskader ADR schrijft voor dat *"bij het ontwikkelen van het algoritme relevante belanghebbenden moeten worden betrokken"* (SV.10). Deels is dit gedaan. Bij de ontwikkeling zijn de handhavingsautoriteiten en het ministerie betrokken. Niet betrokken zijn de subjecten die aan risicoprofilering onderworpen zouden worden; de studenten. Studenten hadden mogelijk kunnen wijzen op aannames of vooringenomenheden in het risicoprofiel.

Verder schrijft het Onderzoekskader ADR voor dat *"de doelstelling van het algoritme concreet is uitgewerkt tot functionele eisen voor het algoritme. De mate waarin aan deze eisen is voldaan is bepaald"* (DM.1). Er is op verschillende momenten voor en tijdens het gebruik van het CUB-risicoprofiel gemeten of het doel van het risicoprofiel (effectiviteit) werd bereikt [\[A.11-A.15,A.20-21\]](#). Dat gebeurde niet op basis van vooraf opgestelde normen – van tevoren was niet bepaald bij welke effectiviteit het risicoprofiel succesvol is.

Het Onderzoekskader ADR schrijft voor dat *"de keuze voor het model en de hyperparameters zijn beargumenteerd en vastgelegd"* (DM.2) en dat de grenzen van de toepasbaarheid van het model zijn gedocumenteerd (DM.3). Hier is deels aan voldaan [\[A.40\]](#). De keuze voor het risicoprofiel en haar parameters is deels te reconstrueren, maar niet consequent vastgelegd [\[C.42-1\]](#). Een onderbouwing van de wegingsfactoren (ook wel: risicocodering) ontbreekt. Waarom voor de gebruikte selectiecriteria is gekozen is nergens uitvoerig onderbouwd [\[E.9-4\]](#).

Een en ander staat ook op gespannen voet met richtlijn 9 van het Toetsingskader Discriminatie door Risicoprofielen van het College voor de rechten van de Mens: *"Selectiebesli*

singen moeten te allen tijde uitlegbaar zijn”.³¹

Ook is niet gebleken dat “de doelpopulatie is vastgesteld en dat er is gecontroleerd dat de testdata representatief is voor de data van de verschillende subgroepen die in de productiedata voorkomen” (DM.7). Er is niet gebleken dat “de input- en outputdata voldoen qua kwaliteit, volledigheid en betrouwbaarheid” (DM.9). In de selectie van 934 studenten in 2010 is rekening gehouden met de subgroepen onderwijsvorm, leeftijd en afstand tot ouder(s) [A.47]. Van andere subgroepen – voor dit onderzoek met name relevant, de subgroep migratieachtergrond/niet-Nederlandse etniciteit – is de representativiteit niet onderzocht. Uit het onderzoek is niet gebleken dat “tijdens de ontwikkeling van het model is beoordeeld of er een verschil bestaat tussen de prestatie van het model tussen verschillende subgroepen”, wat het Onderzoekskader ADR voorschrijft (DM.20).

Een en ander wordt samengevat in richtlijn SV.4 van het Onderzoekskader ADR: “de impact van het algoritme is geïnventariseerd en beoordeeld”. Dit is gedaan wat betreft de effectiviteit en de kosten van het algoritme. Er is geen oog geweest voor het ontstaan van mogelijke vooringenomenheid jegens bepaalde groepen. Hier moet in de toekomst meer rekening mee worden gehouden; niet alleen wat betreft de subgroepen die in het kwantitatieve onderzoek zijn meegenomen en die in het nieuws zijn geweest, maar ook andere subgroepen.

De IAMA bevestigt in meer algemene termen het belang van het expliciet maken van publieke waarden bij het opstellen van een algoritme. Als die waarden – waaronder mogelijke vooringenomenheid – expliciet worden gemaakt, kan er vervolgens op worden gecontroleerd:

Het is niet altijd gemakkelijk om te weten wat publieke waarden zijn, maar het gaat steeds om manifestaties van het algemeen belang. (...) Voorbeelden van publieke waarden zijn gelijkwaardigheid, (...), rechtszekerheid, verdelende rechtvaardigheid, respect voor kwetsbare groepen, participatie en efficiënte besteding van middelen. Op een concreter niveau kan ook bescherming van grondrechten (...) tot de publieke waarden worden gerekend.

Algoritmen kunnen dienen om bepaalde publieke waarden te vertalen naar concrete besluitvorming. Daarbij kunnen algoritmen bepaalde waarden versterken, maar ze kunnen ook publieke waarden - zoals grondrechten - aantasten. Juist daarom is het belangrijk om in kaart te brengen welke publieke waarden aan de orde kunnen zijn bij de inzet van het algoritme.³²

Dit alles - het oog hebben voor subgroepen, het betrekken van alle belanghebbenden bij de ontwikkeling van een risicoprofiel en het vooraf vastleggen van prestatieindicatoren – zal bij eventueel toekomstig gebruik van risicoprofielen beter moeten worden gedaan.

³¹ p.30

³² IAMA p.14

Onderdeel 2: Processen om het risico op vooringenomenheid tegen te gaan

Het Onderzoekskader ADR geeft richtlijnen voor het tegengaan van vooringenomenheid: *“De definitie van de verschillende groepen en de gewenste prestatie van het model voor deze groepen moeten zijn opgenomen in de functionele eisen”* (DM.16). De subgroepen onderwijsvorm, leeftijd en afstand tot ouder(s) zijn deels vastgesteld. Andere subgroepen – onder meer migratieachtergrond/niet-Nederlandse etniciteit – zijn niet gedefinieerd. Van geen van de groepen is de gewenste prestatie vastgesteld. Ook is *“de mate van geaccepteerde bias in de uitkomst niet bepaald en niet uitgewerkt in meetbare prestatiecriteria”* (DM.17). Door dit alles niet vast te leggen, is het veel lastiger om tijdens het gebruik van het risicoprofiel te monitoren of controleren of er door het risicoprofiel een oververtegenwoordiging van bepaalde subgroepen in het CUB-proces ontstond. Er kunnen processen worden ingericht om dit wel te monitoren, zodat mogelijke onbedoelde vooringenomenheid wordt tegengegaan.

Dergelijke processen worden ook voorgeschreven door het Onderzoekskader ADR: richtlijnen DM.18, DM.19 en DM.21 schrijven voor: *“De methoden om bias te voorkomen, detecteren en corrigeren zijn vastgelegd; de uitkomstbias van productiedata is beoordeeld voor de verschillende subgroepen en voldoet aan de prestatiecriteria; de mate van bias in de data, dataverzameling en het model zijn in kaart gebracht”*. Hoe dit kan worden gemonitord, is besproken in [4. Resultaten kwantitatieve analyse](#).

