

# Herijking risico-indicator onderwijsachterstanden

basisonderwijs - fase 2

2026





Centraal Bureau  
voor de Statistiek

**adres** CBS Den Haag  
Henri Faasdreef 312  
2492 JP Den Haag  
Postbus 24500  
2490 HA Den Haag  
T +31 70 337 38 00

**datum** 11-6-2026

**auteur** CBS

# Inhoudsopgave

<b>Over deze publicatie</b>	<b>5</b>
<b>1. Inleiding</b>	<b>6</b>
1.1. Leeswijzer . . . . .	7
<b>2. Theoretisch kader risico op onderwijsachterstanden</b>	<b>8</b>
2.1. Wat is een onderwijsachterstand? . . . . .	8
2.2. Conceptueel model . . . . .	9
2.3. Meten van leerprestaties . . . . .	9
2.4. Kindkenmerken . . . . .	10
2.4.1. Intelligentie en invloeden op de meting daarvan . . . . .	10
2.4.2. Niet-cognitieve kenmerken van het kind en leerprestaties . . . . .	12
2.5. Gezinsomgeving . . . . .	12
2.5.1. Stress en risico's . . . . .	12
2.5.2. Verschillende vormen van kapitaal . . . . .	13
2.6. Schoolomgeving . . . . .	14
2.7. Buurtomgeving . . . . .	14
2.8. Implicaties voor modelontwikkeling . . . . .	15
<b>3. Data</b>	<b>16</b>
3.1. Databronnen . . . . .	16
3.2. Populatieafbakening . . . . .	16
3.3. Gebruikte achtergrondkenmerken en operationalisaties . . . . .	17
<b>4. Keuzes bij modelontwikkeling</b>	<b>18</b>
4.1. Het gebruik van een intelligentiemeting . . . . .	18
4.2. Omgang met categoriale variabelen . . . . .	21
4.3. Meting van onderwijsprestaties . . . . .	21
4.4. Modellen met meerdere groepen . . . . .	23
4.5. Omgang met ontbrekende waarden . . . . .	25
<b>5. Modelontwikkeling</b>	<b>27</b>
5.1. Bivariate analyses . . . . .	27
5.2. Stapsgewijze procedure . . . . .	30
5.3. Een multigroep-model . . . . .	31
5.4. Een beperkt model . . . . .	35
<b>6. Plausibiliteitsanalyses</b>	<b>38</b>
6.1. Kruisvalidatie . . . . .	38
6.2. LASSO . . . . .	39
6.2.1. Analyse . . . . .	39
6.2.2. Resultaten . . . . .	40
6.3. Vergelijking met een ander cohort . . . . .	41
<b>7. Herverdeeleffecten</b>	<b>45</b>
7.1. Verschillen tussen de herijkte en huidige methode . . . . .	45

7.2. Herverdeeleffecten op schoolniveau . . . . .	46
7.2.1. Verschillen in financieringsstatus tussen de methoden . . . . .	46
7.2.2. Verdeling van verschillen in ontvangen percentage van budget op schoolniveau	46
7.3. Herverdeeleffecten voor gemeenten . . . . .	48
7.3.1. Verdeling van verschillen in ontvangen percentage van budget op gemeen-	
teniveau . . . . .	48
7.4. Verandering in percentage van het ontvangen budget per gemeente . . . . .	49
7.5. Geografische verdeling van herverdeeleffecten . . . . .	50
7.6. Conclusie . . . . .	51
<b>8. Samenvatting en conclusie</b>	<b>52</b>
<b>Referenties</b>	<b>54</b>
<b>Bijlagen</b>	<b>59</b>
<b>A. Operationalisering gebruikte achtergrondkenmerken</b>	<b>59</b>
A.0.1. Kenmerken van kind en gezin . . . . .	59
A.0.2. Kenmerken van school en buurt . . . . .	66
<b>B. Begeleidingscommissie</b>	<b>67</b>

## Over deze publicatie

Om onderwijsachterstanden in het primair onderwijs (po) te verminderen kent het ministerie van Onderwijs, Cultuur en Wetenschap (OCW) extra middelen toe aan scholen en gemeenten. Het ministerie verdeelt deze middelen aan de hand van de door het Centraal Bureau voor de Statistiek (CBS) ontwikkelde risico-indicator onderwijsachterstanden. Op verzoek van het ministerie doet het CBS onderzoek naar het actualiseren van het model dat voor deze indicator wordt gebruikt.

Deze publicatie bouwt voort op de vooronderzoeken uit fase 1, waarin onder meer de imputatiemethoden, de verwerking van eindtoetsscores en relevante achtergrondkenmerken zijn onderzocht. In fase 2 staat de herziening van het model voor de risico-indicator onderwijsachterstanden centraal. Daarbij wordt ingegaan op het theoretisch kader van onderwijsachterstanden, de selectie van beschikbare achtergrondkenmerken, de ontwikkeling van het analysemodel, de resultaten van plausibiliteitsanalyses en de herverdeeleffecten van de herijkte methode ten opzichte van de huidige indicator.

# 1. Inleiding

Al jarenlang staat het bevorderen van gelijke kansen hoog op de agenda van beleidsmakers, politici en wetenschappers, waarmee dit een belangrijk en relevant thema is in het publieke debat in Nederland. Het uitgangspunt is daarbij veelal dat het niet mag uitmaken waar je bent geboren, in welk gezin je opgroeit of wie je ouders zijn.

Sinds de jaren zeventig voert het ministerie van Onderwijs, Cultuur en Wetenschap (OCW) beleid om de onderwijskansen van kinderen afkomstig uit achterstands- en kwetsbare milieus te bevorderen. Het beleid is daarbij gericht op het voorkomen van onderwijsachterstanden bij leerlingen in het basisonderwijs die het gevolg zijn van sociale, economische of culturele factoren, en op het verminderen van reeds opgelopen achterstanden. De gewichtenregeling en de impulsregeling waren in het verleden belangrijke onderdelen van het beleid ter bestrijding van onderwijsachterstanden. Deze regelingen bepaalden de omvang van de extra financiële middelen die een basisschool ontving voor het aanpakken van onderwijsachterstanden.

In de afgelopen decennia is het beleid verschillende keren van vorm veranderd. Initieel diende het beroep van de vader als indicator voor de achterstanden. Het beleid richtte zich destijds voornamelijk op het extra stimuleren van 'arbeiderskinderen'. Aansluitend werden ook de kenmerken 'opleidingsniveau van de ouders' en 'herkomst' integraal onderdeel van het beleid.

De gewichtenregeling zoals die in zijn laatste vorm tot 2018 van toepassing was, was volledig gebaseerd op het opleidingsniveau van de ouders [Kloprogge and de Wit, 2015]. Scholen moesten voor deze gewichtenregeling aan ouders vragen wat hun behaalde opleidingsniveau was. Daarbij werden twee gewichten onderscheiden, waarbij hogere gewichten correspondeerden met lagere opleidingsniveaus van de ouders. Als beide ouders van leerlingen geen diploma in het voortgezet onderwijs bezitten, kregen scholen 1,2 leerling extra gefinancierd. Als dat gold voor één ouder en de andere ouder heeft maximaal vmbo, dan kregen scholen 0,3 leerling extra gefinancierd. Alle overige leerlingen kregen geen extra financiering vanuit het onderwijsachterstandenbeleid. Daarnaast was de som van de gewichten in een gemeente de basis voor de verdeling van het gemeentelijke onderwijsachterstandenbudget. Scholen die in een postcodegebied lagen waar het gemiddelde inkomensniveau laag is en de werkloosheid hoog (de zogenaamde impulsgebieden), kregen daarnaast voor elke gewichtenleerling extra budget toegekend.

Het bepalen van mogelijke achterstanden aan de hand van het opleidingsniveau van ouders riep steeds meer vragen op. Is uitsluitend het opleidingsniveau van ouders de meest geschikte indicator voor onderwijsachterstanden, of zijn er aanvullende of alternatieve kenmerken die eveneens een significante rol spelen? Zorgt het toenemende opleidingsniveau van ouders mogelijk onterecht voor een vermindering van het budget voor het aanpakken van onderwijsachterstanden? Is de regeling niet onnodig tijdrovend en kostbaar in de uitvoering? De verantwoordelijkheid voor de administratie van de opleidingsniveaus van ouders en de bijbehorende vertaling naar gewichten berustte destijds bij de scholen. Dit bleek voor hen behoorlijk lastig en tijdrovend te zijn. Voor het ministerie waren er eveneens hoge kosten verbonden aan de controle, aangezien deze voor scholen met een groot aantal gewichtenleerlingen arbeidsintensief is.

Tegen deze achtergrond van hoge administratieve lasten voor scholen en hoge kosten voor het ministerie besloot OCW om het beleid te herzien. Om het onderwijsachterstandenbeleid te kunnen verbeteren, heeft OCW het Centraal Bureau voor de Statistiek (CBS) gevraagd om op basis van kenmerken die in registraties bij het CBS aanwezig zijn een indicator samen te stellen die onderwijsachterstanden in het basisonderwijs beter voorspelt dan enkel het opleidingsniveau van ouders.

Het CBS heeft verschillende onderzoeken uitgevoerd om een nieuwe schatting van het risico op onderwijsachterstanden mogelijk te maken [CBS, 2016, 2017a,b, 2019a]. Sinds 2018 maakt OCW gebruik van deze zogenaamde risico-indicator onderwijsachterstanden. De systematiek achter deze indicator is beschreven in een samenvattend rapport [CBS, 2019b]. Door het CBS wordt met deze indicator een onderwijsscore berekend voor alle peuters van 2,5 tot 4 jaar en voor alle basisschoolleerlingen, waarbij deze scores worden opgeteld tot achterstandsscores per school en per gemeente. Deze drukken dan de verwachte onderwijsachterstandsproblematiek op scholen en in gemeenten uit, op basis waarvan OCW het onderwijsachterstandenbudget over de scholen en gemeenten verdeeld.

Om de risico-indicator onderwijsachterstanden actueel te houden, heeft OCW aangegeven de indicator regelmatig te willen evalueren. Hierdoor kunnen nieuwe ontwikkelingen mee worden genomen in de indicator. De eerste evaluatie is uitgevoerd in 2021 en betrof een beperkte herziening waarbij vooral is gekeken naar de actualisatie van de coëfficiënten en schaalwaarden van het model [CBS, 2021]. De resultaten van deze herziening zijn in 2022 voor het eerst toegepast bij het berekenen van de achterstandsscores voor scholen.

In 2023 is het CBS op verzoek van OCW gestart met een grote herziening van de risico-indicator onderwijsachterstanden. Deze herziening bestaat uit een drietal fases. In de eerste fase stonden een aantal voorbereidende onderzoeken voor het nieuw te schatten model voor onderwijsachterstanden centraal [CBS, 2024b]. De tweede fase stond in het teken van dit nieuw te schatten model voor onderwijsachterstanden. Hierbij werden de bevindingen uit de eerste fase verwerkt in het nieuw te schatten model (dit rapport). In de derde en laatste fase zullen de uitkomsten van de eerste en tweede fase verwerkt worden in een gestandaardiseerde productiestatistiek.

In deze rapportage bespreken we het nieuwe model om het risico op onderwijsachterstanden te schatten. Hierbij wordt onderzocht welke variabelen de onderwijskansen van kinderen kunnen beïnvloeden en hoe goed met verschillende (combinaties van) achtergrondkenmerken uit CBS-registraties het risico op onderwijsachterstand daadwerkelijk kan worden ingeschat. Hoewel de meeste registraties van het CBS betrekking hebben op de gehele bevolking, is de registratie van de hoogst behaalde opleidingsniveaus niet volledig. Aangezien het opleidingsniveau van ouders naar verwachting een belangrijk kenmerk is bij het inschatten van het risico op onderwijsachterstand, hebben we eerder uitgebreid aandacht besteed aan mogelijkheden voor het voorspellen en aanvullen van ontbrekende waarden [CBS, 2016, 2024b].

## 1.1. Leeswijzer

In hoofdstuk 2 schetsen we eerst het theoretisch kader met betrekking tot het risico op onderwijsachterstand. Vervolgens beschrijven we in hoofdstuk 3 de data die in dit onderzoek hebben gebruikt en in hoofdstuk 4 welke keuzes we bij de modelontwikkeling hebben gemaakt. De resultaten van de analyses beschrijven we in hoofdstuk 5 (modelontwikkeling) en 6 (plausibiliteitsanalyses). De herverdeeffecten – hoe de uitkomsten van het nieuwe model zich verhouden tot het huidige model – beschrijven we in hoofdstuk 7. We sluiten af met een samenvatting en conclusie.

## 2. Theoretisch kader risico op onderwijssachterstanden

In dit hoofdstuk inventariseren we welke kind-, gezins-, school- en buurtkenmerken volgens eerder onderzoek mogelijk van invloed zijn op leerprestaties en daardoor een indicatie kunnen geven van mogelijke risico's op onderwijssachterstanden. Om relevante omgevingskenmerken te identificeren, maken we gebruik van een mix van typen onderzoeksliteratuur: review- en overzichtsstudies [o.a. [Badou and Day, 2021](#), [CBS, 2024a](#), [Herweijer et al., 2021](#), [Kloprogge and de Wit, 2015](#), [Leseman, 2019](#), [Sauce and Matzel, 2018](#)], empirisch en experimenteel onderzoek [o.a. [Bluemink et al., 2026](#), [Das and Nollet, 2026](#), [Von Stumm and Plomin, 2015](#), [ter Weel et al., 2025](#)] en een expertraadpleging [[CBS, 2024b](#)].

We beginnen echter met een beschrijving van wat we onder een onderwijssachterstand verstaan. Vervolgens geven we een kort overzicht van het conceptuele model aan de hand waarvan we het theoretisch kader beschrijven. Daarna beschrijven we hoe we leerprestaties hebben gemeten alvorens we ingaan op diverse kind- en omgevingskenmerken. Daarbij gaan we ook in op de invloed van de individuele potentie van leerlingen. Omgevingskenmerken verdelen we onder in kenmerken van het gezin, de school en de buurt.

### 2.1. Wat is een onderwijssachterstand?

Onderwijssachterstanden kunnen op verschillende manieren worden gedefinieerd. In navolging van [Kloprogge and de Wit \[2015\]](#) spreekt het CBS van een onderwijssachterstand als leerlingen door ongunstige economische, sociale of culturele omgevingskenmerken slechter presteren op school dan ze bij een gunstigere situatie zouden kunnen. Lage prestaties of een zorgleerling zijn, zijn daarin niet hetzelfde als een onderwijssachterstand. Onderpresteren kan op alle onderwijsniveaus voorkomen en is daarmee niet gelijk aan lage leerprestaties. Het zijn van een zorgleerling is veel meer te verklaren uit een reeks kenmerken die betrekking hebben op het kind zelf, zoals gedragsmatige stoornissen of tekortkomingen op het gebied van intelligentie of aandacht, terwijl een achterstandssituatie in de eerste plaats te maken heeft met de specifieke omgeving waarin het kind zich bevindt.

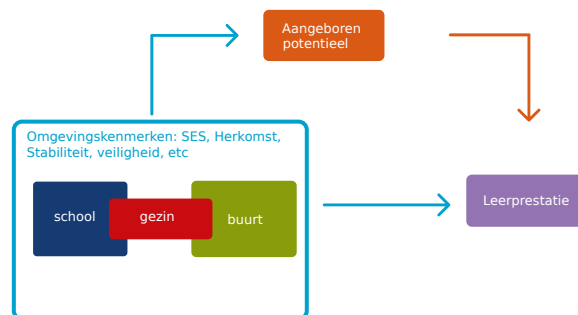
Wat het bespreken en onderzoeken van het risico op onderwijssachterstand lastig maakt, is dat het opgroeien in een omgeving (gezin, school en/of buurt) met ongunstige kenmerken niet automatisch resulteert in een onderwijssachterstand [[Kloprogge and de Wit, 2015](#)]. Niet elk kind uit een minder gunstige omgeving presteert minder dan wat het kind zou kunnen. Het opgroeien in een omgeving met ongunstige kenmerken zorgt er vooral voor dat de kans op een succesvolle schoolcarrière kleiner wordt waardoor kinderen niet op het niveau kunnen komen dat aansluit bij hun potentie.

## 2.2. Conceptueel model

Het risico op onderwijsachterstand wordt dus bepaald door de leerprestatie, potentie en eigenschappen van een kind en kenmerken van de omgeving van het kind. Een onderwijsachterstand is niet direct te meten, maar wordt bepaald door de samenhang van al deze factoren.

In Figuur 2.1 geven we een schematische weergave van de relaties tussen mogelijk voorspellende variabelen van leerprestaties die we in dit hoofdstuk aan de orde laten komen aan de hand van eerdere literatuur. Veel van de kenmerken in deze schematische weergave vertonen onderlinge samenhang. Zo kan bijvoorbeeld de sociaaleconomische status van een gezin op verschillende manieren gerelateerd zijn aan andere achtergrondkenmerken zoals stabiliteit in het gezin of de sociaaleconomische status van de buurt. Het doel van dit onderzoek is niet om de onderlinge relaties tussen alle kenmerken te analyseren, maar om specifiek de relaties tussen deze kenmerken en de leerprestaties van basisschoolleerlingen te onderzoeken. In de uiteindelijke analyses houden we rekening met deze kenmerken door hun invloed op leerprestaties te corrigeren voor de effecten van andere relevante kenmerken.

Figuur 2.1. Conceptueel model effecten op leerprestaties



In de definitie die we voor onderwijsachterstanden gebruiken, staat de invloed van omgevingskenmerken op de leerprestaties centraal. Daarom zullen we eerst ingaan op de wijze waarop we leerprestatie hebben geoperationaliseerd. Voordat we ingaan op de invloed van achtergrondkenmerken van de sociaaleconomische omgeving van een kind, beschrijven we eerst de rol van het aangeboren potentieel van het kind.

## 2.3. Meten van leerprestaties

In het Nederlandse onderwijs worden verschillende toetsen afgenomen om de prestaties en vorderingen van leerlingen in kaart te brengen. Op de meeste basisscholen in Nederland wordt gewerkt met een leerlingvolgsysteem (LVS) om jaarlijks de resultaten van de LVS-toetsen vast te leggen. Daarnaast wordt ieder jaar in groep 8 de eindtoets afgenomen. Deze eindtoets wordt gebruikt om het advies voor vervolgonderwijs vast te stellen. Voor het ontwikkelen van een model dat het risico op onderwijsachterstand beschrijft, is in het verleden vanwege de betrouwbaarheid gekozen voor de eindtoets als maatstaf voor leerprestaties [CBS, 2016]. We zien momenteel geen aanleiding om af te wijken van deze keuze.

Voor de huidige indicator is de Cito Eindtoets gebruikt omdat die door veruit de meeste scholen werd gebruikt [CBS, 2016]. Tegenwoordig maken niet alle kinderen dezelfde eindtoets<sup>1</sup>; iedere school bepaalt zelf welke eindtoets wordt afgenomen. De verschillende eindtoetsen hebben echter niet dezelfde uitkomstmaat. In hoofdstuk 4 beschrijven we op welke wijze wij daarmee omgaan bij de ontwikkeling van het model.

## 2.4. Kindkenmerken

In de eerste plaats worden leerprestaties beïnvloed door de eigenschappen van het kind zelf. Hierbij kan onderscheid gemaakt worden tussen cognitieve en niet-cognitieve (sociaalemotionele) vaardigheden [Boman, 2023, Demange et al., 2022, van Hoorn et al., 2005, Krapohl et al., 2014, Malanchini et al., 2024, Resing and Blok, 2002, Veen et al., 2013]. De cognitieve potentie wordt doorgaans gemeten met behulp van intelligentietesten.

### 2.4.1. Intelligentie en invloeden op de meting daarvan

Volgens Resing and Drenth [2007] kan intelligentie worden omschreven als het geheel van cognitieve of verstandelijke vermogens dat nodig is om kennis te verwerven en daar op een goede wijze gebruik van te maken, teneinde problemen op te lossen die een vast omschreven doel en structuur hebben. Intelligentie wordt gemeten met behulp van een intelligentietest en doorgaans uitgedrukt in intelligentiequotiënt (IQ). Theoretisch gezien kan met de gemeten potentie van een kind, de invloed van achtergrondkenmerken op de leerprestatie worden onderscheiden en daarmee dus ook het risico op het oplopen van onderwijsachterstanden. Het gebruik van intelligentie als maatstaf voor individuele potentie is om meerdere redenen echter problematisch.

Ten eerste is IQ geen statisch gegeven, maar een dynamisch concept dat zich ontwikkelt over tijd en onder invloed van de omgeving. De cognitieve ontwikkeling van de mens is leeftijd-gerelateerd. Internationaal onderzoek laat zien dat intelligentie van kinderen toeneemt met leeftijd [Haworth et al., 2010]. Hoewel de meeste studies op een hoge mate aan erfelijkheid van IQ duiden, is er geen eenduidig antwoord te geven over de mate waarin dit het geval is. Zo laat tweelingonderzoek zien dat erfelijkheid een groot deel van de variantie in intelligentiescores [tot ca 80%: Hanscombe et al., 2012, Hart et al., 2007] of eindtoetsscores verklaart [ca 60%: de Zeeuw et al., 2016]. Daarentegen laten Genetic Wide Association Studies (GWAS) soms echter veel lagere percentages zien van tussen de 20 en 30 procent [Howe et al., 2022]. De ontwikkeling van intelligentie wordt echter ook beïnvloed door de omgeving. Sauce and Matzel [2018] beschrijven verschillende fenomenen die aantonen dat intelligentie ondanks de hoge mate aan erfelijkheid ook sterk wordt beïnvloed door de omgeving en dus kneedbaar is. Zo laten kinderen die vanuit een arm milieu geadopteerd worden in een welvarend gezin vaak een flinke stijging in IQ zien [Van Ijzendoorn et al., 2005]. Ook lijkt in welvarende gezinnen erfelijkheid een veel grotere rol te spelen terwijl in armere gezinnen juist de gedeelde omgeving een grote rol speelt [o.a. Turkheimer et al., 2003]. Verder kan volgens Sauce and Matzel [2018] de wereldwijde stijging van het IQ niet door genetische verandering worden verklaard, maar is dat vooral een gevolg van maatschappelijke ontwikkelingen en interacties tussen genen en omgeving in een dynamisch proces [Dickens and Flynn, 2001, Dickens, 2005].

Ten tweede blijkt uit eerder onderzoek dat groepen met een niet-Nederlandse herkomst gemiddeld lager scoren op gestandaardiseerde IQ-testen [Resing et al., 1986]. Dit kan worden

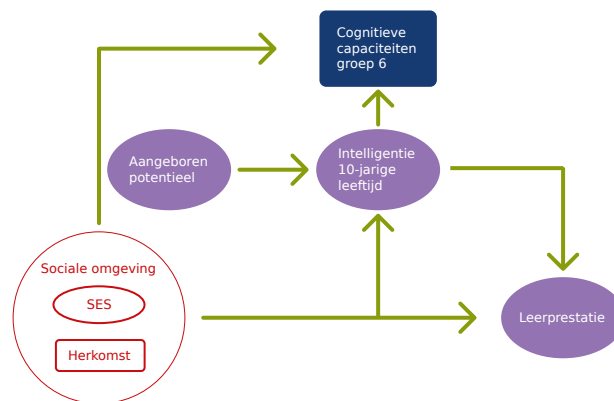
---

<sup>1</sup>De eindtoets is met ingang van schooljaar 2023/'24 vervangen voor de doorstroomtoets

verklaard aan de hand van een aantal verschijnselen. Zo kunnen verschillen in scores tussen bevolkingsgroepen te maken hebben met de gevoeligheid van de vraagstelling in de IQ-tests voor culturele verschillen [Resing and Drenth, 2007]. Variaties in scores tussen verschillende bevolkingsgroepen kunnen verder gevonden worden in accentverschillen in scholing (waarbij specifieke facetten van intelligentie meer geoefend worden) of in factoren zoals cultureel beïnvloede ijver [Kooistra et al., 2008]. De lagere IQ-scores onder minderheidsgroepen kunnen ook te maken hebben met een algeheel lager opleidingsniveau als gevolg van lagere opleidingsniveaus in het land van herkomst waardoor het versterkingseffect zoals beschreven door Sauce and Matzel [2018] minder effect heeft gehad. Daarnaast kunnen de intelligentiescores van minderheidsgroepen ook lager uitvallen als er sprake is geweest van selectieve migratie.

Uit bovenstaande is duidelijk geworden dat intelligentie niet een pure erfelijkheidskwestie is, maar dat de omgeving ook invloed heeft op de mate waarin het aangeboren potentieel tot ontwikkeling komt. Leerprestaties op school zijn hier een uiting van. In Figuur 2.2 hebben we de relaties weergegeven tussen leerprestaties, aangeboren potentieel, de gemeten intelligentie en de invloed van de sociaaleconomische omgeving. Intelligentie is hierin een zogenaamde latente (niet direct waargenomen) variabele waarvoor de uitkomst van een IQ-test een proxy variabele is.

