

Effecten van zittenblijven in de derde kleuterklas op de wiskundegroei:

Een propensityscore-stratificatie-analyse

M. Vandecandelaere, G. Vanlaar, M. Goos, B. De Fraine & J. Van Damme

Samenvatting

In Vlaanderen doet ongeveer 4% van alle kleuters de derde kleuterklas opnieuw. Door middel van een propensityscore-stratificatie-analyse werden de effecten van zittenblijven in de derde kleuterklas op de groei in wiskunde doorheen het lager onderwijs onderzocht. Er werd gebruik gemaakt van de data van een grootschalig, longitudinaal onderzoek in Vlaanderen, namelijk het project 'Schoolloopbanen in het basisonderwijs' (SIBO). Er kan geconcludeerd worden dat zittenblijvers op het einde van het bisjaar beduidend lager scoren dan vergelijkbare leeftijdsgenoten die overgingen naar het eerste leerjaar. Op 10-jarige leeftijd is het prestatieverschil tussen beide groepen verminderd tot minder dan de helft. Toch zouden zittenblijvers gedurende hun ganse loopbaan in het lager onderwijs, gemiddeld genomen, hoger hebben gescoord voor wiskunde, waren ze toch overgegaan naar L1. Uit de resultaten blijkt voorts dat leerlingen die blijven zitten in de derde kleuterklas en equivalente leerlingen die later in hun lagere schoolloopbaan blijven zitten, vijf jaar na de treatment gelijk scoren voor wiskunde. Dat betekent dat het tijdstip van zittenblijven niet uitmaakt voor wat betreft de wiskundescores op 10-jarige leeftijd.

1 Inleiding

De overgang van de derde kleuterklas (groep 2 in Nederland, verder aangeduid met K3) naar het eerste leerjaar (groep 3 in Nederland, verder aangeduid met L1) vormt een belangrijke mijlpaal in de levensloop van kinderen. In vergelijking met K3 wordt L1 onder andere gekenmerkt door een breder en dieper curriculum en meer instructietijd voor lezen, schrijven en rekenen (Bossaert, Doumen, Buyse, & Verschuere, 2011; La Paro, Rimm-Kaufman, & Pianta, 2006). Bij

de overgang naar L1 wordt vaak de vraag gesteld of een kind al dan niet 'schoolrijp' is voor deze stap. Echter, de term schoolrijpheid is controversieel en verschillende visies vullen het begrip anders in. Enerzijds houdt de *biologische maturiteitsvisie* in dat leren pas mogelijk is als het kind er biologisch rijp voor is (Smith & Shepard, 1988). Vanuit dit perspectief is schoolonrijpheid te wijten aan onvoldoende rijping en moet het kind de nodige tijd krijgen om zich verder te ontwikkelen. Anderzijds wordt schoolonrijpheid vanuit de *remediërende visie* gezien als een gevolg van ervaring en de leeromgeving. Extra inspanningen van ouders en leerkrachten kunnen het verschil maken (Smith & Shepard, 1988). Algemeen wordt schoolrijpheid gezien als een compromis tussen beide visies en dus als het resultaat van de interactie of wisselwerking tussen het kind en zijn omgeving (Mashburn & Pianta, 2006). Oordelen of een kind al dan niet klaar is om de lagere school aan te vatten is dan ook geen sinecure; zowel kind- als omgevingsfactoren vormen belangrijke elementen bij dit oordeel. Bovendien zijn er geen formele regels die de toepassing van zittenblijven in K3 reguleren. In tegenstelling tot vele andere Europese landen, waar de beslissing gebeurt op basis van vastliggende regels en resultaten op gestandaardiseerde toetsen, is de beslissing of een kind al dan niet blijft doorkleuteren in Vlaanderen het resultaat van overleg tussen ouders, leerkrachten, directie en het CLB (Centrum voor Leerlingenbegeleiding)¹.

Wanneer een kind niet blijkt te voldoen aan de vereisten van de lagere school staan leerkrachten, ouders en andere betrokkenen dus voor een moeilijke beslissing. Vaak wordt geopteerd om de stap naar L1 uit te stellen en het kind nog een jaartje te laten 'doorkleuteren'. Deze praktijk is populair in onder meer Frankrijk, Spanje, Nederland, Duitsland en België (Ikeda, 2011). In Vlaanderen bleef in schooljaar 2003-2004 4%

van de kinderen zitten in K3². In Nederland bedraagt het aantal kinderen dat blijft zitten in groep 2 zo'n 10% (Van der Grift, 2005).

Uit verschillende internationale studies blijkt dat vooral leerlingen die zwakker presteren voor taal en wiskunde (rekenen), leerlingen met gedrags- en/of emotionele problemen, jonge kinderen³, allochtone leerlingen en leerlingen van wie de moeder geen diploma hoger onderwijs heeft behaald het advies krijgen om de derde kleuterklas over te doen (Byrd & Weitzman, 1994; Cosden, Zimmer, & Tuss, 1993; Elder & Lubotsky, 2009; Gadeyne, Onghena, & Ghesquiere, 2008; Hauser, Pager, & Simmons, 2001; Mantzicopoulos & Neuharth-Pritchett, 1998).

Zoals ook uit de andere artikels in dit nummer blijkt, is de praktijk van zittenblijven ondanks zijn populariteit niet onomstreden. De controverse vraagt om meer en contextspecifiek onderzoek naar de effecten van zittenblijven, waaraan deze studie tegemoet tracht te komen. In dit inleidend hoofdstuk wordt verder nog ingegaan op de Vlaamse onderzoekscontext en bestaand onderzoek naar de effecten van zittenblijven en worden de onderzoeksvragen van deze studie geformuleerd.