Omdat vooringenomenheid niet is gemeten, kan niet aan richtlijn DM.22 worden voldaan: *“bij de geconstateerde bias is beoordeeld of deze op discriminatie duidt”*. Op pagina 29 IAMA wordt het belang van het tegengaan van vooringenomenheid bevestigd. Dat ook ogenschijnlijk neutrale selectiecriteria kunnen discrimineren bevestigt het College voor de Rechten van de Mens.³³

Om te waarborgen dat aan bovengenoemde richtlijnen wordt voldaan, kunnen processen worden ingericht om het risico op vooringenomenheid van het risicoprofiel te minimaliseren. Dit schrijft de ADR ook voor: *“risicomangement vindt gestructureerd plaats vooraf en tijdens de inzet van het algoritme”* (SV.13). Meer specifiek moet *“een proces voor een periodieke evaluatie van de kwaliteit van het algoritme gedocumenteerd en in werking zijn. De resultaten moeten met belanghebbenden worden gedeeld”*. (SV.16) Periodieke evaluatie van het CUB-proces was wel ingericht, maar had geen betrekking op de kwaliteit van het risicoprofiel en het risico op vooringenomenheid [\[A.24\]](#). Als onderdeel van periodieke evaluatie dient de *“verantwoordelijke verantwoording af te leggen over de ontwikkeling, inzet en werking van het algoritme”* (SV.19).

Om evaluatie mogelijk te maken dienen *“de functionele eisen te zijn uitgewerkt tot adequate en meetbare prestatiecriteria”* (DM.4). Dit sluit aan op de reeds besproken aanbeveling om interne normen op te stellen. Hiervoor dienen zowel voor het risicoprofiel relevante prestatiecriteria als voor het CUB-proces relevante evaluatiecriteria te worden opgesteld.

³³ p.18 e.v. *Toetsingskader Discriminatie door Risicoprofielen* van het College voor de Rechten van de Mens.

Onderdeel van de evaluatie dient ook “de analyse te zijn of (interne en externe) klachten en incidenten het gevolg kunnen zijn van het gebruik van het algoritme” (SV.17). Als er al methodes en maatregelen waren om vooringenomenheid tegen te gaan in het CUB-proces – wat uit het onderzoek niet is gebleken³⁴ – dan zijn deze niet gedocumenteerd. Wel verdient opmerking dat er uit het onderzoek geen stelselmatige klachten over het risico-profiel naar boven zijn gekomen.

Zoals besproken is het CUB-proces waarvan het risicoprofiel een onderdeel was wel dege-lijk gemonitord. In M&O-onderzoeken is steeds vastgesteld dat het risico van onrechtma-tig gebruik van de uitwonendenbeurs hoog is. Wat hiervan de invloed is op het risicopro-fiel en het CUB-proces is niet bekend. Mogelijk kan bij deze periodieke M&O-onderzoeken ook een proces worden ingericht om vooringenomenheid tegen te gaan.

Bij het opstellen van risicomanagement kan een rol zijn weggelegd voor CIO Rijk. In het – nog niet in definitieve vorm gepubliceerde – Implementatiekader Verantwoorde Inzet van Algoritmen staat:

Samen met de ministeries geeft CIO Rijk ook vorm aan een interne toezichtstruc-tuur waarin de rollen van de verschillende lines of defense worden vormgegeven. De verschillende lines of defense zorgen voor meerdere controlepunten in het toezicht op algoritmen. De toezichtstructuur zal ook worden beschreven in de volgende fase van het implementatiekader. Daarbij zal ook in worden gegaan op de manier waarop dit binnen gemeentes, provincies en waterschappen kan worden vormgegeven.

Kortom, het strekt tot aanbeveling om risicomanagementsystemen in te richten voor het opstellen en gebruik van risicoprofielen – of algoritmes in het algemeen. Dit verkleint het risico op onbedoelde vooringenomenheid en maakt beslissingen uitlegbaar. Een bijkomend voordeel is.³⁵

Antwoord op onderzoeksvraag 7:

Het risicoprofiel voldeed niet aan verschillende de normen die tegenwoordig worden gesteld aan door de overheid gebruikte algoritmen. Het antwoord op onderzoeksvraag 7 is daarmee negatief.

5.2 Processen om het risico op vooringenomenheid tegen te gaan

Onderzoeksvraag 8:

Als het antwoord op onderzoeksvraag 7 negatief is, wat is er nodig om in de toekomst verantwoord het risicoprofiel te gebruiken?

Antwoord op onderzoeksvraag 8:

Onderzoeksvraag 8 is bij de bespreking van onderzoeksvraag 7 deels beantwoord. In de

³⁴ En wat de ADR overigens ook concludeert, zie B.8 en B.42-4.

³⁵ p.30

aanbevelingen zullen concrete maatregelen worden benoemd voor het in de toekomst (mogelijk) verantwoord gebruiken van het risicoprofiel.

6. Voorbehouden

Dit rapport beoogt geen uitputtend, definitief beeld van het CUB-proces te geven. Er zijn diverse documenten bij DUO opgevraagd en er zijn diverse documenten uit eigen beweging door DUO gedeeld. Mogelijk is er informatie beschikbaar die de onderzoekers niet hebben gezien en de analyse (deels) tegenspreken.

Algorithm Audit heeft verder data opgevraagd bij datawarehousing-experts van DUO. Er hebben controles plaatsgevonden op de correctheid van de data. Algorithm Audit kan niet garanderen dat de data waarop het rapport is gebaseerd, de bijbehorende queries en/of de onderliggende datastructuren volledig zijn en volledig vrij zijn van fouten of onvolkomenheden.

Algorithm Audit is niet verantwoordelijk voor beslissingen die worden genomen naar aanleiding van dit rapport.

Zonder voorafgaande toestemming van Algorithm Audit mag dit rapport – en/of delen van het rapport – niet worden gedeeld met partijen buiten DUO.

Bij het opstellen van dit rapport is de bescherming van persoonsgegevens in ogenschouw genomen.

- > De interne processen van Algorithm Audit voldoen aan de AVG;
- > DUO deelde haar documenten via haar eigen platform. Algorithm Audit heeft van DUO afkomstige documenten derhalve nooit in handen gehad. Bij het ontnemen van toegang tot de systemen van DUO, heeft Algorithm Audit definitief geen toegang meer tot de documenten.

7. Conclusie: bevindingen en aanbevelingen

Naar aanleiding van deze analyse zijn er verbeterpunten aan te wijzen om in de toekomst vooringenomenheid in het CUB-proces te voorkomen. Dit rapport doet daartoe een aanzet in zeven bevindingen (§7.1) en drie aanbevelingen (§7.2).

7.1 Bevindingen

In deze sectie worden de genoemde bevindingen toegelicht. De bevindingen komen voort uit de kwantitatieve en kwalitatieve analyse van het risicoprofiel en van hoe dit is opgesteld en beheerd binnen de organisatie.

Bevinding 1 – Er is jarenlang een regel-gebaseerd algoritme gebruikt, dat op basis van onderwijsvorm, leeftijd en de afstand tussen het adres van de student en het adres van zijn/haar ouder(s) een risicoscore heeft toegekend aan studenten die een uitwonendenbeurs ontvingen. Er is geen sprake van een zelflerend algoritme of van de inzet van kunstmatige intelligentie. Gebruik van risicoprofilering is effectief gebleken.

