



Centraal Bureau  
voor de Statistiek

**Bevolkingstrends 2014**

# **Hoeveel ouderlijk gezang klinkt door in het gepiep van de jongen?**

**Over de samenhang tussen milieu van herkomst  
en schoolprestaties van kinderen**

**Ruben van Gaalen**

**Bart Bakker**

**Jan van der Laan**

**Sue Westerman**

**Sander Scholtus**

**September 2014**

Van de verschillen in schoolprestaties tussen kinderen wordt 42 procent verklaard door verschillen tussen ouderlijk milieus. Uiteindelijk wordt bijna één derde hiervan verklaard door de opleidingsniveaus van beide ouders. Daarnaast spelen het gezinsinkomen, echtscheiding van de ouders en de herkomstgroepering van het kind een kleine rol. De twee derde onverklaarde verschillen zijn terug te voeren op ongemeten kenmerken van het ouderlijk nest zoals het beroepsniveau van de ouders, het cultureel aanbod en genetische factoren. Het ouderlijk nest is dus nog steeds belangrijk voor de schoolprestaties van kinderen

## 1. Inleiding

Bij het verklaren van verschillen in schoolprestaties tussen kinderen worden in de literatuur primaire en secundaire effecten van het milieu van herkomst op schoolprestaties onderscheiden (Boudon, 1974; zie ook Kloosterman, 2010). Bij primaire effecten gaat het om de directe invloed van het milieu van herkomst op leercapaciteiten aan het begin van de schoolloopbaan. Bij secundaire effecten gaat het om de invloed van het milieu van herkomst op het uiteindelijk bereikte onderwijsniveau bij gelijke leercapaciteiten.

In dit onderzoek gaat het er niet om het relatieve belang van beide effecten te bepalen maar worden primaire en secundaire effecten in één keer geschat. Idealiter wordt voor het primaire effect gebruik gemaakt van gemeten leercapaciteiten in het begin van de schoolloopbaan (in groep 2 of 3), maar over die gegevens beschikt het CBS niet. De scores voor de Eindtoets basisonderwijs (hier wordt altijd de Cito Eindtoets mee bedoeld) zijn een goed alternatief, ook omdat bijna ieder kind deze toets maakt. De secundaire effecten worden benaderd met de onderwijspositie in het voortgezet onderwijs drie jaar nadat de Cito-toets werd afgenomen.

## 2. Methode: Multilevel siblingmodellen

In de literatuur wordt de bijdrage van het sociale milieu aan de ontwikkelkansen van kinderen meestal gemodelleerd met zogeheten 'sibling'-modellen. Het idee daarachter is dat de overeenkomst in bijvoorbeeld schoolprestaties van broers en zussen wordt bepaald door het milieu van herkomst. Broers en zussen hebben een overeenkomst in genetische eigenschappen, in opvoeding en andere omgevingsfactoren. Milieu van herkomst kan dan worden opgevat als een latente variabele die de correlatie tussen de schoolprestaties van broertjes en zusjes verklaart (Sieben, 2001).

De daadwerkelijk gemeten indicatoren voor milieu van herkomst, het opleidingsniveau van de vader en de moeder, de voornaamste inkomstenbron van de vader en de moeder, inkomsten van het huishouden, eenoudergezin, etc. worden gerelateerd aan de schoolprestaties