1.1 Vlaamse onderzoekscontext

Dit onderzoek werd uitgevoerd in Vlaanderen. Vlaamse onderwijsstatistieken indiceren dat het aantal kinderen dat met vertraging start in L1 toeneemt. Van de jongens die geboren zijn in 1983, startte 3.1% met vertraging aan het lager onderwijs, dit in vergelijking met 6.5% van de jongens die geboren zijn in 2004. Het aantal vertraagde meisjes aan de start van de lagere school, geboren in 1983 en 2004, steeg respectievelijk van 2.0% naar 4.8% (Vlaams Ministerie van Onderwijs en Vorming, 2012). Deze kinderen zijn met vertraging in de kleuterschool gestart of dubbelden een jaar in het kleuteronderwijs. In het decreet basisonderwijs staat aangegeven dat kinderen maximum acht jaar in het lager onderwijs kunnen doorbrengen (Ministerie van de Vlaamse Gemeenschap, 1997). Dit impliceert dat kinderen die meer dan twee keer blijven zitten in hun lagere schoolloopbaan het diploma basisonderwijs niet op reguliere wijze kunnen behalen. Deze regel kan zittenblijven in K3 in de hand werken omdat kinderen a priori een extra jaar de

tijd krijgen om te ontwikkelen alvorens ze hun lagere schoolloopbaan aanvangen. Een recente aanpassing van het decreet houdt in dat sinds 1 september 2010, vijf- en zesjarige leerlingen pas kunnen starten in het gewoon lager onderwijs van een Nederlandstalige school als ze het schooljaar voordien minstens 220 halve dagen Nederlandstalig kleuteronderwijs gevolgd hebben. Is dit niet het geval dan kunnen ze via een succesvolle taalproef toch nog toegelaten worden (Ministerie van de Vlaamse Gemeenschap, 1997). Hierdoor zal het aantal zittenblijvers in de toekomst vermoedelijk nog stijgen.

Verder blijkt dat zittenblijven in Vlaanderen min of meer maatschappelijk aanvaard wordt. Leraren en directeurs zien zittenblijven eerder als een goede pedagogische maatregel voor leerlingen die niet schoolrijp zijn of leerlingen die de basiskennis nog onvoldoende beheersen. Schoolonrijpheid wordt bovendien als de belangrijkste reden aangegeven voor zittenblijven in K3 (Juchtmans, Goos, Vandenbroucke, & De Fraine, 2012).

1.2 Effecten van zittenblijven in de derde kleuterklas

In internationaal onderzoek is de vraag naar de effecten van zittenblijven regelmatig aan de orde gesteld (voor een recente meta-analyse, zie Allen, Chen, Willson en Hughes (2009); voor frequent geciteerde meta-analyses, zie Holmes (1989) en Jimerson (2001)). Algemeen blijkt uit deze studies dat zittenblijven weinig of geen positieve effecten teweegbrengt, noch op het vlak van de cognitieve, noch op het vlak van de psychosociale ontwikkeling van jonge kinderen (Alet, 2010; Chen, Liu, Zhang, Shi, & Rozelle, 2010; Dong, 2010; Goos, Van Damme, Onghena & Petry 2012; Hofer, Farran, Lipsey, Aydogan, & Bilbrey, 2010; Hong & Yu, 2008; Lorence & Dworkin, 2006; Wu, West, & Hughes, 2008; Wu, West, & Hughes, 2010). De resultaten van deze onderzoeken zijn echter vaak wisselend en de onderzoeken verschillen sterk qua methodologie.

Onderzoeken die een eerlijke vergelijking maken tussen zittenblijvers en een controlegroep maken gebruik van een quasi-experimenteel design. Propensity-matching wordt algemeen aanvaard als de meest accurate

methode om te controleren voor waarneembare verschillen tussen zittenblijvers en controlegroep wanneer een random toewijzing niet mogelijk is (Allen, Chen, Willson, & Hughes, 2009; Raudenbush & Bryk, 2002; Shadish, Cook, & Campbell, 2002). Twee soorten vergelijkingen kunnen hierbij worden onderscheiden: (1) in een *leerjaarvergelijking* worden zittenblijvers vergeleken met leerlingen die in hetzelfde leerjaar zitten maar K3 niet overdeden en die dus een jaar jonger zijn, (2) in een *leeftijdsvergelijking* worden zittenblijvers met hun leeftijdsgenoten vergeleken, ongeacht het leerjaar waarin ze zitten.

Hong en collega's pasten een leeftijdsvergelijking toe en vergeleken de prestaties van zittenblijvers en vergelijkbare leeftijdsgenoten die overgingen naar L1 (Hong & Raudenbush, 2005; 2006; Hong & Yu, 2007). De auteurs stelden vast dat zittenblijvers een jaar later beter zouden hebben gepresteerd voor lezen en wiskunde, waren ze toch overgegaan naar L1. Zittenblijvers lopen tijdens hun bisjaar een achterstand op in vergelijking met vergelijkbare leerlingen die doorstromen en nieuwe leerstof krijgen aangeboden. De auteurs vonden echter geen negatieve effecten twee en vier jaar na het zittenblijven. De opgelopen achterstand wordt met andere woorden ingehaald.

Dong (2010) vergeleek de prestatiegroei van zittenblijvers en vergelijkbare, jongere, *leerjaargenoten* die niet bleven zitten. De auteur stelde vast dat zittenblijvers gedurende de eerste drie jaar na het zittenblijven beter presteren in vergelijking met vergelijkbare leerjaargenoten. De zittenblijvers zijn in deze vergelijking dan ook een jaar ouder dan vergelijkbare leerjaargenoten die niet bleven zitten in K3. Niettemin, het effect verdwijnt eveneens op lange termijn, m.a.w., de voorsprong die zittenblijvers tijdens het bisjaar maken is van beperkte duur.

Naast de effecten voor de zittenblijvers in kwestie is zittenblijven ook een dure interventie. De gemiddelde kost van zittenblijven voor een leeftijdscohorte neemt in België naar schatting 10% van de jaarlijkse nationale uitgave aan basis- en secundair onderwijs in beslag. Dit komt neer op een kost van ongeveer 8.500 EUR per zittenblijver (Ikeda, 2011).

1.3 Onderzoeksvragen

Zittenblijven op jonge leeftijd wordt in Vlaanderen door velen beschouwd als een zinvolle zaak (Juchtmans et al., 2012). Maar is dat wel zo? Het doel van dit onderzoek is door middel van een leeftijdsvergelijking zicht te krijgen op de effecten van zittenblijven in K3 op de groei in wiskunde van leerlingen doorheen het lager onderwijs. Met wiskundegroei wordt de evolutie in wiskundeprestaties bedoeld over een bepaalde periode. Wiskunde (of rekenen) is naast wetenschappen en moedertaal een van de meest onderzochte onderwijsuitkomsten (Scheerens & Bosker, 1997; Van Damme, 2004). Uit onderzoek blijkt immers dat de wiskundeactiviteit in de kleuterklas een belangrijke voorspeller is voor prestaties en leerwinst in wiskunde doorheen het lager onderwijs (Jordan, Kaplan, Ramineni, & Locuniak, 2009). Bovendien is de ontwikkeling van wiskundige inzichten en vaardigheden van cruciaal belang voor een adequaat functioneren in het maatschappelijk en professionele leven.