Tussen 2012 en 2023 is in het CUB-proces een risicoprofiel gebruikt om geautomatiseerd risicoscores toe te kennen aan alle studenten die van DUO een uitwonendenbeurs ontvingen. Het risicoprofiel is een lineair model dat op basis van drie criteria studenten een risicoscore toekent. Deze criteria zijn onderwijsvorm, leeftijd en de afstand tussen het adres waar zij stonden ingeschreven en het adres van hun ouder(s). Ieder criterium is opgedeeld in categorieën, bijvoorbeeld studenten die tussen de 1m-1km van hun ouder(s) wonen, studenten die 50-500km van hun ouder(s) wonen, een mbo 1-2 opleiding volgen etc. Na toepassing van vooraf bepaalde wegingsfactoren per categorie leidt het profiel van een student tot een risicoscore. De risicoscore kan worden verhoogd als de bij DUO bekende leeftijd van de student afwijkt met de leeftijd van de student in de Basis Registratie Personen (BRP). Aan alle studenten zijn risicoscores toegekend. Het toepassen van het CUB-proces, waarvan het risicoprofiel een belangrijk onderdeel uitmaakt, is effectief gebleken. De effectiviteit blijkt uit het feit dat met toepassing van het CUB-proces meer onrechtmatig gebruik van de uitwonendenbeurs is vastgesteld dan bij aselechte steekproeven. Dit betreft respectievelijk 3,6% en 3,8% effectiviteit bij de aselechte steekproeven van 2014 en 2017 en respectievelijk 38,9% en 35,3% effectiviteit bij toepassing van het CUB-proces in 2014 en 2019 waarbij het risicoprofiel is ingezet.

Meer informatie over het risicoprofiel en het CUB-proces wordt gegeven in [2.3 Overzicht CUB-proces](#). Meer informatie over de effectiviteitscijfers kan worden gevonden in [3.1 Kwantitatieve analyse](#).

Bevinding 2 – Uit de aselechte steekproeven uit 2014 en 2017 volgt onvoldoende statistisch verband tussen de selectiecriteria onderwijsvorm en leeftijd en onrechtmatig gebruik van de uitwonendenbeurs. Voor specifieke categorieën binnen het selectie criterium afstand is wel voldoende verband met onrechtmatigheid gevonden. Voor de opdeling in zes risicocategorieën is onvoldoende statistische onderbouwing gevon-

den ten opzichte van een binaire risicoclassificatie.

Per onderwijs-, leeftijds- en afstandscategorie is geturfd hoe vaak studenten die zijn geselecteerd voor onderzoek middels de aselechte steekproef 2014 (n=387) en 2017 (n=293) onrechtmatig gebruik hebben gemaakt van de uitwonendenbeurs. Aan de hand van deze frequenties is getoetst, middels de Z-toets en Fishers exacte toets, of de verschillen in onrechtmatigheidspercentages voor de gebruikte categorieën in het risicoprofiel statistisch significant zijn. Bijvoorbeeld: is er statistisch significant meer onrechtmatigheid vastgesteld in de afstandscategorie 2-5km dan in de categorie 5-10km? Voor leeftijd zijn onvoldoende aanwijzingen aangetroffen voor statistisch significante verschillen tussen de gebruikte categorieën. Voor onderwijsvorm zijn in de steekproef uit 2014 wel en in de steekproef uit 2017 geen significante verschillen aangetroffen. Dit acht Algorithm Audit een onvoldoende consistent statistisch signaal om risicoprofilering op te baseren. Een nieuwe, grotere steekproef zou hier verandering in kunnen brengen. Voor afstand tot ouder(s) leveren vooral de categorieën 1m-1km, 2-5km en 50-500km een statistisch significant signaal. Dit biedt onderbouwing voor gebruik van afstand als selectiecriteria voor risicoprofilering, mits de specifieke categorisering preciezer wordt vastgesteld. Daarnaast is er onvoldoende onderbouwing om de toegekende risicoscores op te delen in zes risicocategorieën. Voor een binaire risicoclassificatie is wel statistische onderbouwing, en deze simplificering geniet de voorkeur.

Meer informatie over de methodologie van de statistische toetsen kan worden gevonden in [3.1 Kwantitatieve analyse](#). De resultaten worden besproken in [4.1 Resultaten aselechte steekproef 2014 en 2017](#).

Bevinding 3 – Studenten die binnen een afstand van 2km tot hun ouders(s) staan ingeschreven zijn aanzienlijk vaker handmatig geselecteerd voor een huisbezoek dan men op basis van de door het risicoprofiel toegekende risicoscores kan verwachten. Voor de hand ligt dat specifieke werkinstructies, die aansporen tot het handmatig selecteren van studenten die nabij het ouderlijk adres staan ingeschreven, hier de oorzaak van zijn.

In de biasmeting is een sterke oververtegenwoordiging waargenomen van studenten die binnen 2km afstand tot hun ouder(s) staan ingeschreven bij handmatige selectie voor huisbezoek. Deze oververtegenwoordiging komt niet voort uit het gehanteerde risicoprofiel.

In 2014 is er voor studenten die 0km van hun ouder(s) staan ingeschreven een disproportionele verhouding tussen de kans op een huisbezoek (8,3x het gemiddelde van alle categorieën) en de toegekende risicoscore (3,7x het gemiddelde). Dat betekent dat studenten uit deze categorie een meer dan 2 keer zo grote kans hebben op een huisbezoek dan vanuit hun risicoscore kan worden verwacht. Voor de 1m-1km categorie is deze verhouding tussen kans op huisbezoek en risicoscore (beide ten opzichte van hun gemiddelden) 6,8x : 3,0x en voor de 1-2km categorie 3,5x : 2,7x. Dit betekent dat de groep met een kleine afstand tot ouder(s) structureel is oververtegenwoordigd in de handmatige selectie voor huisbezoek, ten opzichte van de hun toegekende risicoscore's.

Deze oververtegenwoordiging kan vermoedelijk worden verklaard door de aanwezigheid

van specifieke werkinstructies die medewerkers opdragen om studenten met bijzonderheden in de woonsituatie – zoals de combinatie jonge leeftijd, dichtbij ouder(s) wonend en/of wonend bij familieleden – gericht te selecteren voor huisbezoek. De oververtegenwoordiging van studenten (zoals die op een kleine afstand tot ouder(s) staan ingeschreven) in de handmatige selectie voor huisbezoek zou potentieel een oververtegenwoordiging van andere kenmerken, zoals die van migratieachtergrond, kunnen veroorzaken. Er moet daarom nader onderzoek worden gedaan of er een verband bestaat tussen de groepen die bovenmatig vaak handmatig zijn geselecteerd voor huisbezoek en studenten met een migratieachtergrond. Zie [Aanbeveling 2](#). Daarnaast moet een onderzoek en verbetertraject worden ingesteld voor het proces van handmatige selectie, waaronder de werkinstructies. Zie [Aanbeveling 3](#).

De uitkomsten van de biasmeting, inclusief de genoemde cijfers, kunnen worden teruggevonden [Figuur 8](#) en in [4.2. Resultaten biasmeting 2014 en 2019](#).