Figuur 2.2. Conceptuele weergave invloeden van sociale omgeving en leeftijd op intelligentie en leerprestaties



Deze kennis is ook van invloed op de wijze waarop een model voor het risico op onderwijsachterstand moet worden geschat. Een belangrijk punt daarbij is dat, zoals hierboven duidelijk is geworden, intelligentie onmogelijk kan worden gemeten zonder dat deze meting is beïnvloed door de sociaaleconomische omgeving van het kind. Bij ontwikkeling van het model voor de huidige risico-indicator onderwijsachterstanden is er op basis van de toenmalige inzichten gekozen om alleen het indirecte effect van het opleidingsniveau van de ouders via de intelligentie op leerprestaties mee te nemen in de modelontwikkeling (de pijl in Figuur 2.2 die van sociale omgeving loopt naar intelligentie). Vervolgens is er bij de toepassing van het model voor gekozen om alleen de directe effecten op de leerprestatie te gebruiken. Volgens de huidige inzichten doet die aanpak onvoldoende recht aan de complexiteit van de samenhang tussen leerprestatie, potentie en omgeving. Het huidige inzicht is dan ook dat het beter is om de indicator te baseren op de totale effecten (direct + indirect) van de omgevingskenmerken. Omdat we alleen geïnteresseerd zijn in de totale effecten van de omgevingskenmerken, is het in principe ook niet meer noodzakelijk om intelligentie mee te nemen in de modelontwikkeling.

In hoofdstuk 4 en verder werken we verder uit hoe we de huidige inzichten in de modelontwikkeling hebben meegenomen. Om inzicht te kunnen geven over het effect van de keuze over het wel of niet meenemen van intelligentie in de modelontwikkeling, vergelijken we

daarbij modellen met en zonder het gebruik van intelligentie als latente variabele bij het schatten van het risico op onderwijsachterstand.

## 2.4.2. Niet-cognitieve kenmerken van het kind en leerprestaties

Verschillende onderzoekers onderschrijven dat voor een succesvolle schoolloopbaan niet alleen cognitieve capaciteiten van belang zijn, maar ook niet-cognitieve (sociaalemotionele) vaardigheden van belang zijn [Boman, 2023, Demange et al., 2022, van Hoorn et al., 2005, Krapohl et al., 2014, Malanchini et al., 2024, Veen et al., 2013]. Niet-cognitieve vaardigheden als doorzettingsvermogen, zelfbeheersing en emotieregulatie, het geloof in eigen kunnen zijn van invloed op de mate waarin het aanwezige potentieel wordt benut [Boman, 2023, Demange et al., 2022, Malanchini et al., 2024]. Niet-cognitieve vaardigheden lijken na intelligentie de belangrijkste verklaring voor variatie in leerprestatie, nog voor de invloed van de sociaaleconomische achtergrond [Boman, 2023, Malanchini et al., 2024]. De invloed van deze vaardigheden lijkt ook toe te nemen met de leeftijd [Malanchini et al., 2024]. Dergelijke niet-cognitieve vaardigheden kunnen werken als versterker van cognitieve vaardigheden, maar ook als compensatiemechanisme voor een gemiddelde of lagere intelligentie [Boman, 2023]. Ook hier geldt weer dat de omgeving die ouders creëren voor hun kinderen van invloed is op de ontwikkeling van deze vaardigheden [Demange et al., 2022]. Helaas beschikt het CBS niet over gegevens van niet-cognitieve capaciteiten van leerlingen in het basisonderwijs.

## 2.5. Gezinsomgeving

Uit eerder onderzoek is bekend dat naast individuele kenmerken, zoals de intelligentie van het kind, kenmerken van het gezin van invloed zijn op leerprestaties [voor overzichtsstudies zie o.a.: Badou and Day, 2021, Herweijer et al., 2021, Kloprogge and de Wit, 2015, Leseman, 2019, Mulder et al., 2014]. In deze paragraaf bespreken we gezinskenmerken die in eerder onderzoek gerelateerd worden aan leerprestaties. Hierbij maken we onderscheid tussen enerzijds stress- en risicofactoren en anderzijds ondersteunings- en investeringmogelijkheden en de invloed van deze factoren op de ontwikkeling en het welzijn van het kind.

### 2.5.1. Stress en risico's

In gezinnen waar stress en risico's aanwezig zijn, ligt minder nadruk op en is er minder ruimte voor stimulerende activiteiten en een opvoedprincipe als 'informeel leren' [Mesman, 2011]. Leseman [2019] onderscheidt risico's op drie niveaus die de stress in een gezin verhogen: (1) factoren op kindniveau zoals een laag geboortegewicht, lage cognitieve vaardigheden of een lastig temperament; (2) factoren op gezinsniveau zoals armoede, relationele problemen of ingrijpende levensgebeurtenissen zoals het overlijden van een gezinslid en (3) factoren op buurniveau zoals onveiligheid, armoede of vervuiling. Het *family stress model* van Conger and Donnellan [2007] is daarbij een handzaam theoretisch kader voor de verklaring van de invloed van sociaaleconomische achterstand en economische tegenspoed op de ontwikkeling en het welzijn van kinderen, waarbij dit effect tot stand komt via een proces van gezinsstress. De impact van armoede en tegenspoed op kinderen verloopt volgens dat kader via een causale keten van indirecte mechanismen. Het proces begint bij economische achterstand en tegenspoed, gekenmerkt door factoren als een laag inkomen, schulden of negatieve gebeurtenissen zoals baanverlies of een scheiding [CBS, 2015, de Lange and Dronkers, 2018, Mooi-Reci et al., 2019, Zumbuehl and Dillingh, 2020]. Deze economische druk verhoogt vervolgens het risico op

emotionele problemen bij de ouders zoals depressie, gedragsproblemen, middelengebruik of vervreemding. Direct (als gevolg van de economische tegenslag) of indirect (als gevolg van emotionele problemen) kunnen door economische tegenslag vervolgens conflicten ontstaan in de relationele sfeer. Het family stress model stelt dat ouders die afgeleid worden door hun eigen stress minder betrokken zijn bij de dagelijkse activiteiten van het kind, mogelijk minder genegenheid voor hun kind tonen en vaker geïrriteerd, hard of onvoorspelbaar reageren. Uiteindelijk brengt dit proces de gezonde ontwikkeling van het kind in gevaar. Dit uit zich in een lager welbevinden, lagere cognitieve vaardigheden en leerprestaties en emotionele- of gedragsproblemen. [Leseman \[2019\]](#) benadrukt hierbij ook de rol van risicocumulatie. [Van den Hof et al. \[2025\]](#) identificeren daarnaast ook nog andere risicofactoren die van invloed kunnen zijn op het stressniveau in een gezin zoals mogelijke betrokkenheid bij criminele activiteiten of problemen met de geestelijke gezondheid van de ouders. De aanwezigheid van de hierboven beschreven stress- en risicofactoren is geen garantie dat de ontwikkeling van een kind wordt belemmerd. Tegenover deze risico's staan ook beschermende factoren, zoals ouders die ondanks de stress in staat zijn om een positieve houding te behouden en steun ervaren van hun sociaal netwerk, wat als een buffer voor het kind fungeert [[Leseman, 2019](#)].

## 2.5.2. Verschillende vormen van kapitaal

Waar het *family stress model* zich richt op de negatieve gevolgen van sociaaleconomische achterstand en tegenspoed, belicht het *family investment model* juist hoe ouders hun middelen inzetten om de talenten en ontwikkeling van hun kinderen te bevorderen [[Conger and Donnellan, 2007](#)]. Deze middelen kunnen in verschillende groepen worden gecategoriseerd: economisch, menselijk en sociaal kapitaal [[Badou and Day, 2021](#), [Conger and Donnellan, 2007](#), [Leseman, 2019](#), [Vrooman et al., 2023](#)]. Het economisch kapitaal gaat over het inkomen en vermogen dat ouders op verschillende manieren in kunnen zetten om de ontwikkeling van hun kinderen te bevorderen; dit kan zowel direct als indirect werken. Directe vormen van het inzetten van het economisch kapitaal zijn bijvoorbeeld investeringen in schaduwwonderwijs [[Elffers, 2018](#)], de aanwezigheid van boeken, de beschikbaarheid van IT-apparatuur zoals computers of tablets die nodig zijn om huiswerk te kunnen maken [[Bol, 2020](#), [Haelermans et al., 2022](#), [Zumbuehl and Dillingh, 2020](#)] of het faciliteren van ontwikkelings- en ontspanningsmogelijkheden zoals de aanwezigheid van kinderboeken, bezoek van cultuurevenementen of vakanties [[ter Weel et al., 2025](#)]. Het economisch kapitaal van ouders heeft ook een indirect effect. Welvarende ouders zijn in staat om in veilige buurten te wonen en kunnen meer investeren in een gezonde omgeving door middel van gezonde voeding, kwaliteit van de woning en medische zorg [[Conger and Donnellan, 2007](#)]. Bij het menselijk kapitaal van ouders gaat het om de kennis, vaardigheden, opleidingsniveau en cultureel kapitaal van de ouders. Ouders met een hoger opleidingsniveau zijn bijvoorbeeld beter in staat en vinden het vaak belangrijker om te helpen met het maken van huiswerk [[Bol, 2020](#)], lezen vaker voor aan hun jonge kinderen [[Conger and Donnellan, 2007](#)] en hebben vaak een rijker taalgebruik [[Leseman, 2019](#)]. Cultureel kapitaal, bestaande uit facetten als omgangsvormen, taalgebruik, interesses, kunst- en cultuurvoorkeuren, sociale en culturele opvattingen, wordt van generatie op generatie overgebracht. Kinderen verkrijgen hun cultureel kapitaal in eerste instantie tijdens hun socialisatie binnen het gezin. Aangezien de dominante cultuur ten grondslag ligt aan het curriculum dat de scholen aanbieden, wordt wel verondersteld dat hoe dichterbij het gezin zich bij de dominante cultuur bevindt, hoe makkelijker het wordt voor het kind om zich comfortabel te voelen en mee te komen op school [[Driessen, 2001](#)]. Naast de kwantiteit en kwaliteit van het taalgebruik in het gezin, is meertaligheid een ander aspect van taal dat invloed heeft op de ontwikkeling van kinderen [[Leseman, 2019](#)]. Dit aspect speelt met name in gezinnen waar één of beide ouders een niet-Nederlandse herkomst hebben. Meertaligheid wordt gelieerd aan positieve effecten op cognitieve vaardigheden. Echter, meertaligheid kan er ook voor zorgen dat de woordenschat van tweetaligen – hoewel in totaal groter – in iedere taal apart kleiner is dan

de woordenschat van eentaligen in hun eigen taal. Verder laat onderzoek zien dat hoe groter de taalafstand tussen de herkomsttaal van een migrant en de taal van het bestemmingsland, des te lager de gemiddelde leesvaardigheid is van een leerling met een migratieachtergrond [Kooistra et al., 2008]. Als laatste speelt ook sociaal kapitaal een rol. Het netwerk, de status en invloed die voortvloeien uit de maatschappelijke positie van ouders beïnvloeden de waarde en prioriteit die ouders toekennen binnen de opvoeding van hun kinderen [Conger and Donnellan, 2007].

## 2.6. Schoolomgeving

Het Nederlands onderwijs wordt gekenmerkt door een sterke en aanhoudende segregatie [Leseman, 2019, Muskens et al., 2025]. Hierbij is de scheiding op basis van sociaaleconomische status sterker dan op basis van migratieachtergrond [Leseman, 2019]. Segregatie in het onderwijs heeft via verschillende mechanismen effect op de leerprestaties van kinderen. Ten eerste kan een eenzijdige samenstelling van de school met (bijna) alleen kinderen uit gezinnen met een lage sociaaleconomische status ervoor zorgen dat 'peer-learning' effecten negatief worden beïnvloed [Van Ewijk and Sleegers, 2010]. Ten tweede laat onderzoek zien dat segregatie effect heeft op de verwachtingen en het taalgebruik van leerkrachten [Leseman, 2019]. Leerkrachten passen hun taalgebruik vaak aan op het ingeschatte niveau van de klas. Verder laat onderzoek zien dat leerkrachten lagere verwachtingen hebben bij kinderen uit gezinnen met een lage sociaaleconomische status, zelfs bij vergelijkbare leerprestaties [Entorf, 2015, Fleischmann and de Haas, 2016, Geven et al., 2018]. Verder kan segregatie leiden tot verschil in hulpbronnen en sociaal kapitaal [Leseman, 2019]. Hoewel achterstandsscholen in Nederland extra middelen krijgen, hebben zij vaker last van onderbezetting door het lerarentekort. Verschillen in de sociale en culturele bagage die leerlingen en ouders de school binnenbrengen kunnen deze effecten nog versterken. Als laatste kan differentiatie in de klas verschillen in sociaaleconomische status versterken. Hoewel differentiatie bedoeld is als maatwerk, kan het ongelijkheid juist versterken. Met name *divergente differentiatie* (waarbij meer begaafde leerlingen extra uitdaging krijgen) kan in combinatie met een onderschatting van het leervermogen van achterstandsleerlingen de prestatiekloof vergroten. Differentiatie werkt alleen verkleinend voor de kloof als het gepaard gaat met extra leertijd en een bewuste focus op kansengelijkheid [Leseman, 2019].

## 2.7. Buurtomgeving

De buurt waarin kinderen opgroeien, kan van invloed zijn op de leerprestaties. Kinderen spelen in de buurt en krijgen (onbewust) de situatie in een bepaalde buurt mee. Een woonomgeving die wordt gekenmerkt door een aanzienlijke aanwezigheid van laagopgeleiden, werkloosheid, taalarmoede of een belemmerende dorps- of straatcultuur kan als niet-stimulerend voor het kind worden beschouwd [Badou and Day, 2021, Mulder et al., 2014, ter Weel et al., 2025]. Wanneer een kind opgroeit in een taalarm gebied, is de kans groot dat het kind onvoldoende wordt gestimuleerd, wat nadelig is voor de schoolprestaties [Kuyvenhoven and Boterman, 2021, Nieuwenhuis and Hooimeijer, 2016]. Buurten waar werkloosheid overheerst kunnen de instelling, werkhouding of opvattingen over de samenleving van kinderen beïnvloeden. Een soortgelijke invloed geldt voor buurtveiligheid [Leseman, 2007, ter Weel et al., 2025]. Criminaliteit en onveiligheid in de buurt worden, net als armoede en werkloosheid, beschouwd als negatieve kenmerken, die niet stimulerend zijn voor de ontwikkeling van een kind. Verondersteld wordt dat kinderen omringd door meer criminaliteit en onveiligheid een ander beeld van de samenleving (rechten, plichten, autoriteit) kunnen ontwikkelen, maar ook bijvoorbeeld over discipline, dat uiteindelijk door kan werken in de manier waarop ze zich op

school gedragen, inzetten en presteren. Eerder onderzoek wijst inderdaad uit dat schoolprestaties gemiddeld lager zijn in stedelijke gebieden waar dit soort buurtkenmerken vaker voorkomen [Mulder et al., 2014].

## 2.8. Implicaties voor modelontwikkeling

Aan de hand van het theoretisch kader dat we in dit hoofdstuk hebben beschreven, hebben we de modelontwikkeling ingericht. In paragraaf 2.4 zijn we aan de hand van nieuwe inzichten tot de conclusie gekomen dat corrigeren voor intelligentie in de modelontwikkeling theoretisch wel mogelijk is, maar dat dit in de praktijk niet wenselijk is. Om een goed beeld te krijgen zullen we in hoofdstuk 4 nog wel een aanpak met en zonder intelligentie met elkaar vergelijken. Een andere reden voor het niet meenemen van intelligentie is dat dit wellicht een te beperkte weergave van de potentie van het kind is. Niet-cognitieve vaardigheden als doorzettingsvermogen, zelfbeheersing of emotieregulatie hebben invloed op – of zijn wellicht onderdeel van – de potentie van het kind. Verder hebben we gezien dat een veelvoud van omgevingskenmerken van invloed zijn op de ontwikkeling van een kind. Voor zover deze integraal bekend zijn in de registers van het CBS, nemen we deze kenmerken mee bij de modelontwikkeling. In het volgende hoofdstuk beschrijven we welke data we hebben gebruikt. Vanaf hoofdstuk 4 beschrijven we de modelontwikkeling en de resultaten van de verschillende analyses die we hebben uitgevoerd.

## 3. Data

In dit hoofdstuk bespreken we welke databronnen we hebben gebruikt voor het ontwikkelen van een nieuw model voor de risico-indicator onderwijsachterstanden voor het basisonderwijs.

### 3.1. Databronnen

Bij de oorspronkelijke ontwikkeling van de risico-indicator onderwijsachterstanden is gebruik gemaakt van de aanwezige intelligentiemaatstaf in de COOL-cohorten [CBS, 2016]. Deze intelligentiemaatstaf is gemeten met de Niet-Schoolse Cognitieve Capaciteiten Test (NSCCT). Met het verdwijnen van de COOL-onderzoeken is deze informatie niet meer direct beschikbaar. Om modellen met en zonder het gebruik van intelligentie met elkaar te kunnen vergelijken is in 2021 een leerpotentiëmeting uitgevoerd onder een omvangrijke en representatieve steekproef van leerlingen in groep 6 van het reguliere basisonderwijs [Jenniskens et al., 2022].

De leerlingen uit deze steekproef hebben, op een enkele uitzondering na, in het voorjaar van 2023 deelgenomen aan de eindtoets van het basisonderwijs. Om die reden is het cohort van schooljaar 2022/'23 als basiscohort genomen voor de ontwikkeling van een nieuw model voor de risico-indicator onderwijsachterstanden. De data van dit cohort zijn vervolgens verrijkt met relevante achtergrondkenmerken uit het Stelsel van Sociaal-Statistische Bestanden (SSB) van het CBS. Het SSB bevat een groot aantal microdatabestanden met informatie uit administratieve overheidsregisters over personen en huishoudens. Het CBS ontvangt deze informatie vanwege zijn wettelijke taak. De data bevatten geen namen, geen adressen en geen burgerservicenummers. Om gegevens uit verschillende bronnen aan elkaar te verbinden worden gepseudonimiseerde koppelsleutels gebruikt die buiten het SSB geen betekenis hebben. Individuele personen zijn hierdoor niet direct te identificeren.

### 3.2. Populatieafbakening

Voor het vergelijken van modellen met en zonder intelligentie is de omvang van de leerpotentiëmeting in groep 6 in 2021 bepalend. De omvang van de gerealiseerde steekproef van deze leerpotentiëmeting bedraagt iets meer dan 14 duizend leerlingen. Voor iets meer dan 13 duizend leerlingen konden eindtoetsgegevens en achtergrondkenmerken worden aangekoppeld. Vervolgens hebben we alleen de leerlingen geselecteerd van wie beide ouders in de Basisregistratie Personen (BRP) voorkomen én van wie de achtergrondkenmerken (behalve het opleidingsniveau van de ouders) compleet zijn; er blijven dan uiteindelijk ruim 12 500 leerlingen over.

Voor modellen zonder het gebruik van intelligentie is de steekproefomvang van de leerpotentiëmeting geen beperkende factor. Daardoor kunnen we voor deze modellen in principe beschikken over de gehele populatie van leerlingen groep 8 van het schooljaar 2022/'23 die een eindtoets hebben gemaakt. Uit deze groep hebben we wederom alleen de leerlingen geselecteerd van wie beide ouders in de Basisregistratie Personen (BRP) voorkomen én van wie

de achtergrondkenmerken (met uitzondering van opleidingsniveau) compleet zijn. De totale populatie van deze dataset bedraagt iets meer dan 162 duizend leerlingen.

### 3.3. Gebruikte achtergrondkenmerken en operationalisaties

Om te bepalen welke achtergrondkenmerken we mee hebben genomen in de analyses, hebben we verschillende bronnen gebruikt. Een overzicht daarvan hebben we in het vorige hoofdstuk besproken. Voor alle relevante achtergrondkenmerken hebben we vervolgens gekeken of er in de registraties bij het CBS daar direct informatie over beschikbaar was of een kenmerk dat als proxyvariabele kan dienen. Hierbij hebben we beschikbare informatie uit enquêtes uitgesloten omdat dit geen integrale waarneming van de Nederlandse bevolking is.

Voor de interpretatie van de resultaten die we in de volgende hoofdstukken bespreken, is het belangrijk om te weten hoe de variabelen exact geoperationaliseerd zijn. In bijlage A is daarom een overzicht van alle variabelen en hun operationalisering opgenomen. Zoals ook in tabel 5.1 is te zien, hebben we van veel achtergrondkenmerken verschillende operationalisaties uitgetest. Het doel hiervan was om te kijken welke mate aan detail nodig was om tot een goed model te komen.

Bij de variabele herkomst (als maat voor culturele afstand) en problematische schulden willen we wat uitgebreider stilstaan. Voor de huidige indicator is destijds een op maat gemaakte indeling ontworpen. Het heeft onze voorkeur om voor de operationalisering van herkomst een standaard beschikbare indeling te gebruiken die uitdrukking geeft aan taal- en culturele afstand én die ook binnen Europa onderscheid maakt tussen verschillende groepen landen. In 2021 heeft de Wetenschappelijke Raad voor Regeringsbeleid (WRR) een onderzoek gepubliceerd waarin een indeling wordt gepresenteerd op basis etno-linguïstische afstand [[Jennissen et al., 2021](#)]. In onze analyses hebben we deze indeling getest naast de indeling zoals die voor de huidige indicator is ontwikkeld. In de huidige indicator is het gegeven dat één of beide ouders in een schuldsaneringstraject zitten een verklarende factor. Dit kenmerk vinden we om meerdere redenen minder geschikt om ook nu in de modelontwikkeling mee te nemen. De eerste reden is dat schuldsanering vaak voorafgegaan wordt door een lang traject van schuldenproblematiek en dus eigenlijk een soort eindpunt is. Lang niet iedereen met schuldenproblematiek komt in de schuldsanering terecht. Ten tweede komt niet iedereen in aanmerking voor schuldsanering. Mensen met een schuld bij de belastingdienst worden bijvoorbeeld uitgesloten. Als laatste is de omvang van deze groep vrij klein en als gevolg van een recente wetswijziging zal – met name door de sterk verkorte duur van de schuldsanering – de omvang van de groep ouders in de schuldsanering nog kleiner worden. Het kenmerk problematische schulden heeft deze nadelen niet en is daarom een geschikter alternatief om het effect van schuldenproblematiek in kaart te brengen.

## 4. Keuzes bij modelontwikkeling

In dit hoofdstuk gaan we in op een aantal keuzes die zijn gemaakt bij de ontwikkeling van de nieuwe risico-indicator voor onderwijsachterstanden. In het bijzonder bespreken we hier punten waarop het nieuwe model afwijkt van het model waarop de bestaande indicator is gebaseerd. Paragraaf 4.1 gaat over het gebruik van intelligentie in het oude en nieuwe model. Paragraaf 4.2 gaat over het gebruik van dummyvariabelen om categoriale omgevingskenmerken te coderen. Een belangrijke verandering is dat bij het nieuwe model gebruik wordt gemaakt van eindtoetsen van meerdere toetsaanbieders voor het meten schoolprestaties in groep 8; dit wordt verder toegelicht in paragraaf 4.3 en paragraaf 4.4. Ten slotte beschrijft paragraaf 4.5 kort hoe is omgegaan met ontbrekende waarden in de data.

### 4.1. Het gebruik van een intelligentiemeting

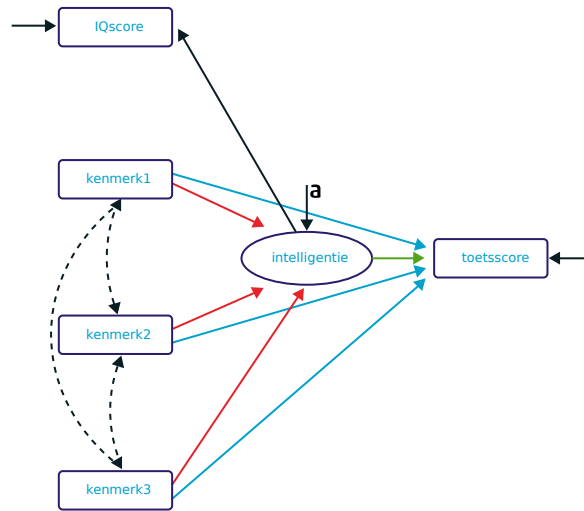
De indicator voor onderwijsachterstanden is gebaseerd op een zogenaamd structureel vergelijkingsmodel, waarin de onderwijsprestaties van leerlingen in groep 8 worden gerelateerd aan kenmerken van hun omgeving. Met zo'n model kunnen systematische verschillen worden gevonden tussen groepen leerlingen, waarbij leerlingen met bepaalde omgevingskenmerken gemiddeld lager presteren dan leerlingen met bepaalde andere omgevingskenmerken. Het idee is dat op deze manier onderwijsachterstanden in beeld worden gebracht.