Aan de hand van een quasi-experimenteel design werd tijdens de analyses gecontroleerd voor een hele reeks leerlingkenmerken en enkele schoolkenmerken, door een beroep te doen op de zogenaamde 'propensityscore' van leerlingen. De propensityscore is de conditionele kans die een bepaalde leerling heeft om te blijven zitten, gegeven alle relevante leerling- en schoolkenmerken samen. Door deze score in rekening te brengen, worden leerlingen vergeleken die equivalent zijn in termen van risicofactoren die zittenblijven voorafgaan. 'Zittenblijvers' worden in dit onderzoek gedefinieerd als de leerlingen die K3 in schooljaar 2003-2004 dubbelden. De 'controlegroep' omvat in dit onderzoek de leerlingen die in vergelijking met de zittenblijversgroep equivalent zijn in termen van risicofactoren maar die in schooljaar 2003-2004 L1 volgden. De wiskundegroei van deze twee groepen werd vergeleken tussen 6-jarige en 10-jarige leeftijd, tussen juni 2004 en juni 2008.

Dit onderzoek onderscheidt zich van vorig onderzoek naar de effecten van zittenblijven in K3 door ook aandacht te besteden aan de schoolloopbaan na de transitie naar L1. Zoals Hong en Yu (2007, 2008) benadrukken, is het waarschijnlijk dat leerlingen

met een grote kans op zittenblijven in K3 maar die toch overgaan naar L1, later in de schoolloopbaan blijven zitten of overstappen naar het buitengewoon onderwijs. In voorliggende studie werd na een (1) leeftijdsvergelijking tussen zittenblijvers en controlegroep (2) de loopbaan na de treatment in kaart gebracht. Op basis van die informatie werd een (3) tweede vergelijking gerealiseerd tussen zittenblijvers en drie subgroepen van de controlegroep, opgedeeld naar hun lagere schoolloopbaan. In totaal werden dus drie onderzoeksvragen voorop gesteld:

Vergelijking tussen zittenblijvers en de controlegroep

Leerlingen die K3 overdoen, hebben een jaar vertraging en staan dus een jaar achter wat betreft de leerstof voor wiskunde. Verwacht wordt dat deze leerlingen tijdens hun bisjaar lager scoren voor wiskunde dan hun vergelijkbare leeftijdsgenoten die in L1 zitten en nieuwe leerstof krijgen aangeboden. Leraren die het advies geven om een leerling K3 te laten overdoen, gaan ervan uit dat deze leerling het tempo van L1 niet zal kunnen volgen. De hypothese in dit onderzoek luidt dan ook dat leeftijdsgenoten die equivalent zijn in termen van risicofactoren die zittenblijven voorafgaan maar die wel overgaan naar L1, op langere termijn problemen zullen hebben met wiskunde en dat zittenblijvers op langere termijn minstens een deel van hun achterstand ten opzichte van deze groep zullen inhalen. De verwachting luidt met andere woorden dat de prestatiekloof tussen zittenblijvers en equivalente doorgestroomde leerlingen op lange termijn afneemt.

Volgende onderzoeksvraag werd beantwoord: 'Wat is de groei in wiskundeprestaties van zittenblijvers tussen 6-jarige (schooljaar 2003-2004) en 10-jarige (schooljaar 2007-2008) leeftijd in vergelijking met de controlegroep?' Deze vraag kan ook als volgt geformuleerd worden: 'Hoe zou de wiskundeontwikkeling van zittenblijvers er uit hebben gezien, waren ze toch overgegaan naar het eerste leerjaar?'

Loopbaanpositie van de zittenblijvers en de controlegroep in schooljaar 2008-2009

Niet alle leerlingen uit de controlegroep volg-

den een modeltraject waarbij ze in schooljaar 2008-2009 in het zesde leerjaar zaten. Zoals Hong en Yu (2007, 2008) benadrukken, is het waarschijnlijk dat leerlingen met een grote kans op zittenblijven in K3 maar die toch overgaan naar L1, later in de schoolloopbaan blijven zitten of overstappen naar het buitengewoon onderwijs. Om een accuraat beeld te vormen van de wiskundegroei van de controlegroep was het dus van belang deze loopbaan in kaart te brengen. Als noodzakelijke tussenstap werd daarom volgende vraag beantwoord: Wat was de loopbaanpositie van de leerlingen in de steekproef in het schooljaar 2008-2009 (zesde leerjaar voor normaalvorderenden)?

Vergelijking tussen zittenblijvers en controlegroep, opgedeeld naar loopbanen

Op basis van de antwoorden op de vorige vraag was het mogelijk om de controlegroep op te delen in drie loopbaangroepen. Tenslotte werd volgende onderzoeksvraag beantwoord: Wat is de groei voor wiskunde van de groep zittenblijvers tussen 6-jarige (schooljaar 2003-2004) en 10-jarige (schooljaar 2007-2008) leeftijd in vergelijking met de controlegroep, opgesplitst naar drie loopbaangroepen?

2 Methode

2.1 Steekproef

Er werd gebruik gemaakt van de data van een grootschalig, longitudinaal onderzoek in Vlaanderen, namelijk het project 'Schoolloopbanen in het basisonderwijs' (SiBO) (Maes et al., 2002). De doelstelling van dit project is het beschrijven en verklaren van de schoolloopbanen en ontwikkeling van Vlaamse leerlingen doorheen het basisonderwijs. Hiertoe werd een cohorte van ongeveer 6000 leerlingen gevolgd vanaf de aanvang van K3 (schooljaar 2002-2003) tot het einde van de eerste graad secundair onderwijs (schooljaar 2010-2011).

De steekproef betrof alle leerlingen in het SiBO-project die in schooljaar 2002-2003 voor het eerst in K3 zaten (N=5624) en in schooljaar 2003-2004 doorstroonden naar L1 (N=5296) of bleven zitten in K3 (N=328).