Bevinding 4 – Er is geen sprake van direct onderscheid op basis van migratieachtergrond (noch op basis van andere beschermde gronden) in het risicoprofiel. Door onvoldoende beschikbare gegevens is het niet mogelijk geweest om kwantitatief onderzoek te doen naar indirecte vooringenomenheid ten aanzien van studenten met een migratieachtergrond.

DUO heeft Algorithm Audit verzocht ook een biasmeting uit te voeren voor mogelijke vooringenomenheid met betrekking tot het bijzondere persoonskenmerk migratieachtergrond. DUO heeft zelf geen data over de migratieachtergrond van studenten. Daarom is aan het CBS verzocht de data van DUO uitsluitend voor het uitvoeren van deze biasmeting te verrijken met data over de migratieachtergrond van studenten. Tot op heden heeft Algorithm Audit deze data niet kunnen verkrijgen. In de toekomst kunnen deze data mogelijk wel beschikbaar worden gesteld voor eventueel vervolgonderzoek. Alternatieve methoden om indirecte vooringenomenheid met betrekking tot migratieachtergrond te meten, zoals met behulp van aggregatiestatistieken per postcodegebied, achtte Algorithm Audit onvoldoende verantwoord in methodologisch opzicht. Algorithm Audit heeft ten tijde van het verschijnen van dit rapport daarom geen kwantitatief onderzoek kunnen uitvoeren naar indirecte vooringenomenheid van het CUB-proces met betrekking tot studenten met een migratieachtergrond. Mogelijke indirecte vooringenomenheid in het risicoprofiel of in het verdere verloop van het CUB-proces valt niet uit te sluiten.

Hoe de biasmeting voor migratieachtergrond zou zijn uitgevoerd in het geval van voldoende beschikbare gegevens wordt toegelicht in [3.1 Kwantitatieve analyse](#).

Bevinding 5 – De selectiecriteria en risicocategorieën die zijn gebruikt in het risicoprofiel zijn naar huidige normen onvoldoende onderbouwd, zeker ten aanzien van mogelijke vooringenomenheid.

De selectiecriteria van het risicoprofiel zijn grotendeels gebaseerd op ervaringsgegevens en zogenaamd gezond verstand van medewerkers. De onderbouwing voor geschiktheid van deze criteria is niet gedocumenteerd. Het ontstaan van het risicoprofiel is getraceerd naar een serie workshops in 2010. Hoewel ervaring en gezond verstand belangrijk en nuttig zijn, vormen *self fulfilling prophecies*, *confirmation biases* en het onbedoeld gebruiken van proxycriteria (zie [Box 1](#)) hier een reëel risico. De wegingsfactoren zijn gebaseerd op een datastudie uitgevoerd in 2010. Documentatie over de oorsprong en gehanteerde methodologie voor het samenstellen van de controlegroep in deze datastudie ontbreekt. Zonder nader onderzoek naar de representativiteit van deze controlegroep kunnen in een handmatig geselecteerde populatie bepaalde groepen over- of ondervertegenwoordigd zijn. Door op die populatie wegingsfactoren te baseren, bestaat het risico op het ontstaan van negatieve *feedback loops*. Dat betekent dat groepen die in de controlegroep oververtegenwoordigd zijn een hogere wegingsfactor krijgen, waardoor deze groepen later door het risicoprofiel systematisch een overmatige risicoscore wordt toebedeeld. Het onderbouwen van de keuze voor bepaalde selectiecriteria en wegingsfactoren en hun mogelijke invloed op vooringenomenheid wordt voorgeschreven door de nu geldende kaders voor algoritmes zoals de Impact Assessment Mensenrechten en Algoritmes (IAMA) en het Onderzoekskader Algoritmes van de ADR.

Een analyse over de onderbouwing van de gebruikte selectiecriteria en categorieën in het risicoprofiel wordt gegeven in [5.1 Resultaten kwalitatieve analyse risicoprofiel](#).

Bevinding 6 – Er is niet gebleken van intern onderzoek naar mogelijke vooringenomenheid van het gebruikte risicoprofiel, noch bij het opstellen van het risicoprofiel noch tijdens gebruik ervan.

Risicoprofilering is een krachtig en door de Centrale Raad van Beroep toegestaan middel voor het vergroten van effectiviteit, efficiëntie en kostenbesparing van de bestrijding van onrechtmatig gebruik van sociale voorzieningen.³⁶ Aan het gebruik zijn echter voorwaarden verbonden, waaronder dat het gemaakte onderscheid geschikt, noodzakelijk en proportioneel dient te zijn. Dit moet worden gewaarborgd bij het opstellen van het risicoprofiel en de inzet daarvan. Er is niet gebleken van bewustzijn binnen DUO (en/of bij betrokken instanties als het Ministerie van OCW, de ADR, de Tweede Kamer na invoering van CUB in 2012, etc.) dat het gebruik van het risicoprofiel in het CUB-proces een risico op (indirecte) ongelijke behandeling met zich meebrengt. Er is ook niet gebleken dat waarborgen ingesteld zijn om dergelijke ongelijke behandeling te monitoren of tegen te gaan. Er is geen vooringenomenheid gemeten en er is niet onderzocht of overwogen of de gebruikte criteria en wegingsfactoren wellicht proxykenmerken waren voor bepaalde demografieën.

Een toelichting op het ontbreken van intern onderzoek naar vooringenomenheid kan worden gevonden in [5.2 Processen om het risico op vooringenomenheid tegen te gaan](#).

³⁶ Onder meer CRvB 8 september 2015, ECLI:NL:CRVB:2015:3249, r.o. 4.5.; zie ook Toetsingskader Discriminatie door Risicoprofielen van het College voor de Rechten van de Mens (2021) Richtlijn 1.

Bevinding 7 – Er is niet gebleken van interne normen voor het gebruik van algoritmes.

In het onderzoek zijn geen richtlijnen gevonden die DUO intern gebruikt voor het tegengaan van vooringenomenheid of het mitigeren van andere risico's bij het gebruik van algoritmes. Er lijkt niet te zijn vastgelegd hoe DUO invulling geeft aan algemene wettelijke kaders en richtlijnen van andere instanties voor de samenstelling van risicoprofielen. Als die richtlijnen er wel zijn, is in ieder geval niet de hele organisatie zich daarvan bewust. Richtlijnen ontbreken op ten minste twee onderdelen die relevant zijn voor het tegengaan van (indirecte) discriminatie. Ten eerste voor de vaststelling bij welke mate van vooringenomenheid sprake is van indirecte discriminatie. Ten tweede om vast te stellen onder welke omstandigheden indirecte discriminatie naar het oordeel van DUO gerechtvaardigd is, omdat wordt voldaan aan de juridische eisen van geschiktheid, noodzakelijkheid en proportionaliteit. In de periode 2012-2022 hebben er wel jaarlijkse effectiviteitsrapportages, privacy-audits en misbruik-en-oneigenlijk-gebruik (M&O) controles plaatsgevonden. Aselecte steekproeven hebben plaatsgevonden in 2010, 2014 en 2017, primair met het doel om de effectiviteit van het CUB-proces en het financieel 'restrisico' in de CUB-populatie vast te stellen. Het gebrek aan richtlijnen voor de omgang met (algoritmische) risicoprofilering heeft eraan bijgedragen dat er bij de controles geen oog is geweest voor mogelijke vooringenomenheid.