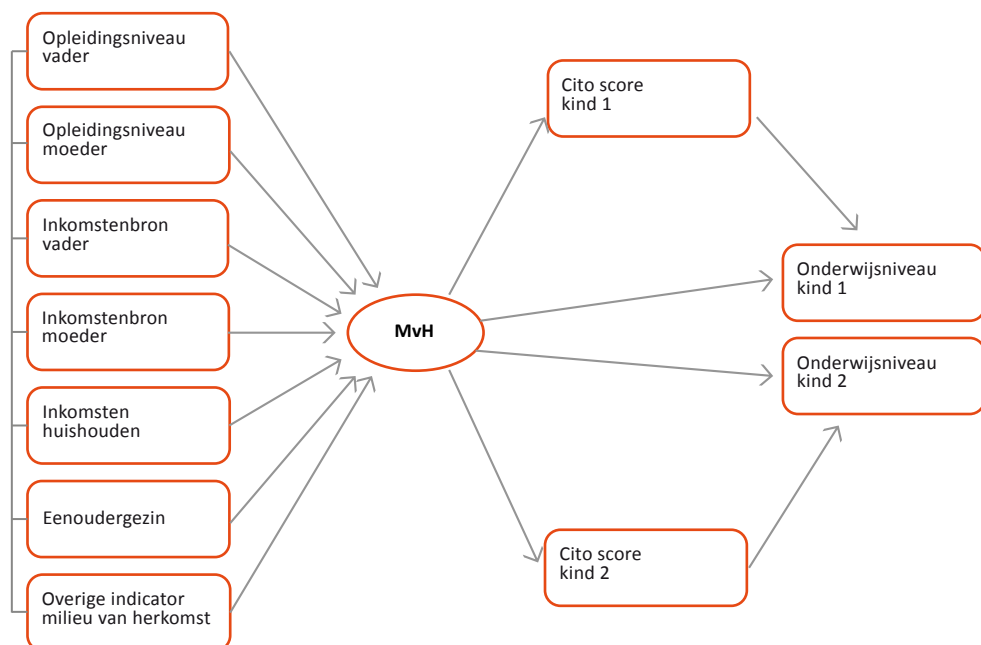
(schema 3.1). Doordat er broers en zussen uit hetzelfde gezin in de data zitten, ontstaan er clustereffecten waarvoor moet worden gecorrigeerd met een multilevel-model.

### 3. Data: Stelsel van Sociaal-statistische Bestanden (SSB)

Het CBS heeft in opdracht van het Ministerie van Onderwijs, Cultuur en Wetenschappen (OCW) onderzocht of het haalbaar is het milieu van herkomst van leerlingen te relateren aan hun schoolprestaties, op basis van het Stelsel van Sociaal-statistische Bestanden (SSB; Bakker, Van Rooijen en Van Toor, 2014). De volgende vragen werden getracht te beantwoorden: (1) Welke kenmerken van het milieu van herkomst van leerlingen zijn via het SSB beschikbaar? (2) Welke van deze kenmerken zijn het meest bruikbaar voor het verklaren van schoolprestaties?

Het SSB is een stelsel van koppelbare registers die onderling zijn afgestemd en consistent gemaakt. Het SSB bevat onder andere gegevens over demografie, arbeidsmarkt, criminaliteit en onderwijs. Op dit moment zijn de meeste gegevens vanaf 1999 beschikbaar voor de gehele Nederlandse bevolking. De Cito-gegevens die in de analyse zijn gebruikt bevatten alleen scores van scholen die toestemming hebben verleend voor het gebruik van deze gegevens door het CBS.

#### 3.1 Multilevel siblingmodel voor het bepalen van effect van milieu van herkomst op schoolprestatie



### 3.1 Combineren opleidingsniveaus ouders en kinderen

Binnen het SSB is het mogelijk om ouders aan kinderen te koppelen. De koppeling van kinderen aan ouders maakt het ook mogelijk om broers en zussen te onderscheiden. Stiefouders, mogelijke halfbroers en –zussen en pleegkinderen worden niet tot de populatie gerekend. Wel selecteren we volle broers en zussen (met dezelfde vader én moeder) als deze niet op hetzelfde adres wonen. De ouder-kind koppeling is voor ouders die na 1945 zijn geboren goed te noemen. Voor de jonge populatie uit ons onderzoek is de koppeling dan ook nagenoeg perfect.

Het opleidingsniveau in het SSB is uitermate geschikt om informatie naar opleidingsniveau uit te splitsen voor ruime dwarsdoorsnedes van de bevolking (Linder, Van Roon en Bakker, 2011). Voor het maken van dergelijke schattingen is een weging ontwikkeld op persoonsniveau, aangezien op dit moment het opleidingsniveau van slechts ongeveer de helft van de Nederlandse bevolking is vastgelegd in het SSB.

Om de opleidingsinformatie over ouders in het SSB te combineren met het SSB Ouder-kind-bestand is een aanvullende weging nodig. Deze is ontwikkeld en noemen we, gezien de context, de intergenerationele weging (zie verder Bakker et al, 2014). Deze wordt gebruikt om de resultaten van de beschrijving en de siblingmodellen te wegen.

### 3.2 Kenmerken populatie

De selectie die werd gemaakt uit het ouder-kind bestand bevat alle kinderen van wie beide ouders bekend zijn, die in de periode 2003-2008 een Cito-toets maakten en die minimaal één broer of zus hadden die ook aan de toets meedeed in die periode. Dit zijn ruim 37 duizend kinderen uit 18 duizend gezinnen.