Tabel 1

Aantallen in de steekproef en geselecteerde leerlingengroepen

	Aantal leerlingen		Totaal	Aantal klassen Totaal	Aantal scholen ⁵ Totaal
	Zittenblijvers K3	Niet-zittenblijvers K3			
Steekproef	328	5296	5624	365	259
Selectie	315	1389	1704	348	250

Leerlingen die na K3 overstapten naar het buitengewoon onderwijs werden niet meege-
nomen in de steekproef. In schooljaar 2002-
2003 had 1.5% van de kleuters in deze steek-
proef reeds een jaar achterstand (geboren in
1996). Daarbij gaat het vaak om anderstalige
kinderen en/of kinderen die de tweede kleu-
terklas dubbelden. In schooljaar 2002-2003
had 0.8% van de kleuters een jaar voorsprong
(geboren in 1998).

Van deze steekproef stroomde in school-
jaar 2003-2004 94.17% door naar L1 en bleef
5.83% van de leerlingen zitten in K3 (zie
Tabel 1)⁴.

2.2 Analyses

Om een afdoend antwoord te formuleren op
de onderzoeksvragen werd gebruik gemaakt
van 'propensityscore stratificatie'. Deze qua-
si-experimentele methode maakt het moge-
lijk om zittenblijvers te vergelijken met leer-
lingen die hetzelfde profiel en dus dezelfde
kans hebben om te blijven zitten, maar toch
overgaan naar L1. Met andere woorden, er
worden twee leerlinggroepen geselecteerd
die vergelijkbaar zijn op tal van kenmer-
ken en enkel verschillen in het feit dat de
ene groep blijft zitten in K3 (experimentele
groep) en de andere groep overgaat naar L1
(controlegroep). Aangezien de leerlingen niet
verschillen op belangrijke kenmerken die zit-
tenblijven voorafgaan, kunnen verschillen in
de groei in wiskundepreraties toegeschreven
worden aan het al dan niet blijven zitten (de
zgn. 'treatment').

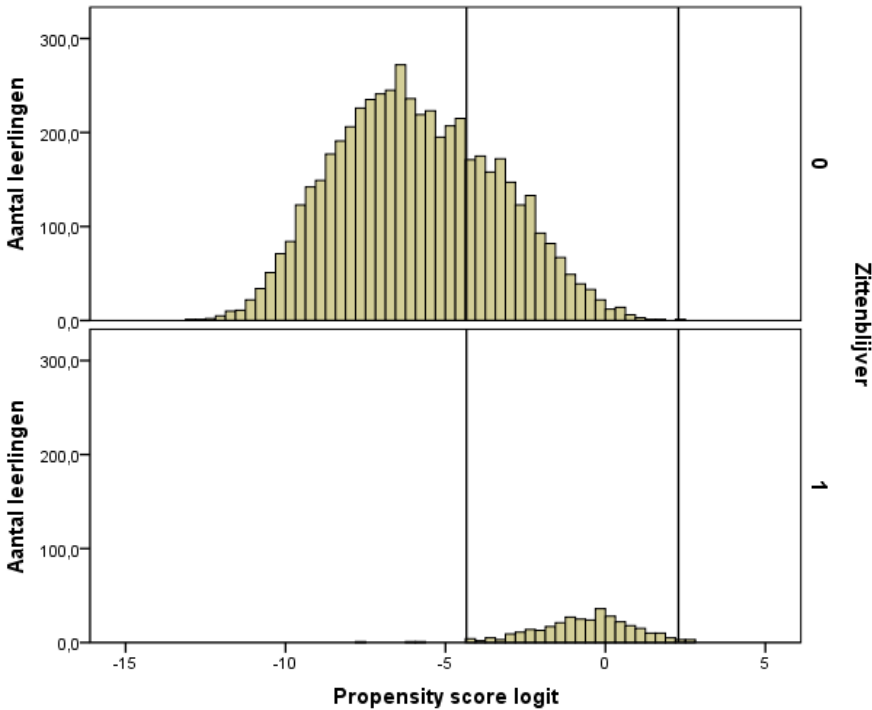
Propensityscore analyse heeft een aantal
voordelen in vergelijking met lineaire regres-
sie (zie o.a. Hill, Reiter en Zunutto (2004)).
Zo veronderstelt lineaire regressie dat de ver-
delingen van de covariaten dezelfde zijn voor
de controle- en experimentele groep. Wan-
neer dit niet zo is, zal het model schattingen
gaan extrapoleren. Propensityscores laten toe

om de leerlingen te selecteren waarvoor er
een match is in de andere conditie, waarmee
extrapolatie wordt vermeden. Daarenboven
spreekt propensityscore analyse intuïtief
meer aan aangezien vergelijkbare leerlingen
in de controle en experimentele conditie ver-
geleken worden. Met lineaire regressie is dit
minder transparant.

Concreet werd een vierstappenprocedure
toegepast die in grote mate gelijklopend is
aan de procedure van Hong en Yu (2008).
In een eerste stap werden de zogenaamde
'true confounders' geïdentificeerd. Dat zijn
de geobserveerde pre-treatment kenmerken
die zowel met de *treatment* als met de *out-
come* (wiskundepreraties in de daaropvol-
gende jaren) samenhangen (Brookhart et al.,
2006; Judkins et al., 2007). Vervolgens werd
voor elke leerling een propensityscore bere-
kend. Dit is de voorwaardelijke kans om te
blijven zitten in K3, gegeven de geïdentifi-
ceerde 'true confounders'. In de derde stap
werden de geselecteerde leerlingen op basis
van hun propensityscore opgedeeld in tien
deelgroepen of strata. In de vierde en laatste
stap werden aan de hand van een reeks drie-
niveaugroei-curve-modellen (metingen binnen
leerlingen binnen scholen) de gemiddelde
effecten van zittenblijven in K3 op de groei
in wiskundepreraties berekend. Daarbij wer-
den de strata waartoe elke leerling behoort en
de treatment opgenomen als covariaten. De
vier stappen worden in wat volgt meer in
detail besproken.

Identificatie van de 'true confounders'

De SiBO-databank bevat naast toetsscores
voor wiskunde en taal (Verachtert, 2003)
ook gegevens die verzameld zijn via een
leerkrachtvragenlijst over de kinderen (Maes,
2003a), een leerkrachtvragenlijst over de
klaspraktijken en didactische overtuigingen
(Gadeyne, 2003), een oudervragenlijst



Figuur 1. Verdeling van de propensityscore logit voor zittenblijvers (beneden) en niet zittenblijvers (boven).