Een toelichting op het ontbreken van interne normen voor gebruik van algoritmes wordt gegeven in [5.2 Processen om het risico op vooringenomenheid tegen te gaan](#).

7.2 Aanbevelingen

Uit de genoemde bevindingen volgen de onderstaande aanbevelingen.

Aanbeveling 1 – Zorg voor een meer omvattende onderbouwing van het risicoprofiel voordat het mogelijk weer in gebruik wordt genomen, onder meer met behulp van een normatief kader.

Inherent aan het gebruik van risicoprofilering in het CUB-proces is dat er onderscheid wordt gemaakt tussen groepen uitwonende studenten. Deels is dit de bedoeling: geen *bug*, maar een *feature*. Het maken van onderscheid kan echter ook juridische en maatschappelijke normen overschrijden. Dit is bijvoorbeeld het geval wanneer het onderscheid niet geschikt, noodzakelijk en proportioneel is. Tegelijkertijd kan onderscheid dat juridisch geoorloofd is toch als maatschappelijk onwenselijk worden beschouwd. Hierbij kan gedacht worden aan onderscheid dat zich niet richt op een juridisch beschermd (persoons) kenmerk, maar op een ander kenmerk zoals opleidingsniveau. Het is aan te bevelen om op voorhand kaders op te stellen waaraan getoetst kan worden voor welke *vormen* en in welke *mate* van onderscheid als (on)wenselijk worden beschouwd. Deze kaders hebben specifiek betrekking op het toetsen van criteria die in het risicoprofiel worden opgeno-

men. Het voorgestelde Fase-2 onderzoek van Algorithm Audit biedt aanknopingspunten om dergelijke normatieve afwegingen te maken. Het toetsen van een risicoprofiel aan een kader moet ook een kwantitatieve onderbouwing bevatten van de gebruikte selectiecriteria, waarvoor de analyses in dit rapport een voorzet doen. Om het selectieprofiel kwantitatief te onderbouwen, en niet alleen de effectiviteit van het CUB-proces te meten, wordt aanbevolen om een grotere aselecte steekproef te trekken dan de aselecte steekproeven uit 2014 en 2017.

Aanbeveling 2 – Er moet nader onderzoek worden gedaan of er een verband bestaat tussen de groepen die bovenmatig vaak handmatig zijn geselecteerd voor huisbezoek en studenten met een migratieachtergrond.

Of groepen die tijdens de handmatige selectie bovenmatig vaak zijn geselecteerd ook bovenmatig vaak een migratieachtergrond hebben, is niet op voorhand duidelijk en moet nader worden onderzocht. De gehanteerde werkinstructies, bijvoorbeeld de richtlijn om bijzonderheden in de woonsituatie nader te onderzoeken (denk aan: combinatie jong en bij familie inwonend), kunnen bijdragen aan mogelijke oververtegenwoordiging van demografische groepen, waaronder migratieachtergrond. Daarnaast is het proces van handmatige selectie vatbaar voor onderscheid op basis van latente kenmerken. Bij de lijsten die medewerkers gepresenteerd krijgen tijdens handmatige selectie zijn, naast de criteria uit het risicoprofiel en de risicoscore, ook andere kenmerken zichtbaar van studenten, zoals naam, adres en geboortedatum. Hier bevindt zich een risico voor (onbewuste) vooringenomenheid ten opzichte van migratieachtergrond of andere kenmerken in handmatige selectie. De invloed die deze en andere kenmerken spelen tijdens de handmatige selectie voor huisbezoek dient te worden onderzocht. Daarnaast moet een verbetertraject worden ingesteld om werkinstructies voor handmatige selectie te verbeteren. Zie [Aanbeveling 3](#).

Aanbeveling 3 – Voorkom het gebruik van hetzelfde selectie criterium in verschillende stappen van het CUB-proces. Stel daartoe een onderzoek en verbetertraject in voor het proces van handmatige selectie, waaronder de werkinstructies.

Er is vastgesteld dat studenten die binnen een afstand van 2km tot hun ouders(s) staan ingeschreven aanzienlijk vaker handmatig zijn geselecteerd voor een huisbezoek dan men op basis van de door het risicoprofiel toegekende risicoscores kan verwachten. Voor de hand ligt dat specifieke werkinstructies die aansporen tot het handmatig selecteren van studenten die nabij het ouderlijk adres staan ingeschreven hier de oorzaak van zijn. Deze bevindingen geven aanleiding om de werkinstructies voor medewerkers in het selectieproces tegen het licht te houden. Als de door het model toegekende risicoscore al kenmerken zoals kleine afstand tot ouder(s) meeweegt, kan een nadere instructie om diezelfde kenmerken handmatig vaker te selecteren een overreactie veroorzaken. Daardoor worden studenten met bijvoorbeeld een kleine afstand tot ouder(s) onterecht veel vaker geselecteerd voor huisbezoek. Het algemene proces van handmatige selectie moet nader worden onderzocht om verbeterpunten vast te stellen. Het huidige proces is ondoorzichtig, omdat de procedure om handmatig studenten in en uit te sluiten van huisbezoek op basis van hun risicoscore's en relevante kenmerken niet een eenduidig protocol is. Het is lastig

te achterhalen hoe medewerkers precies te werk gaan en het proces is daardoor slecht controleerbaar. Daarnaast kunnen latente persoonskenmerken die zichtbaar zijn voor medewerkers (zoals naam, adres, geboortedatum) een risico vormen voor (onbewuste) vooringenomenheid. Het anonimiseren van studenten en het exclusief tonen van alleen de relevante kenmerken en de risicoscore zou een mogelijkheid zijn. Nader onderzoek moet uitwijzen welke verbeteringen kunnen worden aangebracht in het handmatige selectieproces.

Aanbeveling 4 – Stel een algoritmemanagementbeleid op om de risico's te verkleinen die de inzet van algoritmes met zich meebrengt.

Stel beleid op ten behoeve van een consistente en verantwoorde omgang met algoritmes en risicoprofilering. Dergelijk beleid kan opgesteld worden in verschillende dimensies. Hierbij kan gedacht worden aan:

- > **Bestuur:** centralisatie van belangrijke besluitvorming in comités, het afstemmen taken en verantwoordelijkheden binnen de bestaande organisatiestructuur, de implementatie van 3 lines-of-defence (3LoD) risicobeheersingsraamwerk.
- > **Documentatie:** het opstellen van gestandaardiseerde documentatievereisten en werkinstructies.
- > **Algoritme-inventarisatie:** centraal overzicht van alle algoritmes³⁷ binnen de organisatie ten behoeve van informatieverzoeken, monitoring en evaluatie.
- > **Processen:** Opstellen van ontwikkelings- en valideringsprocessen voor ieder algoritme.
- > **Toezicht en rapportage:** Toezien op rapportage over algoritmerisico's bevordert risico-georiënteerde werkwijzen.