Figuur 4.1 toont een schematische weergave (padmodel) van een structureel vergelijkingsmodel voor onderwijsachterstanden waarin intelligentie is opgenomen samen met drie omgevingskenmerken. Intelligentie is in dit model een latente (niet direct waargenomen) variabele. De waargenomen variabele IQ-score is in het model opgenomen als een meting van deze latente variabele. De andere variabelen in het model worden beschouwd als direct waargenomen. Het model in Figuur 4.1 is een technische uitwerking van het conceptuele model uit Figuur 2.2.

In het model zoals weergegeven in Figuur 4.1 heeft elk omgevingskenmerk op twee manieren een effect op de eindtoetsscore. Ten eerste heeft elk omgevingskenmerk een *direct effect*; dit zijn de blauwe pijlen in het paddiagram. Ten tweede heeft elk omgevingskenmerk een *indirect effect* dat loopt via intelligentie. Dit indirecte effect is het product van het directe effect dat het omgevingskenmerk heeft op intelligentie (rode pijl) en het directe effect dat intelligentie heeft op de toetsscore (de groene pijl tussen deze twee variabelen). Het *totale effect* van een omgevingskenmerk op de toetsscore is gelijk aan de som van zijn directe en indirecte effect. Zie bijvoorbeeld Bollen [1989] voor meer uitleg over structurele vergelijkingsmodellen en directe, indirecte en totale effecten.

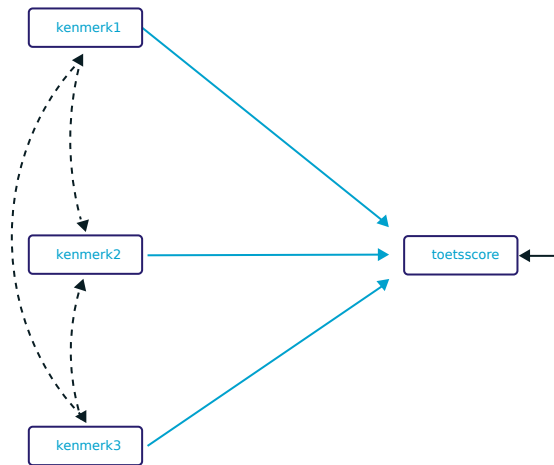
De huidige indicator voor onderwijsachterstanden is gebaseerd op een model zoals in Figuur 4.1, waarbij alleen de blauwe pijlen uiteindelijk zijn opgenomen in de indicator. De rest van het model is gebruikt om deze blauwe pijlen goed te kunnen schatten. Verder waren in dat model alleen rode pijlen opgenomen voor de kenmerken opleidingsniveau moeder en opleidingsniveau vader. Van de overige kenmerken in de indicator is destijds aangenomen dat ze (althans na correctie voor culturele vertekening in de IQ-score; zie [CBS, 2016] voor meer details) geen direct effect hebben op intelligentie. Deze aanname is verder onderzocht in [CBS, 2021].

Figuur 4.1. Pad diagram van een structureel vergelijkingmodel inclusief intelligentie



Zoals beschreven in hoofdstuk 2 is het in de praktijk onmogelijk om intelligentie te meten zonder beïnvloeding door de omgeving. Het huidige inzicht is dat het daarom beter lijkt om de indicator voor onderwijsachterstanden niet te baseren op de directe effecten van de omgevingskenmerken, maar op hun totale effecten, inclusief de indirecte effecten die lopen via intelligentie. Indien men alleen geïnteresseerd is in deze totale effecten, kan ook gewerkt worden met een eenvoudiger structureel vergelijkingmodel zonder intelligentie. Dit vereenvoudigde model is als padmodel weergegeven in Figuur 4.2. In dit model hebben alle omgevingskenmerken alleen een direct effect op de toetscore, dat tevens gelijk is aan het bijbehorende totale effect. In feite is dit een gewoon lineair regressiemodel, dat echter ook kan worden opgevat als een speciaal geval van een structureel vergelijkingmodel.

Figuur 4.2. Paddiagram van een structureel vergelijkingmodel zonder intelligentie



De directe/totale effecten van de omgevingskenmerken uit Figuur 4.2 zijn in theorie exact gelijk aan de totale effecten van de omgevingskenmerken uit Figuur 4.1, mits in dit laatste model alle mogelijke rode en blauwe pijlen zijn opgenomen. Voor de volledigheid hebben we dit in de praktijk geverifieerd door beide modellen te schatten op data van een steekproef van leerlingen die in schooljaar 2022/'23 in groep 8 zaten en waarvoor een intelligentiemeting uit groep 6 beschikbaar was. (Zie hoofdstuk 3 en paragraaf 4.4 hieronder voor meer informatie over deze data.) De opgenomen omgevingskenmerken waren vergelijkbaar met de kenmerken uit de huidige indicator. Zoals verwacht bleken de geschatte directe effecten voor het model uit Figuur 4.2 op afrondingsverschillen na gelijk te zijn aan de geschatte totale effecten voor het model uit Figuur 4.1.

Hoewel de twee modellen equivalent zijn voor het schatten van de totale effecten van de omgevingskenmerken op de toetscore, verschillen de modellen wel in de verklaarde variantie van de toetscore. De verklaarde variantie van het model in Figuur 4.2 is een stuk lager dan de verklaarde variantie van het model in Figuur 4.1; in de bovengenoemde toepassing op data uit 2022/'23 waren deze verklaarde varianties respectievelijk ongeveer 14 procent en ongeveer 43 procent. Dit verschil komt door de bijdrage van de residuele variantie van intelligentie in het model uit Figuur 4.1, aangeduid door de zwarte pijl met bijschrift 'a' in deze figuur. Deze pijl vertegenwoordigt het deel van intelligentie dat niet direct verklaard wordt door de omgevingskenmerken (wat in de praktijk het overgrote deel van intelligentie is). In het model uit Figuur 4.1 draagt het volledige effect van intelligentie op de toetscore bij aan de verklaarde variantie van de toetscore, inclusief het deel dat hoort bij pijl 'a'. In het model uit Figuur 4.2 vormt pijl 'a' geen bron van verklaarde variantie.

Figuur 4.2 heeft dus als theoretisch nadeel dat de verklaarde variantie lager is dan in Figuur 4.1, al maakt dit zoals eerder gezegd geen verschil voor de geschatte totale effecten die relevant zijn voor de indicator. Daar staat als groot praktisch voordeel tegenover dat er geen intelligentiemeting nodig is om het model uit Figuur 4.2 te schatten. In de praktijk beschikt het CBS standaard over eindtoetscores voor vrijwel alle leerlingen in groep 8, terwijl een intelligentiemeting speciaal voor dit onderzoek uitgevoerd moet worden en dan alleen

beschikbaar is voor een kleine steekproef. Dit betekent dat het model uit Figuur 4.2 geschat kan worden op een veel grotere dataset dan het model uit Figuur 4.1, waarmee de totale effecten nauwkeuriger bepaald kunnen worden. Een tweede voordeel is dat het, door niet meer afhankelijk te zijn van een intelligentiemeting, eenvoudiger en goedkoper wordt om het model regelmatig opnieuw te schatten. De indicator zou daarmee vaker herijkt kunnen worden dan in het verleden. In hoofdstuk 5 worden daarom resultaten gepresenteerd van modellen van de vorm uit Figuur 4.2.

## 4.2. Omgang met categoriale variabelen

In het model waarop de huidige indicator is gebaseerd zijn categoriale variabelen op twee verschillende manieren meegenomen. Bij de verblijfsduur van de moeder in Nederland, ingedeeld in drie categorieën, is gewerkt met dummyvariabelen. Dit houdt in dat binnen het structurele vergelijkingsmodel aparte coëfficiënten zijn geschat per categorie. In een paddiagram zoals Figuur 4.1 of Figuur 4.2 betekent dit dat elke categorie is opgenomen als apart kenmerk. Omwille van de identificeerbaarheid van het model is hierbij een van de categorieën aangewezen als referentiecategorie; deze categorie krijgt een coëfficiënt van 0 en de coëfficiënten van de andere categorieën geven de afwijking weer ten opzichte van deze referentiecategorie. (De keuze van de referentiecategorie is hierbij vrij en heeft geen invloed op de verwachte onderwijsscores die volgen uit het model.)

Voor de categoriale kenmerken opleidingsniveau moeder (acht categorieën), opleidingsniveau vader (acht categorieën) en herkomstgroepering (in de huidige indicator acht categorieën) is destijds echter een andere aanpak gevolgd. De categorieën van deze variabelen zijn eerst, via een apart regressiemodel, afgebeeld op een continue schaal. (Hierbij geldt nog steeds dat een van de categorieën wordt aangewezen als referentiecategorie; deze krijgt de schaalwaarde 0.) Vervolgens zijn opleidingsniveau en herkomst in het structurele vergelijkingsmodel opgenomen als continue variabelen op deze schaal, wat inhoudt dat ze in het paddiagram zijn opgenomen als één kenmerk, met één gezamenlijke coëfficiënt [zie CBS, 2016, voor een verdere toelichting op deze aanpak].

Deze twee-stapsaanpak is destijds gekozen met de bedoeling om het eenvoudiger te maken om variabelen met relatief veel categorieën, zoals opleidingsniveau en herkomst, in een structureel vergelijkingsmodel op te nemen. In de praktijk is echter gebleken dat het meer werk met zich meebrengt om in het productiesysteem van de indicator twee aanpakken voor categoriale variabelen naast elkaar te gebruiken (een voor verblijfsduur en een voor opleidingsniveau en herkomst), zonder dat daar een duidelijk voordeel tegenover staat. Er is daarom besloten om bij de ontwikkeling van de nieuwe indicator geen onderscheid meer te maken tussen enerzijds opleidingsniveau en herkomst en anderzijds alle andere categoriale variabelen. In plaats daarvan worden in het vervolg alle categoriale variabelen in het model op dezelfde manier behandeld, namelijk als dummyvariabelen.

## 4.3. Meting van onderwijsprestaties

Bij de ontwikkeling van de oorspronkelijke risico-indicator onderwijsachterstanden in het basisonderwijs en ook bij de kleine herijking van de indicator is alleen de Cito-eindtoets gebruikt als meting van onderwijsprestaties in groep 8. Van de cohorten leerlingen die daarbij zijn meegenomen maakte meer dan 80 procent de Cito-toets. Dit aandeel is inmiddels flink gedaald, omdat scholen vanaf het schooljaar 2014/'15 kunnen kiezen uit verschillende toetsaanbieders.

In de schooljaren 2021/'22 en 2022/'23 die in de huidige analyse centraal staan waren er vijf verschillende toetsaanbieders: naast de Centrale Eindtoets (CET) van Cito zijn dit A-VISION (Route 8), Bureau ICE (IEP), Diataal (DIA) en AMN (AMN). De scores van de vijf eindtoetsen zijn door elke toetsaanbieder uitgedrukt op een eigen schaal en kunnen daarom niet direct worden vergeleken. Wel bevatten de vijf toetsen deels dezelfde vragen, de zogenaamde ankeritems. Met behulp van deze ankeritems maakt de Expertgroep Toetsen PO sinds het schooljaar 2018/'19 jaarlijks een vertaling van de scores op de vijf eindtoetsen naar zes mogelijke toetsadviezen. Tabel 4.1 toont voor het schooljaar 2022/'23 per toets de verdeling van de scores over de toetsadviezen.

Tabel 4.1. Toetsscores per toetsadvies en toetsaanbieder (schooljaar 2022/'23)

Toetsadvies	AMN	CET	DIA	IEP	Route 8
Pro/vmbo-b	301-312	501-510	321-335	50-51	100-112
Vmbo-b/k	313-356	511-524	336-350	52-68	113-161
Vmbo-k/gt	357-395	525-532	351-357	69-76	162-188
Vmbo-gt/havo	396-442	533-539	358-365	77-84	189-214
Havo/vwo	443-475	540-544	366-371	85-91	215-238
Vwo	476-500	545-550	372-390	92-100	239-300

Bij de ontwikkeling van de nieuwe indicator voor onderwijsachterstanden wil het CBS in principe toetsscores van alle vijf de aanbieders meenemen. Door het onderzoek te beperken tot een deel van de toetsaanbieders zou er namelijk een risico zijn dat een selectief deel van de populatie leerlingen wordt uitgesloten bij de ontwikkeling van het model voor onderwijsachterstanden. Om de verschillende toetsen naast elkaar te kunnen gebruiken in hetzelfde model, is het nodig om de scores eerst af te beelden op dezelfde schaal. Een grote beperking is hierbij dat het CBS alleen beschikt over de toetsscores per leerling en geen toegang heeft tot (bijvoorbeeld) de losse antwoorden op de ankeritems. Gegeven deze beperking zijn in een verkennend onderzoek verschillende mogelijke aanpakken uitgewerkt om de toetsscores toch te uniformeren [CBS, 2024b].

Uiteindelijk is gekozen om hier de volgende aanpak te gebruiken. De toetsscores worden herschaald naar een geüniformeerde score die altijd ligt tussen 0 en 6. Een herschaalde score tussen de gehele getallen  $i - 1$  en  $i$  hoort altijd bij een toetsadvies in rij  $i$  van Tabel 4.1; een herschaalde score tussen 1 en 2 hoort dus bijvoorbeeld bij het advies Vmbo-b/k en een herschaalde score tussen 5 en 6 bij het advies Vwo. Per toetsadvies worden de mogelijke scores van een bepaalde toetsaanbieder gelijkmatig verdeeld over het interval dat loopt van  $i - 1$  tot  $i$ . De formule om een score  $x_{ti}$  van toetsaanbieder  $t$  die hoort bij toetsadvies  $i$  te herschalen is:

$$x_{ti, \text{herschaald}} = (i - 1) + \frac{x_{ti} - (\min_{ti} - 0,5)}{(\max_{ti} + 0,5) - (\min_{ti} - 0,5)},$$

waarbij  $\min_{ti}$  en  $\max_{ti}$  respectievelijk de laagste en hoogste score zijn die voor toetsaanbieder  $t$  passen bij toetsadvies  $i$ . In bovenstaande formule is een correctie van 0,5 gebruikt om te voorkomen dat de hoogste score die hoort bij toetsadvies  $i - 1$  en de laagste score die hoort bij toetsadvies  $i$  precies dezelfde herschaalde score zouden krijgen.

Een voorbeeld: gebruikmakend van de informatie uit Tabel 4.1 is te zien dat in het schooljaar 2022/'23 een CET-score van 530 na uniformering overgaat in een herschaalde score  $2 + (5, 5/8) \approx 2,688$ . Dit is vergelijkbaar met een Route 8-score van 180, die wordt herschaald tot  $2 + (18, 5/27) \approx 2,685$ .

Voor meer details, waaronder een vergelijking van deze uniformeringsmethode met andere mogelijke aanpakken, wordt verwezen naar CBS [2024b].

## 4.4. Modellen met meerdere groepen

Na de herschaling uit de vorige paragraaf liggen de scores voor alle toetsaanbieders op dezelfde schaal, zodat deze in principe naast elkaar gebruikt kunnen worden in één model voor onderwijsachterstanden. Een vraag blijft wel of de herschaalde scores van de toetsaanbieders volledig vergelijkbaar zijn: meten de vijf toetsen, nadat de verschillende schalen zijn rechtgetrokken via de toetsadviezen uit Tabel 4.1, echt hetzelfde of zijn er toch inhoudelijke verschillen? Deze vraag is niet direct te beantwoorden met de data waarover het CBS beschikt. Er zijn zeker verschillen te zien in de scoreverdelingen bij verschillende aanbieders [zie CBS, 2024b], maar deze zijn niet per se het gevolg van meetverschillen. Ze zouden bijvoorbeeld ook verklaard kunnen worden door selectie, wanneer leerlingen met een bepaalde achtergrond vaker een bepaalde toets maken. Dergelijke selectieverschillen hoeven geen probleem te zijn bij het modelleren van onderwijsachterstanden, mits de relevante achtergrondkenmerken worden opgenomen in het model. In de praktijk blijft het wel mogelijk dat er ook ongemeten relevante achtergrondkenmerken zijn (confounders), met als gevolg een lagere verklaarde variantie.

In CBS [2026] is met een regressieanalyse onderzocht of de herschaalde scores van verschillende toetsaanbieders nog systematisch van elkaar verschillen, nadat rekening is gehouden met andere mogelijke verklaringen voor verschillen op leerlingniveau, zoals het voorlopige schooladvies en de intelligentiescore van een leerling. In die analyse zijn geen significante verschillen gevonden tussen de toetsaanbieders. Hierbij moeten wel twee kanttekeningen worden gemaakt. Ten eerste zijn de referentieniveaus voor rekenen, taalverzorging en leesvaardigheid in het genoemde regressiemodel opgenomen en dit blijken belangrijke verklarende kenmerken te zijn. Deze referentieniveaus zijn echter afgeleid uit de (oorspronkelijke) toetsscores. Ze zijn daarom alleen geldig als alternatieve verklaring voor verschillen onder de aanname dat de toetsen in elk geval deze referentieniveaus op dezelfde manier meten; ook deze aanname kan niet worden getoetst met de data die beschikbaar zijn. Ten tweede is in het regressiemodel een aantal kenmerken opgenomen – zoals de genoemde referentieniveaus – die zeker niet zullen worden opgenomen in de uiteindelijke indicator voor onderwijsachterstanden, omdat ze niet voor alle leerlingen bekend zijn. In het uiteindelijke model voor onderwijsachterstanden kunnen deze kenmerken daarom niet gebruikt worden om te corrigeren voor eventuele meetverschillen tussen toetsaanbieders. Een relevante vraag blijft daarom of het nuttig is om in het model voor onderwijsachterstanden rekening te houden met mogelijke verschillen tussen toetsaanbieders.

Om deze vraag te beantwoorden is een verkennende analyse gedaan. Uitgangspunt bij deze analyse is het structurele vergelijkingsmodel uit de huidige indicator, zoals in Figuur 4.1. Wel zijn, vooruitlopend op het nieuwe model, de volgende wijzigingen aangebracht om rekening te houden met punten die eerder in dit rapport zijn genoemd:

- Alle kenmerken (dus niet alleen opleidingsniveau en herkomst) hebben in het model zowel een direct effect op de eindtoetsscore als een indirect effect dat loopt via intelligentie.
- Opleidingsniveau en herkomst worden als dummyvariabelen meegenomen in het model en niet eerst afgebeeld op een continue schaal via een apart regressiemodel.
- Het kenmerk schuldsanering is vervangen door het kenmerk problematische schulden (zie paragraaf 3.3).

Er zijn drie varianten geschat van het model voor onderwijsachterstanden:

1. Een gezamenlijk model zonder onderscheid naar toetsaanbieder.
2. Een model per toetsaanbieder waarbij alle parameters per toetsaanbieder mogen verschillen.

3. Een model per toetsaanbieder waarbij alleen de constante term uit de regressievergelijking voor de eindtoetsscore per toetsaanbieder mag verschillen; alle andere parameters zijn gelijk voor alle toetsaanbieders.

Model 2 en 3 zijn voorbeelden van een structureel vergelijkingsmodel met een multigroep-component, waarbij de toetsaanbieder als groep is gebruikt. In een multigroep-model kunnen restricties worden opgelegd die bepaalde parameters over groepen heen aan elkaar gelijkstellen. In model 2 zijn er helemaal geen restricties tussen de groepen, zodat in feite een onafhankelijk model per toetsaanbieder wordt geschat. In model 3 stellen de restricties dat alle omgevingskenmerken (opleidingsniveau, herkomst, etc.) voor alle toetsaanbieders hetzelfde effect moeten hebben, net zoals in model 1. Anders dan in model 1 mag er in model 3 echter wel een verschil zijn in de gemiddelde toetsscore per groep (toetsaanbieder) bij een gegeven combinatie van omgevingskenmerken.

De analyses zijn uitgevoerd op data van leerlingen die in het schooljaar 2022/'23 in groep 8 van het basisonderwijs zaten, een eindtoets hebben gemaakt en geen ontheffing hebben. De eindtoetsscores zijn herschaald met de formule uit de vorige paragraaf. Verder is een intelligentiemeting gebruikt die via het Nationaal Regieorgaan Onderwijsonderzoek in het schooljaar 2020/'21 is uitgevoerd onder een representatieve steekproef van basisschoolleerlingen uit groep 6 [Jenniskens et al., 2022]. Uiteindelijk bleef een steekproef van 12553 leerlingen over met een bekende intelligentiescore uit groep 6 en eindtoetsscore uit groep 8. De verdeling over de vijf toetsaanbieders was als volgt: 5962 leerlingen met CET, 4310 met IEP, 1498 met Route 8, 591 met DIA en 192 met AMN. De steekproef onder leerlingen met een AMN-score bleek echter te klein te zijn om model 2 voor deze groep te kunnen schatten. In deze analyse is AMN daarom buiten beschouwing gelaten.

Alle structurele vergelijkingsmodellen in dit rapport zijn geschat met het R-pakket **lavaan** [Rosseel, 2012]. Tabel 4.2 toont voor de drie geschatte modellen een aantal fitmaten. AIC en BIC zijn maten die een afweging maken tussen enerzijds hoe goed een model bij de data past en anderzijds hoe groot het model is. Voor beide maten geldt dat een lagere waarde wijst op een beter model. Bij een steekproefomvang van ongeveer 12000 zoals hier is BIC strenger voor grotere modellen dan AIC. Verder bevatten de laatste drie kolommen de toetsingsgrootte ( $X^2$ ), het aantal vrijheidsgraden (df) en de bijbehorende p-waarde van een likelihood-ratio-toets van het model bij de gegeven data. Deze statistische toets is in dit geval alleen zinvol bij model 3, omdat model 1 en 2 verzadigde modellen zijn.

Tabel 4.2. Fitmaten voor modellen met en zonder onderscheid tussen toetsaanbieders

Model	Parameters	AIC	BIC	X <sup>2</sup>	df	p-waarde
Model 1	55	134126	134534	0,0	0	
Model 2	220	134203	135836	0,0	0	
Model 3	67	134093	134590	165,3	153	0,235

Te zien is dat model 2 veel groter is dan model 1 (het bevat 165 parameters meer), maar deze extra flexibiliteit leidt volgens zowel AIC als BIC niet tot een beter model. In model 3 zijn 153 parameters uit model 2 weggelaten, maar volgens de likelihood-ratio-toets past dit eenvoudigere model niet significant slechter bij de data dan model 2. Verder is model 3 volgens AIC een beter model dan model 1, terwijl dit volgens BIC net andersom is.

In model 3 wordt per toetsaanbieder een aparte constante term geschat voor de regressievergelijking van de herschaalde eindtoetsscore. Model 1 bevat één constante term die voor alle toetsaanbieders hetzelfde is. Het verschil tussen de kleinste en grootste geschatte

constante term in model 3 is ongeveer 0,16 punt. De verschillen tussen de constante termen van de toetsaanbieders zijn, gegeven de steekproefomvang, statistisch niet significant. Inhoudelijk lijkt een systematisch verschil tussen twee toetsaanbieders van 0,16 punt op de schaal van de geüniformeerde toetsscore echter niet verwaarloosbaar: bij een toetsadvies Vmbo-b/k is dit bijvoorbeeld vergelijkbaar met een verschil van ruim 2 punten op de schaal van de CET.

Tabel 4.3 toont voor de drie modellen de fractie verklaarde variantie ( $R^2$ ) in de herschaalde eindtoetsscore. De eerste kolom bevat de totale verklaarde variantie voor de steekproef, de andere kolommen bevatten de verklaarde variantie per toetsaanbieder. (Bij model 1 is alleen de totale verklaarde variantie beschikbaar.) De totale verklaarde variantie is ruim 43 procent, wat twee procentpunten hoger is dan bij het model waarop de huidige indicator is gebaseerd (zie [CBS, 2016]). Verder zijn soms vrij grote verschillen te zien tussen de toetsaanbieders: de verklaarde variantie is het hoogst bij CET, terwijl de verklaarde variantie bij DIA duidelijk lager ligt dan bij de andere groepen. Vermoedelijk speelt hierbij mee dat CET en DIA respectievelijk de grootste en kleinste groep zijn in de steekproef.

Tabel 4.3. Verklaarde variantie van modellen met en zonder onderscheid tussen toetsaanbieders

Model	Totaal	CET	DIA	IEP	Route 8
Model 1	0,434				
Model 2	0,437	0,455	0,354	0,425	0,433
Model 3	0,431	0,450	0,370	0,416	0,424

Gezien al deze resultaten is besloten om bij de verdere ontwikkeling van het model voor onderwijsachterstanden uit te gaan van een model met een multigroep-component zoals in model 3. Dit is volgens AIC het meest geschikte model van de modellen die hier zijn onderzocht, en volgens zowel AIC als BIC een beter model dan het veel grotere model 2. Verder lijken de verschillen tussen de geschatte constante termen in model 3 te groot om de rol van de toetsaanbieder helemaal te negeren, zoals in model 1. De verschillen in verklaarde variantie uit Tabel 4.3 bevestigen dit beeld.