Onder een (ouderlijk) gezin wordt hier het gedeelde herkomstgezin bedoeld van twee of meer kinderen in de populatie. Dus kinderen die dezelfde vader en moeder hebben maken in onze studie deel uit van hetzelfde ouderlijke gezin. Van de kinderen heeft 86 procent één broer of zus, de overige 14 procent van de kinderen heeft meerdere broers of zussen. Van de vaders heeft 32 procent een hoog opleidingsniveau tegen 22 procent van de moeders. Van de vaders is 81 procent werknemer en 13 procent zelfstandige, voor de moeders was dit 67 procent en 8 procent. Naast het opleidingsniveau, de inkomstenbron (werknemer, zelfstandige, uitkeringsontvanger of geen inkomsten) en of de kinderen in een eenoudergezin wonen, worden ook de inkomsten zelf meegenomen in de analyse. Voor de bepaling van de totale inkomsten zijn de persoonlijke inkomens van beide ouders opgeteld en verdeeld in twintig gelijke percentielgroepen.

De herkomstgroepering van kinderen is ook een van de variabelen in de analyse: 89 procent van de kinderen is van autochtone afkomst, dus van wie de beide ouders in Nederland zijn geboren. De verdeling van het hoogst behaalde opleidingsniveau van de ouders verschilt met de etnische groep waartoe het kind behoort. In vergelijking met autochtone kinderen hebben kinderen van Turkse of Marokkaanse herkomst drie tot vier keer vaker een vader van wie het hoogst behaald opleidingsniveau laag is. Opvallend is dat meer dan de helft van de bij het CBS bekende vaders van kinderen van Antilliaanse of Arubaanse herkomst een hoog

opleidingsniveau heeft. Dit is blijkbaar een selectieve groep, want van de meeste kinderen van Antilliaanse of Arubaanse herkomst is niet geregistreerd wie de vader is.

### 3.2.1 Hoogst behaald opleidingsniveau vader naar herkomst-groepering kind (gewogen)

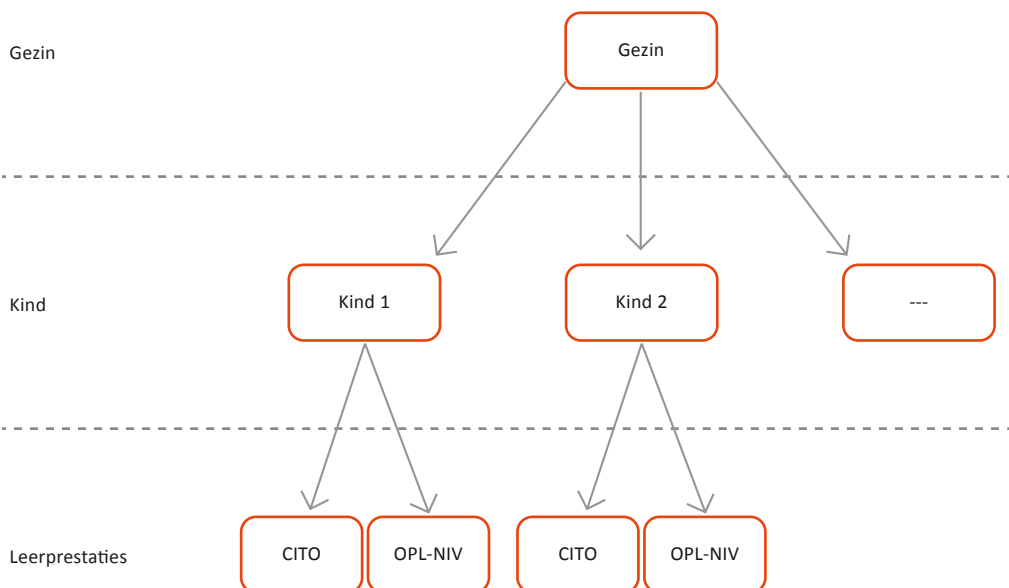
Herkomstgroepering	Hoogst behaald opleidingsniveau vader		
	Laag	Middelbaar	Hoog
	%		
Autochtoon	21,3	46,0	32,7
Marokko	79,5	16,4	4,1
(Voormalig) Nederlandse Antillen en Aruba	17,8	30,8	51,3
Overige niet-westerse landen	28,3	39,5	32,2
Overige westerse landen	21,0	39,4	39,6
Suriname	37,4	43,2	19,4
Turkije	63,1	30,3	6,6

Bron: CBS.