Voor een andere opdeling kan ook inspiratie worden opgedaan uit bestaande kaders voor verantwoorde omgang met algoritmes in de publieke sector.³⁸

³⁷ Voor een definitie van een algoritme zie: <https://algoritmes.overheid.nl/nl/footer/over-algoritmes>.

³⁸ Zie bijvoorbeeld het Implementatiekader 'Verantwoorde inzet van algoritmes' van de Rijksoverheid <https://www.rijksoverheid.nl/documenten/rapporten/2023/06/30/implementatiekader-verantwoorde-inzet-van-algoritmen>, Onderzoekskader algoritme van de Audit Dienst Rijk (ADR) <https://www.rijksoverheid.nl/documenten/rapporten/2023/07/11/onderzoekskader-algoritmes-adr-2023> en het Toetsingskader algoritmes van de Algemene Rekenkamer <https://www.rekenkamer.nl/onderwerpen/algoritmes-digitaal-toetsingskader>.

Appendix A – Verdere verwijzingen

Algoritmeregister

- > Documentatie 'Slimme bijstandcheck levensonderhoud' <https://algoritmeregister.amsterdam.nl/ai-system/onderzoekswaardigheid-slimme-check-levensonderhoud/1086/>.

Maatschappelijke aandacht voor algoritmes

Verder zijn er verschillende (onderzoeks)journalistieke artikelen verschenen over vooringenomenheid in algoritmes. Onder meer relevant zijn:

- > <https://www.vpro.nl/argos/media/luister/argos-radio/onderwerpen/2021/In-het-vizier-van-het-algoritme-.html>
- > <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- > <https://www.lighthousereports.com/methodology/suspicion-machine/>.

Datawetenschappen

- > Kwalificatie van soorten proxydiscriminatie: Michael Carl Tschantz. 2022. *What is Proxy Discrimination?* In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22). Association for Computing Machinery, New York, NY, USA, 1993–2003. <https://doi.org/10.1145/3531146.3533242>.
- > Spanning tussen AI en Europees non-discriminatierecht: Wachter, Sandra, Brent Mittelstadt, and Chris Russell. "Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI." *Computer Law & Security Review* 41 (2021): 105567.

Appendix B – Datawarehouse query

Een voorbeeld van een SQL-query om voor de peildatum 01-01-2014 de volgende gegevens van uitwonende studenten uit de datawarehouse op te vragen:

- > Persoonsnummer (gepseudoanonymiseerd)
- > Risicofactor
- > Afstand tot ouder(s)
- > Resultaat van de controle (indien aanwezig).

```
select distinct PEILDATUM, PERSOONID, RISICOFACOR, AFSTANDCATEGORIE, RESULTAAT_CONTROLE_VERTAALD as RESULTAAT_CONTROLE
from F_WSF_STUDENT_BASISGEGEVENS_HISTORIE
where I=1
and WOONSITUATIE = 'U'
and GEMEENTECODE is not null
and PEILDATUM in ('2014-01-01 00:00:00.000', '2019-01-01 00:00:00.000')
order by PEILDATUM, PERSOONID
```

Figuur 10 – Voorbeeld van een query uit DUO datawarehouse om persoonsnummer (gepseudoanonymiseerd), risicofactor, afstand tot ouder(s) en resultaat van de controle (indien aanwezig) op te halen.

Appendix C – Aanvullende informatie statistische analyse

Hieronder wordt de Z-toets en Fishers exacte toets toegelicht waarvan gebruik wordt gemaakt in 3.1 Kwantitatieve analyse.

Allereerst wordt gebruik gemaakt van een zogenoemde Z-toets voor het vergelijken van twee percentages. We definiëren de test statistiek Z als volgt:

$$Z = \frac{p_A - p_B}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{N_A} + \frac{1}{N_B}\right)}}, \text{ waar } \hat{p} = \frac{N_{A, \text{ onrechtmatig}} + N_{B, \text{ onrechtmatig}}}{N_A + N_B}.$$

Wanneer de grootte van een steekproef groot genoeg is, en $H_0: p_A = p_B$ van kracht is – dan volgt Z een standaard normale distributie – dit is een normale distributie met een gemiddelde van 0 en variatie van 1.

Vervolgens kan een p-waarde worden afgeleid uit de distributie van een standaard normale distributie. Deze p-waarde vertelt ons wat de kans dat we toevallig p_A en p_B observeren, gegeven dat H_0 waar is. Bij het bereken van deze p-waarde wordt rekening gehouden met de groepsgrootte van A en B. Het significantieniveau van de hypothesetest wordt vastgesteld op 5%, wat betekent dat als de p-waarde kleiner is dan 0.05, de nulhypothese wordt verworpen. Dit betekent dat $p_A > p_B$, oftewel: het onrechtmatig gebruik-percentages voor de verschillende groepen zijn niet gelijk. Als de p-waarde groter is dan 0.05 is er geen aanleiding om te vermoeden dat de onrechtmatig gebruik-percentages verschillen voor de verschillende groepen

In sommige gevallen, wanneer de steekproefgrootte erg klein is, kan het gebruik van de Z-toets tot verkeerde conclusies leiden. Om dit te voorkomen wordt Fishers exacte toets ter controle toegepast. De toets is ook een statistische toets voor het vergelijken van twee percentages, en wordt vaak aangeraden wanneer men werkt met kleine steekproeven. Dit is mogelijk relevant voor een aantal van de vergelijkingen – bijvoorbeeld, in de steekproeven zijn er vaak relatief weinig observaties van bepaalde onderwijsvormen. De uitkomsten van Fisher' exacte toets zijn beschikbaar bij de onderzoekers.

Omdat voor iedere variabele onderwijs, leeftijd en afstand de subcategorieën tegen elkaar worden vergeleken ontstaat het risico van de onterechte detectie van statistische (in)significantie. De Bejamini-Yekutieli procedure kan in dit geval een uitkomst bieden. De resultaten van deze procedure zijn beschikbaar bij de onderzoekers.

Appendix D – Documentlijst

A. L: .../Algorithm Audit/2023-10-10

Referentie	Werknaam	Documentnaam
[A.1]	Woo-besluit 2023	1 - Getekend Woo-besluit 28-6-23 Woo-verzoek Investico
[A.2]	Bijlagen Woo-verzoek 2023	2 - Bijlage 2 Inventarisatielijst Woo-verzoek Investico
[A.3]	Beleid Controle Uitwonendenbeurs	3 - Beleid misbruik uitwonende beurs passages m.b.t. fraude selectie, onderzoek, vaststellen fraude, huis- buurtbezoek
[A.4]	Procesbeschrijving CUB	4 - Procesbeschrijving Uitvoeren controle Misbruik Uitwonendenbeurs
[A.5]	Richtlijnen voor controleurs 2022	5 - Richtlijnen voor controleurs 2022
[A.6]	Werkinstructies afhandelen & verwerken CUB	6 - Werkinstructie WI MUB afhandelen & verwerken
[A.7]	Onderzoeksopzet CUB 2023	7 - Onderzoeksopzet CUB versie 02 28062023 CONCEPT
[A.8]	Opdrachtbevestiging onderzoek privacyaspecten CUB aan de ADR 2023	8 - Opdrachtbevestiging onderzoek Beheersing privacyaspecten proces uitwonendencontrole 21122022
[A.9]	ontbreekt	ontbreekt
[A.10]	Organogram 2023	10 - Organogram OVG 2023 (versie 1.2)
[A.11]	Rapportage CUB 2013	11 - Rapportage MUB 2013 v0.2
[A.12]	Rapportage CUB 2014	12 - Rapportage MUB 2014 v1.0
[A.13]	Rapportage CUB 2015	13 - Rapportage MUB 2015 v1.0
[A.14]	Rapportage CUB 2016	14 - Rapportage MUB 2016 v1.0
[A.15]	Rapportage CUB 2017	15 - Rapportage MUB 2017 v1.0
[A.16]	Kamervragen Spekman 2009	16 - 2009 Kamervragen Spekman cs ah-tk-20082009-3686
[A.17]	Actieplan CUB 2009	17 - 200911 Actieplan misbruik uitwonendenbeurs
[A.18]	Voortgangsrapportage 2018	18 - 20100701 Voortgangsrapportage inzake de uitvoering van het Actieplan misbruik uitwonendenbeurs