## 4.5. Omgang met ontbrekende waarden

Een laatste complicatie die in dit hoofdstuk wordt besproken is dat er soms ontbrekende waarden voorkomen in de data waarop het model voor onderwijsachterstanden wordt geschat en later de indicator wordt berekend. We maken hierbij onderscheid tussen drie situaties:

1. Bij een aantal variabelen die we mogelijk willen opnemen in de indicator voor onderwijsachterstanden (en/of als hulpinformatie willen gebruiken om te corrigeren voor ontbrekende waarden in situaties 2 en 3 hieronder) komen ontbrekende waarden voor bij een beperkt aantal kinderen.
2. Bij een grotere groep kinderen is het hoogst behaalde opleidingsniveau van ten minste een van de ouders onbekend in het opleidingsregister.
3. Bij een aantal kinderen ontbreekt bijna alle informatie over een of beide ouders, omdat de ouders niet voorkomen in de Basisregistratie Persoonsgegevens (BRP), of omdat de relatie tussen kinderen en ouders niet volledig bekend is. Bij een kleine groep kinderen die zijn ingeschreven in het basisonderwijs ontbreekt bijna alle informatie omdat het kind zelf niet is ingeschreven in de BRP.

Bij de ontwikkeling van het model voor de indicator zijn alleen situaties 1 en 2 relevant; kinderen met ontbrekende informatie in de BRP worden hierbij uitgesloten. Bij het toepassen van de indicator in de praktijk zijn alle drie de situaties relevant.

De aanpak voor ontbrekende waarden in situaties 1 en 2 is beschreven in [CBS, 2024b]. In dat rapport is meer in detail uitgelegd waarom ontbrekende waarden soms voorkomen. Het is belangrijk om op te merken dat de gekozen aanpakken bij situaties 1 en 2 alleen worden toegepast op kinderen die niet vallen onder situatie 3. Dat wil zeggen: de aanpakken bij situaties 1 en 2 gaan ervan uit dat informatie uit de BRP beschikbaar is over het kind en beide ouders.

In situatie 1 gaat het specifiek om ontbrekende waarden op de kenmerken burgerlijke staat en sociaal-economische categorie voor beide ouders. In totaal betreft dit enkele tienduizenden ontbrekende waarden, op een populatie van circa 1,4 miljoen basisschoolleerlingen. Ontbrekende waarden worden in samenhang geïmputeerd met het R-pakket MICE (multiple imputation with chained equations), waarbij de kenmerken herkomstgroepering, leeftijd, verblijfsduur en opleidingsniveau (met een extra categorie 'onbekend') van beide ouders als hulpvariabelen worden gebruikt. Hoewel MICE geschikt is voor multipale imputatie wordt hier slechts één imputatieronde uitgevoerd, gezien de beperkte fractie ontbrekende informatie.

In situatie 2 worden de ontbrekende waarden van opleidingsniveau van beide ouders geïmputeerd met een multinomiaal logistisch regressiemodel dat gebruikmaakt van een groot aantal achtergrondkenmerken van de ouders en het huishouden. Om rekening te houden met de onzekerheid in de bijgeschatte ontbrekende opleidingsniveaus worden beide ouders tien keer geïmputeerd. Zie [CBS, 2024b] voor meer details. Bij het imputeren wordt onderscheid gemaakt tussen vier deelpopulaties, afhankelijk van welke informatie over opleidingsniveau wel is waargenomen. Tabel 4.4 toont de verdeling van basisschoolleerlingen over de vier deelpopulaties, zowel voor een cohort ten tijde van de ontwikkeling van de oorspronkelijke indicator (2013/'14) als voor het cohort waarop het nieuwe model wordt ontwikkeld (2022/'23).

Tabel 4.4. Deelpopulaties bij het imputeren van opleidingsniveau

Deelpopulatie	Opleidingsniveau		Percentage	
	Moeder	Vader	2013/'14	2022/'23
A	onbekend	onbekend	25	10
B	bekend	onbekend	20	16
C	onbekend	bekend	15	7
D	bekend	bekend	40	67

In de loop van de tijd is meer registerinformatie beschikbaar gekomen over opleidingsniveau, zodat het percentage kinderen met een onbekend opleidingsniveau voor een of beide ouders is afgenomen. Inmiddels is voor ruim twee derde van alle basisschoolleerlingen het opleidingsniveau van beide ouders bekend. Voor ongeveer 10 procent van de leerlingen zijn beide opleidingsniveaus onbekend.

Ten slotte wordt voor de kinderen die vallen onder situatie 3 in de huidige indicator direct een onderwijsscore geïmputeerd, op basis van de zeer beperkte informatie die wel beschikbaar is. Als onderdeel van de herziening van de indicator is onderzoek gedaan om deze imputatiemethode te verbeteren [CBS, 2024b, 2026]. Bij het in kaart brengen van de herverdeeffecten voor de nieuwe indicator in hoofdstuk 7 wordt gebruikgemaakt van de verbeterde imputatiemethode die is voorgesteld.

## 5. Modelontwikkeling

Dit hoofdstuk beschrijft de resultaten van het ontwikkelen van een nieuw model voor onderwijsachterstanden. Hierbij zijn de volgende stappen doorlopen:

- De longlist van beschikbare en mogelijk relevante kenmerken uit Bijlage A is ingekort door voor elk kenmerk de bivariate relatie met schoolprestaties te analyseren (paragraaf 5.1).
- Met de overgebleven kenmerken is via een stapsgewijze procedure gezocht naar een optimaal regressiemodel voor schoolprestaties (paragraaf 5.2).
- Het optimale regressiemodel uit de stapsgewijze procedure is uitgebreid tot een multigroep-model zoals beschreven in hoofdstuk 4, om rekening te houden met resterende verschillen tussen toetsaanbieders (paragraaf 5.3).
- Ter vergelijking is ook een multigroep-versie geschat van een van de beperktere modellen uit de stapsgewijze procedure (paragraaf 5.4).

Bij de analyses in dit hoofdstuk zijn de keuzes die zijn gemaakt in hoofdstuk 4 als uitgangspunt genomen. In het bijzonder is de uitkomstmaat in alle modellen de geüniformeerde eindtoetsscore die is beschreven in paragraaf 4.3.

### 5.1. Bivariate analyses

Voor alle mogelijk relevante kenmerken die beschikbaar zijn op de longlist uit Bijlage A is een eenvoudig lineair regressiemodel voor schoolprestaties geschat met één kenmerk tegelijk als verklarende variabele, plus een constante term. Bij categoriale kenmerken is steeds één van de categorieën als referentiecategorie gebruikt, zoals genoemd in paragraaf 4.2. Hiervoor zijn de data gebruikt van alle leerlingen die in 2022/'23 in groep 8 van het basisonderwijs zaten en voor wie alle kenmerken beschikbaar zijn, inclusief de eindtoetsscore.

Voor elk model is een klassieke F-toets uitgevoerd die toetst of de geschatte coëfficiënt(en) van het opgenomen kenmerk significant van 0 verschillen. Besloten is om in het vervolg van de analyse geen kenmerken mee te nemen waarbij de p-waarde van deze F-toets groter was dan 0,05: een kenmerk komt alleen in aanmerking voor de uiteindelijke indicator als het op zichzelf al gecorreleerd is met schoolprestaties. Bij kenmerken met verschillende beschikbare operationalisaties is besloten om versies van kenmerken te laten afvallen die een duidelijk lagere verklaarde variantie hadden dan andere versies van hetzelfde kenmerk.

Verder zijn enkele kenmerken op de longlist in dit stadium afgefallen om inhoudelijke redenen. Bij het kenmerk 'land van herkomst' is uiteindelijk besloten om de indeling van de WRR te volgen (zie ook hoofdstuk 3). Hiervan zijn twee varianten meegenomen in het vervolg: een met 7 categorieën en een met 13 categorieën. De hier gebruikte indeling wijkt op twee punten af van de oorspronkelijke indeling van de WRR. Ten eerste is besloten om het land van herkomst van de ouders ten behoeve van de indicator op dezelfde manier te definiëren als in de huidige indicator:

- Als ten minste één van de ouders Nederland als land van herkomst heeft, wordt dit het land van herkomst.

- Daarna wordt het land van herkomst van de moeder gevolgd, tenzij dit onbekend is.
- Daarna wordt het land van herkomst van de vader gevolgd, tenzij dit onbekend is.
- Ten slotte wordt het land van herkomst van het kind gevolgd.

Ten tweede is, in overleg met de WRR, besloten om Indonesië in te delen bij de categorie 'Zuidoost-Azië en de Pacific', in plaats van de categorie 'Continentaal-Germaanstalige landen' waar dit land oorspronkelijk was ingedeeld.

Tabel 5.1 geeft een overzicht van de 43 resterende kenmerken na deze stap. Indien er achter de omschrijving van een kenmerk een getal tussen haakjes staat, is dit een categoriaal kenmerk waarbij het getal het aantal categorieën aangeeft. Kenmerken met "(wel/niet)" achter de omschrijving zijn gecodeerd als binaire variabelen, waarbij de "wel"-categorie altijd de waarde 1 krijgt en de "niet"-categorie de waarde 0. De overige kenmerken zijn gedefinieerd als numerieke variabelen.

Tabel 5.1. Geselecteerde kenmerken als input voor de stapsgewijze analyse

Kenmerk
Aantal kinderen in huishouden (3)
Aantal kinderen in huishouden (6)
Asielzoeker kind (wel/niet)
Besteedbaar inkomen huishouden t.o.v. armoedegrens (4)
Besteedbaar inkomen huishouden t.o.v. armoedegrens in de laatste drie jaren (4)
Betaald werk moeder in de laatste vier jaren (12)
Betaald werk vader in de laatste vier jaren (12)
Bruto inkomen huishouden in decielen (11)
Deelpopulatie opleidingsniveau ouders (4)
Deeltijdfactor werkzame personen op huishoudniveau (10)
GGZ-kosten moeder (wel/niet)
GGZ-kosten ouders (wel/niet)
GGZ-kosten vader (wel/niet)
Land van herkomst ouders (13)
Land van herkomst ouders (8)
Leeftijd moeder (6)
Leeftijd moeder (9)
Leeftijd vader (6)
Leeftijd vader (9)
Opleidingsniveau moeder (8)
Opleidingsniveau vader (8)
Ouderlijke structuur (7)
Problematische schulden huishouden (4)
Sociaaleconomische categorie moeder (3)
Sociaaleconomische categorie vader (3)
Sociaaleconomische status buurt
Sociaaleconomische status gemeente
Sociaaleconomische status school
Sociaaleconomische status wijk
Stedelijkheid van de buurt (6)
Verblijfsduur moeder in Nederland (2)
Verblijfsduur moeder in Nederland (3)
Verblijfsduur moeder in Nederland (4)
Verblijfsduur vader in Nederland (3)
Verblijfsduur vader in Nederland (4)
Verdacht van misdrijf laatste jaar moeder (wel/niet)
Verdacht van misdrijf laatste jaar ouders (wel/niet)
Verdacht van misdrijf laatste jaar vader (wel/niet)
Verdacht van misdrijf ooit moeder (wel/niet)
Verdacht van misdrijf ooit ouders (wel/niet)
Verdacht van misdrijf ooit vader (wel/niet)
Vermogen huishouden in decielen (11)
Welvaart huishouden in decielen (11)

Aan de inhoudelijk relevante kenmerken is in Tabel 5.1 nog het kenmerk 'deelpopulatie opleidingsniveau ouders' toegevoegd. Hiermee wordt de indeling bedoeld in deelpopulaties A t/m D uit Tabel 4.4 die aangeeft welke informatie beschikbaar is over het opleidingsniveau van de ouders. Het zou nuttig kunnen zijn om dit kenmerk op te nemen in een model voor onderwijsachterstanden, om te corrigeren voor eventuele resterende selectiviteit in de

ontbrekende informatie over opleidingsniveau die niet kan worden opgevangen door andere inhoudelijke kenmerken in het model. Dat wil zeggen: het zou kunnen dat kinderen in deelpopulatie A (waar beide opleidingsniveaus zijn geïmputeerd) gemiddeld lagere toetsscores halen dan kinderen in deelpopulatie D (waar beide opleidingsniveaus zijn waargenomen), omdat er factoren zijn die het CBS niet heeft waargenomen maar die samenhangen met zowel het ontbreken van opleidingsinformatie als de schoolprestaties van leerlingen. Een dergelijk structureel verschil zou dan geschat kunnen worden via de coëfficiënten van het kenmerk 'deelpopulatie opleidingsniveau ouders'.

## 5.2. Stapsgewijze procedure

Met de lijst kenmerken uit Tabel 5.1 is een stapsgewijze procedure uitgevoerd om het best passende lineaire regressiemodel te vinden voor de eindtoetsscore. Deze procedure is uitgevoerd op dezelfde data van leerlingen in groep 8 uit schooljaar 2022/'23 die eerder zijn gebruikt in paragraaf 5.1. Deze dataset bevat 161925 leerlingen.

In elke stap wordt onderzocht of het mogelijk is om het huidige regressiemodel te verbeteren door ofwel een kenmerk toe te voegen dat nog niet in het huidige model zit, ofwel een kenmerk weg te laten uit het huidige model. Hierbij is als criterium voor 'een beter model' gebruikt: een model met een lagere BIC. In elke stap wordt het kenmerk toegevoegd of weggelaten dat leidt tot de grootste verbetering van BIC. De procedure begint met een 'leeg' model met alleen een constante term. De procedure stopt op het moment dat het huidige model een lagere BIC heeft dan elk ander model dat precies één kenmerk meer of minder bevat.

Bovenstaande aanpak is niet de enige mogelijke manier om een stapsgewijze regressie uit te voeren. In plaats van BIC kan ook een ander criterium worden gebruikt, zoals AIC of een F-toets voor geneste regressiemodellen. Er is hier gekozen voor BIC omdat de dataset groot is. Het is bekend dat de genoemde alternatieve criteria bij grote datasets vaak leiden tot grote modellen, terwijl BIC strenger is voor modellen met veel parameters [Wagenmakers and Farrell, 2004]. Ter vergelijking is ook een vereenvoudigde versie van de procedure toegepast waarbij in elke stap alleen kenmerken mogen worden toegevoegd; dit leidde in dit geval tot dezelfde uitkomsten. Er bestaat ook een variant waarbij in elke stap alleen kenmerken mogen worden verwijderd (beginnend bij een model dat alle kenmerken bevat). Deze laatste variant is hier niet toegepast, omdat het niet realistisch is om alle kenmerken uit Tabel 5.1 tegelijk op te nemen in één (standaard) lineair regressiemodel.

Om rekening te houden met de extra onzekerheid door het imputeren van opleidingsniveau, is de stapsgewijze procedure herhaald voor elk van de tien geïmputeerde versies van opleidingsniveau. Bij elke geïmputeerde versie werden exact dezelfde tien kenmerken geselecteerd in het model. De eerste acht kenmerken werden bovendien altijd geselecteerd in dezelfde volgorde; de volgorde van het negende en tiende kenmerk verschilde per geïmputeerde versie. In alle stappen van de procedure werden alleen nieuwe kenmerken toegevoegd en nooit kenmerken verwijderd. Na het toevoegen van de tiende variabele stopte de procedure omdat BIC niet verder verbeterd kon worden door het toevoegen van een elfde variabele.

Tabel 5.2 toont de tien geselecteerde variabelen, in de volgorde waarin ze zijn geselecteerd; het negende en tiende kenmerk staan hier in de volgorde die het vaakst voorkwam. Verder bevatten de laatste twee kolommen de BIC en verklaarde variantie ( $R^2$ ) na elke stap van de procedure. Gemakshalve zijn deze gerapporteerde BIC en  $R^2$  berekend op alleen de eerste van de tien geïmputeerde versies van de data.

Tabel 5.2. Resultaten van de stapsgewijze analyse

Stap	Kenmerk toegevoegd	BIC	R2
1	Opleidingsniveau moeder (8)	520338	0,108
2	Opleidingsniveau vader (8)	514577	0,140
3	Vermogen huishouden in decielen (11)	512892	0,149
4	Problematische schulden huishouden (4)	512378	0,152
5	Land van herkomst ouders (13)	511963	0,155
6	Deelpopulatie opleidingsniveau ouders (4)	511592	0,157
7	Ouderlijke structuur (7)	511275	0,159
8	Verblijfsduur moeder in Nederland (2)	511078	0,160
9	Verdacht van misdrijf laatste jaar ouders (wel/niet)	511053	0,160
10	Sociaaleconomische status school	511030	0,161

Het kenmerk dat aangeeft of een leerling geregistreerd is als asielzoeker is niet geselecteerd door de stapsgewijze procedure. We kiezen er daarom voor om dit kenmerk niet op te nemen in het model voor de herziene indicator. Niettemin is uit de ontwikkeling van de huidige indicator wel bekend dat asielzoekers een relatief hoog risico op onderwijsachterstanden hebben, en ook uit de bivariate analyse in paragraaf 5.1 bleek dat de toetscores van leerlingen met een asielstatus gemiddeld duidelijk lager zijn dan die van leerlingen zonder asielstatus. Omdat dit een relatief kleine groep betreft, komt dit echter niet tot uitdrukking in een 'optimaal' model voor de hele populatie.

In plaats daarvan stellen we voor om voor deze groep de aanpak uit de huidige indicator te handhaven. Hierbij krijgt elke geregistreerde asielzoeker in het basisonderwijs dezelfde, lage score, zodat deze bijdraagt aan de achterstandsscore van de betreffende school. Dit betekent dat het model voor onderwijsachterstanden dat hier ontwikkeld wordt alleen zal worden toegepast op niet-asielzoekers. Bij het schatten van de verdere modellen in dit hoofdstuk worden asielzoekers daarom niet meer meegenomen. Voor de volledigheid is de stapsgewijze procedure herhaald met dezelfde dataset nadat de geregistreerde asielzoekers uit het bestand verwijderd waren. Dit leidde tot dezelfde selectie van kenmerken als in Tabel 5.2.

### 5.3. Een multigroep-model

Na de analyses uit de vorige paragraaf is een laatste wijziging aangebracht in de operationalisering van het kenmerk 'Sociaaleconomische status school'. Oorspronkelijk werd dit kenmerk afgeleid door het gemiddelde te nemen van de sociaaleconomische status-score per huishouden van alle leerlingen op een school. Een probleem hierbij is dat de sociaaleconomische status niet voor alle huishoudens bekend is. Deze ontbrekende scores werden tot nu toe genegeerd, met als risico dat de gemiddelde sociaaleconomische status voor sommige scholen erg onnauwkeurig was bepaald. Het lijkt daarom onwenselijk om dit kenmerk in deze vorm op te nemen in de uiteindelijke indicator. Het kenmerk is daarom anders geoperationaliseerd, als het gemiddelde van de sociaaleconomische status-score op buurtniveau van alle leerlingen op een school. De sociaaleconomische status op buurtniveau is bijna altijd bekend en wordt bovendien regulier door het CBS gepubliceerd. De nieuwe operationalisering is daarom over de tijd heen stabiel en over scholen heen beter vergelijkbaar. Inhoudelijk is een nadeel dat de relatie met de omgeving van de leerling nu zwakker is dan wanneer de score op huishoudensniveau wordt gebruikt. Tussen de oude en nieuwe scores blijkt echter wel veel samenhang te bestaan: in 2022/'23 was de correlatie tussen de oude en nieuwe score 0,82 op schoolvestigingsniveau en 0,83 op leerlingniveau.

Na verwijdering van de leerlingen met een asielstatus blijven er 162992 leerlingen over die in schooljaar 2022/'23 in groep 8 zaten en een eindtoetsscore hebben. (N.B. Ten opzichte van de vorige paragraaf zijn enkele honderden eerder verwijderde leerlingen weer toegevoegd aan de dataset, indien ze uitsluitend ontbrekende waarden hadden op kenmerken uit Tabel 5.1 die niet geselecteerd zijn door de stapsgewijze procedure.) Op deze data is een structureel vergelijkingsmodel geschat zoals in Figuur 4.2, gebruikmakend van de tien kenmerken uit Tabel 5.2. Dit model bevat een multigroep-component voor toetsaanbieders zoals beschreven in paragraaf 4.4.

Tabel 5.3 toont de verdeling van deze leerlingen over de vijf toetsaanbieders. Verder bevat deze tabel de verklaarde variantie [ $R^2$  (volledig)] in de toetsscore door het geschatte model, zowel in totaal als per toetsaanbieder. De laatste kolom [ $R^2$  (beperkt)] wordt besproken in paragraaf 5.4.

Tabel 5.3. Aantal waarnemingen en verklaarde variantie bij multigroep-modellen

Toetsaanbieder	Aantal leerlingen	R2 (volledig)	R2 (beperkt)
Totaal	162992	0,160	0,157
CET	72105	0,171	0,168
IEP	59847	0,151	0,148
Route 8	20625	0,150	0,148
DIA	7335	0,164	0,161
AMN	3080	0,130	0,130

Tabel 5.4 bevat de geschatte coëfficiënten uit het model, met daarbij de standaardfouten. Deze uitkomsten zijn gemiddeld over tien imputatieronden voor het opleidingsniveau van beide ouders. De standaardfouten zijn gebaseerd op de totale variantie volgens de regels bij multi-pele imputatie uit [Rubin, 1987]; in deze standaardfouten is daarom rekening gehouden met de extra onzekerheid als gevolg van het ontbreken van een deel van de opleidingsniveaus.

Tabel 5.4. Geschatte coëfficiënten in het voorgestelde structurele vergelijkingsmodel

Kenmerk	Categorie	Coëfficiënt	Standaardfout
Constante		2,99	0,01
Opleidingsniveau moeder (8)	Basisonderwijs	0,00	0,00
Opleidingsniveau moeder (8)	Vmbo-b/k, mbo1	-0,06	0,02
Opleidingsniveau moeder (8)	Vmbo-g/t, havo-, vwo-onderbouw	0,11	0,02
Opleidingsniveau moeder (8)	Mbo2 en mbo3	-0,00	0,02
Opleidingsniveau moeder (8)	Mbo4	0,15	0,02
Opleidingsniveau moeder (8)	Havo, vwo	0,34	0,02
Opleidingsniveau moeder (8)	Hbo-, wo-bachelor	0,39	0,02
Opleidingsniveau moeder (8)	Hbo-, wo-master, doctor	0,63	0,02
Opleidingsniveau vader (8)	Basisonderwijs	0,00	0,00
Opleidingsniveau vader (8)	Vmbo-b/k, mbo1	0,01	0,02
Opleidingsniveau vader (8)	Vmbo-g/t, havo-, vwo-onderbouw	0,16	0,02
Opleidingsniveau vader (8)	Mbo2 en mbo3	0,04	0,02
Opleidingsniveau vader (8)	Mbo4	0,16	0,02
Opleidingsniveau vader (8)	Havo, vwo	0,39	0,02
Opleidingsniveau vader (8)	Hbo-, wo-bachelor	0,38	0,02
Opleidingsniveau vader (8)	Hbo-, wo-master, doctor	0,62	0,02
Vermogen huishouden in decielen (11)	Eerste deciel	0,00	0,00
Vermogen huishouden in decielen (11)	Tweede deciel	-0,13	0,02
Vermogen huishouden in decielen (11)	Derde deciel	-0,03	0,02
Vermogen huishouden in decielen (11)	Vierde deciel	0,05	0,02
Vermogen huishouden in decielen (11)	Vijfde deciel	0,06	0,01
Vermogen huishouden in decielen (11)	Zesde deciel	0,10	0,01
Vermogen huishouden in decielen (11)	Zevende deciel	0,16	0,01
Vermogen huishouden in decielen (11)	Achtste deciel	0,21	0,02
Vermogen huishouden in decielen (11)	Negende deciel	0,26	0,02
Vermogen huishouden in decielen (11)	Tiende deciel	0,30	0,02
Vermogen huishouden in decielen (11)	Onbekend	0,04	0,04
Problematische schulden huishouden (4)	Geen schulden in afgelopen 3 jaar	0,00	0,00
Problematische schulden huishouden (4)	1 jaar schulden in afgelopen 3 jaar	-0,15	0,01
Problematische schulden huishouden (4)	2 jaar schulden in afgelopen 3 jaar	-0,24	0,02
Problematische schulden huishouden (4)	3 jaar schulden in afgelopen 3 jaar	-0,30	0,02
Land van herkomst ouders	Nederland	0,00	0,00

Hieronder geven we een korte toelichting bij elk van de kenmerken in het model. Hierbij geldt steeds de kanttekening dat de coëfficiënten in het model rekening houden met alle kenmerken tegelijk die zijn opgenomen in het model. Bijvoorbeeld: een verschil van  $\delta$  punten tussen de coëfficiënten van categorie X en categorie Y van een bepaald kenmerk betekent dat de toetsscore van een kind in categorie X gemiddeld  $\delta$  punten afwijkt van die van een kind in categorie Y, in de hypothetische situatie dat deze kinderen op alle andere kenmerken in het model vergelijkbaar zijn.