## Methode: Multilevel siblingmodellen

Voor de analyse van schoolprestaties is gebruik gemaakt van multilevel siblingmodellen. Deze modellen kunnen gebruikt worden voor het beschrijven van gegevens met een hiërarchische structuur. De gegevens die hier gebruikt worden hebben zo'n structuur met drie niveaus. Op het onderste niveau hebben we de schoolprestaties. Deze horen bij kinderen, die op hun beurt weer binnen gezinnen vallen. Dit is grafisch weergegeven in onderstaand schema. Multilevelmodellen splitsen de variatie die waargenomen wordt in de schoolprestaties in drie termen: één voor ieder niveau. Op ieder niveau is de variantie te splitsen in een tussenvariantie en een binnenvariantie. Op het gezinsniveau bijvoorbeeld geeft de

## Multilevelmodel van schoolprestaties



binnenvariantie de verschillen in de scores van de kinderen binnen het gezin weer. De tussenvariantie meet de verschillen tussen de gezinnen. Deze is daarmee een maat voor het effect van het gezin op de schoolprestaties. Door de tussenvariantie door de totale variantie te delen wordt de intraklassecorrelatie (ICCgezin) bepaald. Deze intraklassecorrelatie geeft de samenhang tussen kinderen binnen een gezin weer en is daarmee een maat voor de invloed van het milieu van herkomst. Een waarde van 1 (één) betekent dat verschillen in schoolprestaties volledig zijn toe te schrijven aan verschillen in het milieu van herkomst. Een waarde van 0 (nul) betekent dat het milieu van herkomst geen invloed heeft op de schoolprestaties.

Ook op het niveau van het kind kan een intraklassecorrelatie (ICCKind) uitgerekend worden. Deze geeft aan in welke mate de Cito-toets en het onderwijsniveau drie jaar later samenhangen.

Het doel is uiteindelijk om zoveel mogelijk van het gezinseffect te verklaren aan de hand van de eerder genoemde gezinskenmerken. Op het moment dat het gezinseffect gedeeltelijk verklaard wordt door de gezinskenmerken, zal de intraklassecorrelatie voor gezin afnemen aangezien een gedeelte van de variatie in de schoolprestaties op gezinsniveau nu verklaard wordt. De mate van afname geeft aan in hoeverre de variabelen in staat zijn het gezinseffect te verklaren.

### **Voordeel onderscheid gezinseffect en individueel effect**

We zouden eventueel ook direct kunnen proberen het effect van de gezinskenmerken op de prestaties van het kind te schatten. Dit heeft echter twee nadelen. Het eerste nadeel is dat we de correlaties die aanwezig zijn in de dataset (tussen de prestaties van één kind en tussen de kinderen in een gezin) negeren en daarmee mogelijk onzuivere schatters krijgen van de parameters en de varianties op de geschatte parameters. Het tweede belangrijkere nadeel is dat het dan niet mogelijk is om een schatting te krijgen van de invloed van het milieu van herkomst op de schoolprestaties en het is ook niet mogelijk een schatting te krijgen welk deel van het milieu van herkomst verklaard wordt met gebruikte gezinskenmerken.

### **Schatting van het model**

Het analysebestand bevat gewichten om ervoor te zorgen dat de populatie correct beschreven wordt. Bij het schatten van het model moeten deze gewichten meegenomen worden om ervoor te zorgen dat de schatters zuiver zijn. In multilevel-modellen is het niet mogelijk om de gewichten direct mee te nemen in de schatting. Er is daarom gekozen om de gewichten op een andere manier te verwerken. Allereerst is er per gezin een gewicht uitgerekend door de gewichten van alle kinderen in het gezin te middelen<sup>1)</sup>. Vervolgens zijn de gezinnen evenveel keer herhaald in het bestand als dit gemiddelde gewicht. Hiermee wordt in feite de totale populatie van kinderen gecreëerd. De schattingen die met dit bestand worden verkregen zijn zuiver. Echter, doordat gezinnen meerdere keren herhaald zijn, is kunstmatig een groter aantal waarnemingen gecreëerd waardoor schattingen van de betrouwbaarheid die door de programmatuur geproduceerd worden niet kloppen. Om een goede schatting van de kwaliteit van het model te krijgen, is ervoor gekozen om de dataset in tweeën te splitsen. De ene helft wordt gebruikt om de modellen te schatten, de andere helft om het model te valideren.