(Reynders, Van Heddegem, Nicaise, & Van Damme, 2004), een schoolteamvragenlijst (Maes, 2003b) en een directievragenlijst (Verhaeghe, 2004). In totaal waren 258 variabelen op leerlingniveau, 123 variabelen op klasniveau en 105 variabelen op schoolniveau beschikbaar. Alle variabelen met een significante samenhang ($p < .001$) en met een effectgrootte van $r > .10$ of $\eta^2 > .01$ werden beschouwd als voldoende samenhangend met zowel treatment als outcome om als ‘true confounder’ geselecteerd te worden.

Van deze 486 variabelen die werden gemeten vóór de treatment (nl. blijven zitten in K3), werden 37 leerlingkenmerken en twee schoolkenmerken geïdentificeerd als ‘true confounders’⁶. Zowel kenmerken die eigen zijn aan het kind zelf als kenmerken die eigen zijn aan het kind in zijn omgeving werden geïdentificeerd, wat de twee eerder genoemde facetten in de definitie van schoolrijpheid weerspiegelt. Voorbeelden van belangrijke leerlingkenmerken zijn de geboortemaand en de prestaties voor taal en wiskunde in K3. Kenmerken van het kind in zijn omgeving zijn bijvoorbeeld de socio-

economische status van het gezin en het opleidingsniveau van de moeder.

De ontbrekende waarden van deze 39 variabelen werden geïmputeerd⁷ met behulp van het EM-algoritme in SPSS (SPSS Inc, 2010). Daarbij werden de overige 447 variabelen gebruikt als hulpvariabelen. Het percentage ontbrekende waarden varieerde tussen 0% en 17% met een gemiddelde van 7%.

Schatting van de propensityscores

In een tweede stap werd voor elke leerling een propensityscore berekend. Dit is de voorwaardelijke kans om te blijven zitten in K3, gegeven de 39 geïdentificeerde ‘true confounders’. De berekening gebeurde door middel van een drielniveau logistisch regressiemodel (5624 leerlingen in 365 klassen in 259 scholen).

Figuur 1 geeft een visuele voorstelling van de verdelingen. De bovenste verdeling geeft de logit van de geschatte propensityscores weer van de leerlingen die doorstroomden naar L1. De onderste verdeling geeft de logit van de geschatte propensityscores weer van de leerlingen die bleven zitten in K3. Logi-

Tabel 2

Aantal zittenblijvers en niet-zittenblijvers per stratum

Stratum	1	2	3	4	5	6	7	8	9	10	Totaal
Niet-zittenblijver	167	169	165	168	160	153	148	130	90	39	1389
Zittenblijver	4	2	6	3	10	17	22	40	80	131	315

scherwijs hebben de leerlingen die bleven zitten gemiddeld genomen een hogere propensiteitscore logit dan leerlingen die doorstroomden naar L1. Omwille van de focus op de wiskundegroei van zittenblijvers in vergelijking met die van leerlingen die niet zijn blijven zitten in K3 maar toch equivalent zijn in termen van risicofactoren (true confounders), werden enkel die leerlingen geselecteerd die zich bevinden in de zone waar de twee verdelingen elkaar overlappen. De deelverzamelingen tussen de twee verticale lijnen werden geselecteerd voor het vervolg van het onderzoek. In Tabel 1 worden de aantallen weergegeven van de geselecteerde leerlingengroepen.

Stratificatie op basis van de geschatte propensiteitscores

In de derde stap werden de geselecteerde leerlingen opgedeeld in tien deelgroepen of strata. De propensiteitscores zijn binnen elk stratum gebalanceerd wat betekent dat de scores van zittenblijvers en niet-zittenblijvers binnen eenzelfde stratum niet significant van elkaar verschillen. Daarenboven is 92% van de 'true confounders' gebalanceerd na decielstratificatie. Dat betekent onder andere dat de zittenblijversgroep en de controlegroep niet verschillen voor wat betreft de wiskundescore op het einde van K3 in schooljaar 2002-2003. In Tabel 2 worden de aantallen per stratum weergegeven.

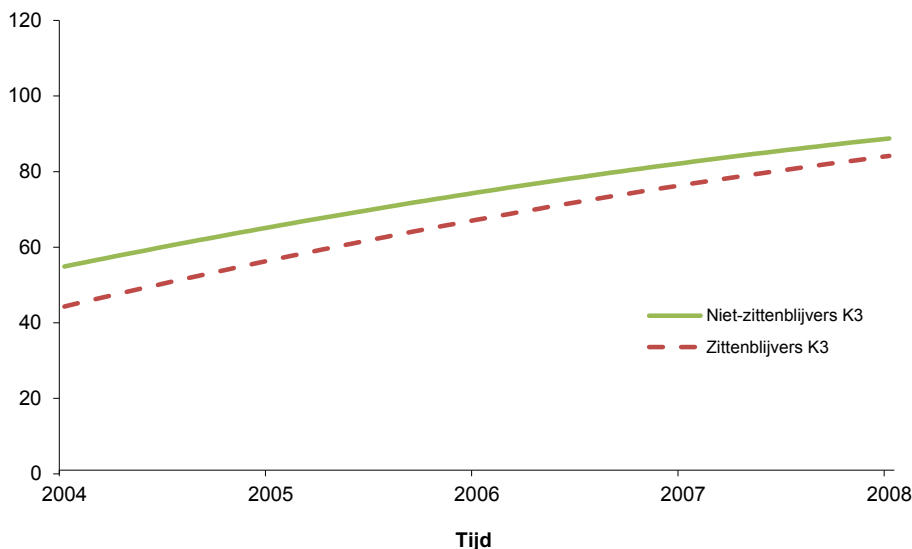
Schatting van de gemiddelde effecten van zittenblijven op de groei in wiskunde van leerlingen

In het SiBO-project werd op het einde van elk schooljaar een wiskundetoets afgenomen die peilt naar het wiskundig inzicht en de rekenvaardigheid van de leerlingen. De wiskundetoetsen hebben tot doel een betrouwbare en juiste beoordeling te geven van het

wiskundig inzicht en de rekenvaardigheid van kinderen. Per leerjaar werd een wiskundetoets ontwikkeld. Vertraagde leerlingen kregen de toets voorgelegd die aansloot bij het leerjaar waarin ze zaten. Met een statistische techniek die gebaseerd is op de zogenaamde Item Response Theory (IRT), werden de toetsen gekalibreerd. De scores van de verschillende toetsen werden hiermee op eenzelfde meetschaal geplaatst en daardoor onderling vergelijkbaar gemaakt.