*Constante term.* Zoals voorgesteld in het vorige hoofdstuk bevat het model een aparte constante term per toetsaanbieder. De constante term die is opgenomen in Tabel 5.4 betreft een gewogen gemiddelde van deze vijf constante termen, waarbij de toetsaanbieders meewegen volgens hun aandeel in de dataset zoals getoond in Tabel 5.3. De afstand tussen de grootste en kleinste constante term was ongeveer 0,08 punt op de schaal van de geüniformeerde toetsscore. De constante termen van de verschillende toetsaanbieders liggen in dit model dus dicht bij elkaar dan in model 3 uit hoofdstuk 4, waarschijnlijk omdat het huidige model meer achtergrondkenmerken bevat.

*Opleidingsniveau moeder en vader.* Net als in de huidige indicator zijn dit de twee kenmerken die het sterkst samenhangen met schoolprestaties. Uit de geschatte coëfficiënten blijkt, zoals verwacht, dat kinderen van ouders met een hoger opleidingsniveau gemiddeld hogere onderwijsprestaties behalen.

*Vermogen huishouden.* Volgens het model halen kinderen uit huishoudens met meer vermogen gemiddeld hogere onderwijsprestaties, zoals verwacht. De geschatte coëfficiënten laten zien dat dit effect vooral zit in een hogere verwachte toetsscore bij de hoogste decielen; in de onderste helft van de vermogensverdeling zijn de verschillen in onderwijsprestaties tussen de decielen klein.

*Problematische schulden huishouden.* Kinderen uit gezinnen die kampen met betalingsachterstanden presteren gemiddeld minder goed in het onderwijs. Volgens de geschatte coëfficiënten is dit negatieve effect groter naarmate de schulden langer duren.

*Land van herkomst ouders.* De rol van herkomst in het model lijkt iets minder belangrijk te zijn geworden ten opzichte van de huidige indicator, door het toevoegen van meer diverse andere kenmerken. Herkomst wordt nu als vijfde kenmerk geselecteerd. Niet alle geschatte coëfficiënten verschillen significant van de referentiecategorie (Nederland). Kinderen met Nederlandse ouders, of ouders uit landen die op cultureel en linguïstisch vlak verwant zijn aan Nederland, halen gemiddeld hogere onderwijsprestaties dan kinderen met ouders uit andere landen. Verder halen kinderen met ouders uit landen in Oost-Azië en in mindere mate Zuid-Azië en Zuidoost-Azië en de Pacific gemiddeld betere toetsscores dan kinderen met Nederlandse ouders die verder vergelijkbaar zijn wat de andere kenmerken in het model betreft.

*Deelpopulatie opleidingsniveau ouders.* Kinderen van wie het opleidingsniveau van één of beide ouders onbekend is bij het CBS hebben volgens het geschatte model gemiddeld een lagere toetsscore dan vergelijkbare kinderen van wie het opleidingsniveau van beide ouders bekend is (deelpopulatie D). Deze gemiddelde afwijking is het grootst voor kinderen van wie het opleidingsniveau van beide ouders onbekend is (deelpopulatie A). Zoals eerder is opgemerkt kunnen we, door dit kenmerk op te nemen in de indicator, voorkomen dat de onderwijsachterstanden in deelpopulatie A (en in mindere mate in deelpopulaties B en C) worden onderschat.

*Ouderlijke structuur.* Volgens het geschatte model presteren kinderen uit eenoudergezinnen gemiddeld minder goed op school dan kinderen die bij beide juridische ouders wonen. Voor kinderen die zonder ouders wonen is dit negatieve effect nog groter.

*Verblijfsduur moeder in Nederland.* Dit kenmerk zit ook in de huidige indicator, waarbij drie categorieën worden onderscheiden. In het nieuwe model wordt alleen nog onderscheid gemaakt tussen kinderen van moeders die korter of langer dan 15 jaar in Nederland verblijven (waarbij moeders met Nederland als herkomstland behoren tot de categorie langer dan 15 jaar). Kinderen van moeders met een verblijfsduur van 15 jaar of korter presteren volgens het geschatte model gemiddeld beter op school, nadat rekening is gehouden met de andere kenmerken die in het model zitten. In de bivariate analyse uit paragraaf 5.1 bleek dat de gemiddelde toetscore bij een verblijfsduur van 15 jaar of korter juist iets lager is dan bij een verblijfsduur van langer dan 15 jaar. Die bivariate analyse houdt echter geen rekening met verschillen tussen de twee categorieën op andere kenmerken, zoals opleidingsniveau en land van herkomst.

*Ouders in het laatste jaar verdacht van een misdrijf.* Volgens de geschatte coëfficiënt presteren kinderen met ouders die recent verdacht zijn van een misdrijf gemiddeld minder goed in het onderwijs.

*Sociaaleconomische status, gemiddelde op schoolniveau.* Dit kenmerk is te zien als een bredere vervanger van het kenmerk 'gemiddeld opleidingsniveau moeders op school' dat in de huidige indicator is opgenomen. Net als bij dat kenmerk in de huidige indicator geldt voor de sociaaleconomische status: naarmate op een school de gemiddelde sociaaleconomische status van de buurt waar kinderen wonen hoger is, zijn ook de gemiddelde onderwijsprestaties van kinderen op die school beter. Op leerlingniveau is dit effect niet groot, maar het zal relatief zwaar meewegen in een indicator voor onderwijsachterstanden op schoolniveau, omdat de bijdrage van dit kenmerk voor alle kinderen op dezelfde school gelijk is.

## 5.4. Een beperkt model

Het model uit de vorige paragraaf bevat tien kenmerken en is daarmee een stuk groter dan het model waarop de huidige indicator is gebaseerd. Het is interessant om te onderzoeken of deze uitbreiding nuttig is, of dat een eenvoudiger model zou leiden tot vergelijkbare uitkomsten. Ter vergelijking willen we daarom ook een beperkter model analyseren.

De fitmaten in Tabel 5.2 bieden geen duidelijke aanknopingspunten voor het kiezen van een beperkter model. Volgens de BIC-waarden is het voorgestelde model met tien kenmerken duidelijk beter dan alle kleinere modellen in de tabel. Het enige model dat op basis van BIC eventueel als alternatief zou kunnen worden gezien is het model met negen kenmerken. Dit model is echter nauwelijks minder groot; het bevat één coëfficiënt minder. Bovendien is het enige kenmerk dat dan zou worden weggelaten de sociaaleconomische status op schoolniveau, waarvan we juist verwachten dat deze relatief zwaar meetelt in een indicator op schoolniveau, zoals boven aangegeven.

Wat de verklaarde variantie betreft zijn alle modellen met zes of meer kenmerken uit de tabel vergelijkbaar. Aangezien er op puur statistische gronden geen duidelijke voorkeur is voor een bepaald beperkt model, stellen we voor om uit te gaan van een beperkt model met zes kenmerken. Dit model is dan qua grootte vergelijkbaar met het model uit de huidige indicator.

Tabel 5.5 bevat de geschatte coëfficiënten uit een beperkt model waarbij alleen de eerste zes achtergrondkenmerken uit Tabel 5.2 zijn geselecteerd.

Tabel 5.5. Geschatte coëfficiënten in een beperkt structureel vergelijkingsmodel

Kenmerk	Categorie	Coëfficiënt	Standaardfout
Constante		2,95	0,01
Opleidingsniveau moeder (8)	Basisonderwijs	0,00	0,00
Opleidingsniveau moeder (8)	Vmbo-b/k, mbo1	-0,08	0,02
Opleidingsniveau moeder (8)	Vmbo-g/t, havo-, vwo-onderbouw	0,09	0,02
Opleidingsniveau moeder (8)	Mbo2 en mbo3	-0,03	0,02
Opleidingsniveau moeder (8)	Mbo4	0,13	0,02
Opleidingsniveau moeder (8)	Havo, vwo	0,33	0,02
Opleidingsniveau moeder (8)	Hbo-, wo-bachelor	0,38	0,02
Opleidingsniveau moeder (8)	Hbo-, wo-master, doctor	0,62	0,02
Opleidingsniveau vader (8)	Basisonderwijs	0,00	0,00
Opleidingsniveau vader (8)	Vmbo-b/k, mbo1	0,00	0,02
Opleidingsniveau vader (8)	Vmbo-g/t, havo-, vwo-onderbouw	0,15	0,02
Opleidingsniveau vader (8)	Mbo2 en mbo3	0,04	0,02
Opleidingsniveau vader (8)	Mbo4	0,16	0,02
Opleidingsniveau vader (8)	Havo, vwo	0,39	0,02
Opleidingsniveau vader (8)	Hbo-, wo-bachelor	0,39	0,02
Opleidingsniveau vader (8)	Hbo-, wo-master, doctor	0,64	0,02
Vermogen huishouden in decielen (11)	Eerste deciel	0,00	0,00
Vermogen huishouden in decielen (11)	Tweede deciel	-0,15	0,02
Vermogen huishouden in decielen (11)	Derde deciel	-0,05	0,02
Vermogen huishouden in decielen (11)	Vierde deciel	0,06	0,02
Vermogen huishouden in decielen (11)	Vijfde deciel	0,09	0,01
Vermogen huishouden in decielen (11)	Zesde deciel	0,14	0,01
Vermogen huishouden in decielen (11)	Zevende deciel	0,20	0,01
Vermogen huishouden in decielen (11)	Achtste deciel	0,26	0,02
Vermogen huishouden in decielen (11)	Negende deciel	0,30	0,02
Vermogen huishouden in decielen (11)	Tiende deciel	0,35	0,02
Vermogen huishouden in decielen (11)	Onbekend	-0,04	0,04
Problematische schulden huishouden (4)	Geen schulden in afgelopen 3 jaar	0,00	0,00
Problematische schulden huishouden (4)	1 jaar schulden in afgelopen 3 jaar	-0,16	0,01
Problematische schulden huishouden (4)	2 jaar schulden in afgelopen 3 jaar	-0,25	0,02
Problematische schulden huishouden (4)	3 jaar schulden in afgelopen 3 jaar	-0,31	0,02
Land van herkomst ouders	Nederland	0,00	0,00

In de laatste kolom van Tabel 5.3 is te zien dat de verklaarde variantie door dit beperkte model nauwelijks lager is dan die van het volledige model; dit geldt ook voor de uitsplitsing naar toetsaanbieder. Verder zijn in Tabel 5.5 geen grote verschuivingen te zien in de coëfficiënten van de overlappende kenmerken ten opzichte van het volledige model uit Tabel 5.4. Dit laatste suggereert dat de aanvullende kenmerken in het volledige model een bijdrage hebben die grotendeels onafhankelijk is van de eerste zes kenmerken. Aangezien we voor deze aanvullende kenmerken ook een inhoudelijke onderbouwing hebben, hebben we gezien deze uitkomsten een voorkeur voor het volledige model met tien kenmerken. Wel zullen in het volgende hoofdstuk enkele aanvullende plausibiliteitsanalyses worden gedaan op dit model.

## 6. Plausibiliteitsanalyses

Om het model voor onderwijssachterstanden dat is gevonden in paragraaf 5.3 te valideren is een aantal vervolganalyses gedaan, die in dit hoofdstuk worden besproken. De volgende analyses worden achtereenvolgens behandeld:

- Er is een kruisvalidatie uitgevoerd op de stapsgewijze procedure uit paragraaf 5.2, om te zien hoe robuust de selectie van kenmerken in het model is (paragraaf 6.1).
- Als tweede controle op de robuustheid van de selectie van kenmerken is een alternatieve modelselectie-procedure toegepast, de zogenaamde LASSO (paragraaf 6.2).
- Om te zien hoe stabiel de geschatte coëfficiënten bij de geselecteerde kenmerken zijn, is het multigroep-model uit paragraaf 5.3 nogmaals geschat op data van een ander schooljaar (paragraaf 6.3).

### 6.1. Kruisvalidatie

Kruisvalidatie is een techniek waarbij een model meerdere keren geschat wordt, waarbij steeds een ander deel van de beschikbare data wordt weggelaten. In veel toepassingen is het doel van kruisvalidatie om elk geschat model vervolgens te evalueren op het weggelaten deel van de data, om zo een 'optimale' keuze te maken voor bepaalde onbekende parameters van het model. In dit geval beperken we ons ertoe om via kruisvalidatie te evalueren hoeveel variatie er zit in de selectie van kenmerken door de stapsgewijze procedure.

De beschikbare data over leerlingen die in schooljaar 2022/'23 in groep 8 zaten zijn hiertoe verdeeld in vijf willekeurige groepen van ongeveer dezelfde omvang. Vervolgens is de stapsgewijze procedure uit paragraaf 5.2 vijf keer uitgevoerd op alle data van vier van de vijf groepen, waarbij dus steeds één groep tegelijk is weggelaten. Net als in paragraaf 5.2 is binnen de stapsgewijze procedure ook weer rekening gehouden met de tien imputatieronden voor opleidingsniveau. In totaal leverde deze kruisvalidatie daarom ( $5 \times 10 =$ ) 50 selecties van kenmerken op.

De uitkomsten waren zeer eenduidig: in alle 50 gevallen heeft de stapsgewijze procedure exact dezelfde kenmerken geselecteerd, namelijk de tien kenmerken uit Tabel 5.2 die ook geselecteerd werden bij gebruik van de hele dataset. Deze selectie lijkt dus zeer robuust tegen kleine willekeurige veranderingen in de dataset. Het enige kleine verschil dat werd gevonden zat in de volgorde waarin de kenmerken zijn geselecteerd: bij vijf van de 50 gevallen werd het opleidingsniveau van de vader als eerste geselecteerd en het opleidingsniveau van de moeder als tweede, in plaats van andersom zoals in Tabel 5.2. Het belang van deze twee kenmerken ligt blijkbaar dicht bij elkaar. Verder wisselde de volgorde van het negende en tiende kenmerk, zoals ook gebeurde bij de oorspronkelijke analyse uit paragraaf 5.2; het kenmerk 'Ouders in het laatste jaar verdacht van een misdrijf' werd in 30 van de 50 gevallen als negende geselecteerd, het kenmerk 'Sociaaleconomische status, gemiddelde op schoolniveau' in de resterende 20 gevallen. Gegeven de gemaakte selectie van kenmerken heeft de volgorde waarin ze zijn geselecteerd overigens geen invloed op de geschatte coëfficiënten in Tabel 5.4 en dus ook niet op de uiteindelijke indicator voor onderwijssachterstanden. Ook voor de keuze van het beperkte model

uit paragraaf 5.4 heeft de kruisvalidatie geen gevolgen, aangezien ook alleen de eerste zes geselecteerde kenmerken in alle 50 gevallen hetzelfde waren.

## 6.2. LASSO

In een reguliere lineaire regressie worden voor alle variabelen coëfficiënten geschat. In dit onderzoek is een stapsgewijze methode toegepast om de kenmerken te selecteren die tot het best passende model leiden. Om te onderzoeken of een andere methode een vergelijkbaar model zou samenstellen is een LASSO lineaire regressie gebruikt.

Een LASSO (least absolute shrinkage and selection operator) regressie is een alternatieve methode om op een selectie van kenmerken een regressiemodel te schatten [Tibshirani, 1996]. In deze methode krijgen, door het toevoegen van een strafterm bij het schatten van de coëfficiënten, coëfficiënten van variabelen die weinig toevoegen een coëfficiënt van nul. Een kenmerk met een coëfficiënt van nul kan vervolgens uit het model gehaald worden. Door te variëren in de hoogte van de strafterm ontstaan er dan varianten van de regressiemodellen waarbij géén tot álle variabelen worden meegenomen.

Bij een LASSO regressie worden dus alle kenmerken samengenomen en vindt er eliminatie van kenmerken plaats door de invloed van de strafterm op de coëfficiënten. Dit verschilt van een stapsgewijze regressie waar het model in rondes één kenmerk toevoegt (of verwijdert). Per ronde worden dan alternatieve modellen op slechts een deel van de kenmerken geschat en wordt daarmee samenhang tussen niet-gekozen kenmerken niet meegenomen. Een risico is dat het hierdoor kan voorkomen dat individueel een kenmerk een kleinere toevoeging geeft op de modelkwaliteit, maar in samenhang met een ander, nog niet geselecteerd kenmerk, juist een grotere bijdrage zou leveren. Of dit in de praktijk optreedt is afhankelijk van de onderliggende verbanden in de data.

Doel van deze plausibiliteitsanalyse is om te bepalen of een LASSO regressie tot een wezenlijk andere selectie komt dan de stapsgewijze regressie, en zo ja, of de modelkwaliteit dan ook verschilt. Twee verschillende modellen kunnen namelijk vergelijkbaar zijn in de hoeveelheid variantie die zij verklaren.

### 6.2.1. Analyse

Voor deze analyse is gebruik gemaakt van een gegroepeerde LASSO [Yang and Zou, 2015]<sup>1</sup>. Bij een reguliere LASSO worden de verschillende niveaus van een categoriaal kenmerk als losse kenmerken meegenomen. Dan kan het voorkomen dat van een variabele als opleidingsniveau niet alle verschillende niveaus worden geselecteerd. Bij een gegroepeerde LASSO kan de onderlinge structuur van de bij elkaar horende niveaus van een kenmerk worden opgegeven, zodat deze samen wel of niet geselecteerd worden. Dit is wenselijk omdat de stapsgewijze regressie deze niveaus ook samen wel of niet selecteert en daarmee aan blijft sluiten bij de standaard uitsplitsingen van kenmerken van het CBS.

Voor de gegroepeerde LASSO zijn in de basis dezelfde kenmerken gebruikt als bij de stapsgewijze procedure, zoals beschreven in paragraaf 5.2. Hierbij zijn ook de wijzigingen meegenomen die beschreven worden in paragraaf 5.2 en 5.3, namelijk de aanpassing van 'Sociaaleconomische status school' en de uitsluiting van leerlingen met een asielstatus. Voor het uitvoeren van de LASSO analyse is met behulp van kruisvalidatie een reeks van toenemende straftermen toegepast.

---

<sup>1</sup>Voor het schatten van de gegroepeerde LASSO regressie is het R-package `gglasso` gebruikt.

De uitkomst van deze analyse zijn verschillende lineaire regressiemodellen met een aflopend aantal kenmerken. Het aantal geselecteerde kenmerken kan dan worden afgezet tegen een (kruisgevalideerde) kwaliteitsmaat (*mean square error*) om een model te selecteren.

Voor de vergelijking met de stapsgewijze regressie kijken we naar twee van deze modellen. Dit zijn het model waarvoor de kruisgevalideerde *mean square error* maximaal 1 standaardfout verschilt van het LASSO model met de laagste error (aangeduid als *1se*), en het model dat evenveel kenmerken selecteert als de stapsgewijze regressie (10 excl. het intercept).

De LASSO analyses zijn net als de stapsgewijze regressie 10 keer gedraaid, elk met een andere imputatie van de opleidingsniveauvariabelen.

## 6.2.2. Resultaten

In Tabel 6.1 zijn de geselecteerde kenmerken van zowel de stapsgewijze als de LASSO regressiemodellen opgenomen. In de kolommen wordt aangegeven hoe vaak een kenmerk geselecteerd is over de 10 imputaties van de opleidingsniveauvariabelen.

Tabel 6.1. Kenmerken geselecteerd in de lasso en stapsgewijze analyse

Kenmerk	Stapsgewijze regressie	LASSO regressie (1se)	LASSO regressie (10)
Constante	10	10	10
Opleidingsniveau moeder (8)	10	10	10
Opleidingsniveau vader (8)	10	10	10
Vermogen huishouden in decielen (11)	10	10	6
Problematische schulden huishouden (4)	10	10	10
Land van herkomst ouders (13)	10	10	
Deelpopulatie opleidingsniveau ouders (4)	10	10	10
Ouderlijke structuur (7)	10	10	
Verblijfsduur moeder in Nederland (2)	10	10	
Sociaaleconomische status school	10	10	10
Verdacht van misdrijf laatste jaar ouders (wel/niet)	10	10	
Verdacht van misdrijf ooit ouders (wel/niet)		10	
GGZ-kosten ouders (wel/niet)		10	10
GGZ-kosten vader (wel/niet)		9	
Aantal kinderen in huishouden (3)		10	
Aantal kinderen in huishouden (6)		10	
Stedelijkheid van de buurt (6)		10	
Besteedbaar inkomen huishouden t.o.v. armoedegrens (4)		10	10
Land van herkomst ouders (8)		10	
Bruto inkomen huishouden in decielen (11)		10	10
Deeltijdfactor werkzame personen op huishoudniveau (10)		10	
Sociaaleconomische categorie vader (3)		10	10
Leeftijd moeder (6)		10	
Leeftijd vader (9)		10	
Leeftijd moeder (9)		10	
R <sup>2</sup>	0.162 (0.000)	0.161 (0.000)	0.140 (0.000)

In de tabel zien we dat het beste model dat is samengesteld door de LASSO (1se) alle kenmerken selecteert die ook door de stapsgewijze regressie geselecteerd zijn. Wel kiest de LASSO daarnaast nog een reeks andere kenmerken. Als we kijken naar het LASSO model dat slechts 10 kenmerken kiest, zien we dat 5 of 6 dezelfde kenmerken worden gekozen<sup>2</sup>. Land van herkomst ouders (13), Ouderlijke structuur, Verblijfsduur moeder in Nederland en Verdacht van misdrijf laatste jaar ouders (wel/niet) worden niet gekozen. In plaats daarvan worden de kenmerken Sociaaleconomische categorie vader, Bruto inkomen huishouden in decielen, GGZ-kosten ouders (wel/niet) en Besteedbaar inkomen huishouden t.o.v. armoedegrens geselecteerd.

Onderaan in de tabel is de verklaarde variantie, de  $R^2$ , van de modellen opgenomen. Hier zien we dat de verklaarde variantie van de stapsgewijze regressie en het beste LASSO model gelijk zijn. De verklaarde variantie van het LASSO model met 10 (of 11) kenmerken is lager dan van de stapsgewijze analyse die 10 kenmerken selecteert. Dat suggereert dat de andere selectie uit dat LASSO model de data minder goed modelleert dan de 10 kenmerken van de stapsgewijze regressie.

Aan de set van kenmerken waaruit geselecteerd wordt zijn in bepaalde gevallen verschillende indikkingen van hetzelfde kenmerk toegevoegd, zoals leeftijd, herkomst en aantal kinderen in huishouden. Opvallend is dat de uitgebreidere LASSO bij deze kenmerken beide indikkingen selecteert. Dit suggereert dat specifieke subcategorieën van de verschillende indikkingen relevante informatie bevatten bij het schatten van de eindtoetscores. Dit kan erop duiden dat een andere indikking optimaal zou zijn. Het heeft echter over het algemeen de voorkeur om aan te sluiten bij standaard CBS indelingen.

Doordat de selectie van kenmerken in de LASSO plaatsvindt bij het schatten van de coëfficiënten zijn deze niet direct vergelijkbaar met die van een reguliere of stapsgewijze regressie. Deze hebben we daarom hier niet naast elkaar gezet.

Op basis van de literatuur zouden we verwachten dat de LASSO geen slechter model oplevert dan de stapsgewijze regressie, zoals hier het geval voor het LASSO model met 10 kenmerken [McNeish, 2015], al is dat dan niet gebaseerd op een gegroepeerde LASSO. Het is mogelijk dat voor de categoriale kenmerken zoals hier gebruikt de stapsgewijze regressie geschikter is.

De belangrijkste conclusie van deze robuustheidsanalyse is dat op basis van de uitkomstmaat, in vorm van de verklaarde variantie, er geen aanleiding is om aan te nemen dat er met de selectie van de stapsgewijze methode nog wezenlijke (lineaire) verbanden in de onderliggende data worden gemist.

### 6.3. Vergelijking met een ander cohort

Bij wijze van laatste plausibiliteitsanalyse is het model uit paragraaf 5.3 opnieuw geschat op data van een ander cohort leerlingen, zodat de geschatte coëfficiënten kunnen worden vergeleken. Het oorspronkelijke model is geschat op eindtoetsgegevens uit schooljaar 2022/'23. Bij de keuze voor een ander schooljaar als vergelijking speelt een aantal praktische beperkingen:

- Vanaf het schooljaar 2023/'24 is de eindtoets vervangen door een doorstroomtoets. Omwille van vergelijkbaarheid van de data vallen alle schooljaren na 2022/'23 daarom af.

---

<sup>2</sup>Bij de LASSO kan niet worden aangegeven hoeveel kenmerken er worden geselecteerd, dus het is mogelijk dat er geen model is met exact 10 kenmerken.

- Om een vergelijkbare reden vallen ook de schooljaren vóór 2018/'19 af, omdat pas vanaf dat schooljaar een tabel zoals Tabel 4.1 beschikbaar is om de toetsen van verschillende aanbieders te vergelijken. De geüniformeerde toetsscore uit ons model kan dus alleen vanaf dat schooljaar berekend worden.
- Opnieuw omwille van de vergelijkbaarheid willen we geen schooljaar kiezen waarin de eindtoets niet op de gebruikelijke manier kon worden afgenomen door de Covid19-pandemie. De schooljaren 2019/'20 en 2020/'21 vallen daarom af.