<sup>1)</sup> In het algemeen zijn de gewichten van kinderen in een gezin gelijk aangezien verschillen in gewichten voornamelijk veroorzaakt worden door het ontbreken van opleidingsniveaus van ouders. De opleidingsniveaus van kinderen zijn meestal bekend.

Voor het schatten van de modellen is de volgende procedure doorlopen:

1. Er is een model geschat zonder verklarende variabelen. Hiermee wordt een schatting verkregen van het totale gezinseffect.
2. Er zijn modellen geschat met telkens één covariaat. Met behulp van de referentiedataset is gekeken hoeveel ieder model verklaart.
3. Beginnend met de meest verklarende covariaat wordt er telkens één covariaat toegevoegd aan het model en wordt er gekeken hoeveel dit model verklaart ten opzichte van het voorgaande model (zonder deze covariaat).

Bovenstaande procedure is zowel voor het complete model waarin tegelijkertijd de Cito-score en het onderwijsniveau drie jaar later gemodelleerd wordt, als voor beide apart uitgevoerd. Voordat de schattingen zijn uitgevoerd zijn de twee metingen van schoolprestaties (Cito-score en onderwijsniveau drie jaar later) eerst geschaald zodat de scores gemiddeld nul zijn en een variantie van één hebben. Zo volgen de twee metingen van schoolprestaties meer een normaalverdeling.

De schattingen zijn uitgevoerd in het statistische programma R (R Core Team, 2012) met behulp van het pakket voor multilevelmodellen lme4 (Bates et al., 2012).

---

## 4. Resultaten

### 4.1 Beschrijvende resultaten

In deze paragraaf wordt ingegaan op de wijze waarop de schoolprestaties van de kinderen zijn gemeten. Aan het einde van groep 8 maken de meeste basisschoolleerlingen de Cito-toets die medebepalend is voor het gevolgde onderwijstype na de basisschool. In de voorkolom van tabel 4.1.1 is aangegeven wat een gebruikelijk schooladvies is bij een bepaalde Cito-score. De Cito-score fungeert in de siblingmodellen als maat voor het primaire effect van het milieu van herkomst. We hebben gebruik gemaakt van de standaardscore. Standaardcores kunnen waarden aannemen van 501 tot en met 550. Tabel 4.1.1 laat een sterke samenhang zien tussen het opleidingsniveau van de vader en de behaalde Cito-score in groep 8. Deze is vergelijkbaar met de samenhang tussen het opleidingsniveau van moeders en de Cito-score. Kinderen van laagopgeleide ouders scoren veel lager en kinderen van hoogopgeleide ouders veel hoger dan gemiddeld.

Een andere manier waarop schoolprestaties zijn gemeten, is het onderwijstype in het derde leerjaar van het voortgezet onderwijs, dus drie jaar na het afleggen van de Cito-toets (tabel 4.1.2). We onderscheiden verschillende leerwegen binnen het vmbo en de onder- en bovenbouw van havo of vwo. De leerlingen die een andere onderwijssoort volgen, zijn ondergebracht in een van de categorieën met een vergelijkbaar niveau. Kinderen volgen vaker dan gemiddeld de laagste leerweg binnen vmbo als ze een vader hebben met een lager

#### 4.1.1 Cito-schooladvies naar hoogst behaald opleidingsniveau vader (gewogen)

Cito-schooladvies	Hoogst behaald opleidingsniveau vader			
	Totaal	Laag	Middelbaar	Hoog
	%			
501–523 vmbo basisberoepsgerichte leerweg	9,6	18,4	9,5	3,4
524–529 vmbo kaderberoepsgerichte leerweg	11,8	17,1	12,7	6,8
530–536 vmbo gemengde/theoretische leerweg	21,5	25,0	23,1	16,7
537–544 havo	33,6	27,7	34,6	36,5
545–550 vwo	23,5	11,8	20,1	36,6
Totaal	100,0	100,0	100,0	100,0

Bron: CBS.