De toets die werd afgenomen in L1 bestaat uit volgende toetsdelen: 'Getallenkennis', 'Vraagstukken', 'Toepassingsituaties metend rekenen', 'Structureren of splitsen van getallen', 'Getallenreeksen' en 'Eenvoudige bewerkingen met tijdslimiet'. 'Getallenkennis' bijvoorbeeld bevat verbale opgaven, waarbij een beroep gedaan wordt op bewerkingsstermen zoals 'meer dan', 'de helft' of 'het dubbele'. In het toetsdeel 'Vraagstukken' bijvoorbeeld krijgen kinderen meer concrete probleemsituaties voorgelegd waarbij ze moeten rekenen om een oplossing te vinden. De toets die werd afgenomen in het tweede leerjaar bestaat daarenboven ook uit een toetsdeel 'Combinatie-opgaven' en een toetsdeel waarin gespeeld wordt naar de vaardigheid in 'Meetkunde'. In het derde leerjaar werd ook gespeeld naar de vaardigheid en inzicht in 'Hoofdrekenen' en 'Hoofdbewerkingen' en in het vierde leerjaar werd dit verder uitgebreid met de toetsdelen 'Breuken', 'Cijferen' en 'Kommagetallen'. In het vijfde leerjaar bestond de toets ook uit een toetsdeel 'Procenten'.

Aan de hand van een reeks drieniveaugroei-curve modellen (metingen binnen leerlingen binnen scholen) werden de gemiddelde effecten van zittenblijven in K3 op de groei in wiskunde prestaties van de leerlingen berekend. Daarbij werden de strata en de treatment opgenomen als covariaten. De analyses wer-



Figuur 2. Leeftijdvergelijking voor de groei in wiskunde-prestaties van zittenblijvers en niet-zittenblijvers derde kleuterklas (K3).

den uitgevoerd in MLwiN-software (Rasbash, Charlton, Browne, Healy, & Cameron, 2010).

De gemiddelde effecten werden geschat op basis van de beschikbare wiskundescores⁸. Een vergelijking van de kenmerken van de totale steekproef en de kenmerken van de kinderen waarvan een wiskundescore beschikbaar is op het laatste meetmoment (2007-2008) wijst erop dat de uitval niet willekeurig is. De resterende groep betreft kinderen die gemiddeld genomen een gunstiger profiel hebben in vergelijking met de totale steekproef. Zo scoort deze groep bijvoorbeeld gemiddeld genomen hoger voor initiële wiskunde en taalvaardigheid, socio-economische status en prosociaal gedrag. Dit betekent dat de wiskunde-prestaties op latere meetmomenten betrekking hebben op kinderen met een gemiddeld genomen kleinere kans om te blijven zitten in K3. Om zeker te zijn dat op de latere meetmomenten nog sprake was van een eerlijke vergelijking tussen de groep zittenblijvers en de controlegroep, werd de propensityscore van beide groepen vergeleken binnen elk stratum, dit voor de leerlingen met een beschikbare wiskundescore in schooljaar 2007-2008. De propensityscores waren binnen elk stratum gebalanceerd wat betekent dat leerlingen in de treatment en controleconditie binnen elk stratum een gelijke conditionele kans hadden om te blijven zitten.

3 Resultaten

3.1 Vergelijking tussen zittenblijvers en de controlegroep

“Wat is de groei voor wiskunde van zittenblijvers tussen 6-jarige (schooljaar 2003-2004) en 10-jarige (schooljaar 2007-2008) leeftijd in vergelijking met de controlegroep?” is de vraag die centraal staat. In schooljaar 2002-2003 zaten alle leerlingen nog in K3. Zoals aangehaald, werden enkel leerlingen vergeleken die equivalent zijn inzake kenmerken die het zittenblijven voorafgaan. Ook de wiskundescores op het einde van schooljaar 2002-2003 van de leerlingen die het volgende jaar bleven zitten (treatment) en de leerlingen die het volgende jaar naar L1 gingen (controlegroep) verschilden niet significant.

In schooljaar 2003-2004 zat de zittenblijversgroep opnieuw in K3 en de controlegroep in L1. Figuur 2 toont aan dat er op het einde van dit schooljaar, in juni 2004, een kloof is ontstaan tussen beide groepen voor wat betreft de wiskunde-prestaties. Zoals verwacht, liepen zittenblijvers tijdens hun schooljaar een achterstand op voor wiskunde, in vergelijking met vergelijkbare leerlingen die doorstroomden naar L1.

In Figuur 2 wordt verder vastgesteld dat de zittenblijvers vier jaar later hun achterstand grotendeels inhaalden ten aanzien van de

Tabel 3

Loopbaanposities van de zittendblijvers en de controlegroep in schooljaar 2008-2009

	Aantal leerlingen					
	Zittendblijvers K3		Niet zittendblijvers K3		Totaal	
L6	3	1.0 %	678	48.8%	681	40.0%
L5	137	43.5%	454	32.7%	591	34.7%
L4	46	14.6%	60	4.3%	106	6.2%
L3	6	1.9%	1	0.1%	7	0.4%
BuLO	102	32.4%	155	11.2%	257	15.1%
BuSO	0	0.0%	1	0.1%	1	0.1%
1B	0	0.0%	4	0.3%	4	0.2%
Onbekend	21	6.7%	36	2.6%	57	3.3%
Totaal	315	100.00%	1389	100.00%	1704	100.00%

Noot: L3-L6: derde tot zesde leerjaar; BuLO: Buitengewoon Lager Onderwijs; BuSO: Buitengewoon Secundair onderwijs; 1B: Eerste jaar secundair onderwijs, bestemd voor leerlingen die behoefte hebben aan een aangepast onderwijs.

controlegroep, wat de hypothese bij de eerste onderzoeksvraag bevestigt. Op het einde van schooljaar 2007-2008 scoorden zittendblijvers nog steeds significant lager dan de controlegroep maar het verschil daalde van 11 IRT-punten (= 1.17SD) verschil in juni 2004 naar vijf IRT-punten (= 0.52SD) verschil in juni 2008. De vraag kan dan ook gesteld worden hoe de schoolloopbaan van beide groepen gedurende deze vijf volgende jaren er uit ziet en of dit een mogelijke verklaring kan bieden voor de verkleinde achterstand.