De enige schooljaren die overblijven voor een mogelijke vergelijking zijn daarom 2018/'19 en 2021/'22. We kiezen voor schooljaar 2021/'22 als validatiecohort, zodat de afstand in de tijd tussen de twee cohorten zo klein mogelijk is. Voor dit schooljaar zijn, als voorheen, de data geselecteerd van leerlingen in groep 8 die een eindtoets hebben gemaakt. De eindtoetsscores zijn geüniformeerd met de methode uit paragraaf 4.3, gebruikmakend van de versie van Tabel 4.1 die van toepassing was op schooljaar 2021/'22. Vervolgens is het model met de tien kenmerken uit Tabel 5.2 geschat op deze data.

Bij elke geschatte coëfficiënt is het verschil (Delta) berekend tussen de schatting uit 2021/'22 en 2022/'23. Bij dit verschil is ook een standaardfout berekend, onder de aanname dat de schattingen uit de twee jaren onafhankelijk van elkaar zijn. Aangezien het gaat om twee vrijwel volledig verschillende leerlingpopulaties in groep 8, lijkt deze aanname niet onredelijk. De standaardfout van het verschil tussen twee coëfficiënten ( $\Delta = \text{coef1} - \text{coef2}$ ) kan in dat geval worden berekend met de formule

$$SE(\Delta) = \sqrt{[SE(\text{coef1})]^2 + [SE(\text{coef2})]^2}.$$

Tabel 6.2 toont voor de coëfficiënten uit het voorgestelde model met tien kenmerken de gevonden verschillen, hun standaardfouten en de bijbehorende p-waarde van een t-toets dat er geen verschil is.

Tabel 6.2. Verschil tussen geschatte coëfficiënten in het voorgestelde structurele vergelijkingsmodel (Delta = schooljaar 2021/'22 - schooljaar 2022/'23)

Kenmerk	Categorie	Delta	SE (Delta)	p-waarde (Delta)
Constante		0,05	0,02	0,01
Opleidingsniveau moeder (8)	Basisonderwijs	0,00	0,00	
Opleidingsniveau moeder (8)	Vmbo-b/k, mbo1	-0,03	0,03	0,28
Opleidingsniveau moeder (8)	Vmbo-g/t, havo-, vwo-onderbouw	-0,03	0,03	0,39
Opleidingsniveau moeder (8)	Mbo2 en mbo3	-0,00	0,02	0,96
Opleidingsniveau moeder (8)	Mbo4	-0,02	0,02	0,37
Opleidingsniveau moeder (8)	Havo, vwo	-0,02	0,03	0,44
Opleidingsniveau moeder (8)	Hbo-, wo-bachelor	-0,02	0,02	0,52
Opleidingsniveau moeder (8)	Hbo-, wo-master, doctor	-0,04	0,03	0,13
Opleidingsniveau vader (8)	Basisonderwijs	0,00	0,00	
Opleidingsniveau vader (8)	Vmbo-b/k, mbo1	0,01	0,03	0,60
Opleidingsniveau vader (8)	Vmbo-g/t, havo-, vwo-onderbouw	0,01	0,03	0,71
Opleidingsniveau vader (8)	Mbo2 en mbo3	0,02	0,02	0,50
Opleidingsniveau vader (8)	Mbo4	0,02	0,03	0,33
Opleidingsniveau vader (8)	Havo, vwo	-0,01	0,03	0,65
Opleidingsniveau vader (8)	Hbo-, wo-bachelor	0,01	0,02	0,73
Opleidingsniveau vader (8)	Hbo-, wo-master, doctor	0,01	0,03	0,85
Vermogen huishouden in decielen (11)	Eerste deciel	0,00	0,00	
Vermogen huishouden in decielen (11)	Tweede deciel	-0,02	0,02	0,54
Vermogen huishouden in decielen (11)	Derde deciel	0,00	0,02	0,85
Vermogen huishouden in decielen (11)	Vierde deciel	-0,01	0,02	0,77
Vermogen huishouden in decielen (11)	Vijfde deciel	0,02	0,02	0,41
Vermogen huishouden in decielen (11)	Zesde deciel	0,02	0,02	0,30
Vermogen huishouden in decielen (11)	Zevende deciel	0,00	0,02	0,90
Vermogen huishouden in decielen (11)	Achtste deciel	0,02	0,02	0,27
Vermogen huishouden in decielen (11)	Negende deciel	0,01	0,02	0,79
Vermogen huishouden in decielen (11)	Tiende deciel	0,00	0,02	0,83
Vermogen huishouden in decielen (11)	Onbekend	0,04	0,06	0,55
Problematische schulden huishouden (4)	Geen schulden in afgelopen 3 jaar	0,00	0,00	
Problematische schulden huishouden (4)	1 jaar schulden in afgelopen 3 jaar	-0,01	0,02	0,54
Problematische schulden huishouden (4)	2 jaar schulden in afgelopen 3 jaar	-0,00	0,03	0,89

Het is duidelijk dat de meeste coëfficiënten nauwelijks verschillen tussen de twee jaren. Bij een aantal coëfficiënten verschilt de geschatte waarde met enkele honderdste punten, maar in de meeste gevallen zijn ook deze verschillen nog klein in vergelijking met hun standaardfout. Deze verschillen wijzen dus meer op de onzekerheid waarmee de coëfficiënten per jaar geschat zijn, dan op werkelijke verschillen tussen de jaren. De kleinste p-waarde ( $p = 0,01$ ) doet zich voor bij de constante term. Verder komen drie p-waarden kleiner dan 0,05 voor bij de categorie 'Midden- en Oost-Europa' van land van herkomst ( $p = 0,01$ ), bij de deelpopulatie waar het opleidingsniveau van beide ouders geïmputeerd is ( $p = 0,03$ ) en bij de categorie 'Woont bij vader en partner' van ouderlijke structuur ( $p = 0,03$ ). Gezien het grote aantal vergelijkingen dat hier is gemaakt (in totaal 52 coëfficiënten wanneer de referentiecategorieën niet worden meegeteld), is echter te verwachten dat er enkele p-waarden kleiner dan 0,05 voorkomen, ook als er geen werkelijke verschillen zouden bestaan.

We concluderen dat het gevonden model voor schooljaar 2022/'23 ook goed toepasbaar lijkt op een ander jaar. Er zijn geen aanwijzingen gevonden dat het model te uitgebreid is en daarom te specifiek geschat is voor dit ene jaar.

## 7. Herverdeeleffecten

In dit hoofdstuk analyseren we de mogelijke herverdeeleffecten die ontstaan als de herijkte methode wordt gebruikt ten opzichte van de huidige methode. Dit doen we voor schooljaar 2021/'22 eerst op schoolniveau en vervolgens op gemeenteniveau. Onder herverdeeleffecten verstaan we in dit geval verschillen in financiering van scholen op gemeenten puur als gevolg van de verandering van methode. Dat wil zeggen dat we in dit hoofdstuk aannemen dat de beleidskeuzes omtrent de financieringsystematiek verder hetzelfde blijven.

### 7.1. Verschillen tussen de herijkte en huidige methode

De herijking van de methode om het risico op onderwijsachterstanden vast te stellen heeft tot een aantal veranderingen geleid die invloed hebben op de resultaten en daardoor bijdragen aan de herverdeeleffecten. Zoals in hoofdstuk 3 en 4 uitgebreid is toegelicht, is het model om het risico op onderwijsachterstanden te voorspellen aanzienlijk veranderd. Daarnaast is de imputatiemethode van het opleidingsniveau van de ouders) verbeterd. Net als in de huidige methode worden bij de herijkte methode tien iteraties van het opleidingsniveau geïmputeerd. Vervolgens wordt voor elke iteratie een onderwijsscore uitgerekend voor alle kinderen waarvoor de modelvariabelen gevuld zijn. In de huidige methode worden deze tien onderwijsscores direct gemiddeld naar één onderwijsscore, die vervolgens gebruikt wordt om de achterstandsscore op school of gemeenteniveau te bepalen. In de herijkte methode wordt deze middeling uitgesteld: de tien onderwijsscores worden gebruikt om tien achterstandsscores op school of gemeenteniveau uit te rekenen. Die worden vervolgens gemiddeld tot een enkele achterstandsscore. Deze methode houdt beter rekening met de kansverdeling van de imputatie van opleidingsniveau [CBS, 2024b].

Tot slot is de imputatie van de onderwijsscores verbeterd. De vernieuwde imputatiemethode van de onderwijsscores zal uitgebreid worden beschreven in een nog te publiceren rapport. In het kort houdt de verbeterde methode in dat verschillende voorspelmodellen worden gebruikt om de onderwijsscore te voorspellen, afhankelijk van welke gegevens van het kind ontbreken. In heel Nederland wordt op basis van die voorspelmodellen het meest vergelijkbare kind gevonden om de onderwijsscore over te nemen (de zogenaamde "donor"). Als een donor te vaak hergebruikt wordt, worden alle onderwijsscores die op basis van dat kind zijn geïmputeerd opnieuw geïmputeerd op basis van een steekproef van de 25 meest vergelijkbare kinderen getrokken uit de hele populatie. Voor leerlingen die niet in de BRP staan wordt de score bepaald op basis van een willekeurige leerling op dezelfde school. Als dezelfde donor te vaak wordt gebruikt wordt ook hier uitgeweken naar heel Nederland. Net als in de huidige methode krijgen kinderen met een asielstatus de gemiddelde onderwijsscore van een kind onder de grenswaarde geïmputeerd.

## 7.2. Herverdeeleffecten op schoolniveau

In schooljaar 2021/'22 waren er 6198 reguliere basisscholen. In onderstaande analyses worden al die scholen vergeleken in het percentage van het onderwijsachterstandenbudget dat ze kregen onder de huidige methode tegenover wat ze zouden krijgen onder de herijkte methode. Verschillen zijn uitgedrukt in procentpunt. De achterstandsscores zonder verwerkte fusies en bezwaren zijn als referentie gebruikt, zodat de populaties zoveel mogelijk overeenkomen. Dit wijkt iets af van de verdeling van het budget zoals dat in werkelijkheid plaatsvond in 2023, op basis van de berekende achterstandsscores van schooljaar 2021/'22.

### 7.2.1. Verschillen in financieringsstatus tussen de methoden

Tabel 7.1 toont het verschil tussen de oude en de herijkte methode in de financieringsstatus van scholen. Dat wil zeggen: het aantal scholen dat al dan niet gelden ontvangt uit het onderwijsachterstandenbudget. Onder de herijkte methode zouden 3195 van de 6198 scholen geld krijgen. Dat aantal is 564 hoger dan het aantal dat daadwerkelijk geld ontving onder de huidige methode. Dat is een toename van ongeveer 9 procent. Ruim 10 procent van het totale aantal scholen ontving onder de huidige methode geen financiering, en onder de herijkte methode wel. Ongeveer 2,5 procent van het budget gaat onder de herijkte methode naar deze scholen. Daarnaast kreeg ruim 2 procent van de scholen onder de huidige methode wel financiering maar zou dat onder de herijkte methode niet meer krijgen. Deze scholen ontvingen samen minder dan 1 procent van het budget. Samen genomen betekent dit dat meer scholen met de herijkte methode geld ontvangen uit het onderwijsachterstandenbudget, maar dat de hoeveelheid geld die scholen gemiddeld ontvangen iets afneemt doordat het budget verspreid wordt over meer scholen.

Tabel 7.1. Budget analyse huidige en herijkte methode

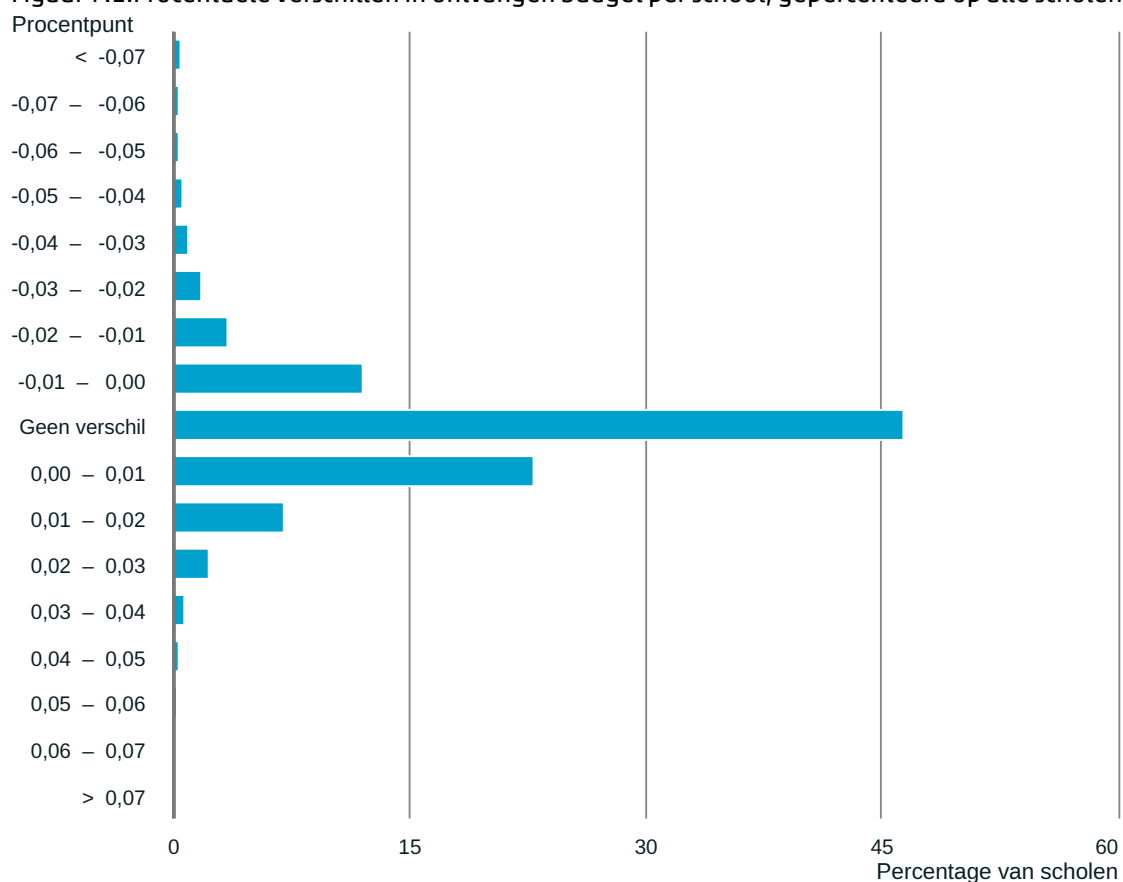
	Scholen	Scholen (%)	Budget (%)
Eerst geen budget, nu wel	693	11,18	2,48
Eerst wel budget, nu niet	129	2,08	0,64

### 7.2.2. Verdeling van verschillen in ontvangen percentage van budget op schoolniveau

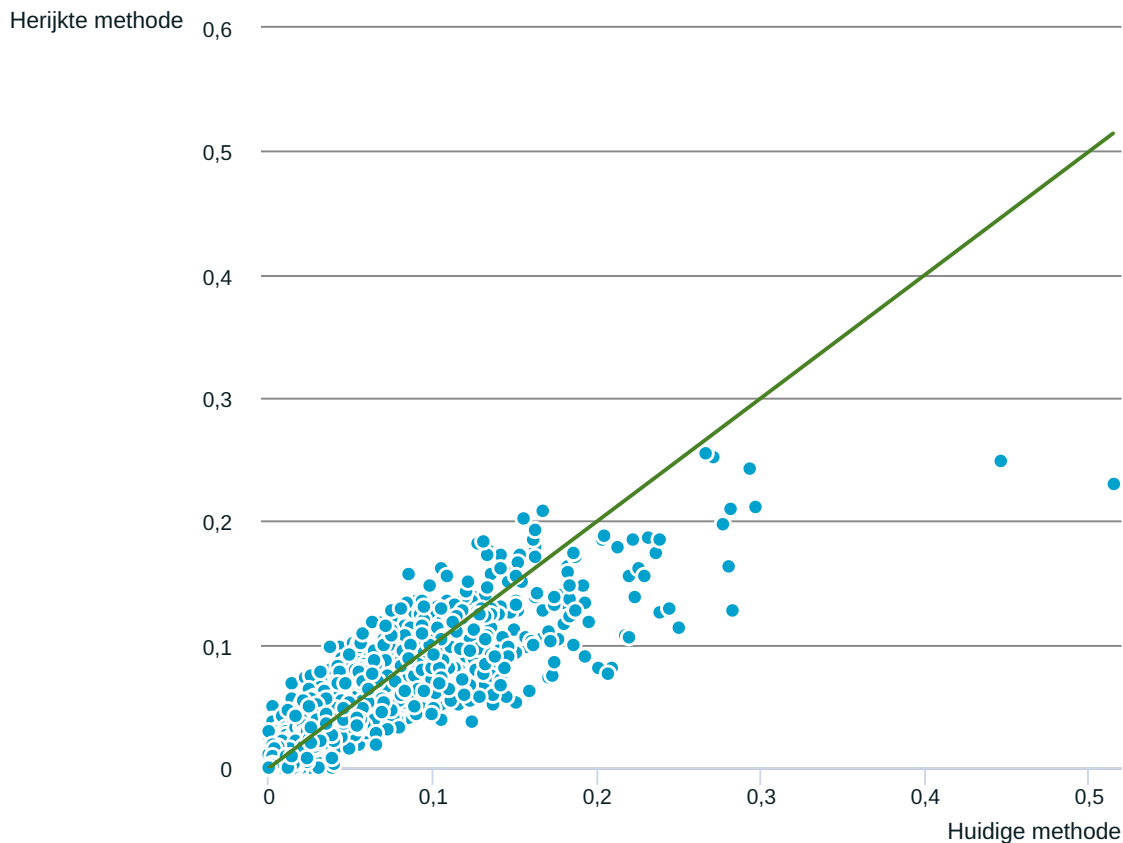
Naast de financieringsstatus is onderzocht in hoeverre het aandeel van het budget dat scholen ontvangen verschilt tussen de methodes. Dat verschil is zo uitgerekend dat een positief verschil betekent dat een school meer financiering ontvangt onder de herijkte methode ten opzichte van de huidige methode. De uitkomst is te zien in Figuur 7.1. Voor bijna de helft van de scholen is er geen verschil tussen de methodes: dit zijn scholen die onder beide methodes geen financiering ontvangen. Daarnaast is te zien dat er meer scholen zijn die een bescheiden (0-0,03 procentpunt) positief verschil tonen dan een bescheiden negatief verschil. Vanaf 0,03 procentpunt draait het verschil om: er zijn in die categorieën meer scholen met een negatief verschil dan een positief verschil. Over het algemeen geldt dat grotere verschillen minder vaak voorkomen dan kleinere verschillen, maar er zijn relatief veel scholen die een relatief groot negatief (negatiever dan 0,07 procentpunt) verschil hebben. Dit is consistent met de versnippering die eerder werd beschreven.

Hetzelfde patroon is in Figuur 7.2 duidelijker te zien. Scholen die precies op de diagonale lijn liggen, zouden onder beide methodes hetzelfde percentage van het budget ontvangen. Als alle scholen precies op de lijn zouden liggen, zou er dus geen verschil zijn tussen de twee methodes. Scholen die boven de lijn liggen, krijgen onder de herijkte methode een hoger percentage van het budget dan onder de huidige methode, en vice versa voor scholen die onder de lijn liggen. In de eerste instantie valt op dat scholen veelal relatief dichtbij de lijn verspreid zijn. Dit wijst op een hoge mate van samenhang tussen de twee methodes. Hoewel de financiering die een school ontvangt dus per methode kan verschillen, geldt over het algemeen dat scholen die bij één methode een groot of klein percentage van het budget ontvangen, dat bij de andere methode ook doen. Uit de grafiek is ook af te lezen dat de sterkste daling in het aandeel van het budget plaatsvindt onder scholen die relatief veel gelden ontvingen. Over het geheel genomen is de spreiding van budgetten onder de herijkte methode dan ook kleiner dan onder de huidige methode.

**Figuur 7.1.** Procentuele verschillen in ontvangen budget per school, gepercenteerd op alle scholen



Figuur 7.2. Percentage budget van scholen, huidige versus herijkte methode



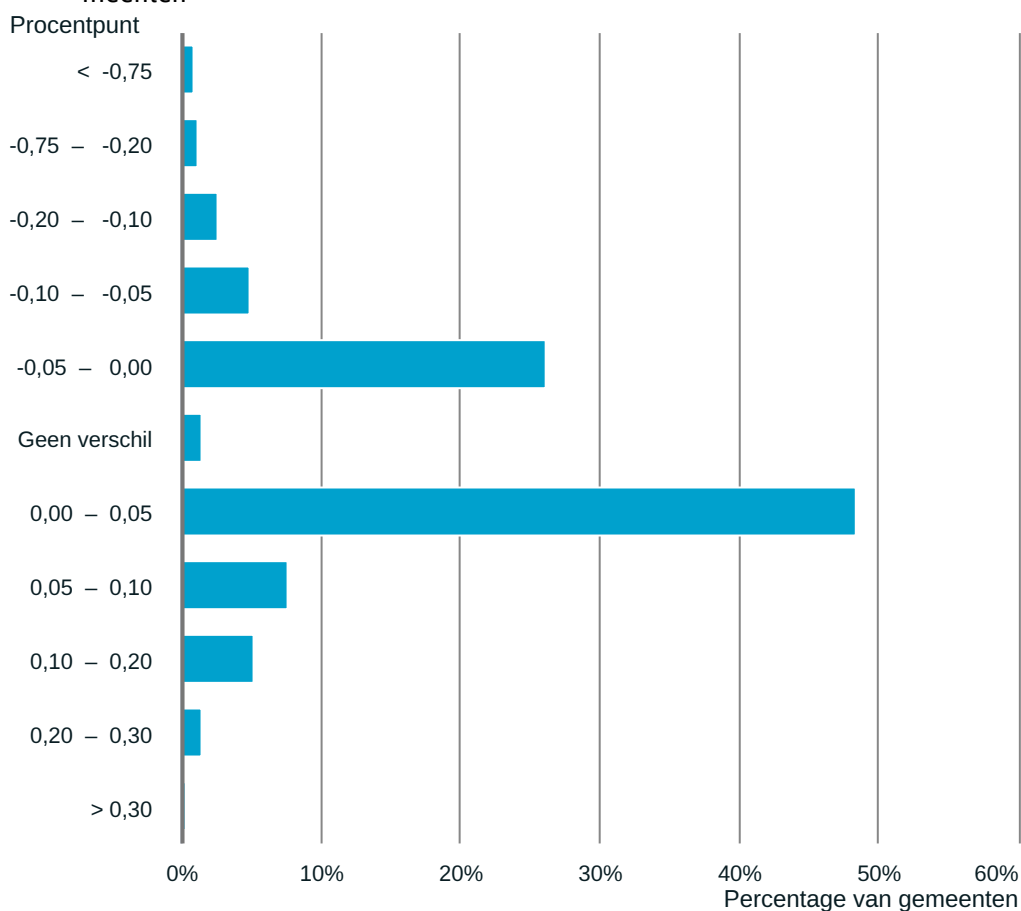
### 7.3. Herverdeeleffecten voor gemeenten

In deze paragraaf worden de herverdeeleffecten op gemeenteniveau geanalyseerd. In 2022 waren er 345 gemeenten die bijna allemaal geld ontvingen vanuit het onderwijsachterstandenbudget. Net als bij de analyses op schoolniveau hebben we de herverdeeleffecten onderzocht in termen van het percentage van het budget dat gemeenten ontvingen, uitgaande van gelijke beleidskeuzes.

#### 7.3.1. Verdeling van verschillen in ontvangen percentage van budget op gemeenteniveau

Figuur 7.3 toont de verdeling van de verschillen in ontvangen budget van gemeenten in procentpunt van het totale budget. Anders dan bij scholen zijn er maar weinig gemeenten waar geen verschil optreedt. Dit is omdat er onder de huidige methode maar 11 gemeenten zijn die niets ontvangen. Onder de herijkte methode zouden dit er 9 zijn. Er zijn 8 gemeenten die onder beide methodes geen financiering ontvangen. Net als bij de scholen zijn er meer gemeenten die onder de herijkte methode meer financiering ontvangen dan onder de huidige methode dan gemeenten die minder ontvangen. Daar staat tegenover dat bij een aantal gemeenten een groter negatief verschil optreedt tussen de methoden. Over het algemeen lijkt het beeld bij de gemeenten dan ook op het beeld van de scholen.