#### 4.1.2 Onderwijstype 3 jaar later naar hoogst behaald opleidingsniveau vader (gewogen)

Onderwijstype	Hoogst behaald opleidingsniveau vader			
	Totaal	Laag	Middelbaar	Hoog
	%			
Vmbo basisberoepsgerichte leerweg	5,8	13,9	5,0	1,3
Vmbo kaderberoepsgerichte leerweg	10,5	17,6	11,7	3,9
Vmbo gemengde leerweg	8,6	10,7	10,0	5,1
Vmbo theoretische leerweg	15,4	18,3	17,5	10,5
Havo	34,3	28,2	35,1	37,4
Vwo, gymnasium	25,4	11,3	20,7	41,8
Totaal	100,0	100,0	100,0	100,0

Bron: CBS.

opleidingsniveau. Hebben kinderen een vader met een middelbaar opleidingsniveau, dan volgen ze vaak een van de andere vmbo-leerwegen of zitten ze op de havo. Kinderen van een vader met een hoger opleidingsniveau gaan het vaakst naar het vwo.

## 4.2 Resultaten multilevel siblingmodellen

Het model dat uiteindelijk de beste verklaring geeft van de schoolprestaties van het kind is het model met de opleidingsniveaus van beide ouders, het gezinsinkomen en herkomstgroepering. Het inkomen van vaders en moeders voegde niet veel toe. Er is duidelijk te zien dat de schoolprestaties van de kinderen toenemen wanneer de opleidingsniveaus van de ouders hoger zijn. Vooral de schoolprestaties van kinderen met een Marokkaanse of Turkse achtergrond zijn minder dan die van autochtone kinderen. Deze uitkomsten komen in hoge mate overeen met de bevindingen van Sieben (2001), die soortgelijke multilevel siblingmodellen gebruikte, al was de uitkomstmaat het uiteindelijk bereikte opleidingsniveau.

De tabel toont de resultaten van het stapsgewijs uitbouwen van het model dat Cito en onderwijsniveau gezamenlijk modelleert. Met een asterisk is aangegeven of het model een betere schatting geeft dan het voorgaande model. Omdat dit model drie niveaus heeft zijn er twee intraklassecorrelaties af te leiden: één op gezinsniveau die weer het effect van het



#### 4.2.1 Multilevel siblingmodel voor Cito en onderwijsniveau

Coëfficiënt	Schatting	Std. fout
Intercept	-0,94	0,01
<b>Opleidingsniveau vader</b>		
Laag	<i>Ref.</i>	<i>Ref.</i>
Middelbaar	0,27	0,01
Hoog	0,53	0,01
<b>Opleidingsniveau moeder</b>		
Laag	<i>Ref.</i>	<i>Ref.</i>
Middelbaar	0,31	0,01
Hoog	0,61	0,01
<b>Inkomen</b>		
1ste quintiel	<i>Ref.</i>	<i>Ref.</i>
2de quintiel	0,12	0,01
3de quintiel	0,17	0,01
4de quintiel	0,23	0,01
5de quintiel	0,33	0,01
<b>Herkomstgroepering</b>		
Autochtoon	<i>Ref.</i>	<i>Ref.</i>
Marokko	-0,32	0,02
(Voormalig) Nederlandse Antillen en Aruba	0,03	0,03
Overig niet westers	0,03	0,02
Overig westers	0,04	0,01
Suriname	-0,15	0,02
Turkije	-0,40	0,02

Bron: CBS.  
Ref.=referentiecategorie.

milieu van herkomst weergeeft en één op kindniveau die de samenhang tussen Cito-score en onderwijsniveau voor een kind weergeeft. Er is te zien dat de samenhang tussen deze twee voor een kind 57 procent is (ICCKind). Het gezinseffect is 42 procent (ICCgezin). Dit is vergelijkbaar met andere studies (Sieben, 2001; Kloosterman, 2010). Deze samenhang neemt af naarmate er meer variabelen aan het model worden toegevoegd tot 29 procent. Uiteindelijk wordt ongeveer 1/3 van het gezinseffect wel en 2/3 niet verklaard door de beschikbare verklarende variabelen.

#### 4.2.2 Samenvatting van modellen voor Cito en onderwijsniveau (3 jaar na Cito)

Model	ICCgezin	ICCKind	Significant
1. (leeg model)	0,42	0,58	
2. Opleidingsniveau vader	0,33	0,58	*
3. Opleidingsniveau moeder	0,30	0,58	*
4. Inkomen	0,30	0,58	*
5. Herkomstgroepering	0,29	0,58	*
6. Inkomstenbron moeder	0,29	0,57	
7. Inkomstenbron vader	0,29	0,57	
8. Type huishouden	0,29	0,57	

Bron: CBS.