3.2 Loopbaanpositie van de zittendblijvers en de controlegroep in schooljaar 2008-2009

De loopbaanposities van de zittendblijvers en de controlegroep in schooljaar 2008-2009 worden weergegeven in Tabel 3. Schooljaar 2008-2009 is het schooljaar waarin normaalvorderende leerlingen in het zesde leerjaar zaten.

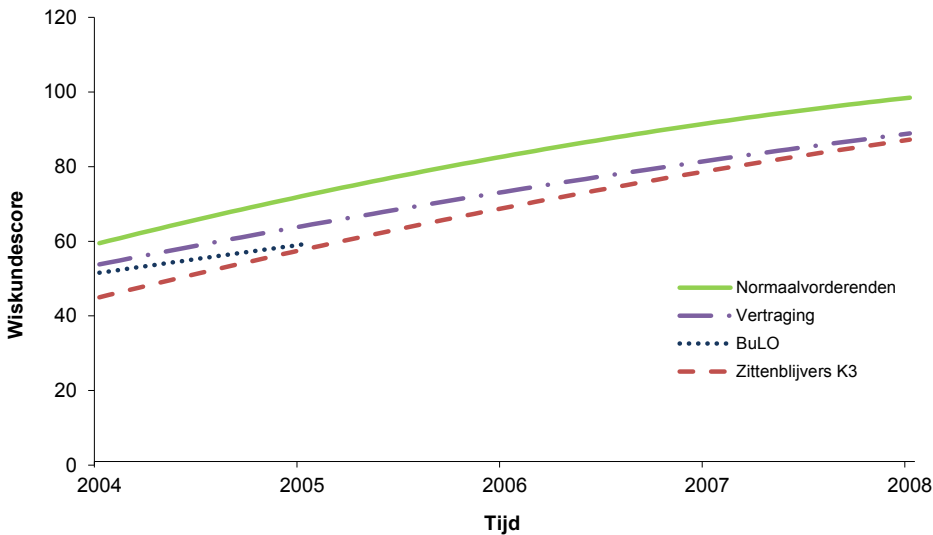
Zoals uit Tabel 3 blijkt, zat bijna 44% van de zittendblijvers in schooljaar 2008-2009 in het vijfde leerjaar. Zij hebben vanaf het eerste leerjaar t.e.m. minstens het vijfde leerjaar een 'normaal' traject gevolgd. Bijna 15% bleef tussen schooljaar 2003-2004 en schooljaar 2008-2009 nog eens zitten en liep dus een vertraging op van twee jaar. Meer dan 32% van de zittendblijvers zette in die periode de stap naar het buitengewoon onderwijs.

Van de controlegroep (N=1389) zat een kleine helft in schooljaar 2008-2009 in het zesde leerjaar. Deze subgroep heeft dus een normaal traject afgelegd. Bijna 33% van de controlegroep had in schooljaar 2008-2009 één jaar vertraging opgelopen en meer dan 4% had zelfs twee of meer jaar vertraging opgelopen. Ruim 11% van deze groep stapte tussen schooljaar 2003-2004 en schooljaar 2008-2009 over naar het buitengewoon onderwijs.

Zoals gespeculeerd door Hong en Yu (2007, 2008), is de controlegroep niet homogeen wat betreft de loopbaan die ze hebben afgelegd vanaf L1. Leerlingen met een grote kans op zittenblijven in K3 maar die toch overgaan naar L1, blijven vaak op een later tijdstip zitten of stappen op een later tijdstip over naar het buitengewoon onderwijs. Om een accuraat beeld te vormen van hun groei in wiskunde-prestaties tussen zesjarige en 10-jarige leeftijd, werd in een volgende analysestap deze controlegroep opgedeeld in drie 'loopbaangroepen'.⁹

3.3 Vergelijking tussen zittendblijvers en controlegroep, opgedeeld naar loopbanen

De controlegroep werd opgedeeld in drie loopbaangroepen. Een eerste groep omvatte de 'normaalvorderende leerlingen', namelijk de leerlingen die vanaf K3 elk jaar overgin-



Figuur 3. Leeftijdvergelijking voor de groei in wiskundeprestaties van zittenblijvers K3 en niet-zittenblijvers K3, met verschillende loopbaangroepen.

gen naar het volgende leerjaar en in schooljaar 2008-2009 in het zesde leerjaar zaten (N= 678). Een tweede groep omvatte de leerlingen met ‘vertraging’, namelijk de leerlingen die na K3 overgingen naar L1 maar in de loop van hun verdere lagere schoolloopbaan minstens één keer bleven zitten (N=515). Een derde groep bestond uit leerlingen die na K3 overgingen naar L1 maar in de loop van hun verdere lagere schoolloopbaan de stap zetten naar het buitengewoon onderwijs, verder aangeduid als ‘BuLO’ (N=155). Aangezien we voor deze groep enkel over wiskundescores beschikken t.e.m. schooljaar 2004-2005, worden in wat volgt geen uitspraken gedaan over de groei van deze groep vanaf schooljaar 2005-2006. Deze drie groepen werden vergeleken met de experimentele groep, namelijk de leerlingen die K3 bisten in schooljaar 2003-2004, ongeacht het traject dat ze vanaf schooljaar 2004-2005 aflegden (N=315).

Hoewel de zittenblijversgroep en de totale controlegroep gemiddeld genomen niet significant verschillend scoorden voor wiskunde op het einde van K3 (juni 2003), worden hiervoor wel significante verschillen gevonden tussen de zittenblijversgroep en de subgroepen van de controlegroep. Op het einde van schooljaar 2002-2003 scoorden de normaalvorderenden gemiddeld genomen hoger

dan de drie andere loopbaangroepen. Meer concreet scoorden de groepen ‘zittenblijvers K3’, ‘BuLO’ en ‘vertraging’ respectievelijk drie IRT-punten (= 0.28SD), drie IRT-punten (= 0.30SD) en twee IRT-punten (= 0.26SD) lager dan de groep ‘normaalvorderenden’.