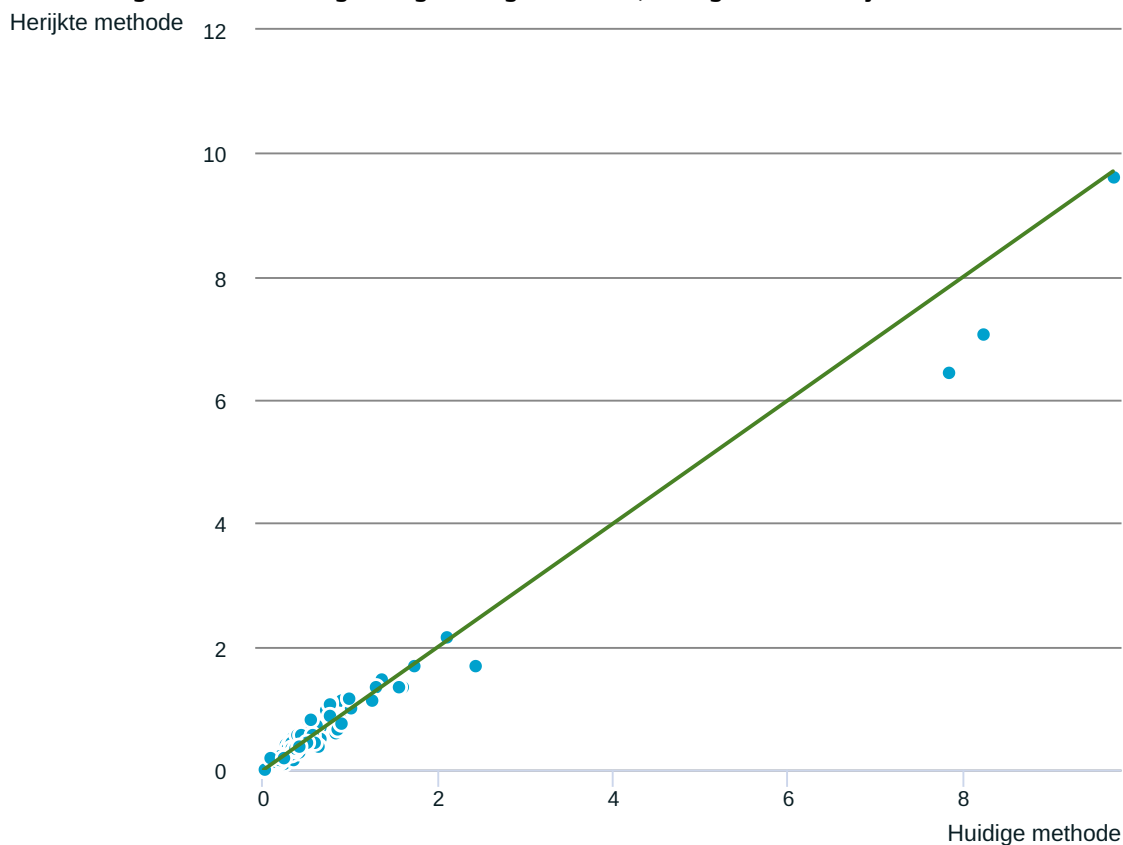
Figuur 7.3. Procentuele verschillen in ontvangen budget per gemeente, gepercenteerd op alle gemeenten



## 7.4. Verandering in percentage van het ontvangen budget per gemeente

In het algemene beeld uit de vorige paragraaf is te zien dat de meeste verschillen tussen de methoden licht positief zijn, maar dat bij een aantal gemeenten grotere negatieve verschillen plaatsvinden. In Figuur 7.4 is die verdeling op meer gedetailleerd niveau zichtbaar. Een aantal gemeenten, die veelal een relatief groot aandeel van het budget ontvingen, krijgen minder financiering op basis van de herijkte methode dan op basis van de huidige methode. Opvallend daarbij is wel dat de gemeente die het grootste aandeel van het budget ontving, ongeveer evenveel ontvangt bij beide methodes. De positieve verschillen die in de vorige paragraaf werden geconstateerd zijn met name te zien onder gemeenten die rond de 1 procent van het budget of minder kregen op basis van de huidige methode. Net als bij de scholen lijkt de spreiding van het ontvangen budget op basis van de herijkte methode wat lager dan op basis van de huidige methode.

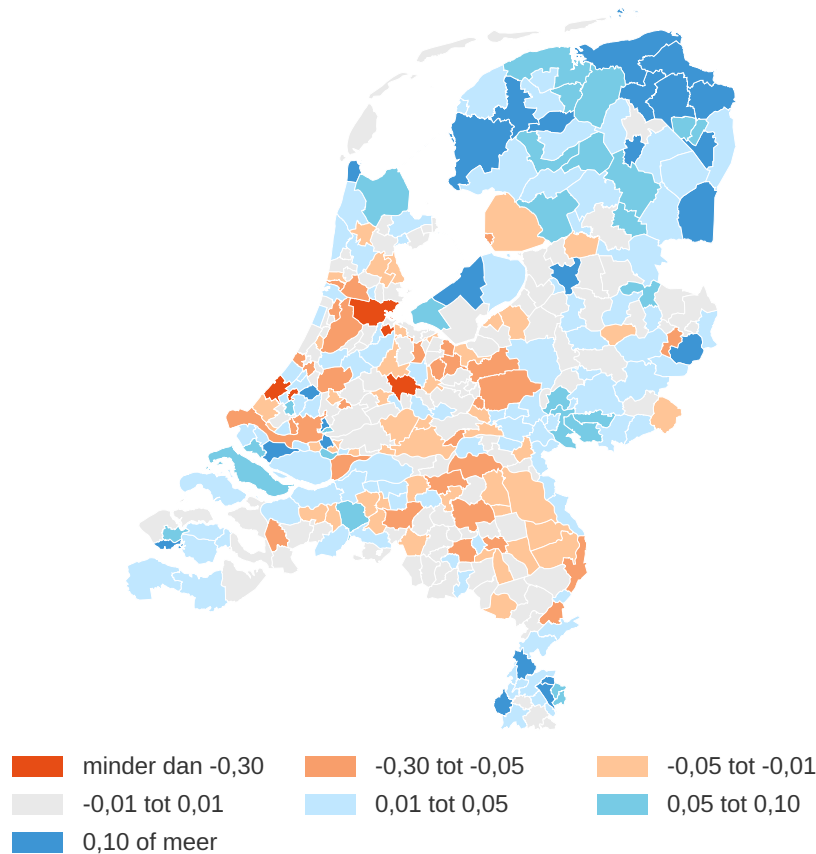
Figuur 7.4. Percentage budget van gemeenten, huidige versus herijkte methode



## 7.5. Geografische verdeling van herverdeeleffecten

Figuur 7.5 toont de verschillen tussen de methodes op gemeenteniveau en illustreert het algemene beeld van de verschuiving van financiering. Over het algemeen is te zien dat de vier gemeenten (de G4) die het meeste geld ontvangen onder de huidige methode, minder geld ontvangen onder de herijkte methode. Rotterdam, de gemeente die het grootste aandeel heeft, is echter nagenoeg gelijk onder beide methodes. Geografisch is verder te zien dat met name het noorden van Nederland, zuid-Limburg en Zeeland onder de herijkte methode een groter aandeel van het budget zouden ontvangen dan onder de huidige methode het geval is. In veel gemeentes in midden en zuid Nederland ligt het verschil tussen de methodes rond de nul of is het verschil negatief. Zoals uit hoofdstuk 5 opgemaakt kan worden is dit waarschijnlijk het gevolg van het feit dat het nieuwe model variabelen bevat die meer verspreid zijn over mensen in Nederland, en die meer verschillende oorzaken van onderwijsachterstanden omvatten. Hierdoor wordt de financiering ook geografisch meer door Nederland verspreid.

Figuur 7.5. Verschil tussen methodes per gemeente in procentpunt



## 7.6. Conclusie

In dit hoofdstuk zijn de herverdeeleffecten op school- en gemeenteniveau beschreven van de herijkte methode om het risico op onderwijsachterstanden te schatten ten opzichte van te huidige methode. Op schoolniveau blijkt dat meer scholen een deel van het budget zouden krijgen onder de herijkte methode dan onder de huidige methode, waardoor het budget uitgesmeerd wordt over meer scholen. Veel scholen krijgen met de herijkte methode een iets groter aandeel van het budget. Daar staat tegenover dat een aantal scholen onder de herijkte methode relatief veel minder financiering krijgen dan onder de huidige methode. Dit zijn met name scholen die onder de huidige methode een relatief groot aandeel van het budget ontvingen. Over het geheel genomen betekent dit dat het budget gelijkmatiger over scholen is verdeeld bij de herijkte methode dan bij de huidige methode. Op gemeenteniveau is ook te zien dat het budget onder de herijkte methode wat gelijkmatiger verspreid is. Dit komt met name doordat een aantal van de grootste gemeenten een kleiner aandeel van het budget ontvangen, en een relatief groot aantal gemeenten een wat groter aandeel ontvangen. Ook verandert de geografische spreiding van het budget: terwijl met name de grootste gemeenten, en gemeenten in het midden en zuiden van Nederland een kleiner aandeel ontvangen, krijgen gemeenten in het noorden van Nederland, in Zeeland en Zuid-Limburg juist meer. Dit is vermoedelijk het gevolg van de variabelen die in het model zitten dat gebruikt wordt voor het schatten van het risico op onderwijsachterstanden. Hierdoor ligt de nadruk op een grotere verscheidenheid van factoren die kunnen bijdragen aan een verhoogd risico op een onderwijsachterstand. Deze factoren zijn ook meer verspreid onder de bevolking van Nederland, en daarom ook over het land.

## 8. Samenvatting en conclusie

Al jarenlang staat het bevorderen van gelijke kansen hoog op de agenda van beleidsmakers, politici en wetenschappers, waarmee dit een belangrijk en relevant thema is in het publieke debat in Nederland. Het uitgangspunt is daarbij veelal dat het niet mag uitmaken waar je bent geboren, in welk gezin je opgroeit of wie je ouders zijn. Sinds de jaren zeventig voert het ministerie van Onderwijs, Cultuur en Wetenschap (OCW) beleid om de onderwijskansen van kinderen afkomstig uit achterstands- en kwetsbare milieus te bevorderen. Ter ondersteuning van dit beleid gebruikt OCW sinds 2018 de risico-indicator onderwijsachterstanden die door het CBS in opdracht van OCW is ontwikkeld. Om de risico-indicator onderwijsachterstanden actueel te houden, heeft OCW te kennen gegeven de indicator regelmatig te willen evalueren. Hierdoor kunnen nieuwe ontwikkelingen mee worden genomen in de indicator.

In voorgaande hoofdstukken en een eerder rapport (CBS, 2024b) hebben we beschreven hoe we tot een nieuwe aanpak voor de risico-indicator onderwijsachterstanden zijn gekomen. Daarbij zijn alle aspecten van de indicator opnieuw bekeken: de gebruikte achtergrondkenmerken, de wijze van modelontwikkeling en de omgang met ontbrekende waarden. In hoofdstuk 2 hebben we stilgestaan bij de wijze waarop we onderwijsachterstanden definiëren en de consequenties daarvan voor het schatten van de invloed van achtergrondkenmerken in relatie tot leerprestaties en de potentie van het kind. Verder hebben we in dit hoofdstuk gekeken naar relevante achtergrondkenmerken vanuit drie verschillende omgevingen: de gezinsomgeving, de schoolomgeving en de fysieke omgeving van de buurt of wijk waar een kind opgroeit. Vervolgens hebben we in hoofdstuk 3 beschreven welke data we hebben gebruikt voor de modelontwikkeling en welke keuzes we daarin hebben gemaakt.

De keuzes die we hebben gemaakt bij de modelontwikkeling en de resultaten hebben we beschreven in de hoofdstukken 4 en 5. Volgens de huidige inzichten is het beter om de model niet te baseren op alleen de direct effecten van de achtergrondkenmerken, maar op de totale effecten inclusief de indirecte effecten die lopen via de potentie van het kind. Omdat we alleen geïnteresseerd zijn in de totale effecten gebruiken we voor het uiteindelijke model een vereenvoudigd structureel vergelijkingsmodel waarbij de potentie van kind niet expliciet is meegenomen. Het bijkomend voordeel van die keuze is dat we daardoor de beschikking hebben over een grotere steekproef. In het uiteindelijke model gebruiken we 10 achtergrondkenmerken om het risico op onderwijsachterstand te schatten: het hoogstbehaalde opleidingsniveau van beide ouders, het vermogen van het huishouden, de aanwezigheid van problematische schulden, herkomst, de ouderlijke structuur, de verblijfsduur van de moeder in Nederland, of de ouders verdacht zijn van een misdrijf en de sociaaleconomische status van de schoolpopulatie. Omdat de ontbrekende waarden bij de opleidingsniveau variabelen niet 'missing at random' zijn, hebben we ook een kenmerk toegevoegd wat hiervoor een indicatie geeft. Dit kenmerk laat zien dat de mate van bekendheid van het hoogstbehaalde opleidingsniveau van de ouders een aanvullende indicatie is van (on)gunstige omgevingskenmerken.

Om de validiteit van de uitkomsten te testen hebben we in hoofdstuk 6 de uitkomsten van een aantal plausibiliteitsanalyses beschreven. Ten eerste hebben we middels kruisvalidaties gekeken of dezelfde achtergrondkenmerken worden geselecteerd als de analyse wordt toegepast op willekeurige steekproeven van de modelpopulatie. In de uitkomsten zien we dat consequent dezelfde achtergrondkenmerken worden geselecteerd. Daarnaast hebben we middels een

'grouped lasso' regressie gekeken of deze techniek dezelfde kenmerken selecteert. De resultaten liet zien dat de 'grouped lasso' in ieder geval dezelfde kenmerken selecteert als de stepwise regressie. De 'grouped lasso' procedure selecteert daarbovenop nog een extra set aan variabelen; dit zorgt echter niet voor een beter model. Als laatste hebben we het structurele vergelijkingsmodel geschat voor een ander cohort. Als we vervolgens dit model vergelijken met het basiscohort, dan zien we geen onredelijk grote verschillen. Al deze uitkomsten geven aan dat het ontwikkelde model robuust is.

Vergeleken met de huidige indicator verandert deze nieuwe indicator aanzienlijk qua gebruikte achtergrondkenmerken. Om die reden hebben we ook gekeken naar herverdeeleffecten van de nieuwe indicator ten opzichte van de oude indicator op school en gemeenteniveau. In die analyses is onderzocht hoe de verdeling van het onderwijsachterstanden budget zou verschillen puur op basis van de gebruikte methode, dus bij gelijke beleidskeuzes wat betreft de financieringsystematiek. Het algemene beeld dat uit die analyses naar voren komt is dat onder de herijkte methode meer scholen en gemeenten financiering krijgen via de onderwijsachterstandenregeling dan bij de oude indicator het geval was.

Verder is er een hoge samenhang tussen de hoeveelheid financiering die scholen en gemeenten ontvangen onder de twee methodes. Ongeveer 10% van de scholen die in 2021/'22 onder de oude methode geen financiering ontving, zou die volgens de herijkte methode wel ontvangen. Daar staat tegenover dat een kleine groep scholen er relatief veel op achteruit zou gaan. De financiering is onder de herijkte methode dan ook gelijkmatiger verdeeld.

Op gemeenteniveau zien we dat met name de G4 onder de herijkte methode minder financiering ontvangen, en gebieden in noordoost Nederland, Zeeland en Zuid-Limburg juist meer. Over het algemeen is de verdeling van het budget onder de herijkte methode dan ook meer verspreid over Nederland. We vermoeden dat dit komt doordat de variabelen die aan het herijkte model zijn toegevoegd meer verspreid zijn over Nederland en ook meer facetten van het risico op een onderwijsachterstand omvatten. De conclusie is dat de herverdeeleffecten zowel in ordergrootte als in spreiding over scholen en gemeenten in de lijn der verwachting vallen.

# Referenties

- Mariam Badou and Mehmet Day. Kansengelijkheid in het onderwijs. *Verwey-Jonker Instituut*, 2021.
- Chris Bluemink, Tessa Jenniskens, Annemarie Van Langen, Bianca Leest, and Maarten H. J. Wolbers. Underachievement among disadvantaged pupils in Dutch primary education. *Irish Educational Studies*, 45(2):301–319, April 2026. ISSN 0332-3315, 1747-4965. doi: 10.1080/03323315.2024.2430291. URL <https://www.tandfonline.com/doi/full/10.1080/03323315.2024.2430291>.
- Thijs Bol. Inequality in homeschooling during the corona crisis in the netherlands. first results from the liss panel, 2020.
- Kenneth A. Bollen. *Structural equations with latent variables*. Wiley, 1 edition, April 1989. ISBN 9780471011712 9781118619179. doi: 10.1002/9781118619179. URL <https://onlinelibrary.wiley.com/doi/book/10.1002/9781118619179>.
- Björn Boman. The influence of ses, cognitive, and non-cognitive abilities on grades: cross-sectional and longitudinal evidence from two swedish cohorts: The influence of ses, cognitive, and non-cognitive abilities on grades: cross-sectional and longitudinal evidence from two swedish cohorts. *European Journal of Psychology of Education*, 38(2):587–603, 2023.
- CBS. Gezinsdynamiek, ouderlijk opleidingsniveau en schoolniveau van vijftienjarigen. Technical report, Centraal Bureau voor de Statistiek, Den Haag, 2015. URL <https://www.cbs.nl/nl-nl/achtergrond/2015/45/gezinsdynamiek-ouderlijk-opleidingsniveau-en-schoolniveau-van-vijftienjarigen>.
- CBS. Herziening gewichtenregeling primair onderwijs - Fase 1. Technical report, Centraal Bureau voor de Statistiek, Den Haag, 2016. URL <https://www.cbs.nl/nl-nl/maatwerk/2016/50/herziening-gewichtenregeling-primair-onderwijs-fase-i>.
- CBS. Herziening gewichtenregeling primair onderwijs - Fase 2. Technical report, Centraal Bureau voor de Statistiek, Den Haag, 2017a. URL <https://www.cbs.nl/nl-nl/maatwerk/2017/04/herziening-gewichtenregeling-primair-onderwijs-fase-ii>.
- CBS. Herziening gewichtenregeling primair onderwijs - Fase 3. Technical report, Centraal Bureau voor de Statistiek, Den Haag, 2017b. URL <https://www.cbs.nl/nl-nl/maatwerk/2017/35/herziening-gewichtenregeling-primair-onderwijs-fase-3>.
- CBS. Herziening gewichtenregeling primair onderwijs - Fase 4. Technical report, Centraal Bureau voor de Statistiek, Den Haag, 2019a. URL <https://www.cbs.nl/nl-nl/maatwerk/2019/07/herz-gewichtenregeling-fase-4-verfijning-imputaties>.
- CBS. Herziening gewichtenregeling primair onderwijs - Samenvatting. Technical report, Centraal Bureau voor de Statistiek, Den Haag, 2019b. URL <https://www.cbs.nl/nl-nl/maatwerk/2019/45/de-nieuwe-onderwijsachterstandenindicator>.

- CBS. Herijking onderwijsachterstandenindicator primair onderwijs 2021 - Vervolgonderzoek naar het actualiseren van het model van de onderwijsachterstandenindicator. Technical report, Centraal Bureau voor de Statistiek, Den Haag, 2021. URL <https://www.cbs.nl/nl-nl/longread/rapportages/2022/herijking-onderwijsachterstandenindicator-primair-onderwijs-2021>.
- CBS. Hoe kan de Landelijke Jeugdmonitor bijdragen aan het monitoren van kansenongelijkheid? Technical report, Centraal Bureau voor de Statistiek, 2024a. URL <https://www.cbs.nl/nl-nl/longread/statistische-trends/2024/hoe-kan-de-landelijke-jeugdmonitor-bijdragen-aan-het-monitoren-van-kansenongelijkheid->.
- CBS. Onderzoek herijking risico-indicator onderwijsachterstanden – fase 1. Technical report, Centraal Bureau voor de Statistiek, October 2024b. URL <https://www.cbs.nl/nl-nl/longread/aanvullende-statistische-diensten/2024/onderzoek-herijking-risico-indicator-onderwijsachterstanden-fase-1>.
- CBS. Onderzoek herijking risico-indicator onderwijsachterstanden – aanvullende analyses fase 2. 2026. URL [Nogtepubliceren](#).
- Rand D Conger and M Brent Donnellan. An interactionist perspective on the socioeconomic context of human development. *Annu. Rev. Psychol.*, 58(1):175–199, 2007.
- Marjolijn Das and Mayke Nollet. Trends in inequality of opportunity: changes in heritability of education over time in the netherlands. *European Sociological Review*, page jcag003, 2026.
- Marloes de Lange and Jaap Dronkers. Single parenthood and children's educational performance: inequality among families and schools. In *The triple bind of single-parent families*, pages 125–144. Policy Press, 2018.
- Eveline L de Zeeuw, Catharina EM van Beijsterveldt, Tina J Glasner, Eco JC de Geus, and Dorret I Boomsma. Arithmetic, reading and writing performance has a strong genetic component: A study in primary school children. *Learning and Individual Differences*, 47:156–166, 2016.
- Perline A Demange, Jouke Jan Hottenga, Abdel Abdellaoui, Espen Moen Eilertsen, Margherita Malanchini, Benjamin W Domingue, Emma Armstrong-Carter, Eveline L de Zeeuw, Kaili Rimfeld, Dorret I Boomsma, et al. Estimating effects of parents' cognitive and non-cognitive skills on offspring education using polygenic scores. *Nature communications*, 13(1):4801, 2022.
- William T Dickens. Genetic differences and school readiness. *The Future of Children*, pages 55–69, 2005.
- William T Dickens and James R Flynn. Heritability estimates versus large environmental effects: the iq paradox resolved. *Psychological review*, 108(2):346, 2001.
- Geert WJM Driessen. Ethnicity, forms of capital, and educational achievement. *International Review of Education*, 47(6):513–538, 2001.
- Louise Elffers. *De bijlesgeneratie: opkomst van de onderwijscompetitie*. Amsterdam University Press, 2018.
- Horst Entorf. Migrants and educational achievement gaps. *IZA World of Labor*, 2015.
- Fenella Fleischmann and Annabel de Haas. Explaining parents' school involvement: The role of ethnicity and gender in the netherlands. *The Journal of Educational Research*, 109(5):554–565, 2016.
- Sara Geven, Anatolia Batruch, Herman van de Werfhorst, et al. Inequality in teacher judgements, expectations and track recommendations: A review study. *Amsterdam: Universiteit van Amsterdam*, 2018.

- Carla Haerlemans, Roxanne Korthals, Madelon Jacobs, Suzanne de Leeuw, Stan Vermeulen, Lynn van Vugt, Bas Aarts, Tijana Prokic-Breuer, Rolf Van der Velden, Sanne van Wetten, et al. Sharp increase in inequality in education in times of the covid-19-pandemic. *Plos one*, 17(2): e0261114, 2022.
- Ken B Hanscombe, Maciej Trzaskowski, Claire MA Haworth, Oliver SP Davis, Philip S Dale, and Robert Plomin. Socioeconomic status (ses) and children's intelligence (iq): In a uk-representative sample ses moderates the environmental, not genetic, effect on iq. *PloS one*, 7(2):e30320, 2012.
- Sara A Hart, Stephen A Petrill, Kirby Deater Deckard, and Lee A Thompson. Ses and chaos as environmental mediators of cognitive ability: A longitudinal genetic analysis. *Intelligence*, 35(3):233–242, 2007.
- Catherine MA Haworth, Margaret J Wright, Michelle Luciano, Nicholas G Martin, Eco JC de Geus, Catharina EM van Beijsterveldt, Meike Bartels, Danielle Posthuma, Dorret I Boomsma, Oliver SP Davis, et al. The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular psychiatry*, 15(11):1112–1120, 2010.
- Lex Herweijer, Ralf Maslowski, and Jesper Rözer. Ongelijke kansen in het onderwijs. In *Kwesties voor het kiezen 2021: Analyse van vijf urgente maatschappelijke vraagstukken voor de komende kabinetsperiode*, pages 20–81. Sociaal en Cultureel Planbureau, 2021.
- Laurence J Howe, Michel G Nivard, Tim T Morris, Ailin F Hansen, Humaira Rasheed, Yoonsu Cho, Geetha Chittoor, Rafael Ahlskog, Penelope A Lind, Teemu Palviainen, et al. Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nature genetics*, 54(5):581–592, 2022.
- T Jenniskens, A van Langen, E Stronkhorst, I van der Veen, A Veen, and M Wolbert. Meting leerpotentie in groep 6. Technical report, KBA Nijmegen / Kohnstamm Instituut, Nijmegen / Amsterdam, 2022.
- Roel Jennissen, Mark Bovens, and GBM Engbersen. Clusteren op basis van cultuur: Etno-linguïstische indelingen van herkomstlanden als alternatieven voor westers/niet-westers in integratieonderzoek. Technical report, Wetenschappelijke Raad voor het Regeringsbeleid, 2021.
- J Klopogge and W de Wit. Het onderwijsachterstandenbeleid na 2015. literatuurstudie t.b.v. expertbijeenkomst oab september 2015. Technical report, Nationaal Regieorgaan Onderwijsonderzoek, 2015.
- Jan-Paul Kooistra, Wout Ultee, and Ben Pelzer. Aap, noot, mustafa...: Het effect van taalafstand en koloniaal verleden op leesprestaties. *Mens & Maatschappij*, 83(3):258, 2008.
- Eva Krapohl, Kaili Rimfeld, Nicholas G Shakeshaft, Maciej Trzaskowski, Andrew McMillan, Jean-Baptiste Pingault, Kathryn Asbury, Nicole Harlaar, Yulia Kovas, Philip S Dale, et al. The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proceedings of the national academy of sciences*, 111(42):15273–15278, 2014.
- Joeke Kuyvenhoven and Willem R Boterman. Neighbourhood and school effects on educational inequalities in the transition from primary to secondary education in amsterdam. *Urban Studies*, 58(13):2660–2682, 2021.
- P. Leseman. *Onderwijsachterstandenbeleid voorbij de voor- en vroegschoolse periode*, pages 113–130. Amsterdam University Press, 2007.
- P Leseman. Ongelijke kansen in het onderwijs: Een literatuuroverzicht. Technical report, Universiteit Utrecht, 2019.