Blijkbaar spelen andere kenmerken die niet in de index zitten een rol (zoals het cultureel aanbod in het gezin; aantal boeken, museumbezoek, overige genetische factoren). Dit maakt twee derde van het totale effect uit, maar we kunnen niet overzien hoe belangrijk deze zaken afzonderlijk zijn.

## 5. Conclusie en discussie

Het gezinseffect is 42 procent, dat wil zeggen dat 42 procent van de schoolprestaties van leerlingen wordt verklaard door het milieu van herkomst. Tegelijkertijd wil dit zeggen dat 58 procent niet door het milieu van herkomst wordt verklaard, maar komt door individuele verschillen tussen leerlingen die niets met het milieu van herkomst te maken hebben. Dit is goed vergelijkbaar met de resultaten van andere studies (Sieben, 2001, Kloosterman, 2010).

Uiteindelijk wordt een derde van het gezinseffect verklaard door de beschikbare verklarende variabelen: dat is dus 14 procent van de totale verschillen in schoolprestaties. Van het gezinseffect blijft twee derde onverklaard. Blijkbaar spelen andere kenmerken een rol. Het model dat uiteindelijk de beste verklaring geeft van de schoolprestaties van het kind is het model met de opleidingsniveaus van beide ouders, het gezinsinkomen en herkomstgroepering. Hoe hoger het opleidingsniveau van de ouders, hoe beter de schoolprestaties. De hoge samenhang tussen het opleidingsniveau van ouders en kinderen die in het verleden is aangetoond geldt dus nog steeds (Dronkers 1993; De Graaf en Ganzeboom 1993; Bakker en Creemers 1994; Blau en Duncan, 1967). De schoolprestaties van kinderen met een Marokkaanse of Turkse achtergrond blijven achter bij die van autochtone kinderen. Hoe vaders en moeders hun inkomsten verdienen en uit welk type huishouden de kinderen komen blijkt geen invloed te hebben.

Er zijn verschillende factoren die we niet hebben gemeten. Kloosterman (2010) liet zien dat naast het opleidingsniveau, ook de taalvaardigheid, de leessocialisatie en de mate van betrokkenheid van ouders bij de schoolactiviteiten van de kinderen (taal en rekenen) sterk samenhangen met de schoolprestaties. Daarnaast spelen - zoals aangegeven - belangrijke andere kenmerken van het milieu van herkomst, zoals het cultureel aanbod in het gezin (het aantal boeken, museumbezoek) en genetische factoren, een rol. Deze ongemeten factoren – die overigens wel samenhangen met het opleidingsniveau maar tevens een zelfstandig bijdrage vormen – liggen aan de basis van een groot deel van het twee derde deel onverklaarde variantie van het gezin.

Verder hebben we geen rekening gehouden met eventuele gebeurtenissen die plaatsvonden tussen het moment van de deelname aan de Cito-toets en de meting drie jaar later, bijvoorbeeld een echtscheiding van de ouders, of juist een rustigere gezinssituatie na de scheiding, een verhuizing naar een betere buurt, huiselijk geweld of werkloosheid van een ouder. Deze kunnen deels samenhangen met een lager of een hoger niveau in het voortgezet onderwijs.

Als laatste houden we in het toegepaste multilevel siblingmodel geen rekening met clustering binnen scholen. Er kan namelijk ook een zogeheten schooleffect bestaan: de ene school is mogelijk beter dan de andere. Echter, omdat kinderen uit één gezin vaak dezelfde school

bezoeken zal een eventueel schooleffect waarschijnlijk niet veel toevoegen of afdoen aan het gezinseffect.

De beginpopulatie bestaat uit kinderen van wie beide juridische ouders bekend zijn. Het is niet duidelijk in hoeverre de uitkomsten representatief zijn voor kinderen van wie één of allebei de ouders onbekend zijn. Vaak betreft het een onbekende vader (ongeveer 5 procent) en is de moeder van een kind met een onbekende vader vaker relatief jong, vaker allochtoon (vooral Antilliaans) en gemiddeld lager opgeleid. Niet-westerse allochtone kinderen zijn relatief vaak goed geïntegreerd, en hebben weinig taalachterstand, maar we missen nogal eens gegevens over de opleidingen die hun ouders in het buitenland genoten hebben.