Referentie	Werknaam	Documentnaam
[A.19]	Kamerbrief 2012	19 - kamerbrief-over-eindrapportage-actieplan-misbruik-uitwonendenbeurs
[A.20]	Rapportage CUB 2012	20 - Rapportage MUB 2012 v1.1
[A.21]	Eindrapportage Intensivering IHS en CUB 2018	21 - Eindrapportage IHS en Mub 2018 1.0
[A.22]	Avg verwerkingsgrond uitwonendecontrole	22 - ZZZ GGL HenI - uitwonendencontrole 000
[A.23]	Avg verwerkingsgrond procedure uitwonendecontrole	23 - ZZZ GGL HenI – uitwonend procedure 000
[A.24]	Informatieverkenning CUB 2020	24 - 20200629 Informatieverkenning H&I proces MUB
[A.25]	Orderingsplan CUB 2022	25 - Kopie van 20221118 Orderingsplan HI versie 0.4 netto lijst
[A.26]	Datawarehouse overzicht	26 - Datawarehouse SFS
[A.27]	Kopie risicocodering 2023	27 - Kopie van 20230704 Risicocodering
[A.28]	Mogelijkheden invoer resultaat controle	28 - Mogelijkheden invoer resultaat controle
[A.29]	Resultaten controle mogelijkheden	29 - resultaten controle mogelijkheden
[A.30]	Resultaten Den Haag 2 2013	30 - resultaten Den Haag 2
[A.31]	Resultaten Den Haag 3 2013	31 - resultaten Den Haag 3
[A.32]	Resultaten Den Haag 2013	32 - resultaten Den Haag
[A.33]	Mailwisseling uitbreiding risicoprofiel 2023	33 - Risicoprofiel MUB
[A.34]	Overzicht Den Haag	34 - -----
[A.35]	Risicocodering 2021	35 - 20210630 Risicocodering
[A.36]	Snippet datalevering #1	36 - access aangeboden ter controle
[A.37]	Basisbestand Utrecht	37 – basisbestand utrecht
[A.38]	Snippet datalevering #2	38 - basisbestand
[A.39]	Informatie en bron analyse rapport CUB 2015	39 - DSF080 Selectie uitwonende fraude v0.2

Referentie	Werknaam	Documentnaam
[A.40]	Rapportage aselecte steekproef 2014	40 - Rapportage a-selecte steekproef MUP 0.1
[A.41]	MUB-AP	41 – MUP-AP
[A.42]	AP 20230705	42 Notulen 8 juni 2023 en 25 mei 2023
[A.43]	Notulen 2021 en 2022	43 - Notulen 2021 2022, Notulen 20211111
[A.44]	Kalender 2022	44 - MO kalender 2023 concept (geclusterd)
[A.45]	Viermaandse rapportage MUB 2023	45 - 20230612 1ste viermaandsrapportage MUB regulier en MUB HBB tijdvak jan-apr 2023 v05TK
[A.46]	Uitvoeren Controle MUB	46 - Uitvoeren controle misbruik Uitwonendenbeurs (2.0)
[A.47]	Risicocodering 2010	47 - Risicocoderingen misbruik uitwonendenbeurs
[A.48]	Kwalitatieve onderbouwing art.11	48 - Kwalitatieve onderbouwing MO-inventarisatie art 11 HOS 2022
[A.49]	Kwalitatieve onderbouwing art.12	49 - Kwalitatieve onderbouwing MO-inventarisatie art 12 HOS 2022
[A.50]	Kwalitatieve onderbouwing art.13	50 - Kwalitatieve onderbouwing MO-inventarisatie art 13 HOS 2022
[A.51]	Aselecte steekproef 2010 verdeling	51 - A-selecte steekproef def
[A.52]	Aselecte steekproef 2010 verdeling fout	52 - A-selecte steekproef fout def
[A.53]	Sjabloon Beheersdocument	53 - IAR Misbruik Uitwonendenbeurs (MUB) v.1.1

B. L: .../Algorithm Audit/ 2023-10-10/41 - MUB-AP/

Referentie	Werknaam	Documentnaam
[B.41-1]	Properties risicofactoren 2018	-----
[B.41-2]	Informatie Analyse Rapport - Misbruik Uitwonendenbeurs	IAR - Misbruik Uitwonendenbeurs (MUB)
[B.41-3]	Inhoud H SWF	-----
[B.41-4]	Jira HENI fiPF3&	-----
[B.41-5]	Query documentatie	-----

C. L: .../Algorithm Audit/2023-10-10/42 - AP 20230705/

Referentie	Werknaam	Documentnaam
[C.42-1]	Notulen 05-01-23 MUB overleg	20230105 Verslag MUB overleg
[C.42-2]	Notulen 02-02-23 MUB overleg	20230202 Verslag MUB overleg
[C.42-3]	Notulen 15-02-23 MUB overleg	20230215 Verslag MUB overleg
[C.42-4]	Notulen 16-03-23 MUB overleg	20230302 Verslag MUB overleg
[C.42-5]	Notulen 30-03-23 MUB overleg	20230316 Verslag MUB overleg
[C.42-5]	Notulen 30-03-23 MUB overleg	20230323 Verslag MUB overleg
[C.42-6]	Notulen 17-04-23 MUB overleg	20230417 Verslag MUB overleg
[C.42-7]	Notulen 11-05-23 MUB overleg	20230511 Verslag MUB overleg
[C.42-8]	Notulen 25-05-23 MUB overleg	20230525 Verslag MUB overleg
[C.42-9]	Notulen 08-06-23 MUB overleg	20230623 Verslag MUB overleg
[C.42-10]	2021 Leidraad selecteren	2021 leidraad selecteren
[C.42-11]	2021 Werkinstructies	2021 Werkinstructie en actielijst MUB
[C.42-12]	Access zoeken	Access zoeken
[C.42-13]	Woonoppervlakte tool	BAG Viewer woonoppervlakte
[C.42-14]	BRP documenten	BRP ...
[C.42-15]	Aanlevering formulier	FORMAT AANLEVERING NIEUWE STIJL 2023
[C.42-16]	Voorbeeld inschrijving	Studie inschrijving
[C.42-17]	overzicht bestand	WI MUB selecteren versie 0.1