- Margherita Malanchini, Andrea G Allegrini, Michel G Nivard, Pietro Biroli, Kaili Rimfeld, Rosa Cheesman, Sophie von Stumm, Perline A Demange, Elsje van Bergen, Andrew D Grotzinger, et al. Genetic associations between non-cognitive skills and academic achievement over development. *Nature human behaviour*, 8(10):2034–2046, 2024.
- Daniel M McNeish. Using lasso for predictor selection and to assuage overfitting: A method long overlooked in behavioral sciences. *Multivariate behavioral research*, 50(5):471–484, 2015.
- Judi Mesman. Oud geleerd, jong gedaan. *Het Jonge Kind*, 11:9–12, 2011.
- Irma Mooi-Reci, Bart Bakker, Matthew Curry, and Mark Wooden. Why parental unemployment matters for children's educational attainment: empirical evidence from the netherlands. *European Sociological Review*, 35(3):394–408, 2019.
- CWJ Mulder, Daan Fettelaar, Ilona Schouwenaars, Guuske Ledoux, Liselotte Dijkers, and Els Kuiper. De achterstand van autochtone doelgroeperlingen oorzaken en aanpak. Technical report, 2014.
- M Muskens, D van Helvoirt, P Aalders, MHJ Wolbers, M Karssen, E Stronkhorst, DP Elshof, R de Loos, and S Lepoeter. Gescheiden werelden: Onderwijssegregatie houdt aan. onderzoek naar segregatie in het primair en voortgezet onderwijs van 2017-2018 tot en met 2023-2024. eindrapport 2025. 2025.
- Jaap Nieuwenhuis and Pieter Hooimeijer. The association between neighbourhoods and educational achievement, a systematic review and meta-analysis. *Journal of Housing and the Built Environment*, 31(2):321–347, 2016.
- WCM Resing and PJD Drenth. Intelligentie: weten en meten: 2e, herziene druk [intelligence: Knowing and measuring. 2007.
- Wilma Resing and Jan B Blok. De classificatie van intelligentiescores: Voorstel voor een eenduidig systeem. *Psycholoog*, 2002.
- Wilma C Resing, Nico Bleichrodt, and Pieter J Drenth. Het gebruik van de {rakit} bij allochtoon etnische groepen. *Nederlands Tijdschrift voor de Psychologie en haar Grensgebieden*, 1986.
- Yves Rosseel. lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2):1–36, 2012. doi: 10.18637/jss.v048.i02.
- Donald B. Rubin. *Multiple imputation for nonresponse in surveys*. Wiley Series in Probability and Statistics. Wiley, 1 edition, June 1987. ISBN 9780471087052 9780470316696. doi: 10.1002/9780470316696. URL <https://onlinelibrary.wiley.com/doi/book/10.1002/9780470316696>.
- Bruno Sauce and Louis D Matzel. The paradox of intelligence: Heritability and malleability coexist in hidden gene-environment interplay. *Psychological bulletin*, 144(1):26, 2018.
- B ter Weel, H Bussink, and M Hartgerink. Kansenongelijkheid in Nederland. Technical report, SEO, Amsterdam, 2025.
- Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1):267–288, 1996.
- Eric Turkheimer, Andreana Haley, Mary Waldron, Brian d'Onofrio, and Irving I Gottesman. Socioeconomic status modifies heritability of iq in young children. *Psychological science*, 14(6): 623–628, 2003.

- Malon Van den Hof, Ilona Veer, Ruben van Gaalen, and Tessa Roseboom. Clustering of circumstances during the first 1000 days after conception and their association with school performance: a population-based cohort study from the netherlands. *BMJ Public Health*, 3(2), 2025.
- Reyn Van Ewijk and Peter Slegers. The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational research review*, 5(2):134–150, 2010.
- W van Hoorn, L van Kamp, W den Brinker, J Ter Laak, and M de Groot. Nederlandse differentiatie testserie: Ndt-sef 2005. vragenlijst voor sociaal-emotioneel-functioneren. 2005.
- Marinus H Van Ijzendoorn, Femmie Juffer, and Caroline W Klein Poelhuis. Adoption and cognitive development: a meta-analytic comparison of adopted and nonadopted children's iq and school performance. *Psychological bulletin*, 131(2):301, 2005.
- AM Veen, M Karssen, M van Daalen, J Roeleveld, B Triesscheijn, and D Elshof. *De aansluiting tussen voor-en vroegschoolse educatie en tussen vroegschoolse educatie en groep 3*. Kohnstamm Instituut, 2013.
- Sophie Von Stumm and Robert Plomin. Socioeconomic status and the growth of intelligence from infancy through adolescence. *Intelligence*, 48:30–36, 2015.
- Cok Vrooman, Jeroen Boelhouwer, Jurjen Iedema, and Ab Van der Torre. Eigentijdse ongelijkheid. de postindustriële klassenstructuur op basis van vier typen kapitaal. *Sociaal en Cultureel Planbureau*, 2023.
- Eric-Jan Wagenmakers and Simon Farrell. AIC model selection using Akaike weights. *Psychonomic Bulletin & Review*, 11(1):192–196, 2004.
- Yi Yang and Hui Zou. A fast unified algorithm for solving group-lasso penalize learning problems. *Statistics and Computing*, 25(6):1129–1141, 2015.
- Maria Zumbuehl and Rik Dillingh. Ongelijkheid van het jonge kind. *Den Haag: Centraal Planbureau (CPB)*, 2020.

# A. Operationalisering gebruikte achtergrondkenmerken

In deze bijlage beschrijven we de achtergrondkenmerken die we in dit onderzoek hebben gebruikt. Bij sommige variabelen zijn verschillende indelingen getest. Verder zijn sommige variabelen apart voor beide ouders meegenomen en/of voor beide ouders als geheel. Bij de variabelen waarvoor dit van toepassing is, hebben we dit tussen haakjes achter de variabele weergegeven.

## A.0.1. Kenmerken van kind en gezin

### Geüniformeerde eindtoetscore

Omdat de eindtoetscores van de verschillende aanbieder verschillende schalen hebben, zijn de eindtoetscores herschaald naar één uniforme schaal [CBS, 2024b, 2026]. De geüniformeerde schaal heeft waarden die tussen de nul en zes liggen.

### Hoogst behaalde opleidingsniveau ouders (vader / moeder)

Van een groot deel van de Nederlandse bevolking is het hoogst behaalde opleidingsniveau bekend uit diplomaregisters of de Enquête Beroepsbevolking. Het opleidingsniveau is niet voor iedereen bekend. Ontbrekende waarden zijn daarom geïmputeerd met behulp van achtergrondkenmerken [CBS, 2024b]. Het hoogst behaalde opleidingsniveau is gecategoriseerd aan de hand van het 3e aggregatieniveau van de [Standaard Onderwijsindeling van het CBS](#):

- Basisonderwijs
- Praktijkonderwijs, vmbo-b/k, mbo entreeopleiding
- Vmbo-g/t, havo-, vwo-onderbouw
- Mbo niveau 2 en 3
- Mbo niveau 4
- Havo-, vwo-bovenbouw
- Hbo-, wo-bachelor
- Hbo-, wo-master, doctor

### Bekendheid opleidingsniveau ouders

Omdat de ontbrekende waarden met betrekking tot het hoogst behaalde opleidingsniveau niet *missing at random* zijn, hebben we ook een correctiefactor meegenomen in de analyses met de volgende indeling:

- Opleidingsniveau beide ouders *onbekend*
- Opleidingsniveau vader *onbekend*
- Opleidingsniveau moeder *onbekend*
- Opleidingsniveau beide ouders *bekend*

## Ouderlijke structuur

Op basis van de BRP is in kaart gebracht of de leerling in hetzelfde huishouden woont als de juridische ouders:

- Woont bij beide juridische ouders
- Woont bij moeder
- Woont bij vader
- Woont bij moeder en partner
- Woont bij vader en partner
- Woont zonder juridische ouders
- Onbekend

## Herkomstgroepering (vader / moeder / kind)

Voor herkomst is de indeling zoals voorgesteld door de WRR gebruikt [Jennissen et al., 2021]. De herkomstgroepering is zowel voor de ouders als voor het kind bekeken. De herkomstgroepering van het kind is daarbij afgeleid van de herkomstgroepering van de ouders op basis van de volgende regels:

- Indien bekend, dan wordt de herkomstindeling van de moeder overgenomen.
- Indien de herkomstindeling niet bekend is, dan wordt de herkomstindeling van de vader gebruikt.
- Als één van beide ouders Nederland als herkomst heeft, dan krijgt het kind ook Nederland als herkomst toegewezen.

Ten opzichte van de door de WRR gepubliceerde indeling zijn - in overleg met de WRR - er een aantal wijzigingen doorgevoerd. De belangrijkste daarvan zijn:

- Nederland is als aparte categorie opgenomen
- Indonesië is ingedeeld bij Zuidoost-Azië en de Pacific in plaats van bij de Germaanstalige landen
- Puerto Rico is ingedeeld bij Latijns-Amerika in plaats van bij de Angelsaksische landen

De gebruikte indeling in 13 categorieën:

- Nederland
- Angelsaksische landen
- Continentaal-Germaanstalige landen
- Midden- en Oost-Europese landen
- Mediterraan Europa
- Latijns-Amerika
- Cariben
- Arabische landen
- Centraal-Azië
- Zuid-Azië
- Oost-Azië
- Zuidoost-Azië en de Pacific
- Sub-Sahara-Afrika
- Onbekend

De gebruikte indeling in 8 categorieën:

- Nederland
- Anglo-Germaanstalige landen

- Midden- en Oost-Europese landen
- Latijnstalige en andere mediterrane landen
- Cariben
- Arabische landen en Centraal-Azië
- Overig Azië en de Pacific
- Sub-Sahara-Afrika
- Onbekend

### **Leeftijd vader**

In 6 categorieën:

- Jonger dan of gelijk aan 35 jaar
- Ouder dan 35 jaar of jonger dan of gelijk aan 40 jaar
- Ouder dan 40 jaar of jonger dan of gelijk aan 45 jaar
- Ouder dan 45 jaar of jonger dan of gelijk aan 50 jaar
- Ouder dan 50 jaar
- Leeftijd onbekend

In 9 categorieën:

- Jonger dan of gelijk aan 30 jaar
- Ouder dan 30 jaar of jonger dan of gelijk aan 35 jaar
- Ouder dan 35 jaar of jonger dan of gelijk aan 40 jaar
- Ouder dan 40 jaar of jonger dan of gelijk aan 45 jaar
- Ouder dan 45 jaar of jonger dan of gelijk aan 50 jaar
- Ouder dan 50 jaar of jonger dan of gelijk aan 55 jaar
- Ouder dan 55 jaar of jonger dan of gelijk aan 60 jaar
- Ouder dan 60 jaar
- Leeftijd onbekend

### **Leeftijd moeder**

In 6 categorieën:

- Jonger dan of gelijk aan 30 jaar
- Ouder dan 30 jaar of jonger dan of gelijk aan 35 jaar
- Ouder dan 35 jaar of jonger dan of gelijk aan 40 jaar
- Ouder dan 40 jaar of jonger dan of gelijk aan 45 jaar
- Ouder dan 45 jaar
- Leeftijd onbekend

In 9 categorieën:

- Jonger dan of gelijk aan 25 jaar
- Ouder dan 25 jaar of jonger dan of gelijk aan 30 jaar
- Ouder dan 30 jaar of jonger dan of gelijk aan 35 jaar
- Ouder dan 35 jaar of jonger dan of gelijk aan 40 jaar
- Ouder dan 40 jaar of jonger dan of gelijk aan 45 jaar
- Ouder dan 45 jaar of jonger dan of gelijk aan 50 jaar
- Ouder dan 50 jaar of jonger dan of gelijk aan 55 jaar
- Ouder dan 55 jaar
- Leeftijd onbekend

## Aantal kinderen in huishouden

In 6 categorieën:

- Eén kind
- Twee kinderen
- Drie kinderen
- Vier kinderen
- Vijf kinderen
- Zes of meer kinderen

In 3 categorieën:

- Eén of twee kinderen
- Drie of vier kinderen
- Vijf of meer kinderen

## Bruto inkomen huishouden in decielen

De inkomens van alle huishoudens in Nederland zijn gerangschikt van laag naar hoog en in 100 groepen van gelijke omvang ingedeeld. De eerste groep bevat dan de 1 procent huishoudens met het laagste inkomen, de honderdste groep bevat de 1 procent huishoudens met het hoogste inkomen. Voor het huidige onderzoek zijn de inkomenspercentielen ingedeeld in 11 categorieën: 10 decielen en een categorie *onbekend*.

## Besteedbaar inkomen huishouden t.o.v. armoedegrens

Het Centraal Bureau voor de Statistiek (CBS), het Nationaal Instituut voor Budgetvoorlichting (Nibud) en het Sociaal en Cultureel Planbureau (SCP) hebben gezamenlijk een methode ontwikkeld om armoede in Nederland te meten. Voor 35 verschillende typen huishoudens is vastgesteld welke bedragen voor welke uitgavenposten minimaal nodig zijn om volwaardig mee te kunnen doen in de samenleving (de zogeheten minimumvoorbeeldbegrotingen). Deze vormen de basis voor de armoedegrenzen. Om vast te stellen of iemand in armoede leeft, wordt ook gekeken naar daadwerkelijke uitgaven van het huishouden en de financiële middelen: het besteedbaar inkomen en eventueel spaargeld of ander direct te besteden bezit, de (vrij opneembare) vermogensbuffer. Voor de meeste huishoudens zijn de maandelijkse uitgaven aan wonen, energie en zorg de grootste vaste lasten. Als er na het betalen van deze vaste lasten te weinig geld overblijft voor de andere basisbehoeften, dan is een huishouden (en de mensen die er deel van uitmaken) arm.

De inkomens zijn uitgedrukt als percentage van de armoedegrens met een maximum van 500%. Deze percentage hebben we ingedeeld in de volgende categorieën:

- Niet arm (meer dan 150% van de armoedegrens).
- Minder of gelijk aan armoedegrens.
- Boven de armoede grens, maar minder of gelijk aan 125% van de armoedegrens.
- Meer dan 125% van de armoedegrens, maar minder of gelijk aan 150% van de armoedegrens.

## Besteedbaar inkomen huishouden t.o.v. armoedegrens in de laatste drie jaren

Idem als voorgaande, maar dan voor de laatste drie jaar voorafgaand aan het peilmoment samengenomen. Deze variabele kan worden gezien als een indicator voor langdurige armoede.

## Vermogen huishouden

De vermogens van alle huishoudens in Nederland zijn gerangschikt van laag naar hoog en in 100 groepen van gelijke omvang ingedeeld. De eerste groep bevat dan de 1 procent huishoudens met het laagste vermogen, de honderdste groep bevat de 1 procent huishoudens met het hoogste vermogen. Voor het huidige onderzoek zijn de vermogenspercentielen ingedeeld in 11 categorieën: 10 decielen en een categorie *onbekend*.

### **Vermogen huishouden exclusief eigen woning**

Van de vermogens van de huishoudens hebben we ook een variant onderzocht waarin de eigen woning niet is meegenomen.

### **Welvaart huishouden**

De financiële welvaart van een huishouden is gebaseerd op zowel het gestandaardiseerd inkomen als het vermogen. Na rangschikking op basis van het inkomen en vermogen is de financiële welvaart van een huishouden bepaald uit de som van het cumulatieve aandeel in het totale inkomen en het cumulatieve aandeel in het totale vermogen. Op grond van de optelling zijn de huishoudens vervolgens gerangschikt van laag naar hoog en in 100 groepen van gelijke omvang verdeeld. De eerste groep bevat dan de 1 procent huishoudens met de laagste financiële welvaart, de honderdste groep bevat de 1 procent huishoudens met de hoogste financiële welvaart. Voor het huidige onderzoek zijn de welvaartspercentielen ingedeeld in 11 categorieën: 10 decielen en een categorie *onbekend*.

### **Problematische schulden huishouden**

Het Centraal Bureau voor de Statistiek (CBS) heeft op basis van registerdata in beeld gebracht hoeveel Nederlandse huishoudens te maken hadden met [geregistreeerde problematische schulden](#).

Met behulp van verschillende registraties is afgeleid welke huishoudens te maken hebben met geregistreeerde problematische schulden. De informatie over schulden verschilt per registratie. Daarom is per registratie bepaald wanneer er sprake is van een geregistreeerde problematische schuldsituatie.

We hebben hierbij gebruik gemaakt van de *smalle definitie* van problematische schulden, waarin in tegenstelling tot de volledige definitie de volgende voorwaarden niet zijn meegenomen:

- Volgt een bij het Bureau Krediet Registratie (BKR) geregistreeerd minnelijk traject;
- Heeft een bij BKR geregistreeerde betalingsachterstand;
- Is in het Centraal Curatele en Bewindsregister (CCBR) opgenomen op grond van verkwisting en/of problematische schulden.

Laatste jaar:

- Geen problematische schulden
- Wel problematische schulden

Laatste drie jaren:

- Geen problematische schulden in de afgelopen drie jaren
- Eén jaar van de afgelopen drie jaren problematische schulden
- Twee jaar van de afgelopen drie jaren problematische schulden
- Drie jaar van de afgelopen drie jaren problematische schulden

## **Deeltijdfactor werkzame personen huishouden**

Het aantal gewerkte uren in een maand in verhouding tot een voltijdsbaan. Wanneer iemand meerdere banen heeft, zijn de deeltijdfactoren van deze banen opgeteld.

Vader / moeder:

- Lager dan of gelijk aan 0,1 fte
- Hoger dan 0,1, maar lager dan of gelijk aan 0,2 fte
- Hoger dan 0,2, maar lager dan of gelijk aan 0,3 fte
- Hoger dan 0,3, maar lager dan of gelijk aan 0,4 fte
- Hoger dan 0,4, maar lager dan of gelijk aan 0,5 fte
- Hoger dan 0,5, maar lager dan of gelijk aan 0,6 fte
- Hoger dan 0,6, maar lager dan of gelijk aan 0,7 fte
- Hoger dan 0,7, maar lager dan of gelijk aan 0,8 fte
- Hoger dan 0,8, maar lager dan of gelijk aan 0,9 fte
- Hoger dan 0,9, maar lager dan of gelijk aan 1,0 fte

Huishouden (beide ouders samen):

- Lager dan of gelijk aan 0,2 fte
- Hoger dan 0,2, maar lager dan of gelijk aan 0,4 fte
- Hoger dan 0,4, maar lager dan of gelijk aan 0,6 fte
- Hoger dan 0,6, maar lager dan of gelijk aan 0,8 fte
- Hoger dan 0,8, maar lager dan of gelijk aan 1,0 fte
- Hoger dan 1,0, maar lager dan of gelijk aan 1,2 fte
- Hoger dan 1,2, maar lager dan of gelijk aan 1,4 fte
- Hoger dan 1,4, maar lager dan of gelijk aan 1,6 fte
- Hoger dan 1,6, maar lager dan of gelijk aan 1,8 fte
- Hoger dan 1,8, maar lager dan of gelijk aan 2,0 fte

## **Economische sector werkomgeving ouders (vader / moeder)**

Economische activiteit van het bedrijf van de voornaamste werkkring waarin werkzame personen werkzaam zijn. Standaardbedrijfsindeling in 10 categorieën:

- A Landbouw, bosbouw en visserij
- B-E Nijverheid (geen bouw) en energie
- F Bouwnijverheid
- G-I Handel, vervoer en horeca
- J Informatie en communicatie
- K Financiële dienstverlening
- L Verhuur en handel van onroerend goed
- M-N Zakelijke dienstverlening
- O-Q Overheid en zorg
- R-U Cultuur, recreatie, overige diensten

## **Inkomstenbron ouders (vader / moeder)**

De wijze waarop de ouders hun inkomen hebben verkregen:

- Werk of pensioen
- Uitkering of geen inkomen
- Onbekend

### **Arbeidsverleden ouders (vader / moeder)**

Het recente arbeidsverleden geeft aan of een persoon in de afgelopen 4 jaren heeft gewerkt.

### **Psychosociale problemen (vader / moeder / ouders)**

Om psychosociale problemen te meten, is gekeken naar GGZ-kosten.

- Persoon (vader, moeder of minimaal één van beide ouders) heeft geen GGZ-kosten gehad in de afgelopen 4 jaar.
- Persoon heeft GGZ-kosten gehad in de afgelopen 4 jaar.

### **Verblijfsduur in Nederland (vader / moeder)**

Het aantal jaren dat de juridische ouders in Nederland verblijven is ingedeeld in categorieën.

Indeling in 2 categorieën:

- Langer dan vijftien jaar in Nederland
- Korter dan vijftien jaar in Nederland

Indeling in 3 categorieën:

- Korter dan vijf jaar in Nederland
- Tussen vijf en tien jaar in Nederland
- Langer dan tien jaar in Nederland

Indeling in 4 categorieën:

- Korter dan vijf jaar in Nederland
- Tussen vijf en tien jaar in Nederland
- Tussen tien en vijftien jaar in Nederland
- Langer dan vijftien jaar in Nederland

### **Asielstatus kind**

Registratie van personen die bekend zijn bij het Centraal Orgaan opvang Asielzoekers (COA) en de Immigratie- en Naturalisatiedienst (IND). De registratie gaat terug tot 2014.

- Leerling is niet bekend in COA/IND registers.
- Leerling is via asielpcedure Nederland binnengekomen.

### **Ouder(s) verdacht van misdrijf (vader / moeder / ouders)**

Verdachten (redelijk vermoeden van schuld) van geregistreerde misdrijven die in een bepaald jaar voorkomen in het registratiesysteem BVI van de Nationale Politie. Er wordt geen onderscheid gemaakt naar type misdrijf.

Laatste jaar:

- Niet verdacht van een misdrijf
- Wel verdacht van een misdrijf

Ooit:

- Niet verdacht van een misdrijf
- Wel verdacht van een misdrijf

## **A.0.2. Kenmerken van school en buurt**

### **Sociaaleconomische status**

De sociaaleconomische status (SES-WOA) wordt beschreven in termen van de financiële welvaart, het opleidingsniveau en het recente arbeidsverleden van particuliere huishoudens. De scores per huishouden liggen tussen de -2 en 1. Voor verschillende aggregatieniveaus is het gemiddelde van deze huishoudscores berekend:

- School
- Buurt woonadres leerling
- Wijk woonadres leerling
- Gemeente woonadres leerling

### **Stedelijkheid van de buurt**

Op basis van de BRP is bekend in welke gemeente, wijk en buurt een leerling woont. Voor ieder adres binnen een gemeente is de adressendichtheid vastgesteld van een gebied met een straal van één kilometer rondom dat adres. De volgende klassen van stedelijkheid worden onderscheiden:

- Zeer sterk (meer dan 2500 omgevingsadressen/km<sup>2</sup>)
- Sterk (1500 tot 2500 omgevingsadressen/km<sup>2</sup>)
- Matig (1000 tot 1500 omgevingsadressen/km<sup>2</sup>)
- Weinig (500 tot 1000 omgevingsadressen/km<sup>2</sup>)
- Niet (minder 500 omgevingsadressen/km<sup>2</sup>)

## **B. Begeleidingscommissie**

- Thijs Bol - Universiteit van Amsterdam
- Jonneke Bolhaar - Centraal Plan Bureau
- Peter van Corler - Vereniging Nederlandse Gemeenten
- Louise Elffers - Universiteit van Amsterdam
- Carla Haelermans - Universiteit Maastricht
- Lex Herwijer - Sociaal Cultureel Planbureau
- Wendy Mol - Ministerie van Onderwijs, Cultuur & Wetenschap
- Marius Lam - Inspectie voor het Onderwijs
- Paul Leseman - Universiteit Utrecht
- Tim Schokker - Ministerie van Onderwijs, Cultuur & Wetenschap
- Reinoud Stoel - Centraal Bureau voor de Statistiek
- Edi Terlaak - Inspectie voor het Onderwijs
- Ellen Timminga - Inspectie voor het Onderwijs
- Janis van Vliet - Ministerie van Onderwijs, Cultuur & Wetenschap
- Peter Vos - PO-raad
- Thomas Westendorp - Inspectie voor het Onderwijs
- Anneke de Wolff - Ministerie van Onderwijs, Cultuur & Wetenschap