De ontwikkelde intergenerationale weegfactor kan worden gebruikt voor verdere analyses. Op dit moment is deze voorhanden voor alle personen die in 2006 jonger waren dan 25 jaar. Het is de bedoeling dit de komende jaren uit te breiden. Omdat ieder jaar weer alle behaalde diploma's integraal via registers worden toegevoegd, wordt het opleidingsniveaubestand ieder jaar completer. Als gevolg daarvan is de variabele 'opleidingsniveau' van vader en/of moeder steeds beter te koppelen en te gebruiken in nadere analyses van de schoolloopbanen en de verdere levensloop van kinderen.

## Referenties

Bakker, B.F.M., J. van Rooijen & L. van Toor, 2014, The system of social statistical datasets of Statistics Netherlands: An integral approach to the production of register-based social statistics. *Journal of the International Association for Official Statistics*, 1–14.

Bakker, B.F.M., R. van Gaalen, J. van der Laan, S. Scholtus & S. Westerman, 2014, Rapportage haalbaarheidsonderzoek sociaaleconomische status index van het gezin gerelateerd aan schoolprestaties van leerlingen met gegevens uit het SSB in opdracht van het Ministerie van OCW, CBS Den Haag/Heerlen.

Bakker, B.F.M., en P.G.J. Creemers, 1994. Gelijke kansen in het onderwijs? Een vergelijking van vier cohorten leerlingen in hun overgang naar het voortgezet onderwijs, *Tijdschrift voor Onderwijsresearch*, vol. 19, 191–203.

Bates, D., M. Maechler & Ben Bolker, 2012. lme4: Linear mixed-effects models using Eigen and Eigen. R package version 0.999999-0. <http://CRAN.R-project.org/package=lme4>

Blau, P.M., & O. D. Duncan, 1967, *The American Occupational Structure* (New York, London, Sydney: Wiley).

Boudon, R., 1974. *Education, opportunity and social inequality* (New York: Wiley).

Dronkers, J., 1993. Educational reform in the Netherlands: did it change the impact of parental occupation and education? In: *Sociology of Education*, vol. 66, 262–277

Graaf, P.M. de, en H.B.G. Ganzeboom, 1993. Family background and educational attainment in the Netherlands for the 1891-1960 birth cohorts, Y. Shavit & H.-P. Blossfeld (eds.), *Persistent inequality: Changing educational attainment in thirteen countries* (Boulder: Westview Press) 75–99

Kloosterman, R., 2010. Social background and children's educational careers. *The primary and secondary effects of social background over transitions and over time in the Netherlands* (Nijmegen: Radboud Universiteit).

Linder, F., D. van Roon, & B.F.M. Bakker, 2011, Combining data from administrative sources and sample surveys; the single-variable case. Case study: educational attainment. Rapport (Work Package 4.2 van het ESSnet-project 'Data Integration'), CBS, Den Haag.

R Core Team 2012, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.

Sieben, I., 2001. *Sibling similarities and social stratification. The impact of family background across countries and cohorts* (Nijmegen: ICS).

## Verklaring van tekens

.	Gegevens ontbreken
*	Voorlopig cijfer
**	Nader voorlopig cijfer
x	Geheim
–	Nihil
–	(Indien voorkomend tussen twee getallen) tot en met
0 (0,0)	Het getal is kleiner dan de helft van de gekozen eenheid
Niets (blank)	Een cijfer kan op logische gronden niet voorkomen
2013–2014	2013 tot en met 2014
2013/2014	Het gemiddelde over de jaren 2013 tot en met 2014
2013/'14	Oogstjaar, boekjaar, schooljaar enz., beginnend in 2013 en eindigend in 2014
2011/'12–2013/'14	Oogstjaar, boekjaar, enz., 2011/'12 tot en met 2013/'14

In geval van afronding kan het voorkomen dat het weergegeven totaal niet overeenstemt met de som van de getallen.

## Colofon

### *Uitgever*

Centraal Bureau voor de Statistiek  
Henri Faasdreef 312, 2492 JP Den Haag  
[www.cbs.nl](http://www.cbs.nl)

### *Inlichtingen*

Tel. 088 570 70 70, fax 070 337 59 94  
Via contactformulier: [www.cbs.nl/infoservice](http://www.cbs.nl/infoservice)

© Centraal Bureau voor de Statistiek, Den Haag/Heerlen, 2014.  
Verveelvoudigen is toegestaan, mits het CBS als bron wordt vermeld.