D. L: .../Algorithm Audit/2023-10-10/43 - Notulen 2021 en 2022/

Referentie	Werknaam	Documentnaam
[D.43-1]	Notulen 23-04-20 MUB overleg	Notulen 23-04-20 MUB overleg
[D.43-2]	Notulen 16-09-21 MUB overleg	Notulen 16-09-21 MUB overleg
[D.43-3]	Notulen 30-09-21 MUB overleg	Notulen 30-09-21 MUB overleg
[D.43-4]	Notulen 14-10-21 MUB overleg	Notulen 14-10-21 MUB overleg
[D.43-5]	Notulen 28-10-21 MUB overleg	Notulen 28-10-21 MUB overleg
[D.43-6]	Notulen 11-11-21 MUB overleg	Notulen 11-11-21 MUB overleg
[D.43-7]	Notulen 25-11-21 MUB overleg	Notulen 25-11-21 MUB overleg
[D.43-8]	Notulen 09-12-21 MUB overleg	Notulen 09-12-21 MUB overleg

E. L: .../Algorithm Audit/2023-10-18/

Referentie	Werknaam	Documentnaam
[E.1]	Mail agenda stuurgroep 22-02-2010	Agenda Stuurgroep uitwonendenbeurs 22022010.doc
[E.1-1]	Notitie Risicoprofielen 22-02-2010	DUO-notitie Risicoprofielen
[E.1-2]	Agenda stuurgroep 22-02-2010	Agenda Stuurgroep uitwonendenbeurs 2202010.doc
[E.1-3]	Besluitenlijst Stuurgroep Misbruik Uitwonendenbeurs 25-01-2010	Concept Besluitenlijst Stuurgroep Misbruik Uitwonendenbeurs 25 jan 2010.doc
[E.1-4]	DUO Opsporing & Handhaving	DUO-opsporing&handhaving.doc
[E.2]	Nota Misbruik controlevolume 2015 en jaarrapportage 2013	-----
[E.3]	Opdrachtbrief uitvoering wetsvoorstel Studievoorschot	-----
[E.4]	Verslag uitvoeringsoverleg 08-07-2016	-----
[E.5]	Mail stuurgroep vergadering MUB 15-03-2010	FW Stuurgroep vergadering 'misbruik uitwonendenbeurs' 15 maart 2010

Referentie	Werknaam	Documentnaam
[E.6]	Mail onderbouwing uitvoering- skosten Studievoorschot	Onderbouwing uitvoeringskosten Studievoor- schot
[E.7]	Agenda, actie- en besluitenlijst uit- voeringsoverleg 08-07-2016	-----
[E.8]	Mail stuurgroep MUB 21-01-2010	Stuurgroep Misbruik Uitwonendenbeurs
[E.8-1]	Agenda Stuurgroep misbruik uit- wonendenbeurs 25-01-2010	Agenda Stuurgroep uitwonendenbeurs 25012010.doc
[E.8-2]	Besluitenlijst Stuurgroep 14-12- 2009	Besluitenlijst Stuurgroep dd 14122009.doc
[E.8-3]	Communicatieplan Misbruik Uit- wonenden Beurs 2010	Communicatieplan fraude berus uitwonen- den_0.2.doc
[E.8-4]	Notitie afspraken over convenanten voor pilots uitwonendenbeurs	Notitie Stuurgroep 25012010 convenanten.doc
[E.8-5]	Raming kosten Misbruik Uitwonen- den Beurs	Raming Project Misbruik Uitwonenden- beurs-21012010.doc
[E.8-6]	Notitie ter voorbereiding voor overleg met BZK (toezichthouder en adresadministratie)	Notitie ter voorbereiding voor overleg met BZK
[E.8-7]	Planning Stuurgroepagenda	Agenda- planning Stuurgroep.doc
[E.9]	Mail stuurgroep MUB 24-06-2010	Stuurgroepvergadering Project Misbruik Uit- wonendenbeurs 28 juni
[E.9-1]	Agenda Stuurgroep MUB 28-06- 2010	Agenda Stuurgroep 28.06.10.doc
[E.9-2]	Besluitenlijst Stuurgroep 19-04- 2010	Besluitenlijst Stuurgroep Misbruik uitwonenden- beurs 19 april 2010
[E.9-3]	Concept artikelen wetsvoorstel aan- pak MUB 24-06-2010	1759-artikelen-24-6-10.doc
[E.9-4]	Overdrachtsdocument 22-06-2010	Overdrachtsdocument 22.06.10.doc
[E.9-5]	Nota voortgangsrapportage Actie- plan uitwonendenbeurs 2010	Ter ondertekening 15.06.10.doc
[E.9-6]	Voortgangsrapportage Actieplan uitwonendenbeurs 2010	Voortgangsrapportage TK 24 06 10 (incl WJZ+DUO+HB) (laatste versie) (2) (2).doc

F: L: .../Algorithm Audit/2023-11-1/

Referentie	Werknaam	Documentnaam
[F.1]	Formulier Uitwonendentoelage 2006	6 formulier uitwonendentoelage 2006
[F.2]	Agenda Workshop Misbruik Uitwonendenbeurs	Agenda Workshop misbruik uitwonendenbeurs
[F.3]	Brief Keupink	Brief melding fraude Keupink
[F.4]	Controlebeleid IB-Groep	Controlebeleid uitwonendheid IB-Groep
[F.5]	PowerPoint Workshop maart 2010	DUO Workshop uitwonendheid 25032010
[F.6]	Gegevens voor ontwikkeling risicoprofiel	Gegevens voor ontwikkeling risicoprofiel studerende met een uitwonende beurs
[F.7]	Nota gesprek 13 november 2007	Nota gesprek over fraude 13 november
[F.8]	Notitie Renk	Notitie 06112007 uitwonendencontrole
[F.9]	Proces uitwisseling gemeenten	Proces uitwisseling gemeenten
[F.10]	Verslag workshop 25 maart 2010	Verslag workshop misbruik uitwonendenbeurs 25 maart 2010
[F.11]	Voorgangsrapportage 28 juni 2010	Voortgangsrapportage 28.06.10 (eindversie)

G: L: .../Algorithm Audit/Gegevenslevering/Werkplan intern DUO/

Referentie	Werknaam	Documentnaam
[G.1]	Steekproef 2014	Tabel 1 – Aselect 2014
[G.2]	Steekproef 2017	Tabel 3 – Aselecte 2017
[G.3]	CUB-2014	Tabel 1 – CUB 2014
[G.4]	CUB-2019	Tabel 2 – CUB 2019

H: Losse overige stukken ontvangen na 14 december 2023

Referentie	Werknaam	Documentnaam
[H.1]	Inventarisatie CUB werkwijze	Inventarisatie CUB werkwijze versie 1.0



www.algorithmaudit.eu



www.github.com/NGO-Algorithm-Audit



info@algorithmaudit.eu

Stichting Algorithm Audit is geregistreerd bij de Kamer
van Koophandel onder nummer 83979212