In Figuur 3 wordt de groei van de vier groepen weergegeven na de treatment. Zoals verwacht, nam de kloof tussen de zittenblijvers en de normaalvorderenden het volgende jaar toe, om dezelfde reden als in de vergelijking met de totale controlegroep: zittenblijvers liepen tijdens het bisjaar een achterstand op voor wiskunde, in vergelijking met vergelijkbare leerlingen die doorstroomden naar L1. Op het einde van schooljaar 2007-2008 scoorden zittenblijvers nog steeds lager dan normaalvorderende leeftijdsgenoten, al haalden de zittenblijvers hun achterstand gedeeltelijk in. Het verschil tussen beide groepen daalde van 15 IRT-punten (= 1.60SD) in schooljaar 2003-2004 naar 11 IRT-punten (= 1.27SD) in schooljaar 2007-2008.

Interessant is de vergelijking tussen de K3 zittenblijvers en leerlingen die later in hun lagere schoolloopbaan minstens een keer bleven zitten. Op het einde van schooljaar 2003-2004 scoorden de K3 zittenblijvers negen IRT-punten (= 0.98SD) lager dan deze groep terwijl er op het einde van schooljaar 2007-2008 geen significant verschil bestond

tussen beide groepen. Dit suggereert dat het voor het prestatieniveau op 10-jarige leeftijd niet uitmaakt of een leerling in K3 of later in de lagere schoolloopbaan blijft zitten. Op het einde van de rit scoren de leerlingen even zwak voor wiskunde.

Omwille van de beperkte beschikbaarheid van de wiskundescores kunnen over de deelgroep leerlingen die in de loop van de lagere schoolloopbaan overstapte naar het buitengewoon onderwijs enkel uitspraken worden gedaan over de groei t.e.m. schooljaar 2004-2005. In juni 2005 scoorde deze groep gemiddeld genomen 13 IRT-punten (= 1.35SD) lager dan de normaalvorderenden en vijf IRT-punten (= 0.51SD) lager dan de deelgroep 'vertraging'. Er was geen significant verschil met de groep 'zittenblijvers K3'.

4 Conclusie en discussie

In Vlaanderen is zittenblijven een populaire en min of meer maatschappelijk aanvaarde praktijk. Van alle kleuters doet ongeveer 4% in Vlaanderen de derde kleuterklas opnieuw. Dit is financieel en tijdsgewijs een grote kost voor zowel de overheid als de kinderen in kwestie en hun omgeving. Door middel van een propensityscore-stratificatie-analyse werden de effecten van zittenblijven in de derde kleuterklas op de groei in wiskunde van leerlingen doorheen het lager onderwijs onderzocht. De resultaten van dit onderzoek stellen de vanzelfsprekendheid van zittenblijven in vraag en geven alvast een aanzet om het gehele plaatje van de effecten van zittenblijven in Vlaanderen in kaart te brengen.

Er kan geconcludeerd worden dat K3-zittenblijvers op het einde van het bisjaar beduidend lager scoren dan vergelijkbare leeftijdsgenoten die overgingen naar L1. Deze laatste groep krijgt gedurende een schooljaar nieuwe leerstof aangeboden en scoort hoger op de toets op het einde van dit schooljaar. De zittenblijvers daarentegen krijgen gedurende het bisjaar vooral dezelfde leerstof als het jaar voordien aangeboden en scoren lager op de toets op het einde van dit bisjaar. Op 10-jarige leeftijd, is het prestatieverschil tussen beide groepen verminderd tot minder dan

de helft. Toch zouden zittenblijvers gedurende hun ganse loopbaan in het lager onderwijs, gemiddeld genomen, hoger hebben gescoord voor wiskunde, waren ze toch overgegaan naar L1. Deze resultaten bevestigen de bevindingen in eerder quasi-experimenteel onderzoek naar de effecten van zittenblijven in K3 voor wat betreft de wiskundeprestaties op het einde van het bisjaar (Hong & Raudenbush, 2005; 2006; Hong & Yu, 2007). In tegenstelling tot voorgaand onderzoek, waar geen significante prestatiekloof werd vastgesteld twee en vier jaar na het bisjaar, tonen de resultaten van deze studie een dalende doch blijvend significante kloof tot vijf jaar na het bisjaar.

Een belangrijke bijdrage van deze studie ten opzichte van eerder onderzoek is de opsplitsing van de controlegroep op basis van de loopbaan na de transitie naar L1. Uit de resultaten blijkt dat leerlingen die blijven zitten in K3 en equivalente leerlingen die later in hun lagere schoolloopbaan blijven zitten, vijf jaar na de treatment gelijk scoren voor wiskunde. Dat betekent dat het tijdstip van zittenblijven niet uitmaakt voor wat betreft de wiskundescores op 10-jarige leeftijd.

Belangrijk bij het interpreteren van de resultaten is het gegeven dat het hier gaat over gemiddeldes. De zittenblijvers en equivalente leerlingen die overgaan naar L1 werden d.m.v. propensityscore-stratificatie gelijkgesteld op basis van kenmerken die zittenblijven voorafgaan. Dit maakte het mogelijk om verschillen in wiskundegroei toe te schrijven aan 'zittenblijven'. Echter, er is ook sprake van gebeurtenissen die zich na het zittenblijven voordoen en de wiskundeprestaties beïnvloeden. Dit werd duidelijk bij het in kaart brengen van het schooltraject van leerlingen die overgaan naar L1 (de controlegroep) en bij het bekijken van de wiskundegroei van verschillende subgroepen opgedeeld naargelang dit schooltraject. Na het zittenblijven of het overgaan naar L1 (1) vorderen de kinderen normaal, (2) blijven ze (nogmaals) zitten of (3) gaan ze naar het buitengewoon onderwijs. In deze studie werd aangetoond dat de wiskundegroei significant verschilt tussen leerlingengroepen, opgedeeld naar deze loopbaan in het lager onderwijs. Verschillen in wiskundegroei

kunnen dan niet louter toegeschreven worden aan het al dan niet blijven zitten in K3 maar minstens deels aan de loopbaanwending na de treatment. Bovendien bleek dat bepaalde pre-treatment kenmerken een voorspellende waarde hebben voor de loopbaangroep waarin leerlingen terechtkomen. Hoewel zittenblijvers en leerlingen in de controlegroep, na propensityscore-stratificatie, gemiddeld genomen gelijk scoorden voor wiskunde in K3, werden in deze studie wel verschillen gevonden tussen leerlingen die normaal vorderden en andere loopbaangroepen. Hiermee wordt meteen een beperking aangegeven van deze studie. Aangezien de loopbaangroepen niet equivalent zijn voor wat betreft kenmerken die zittenblijven voorafgaan, kunnen geen causale uitspraken worden gedaan over het verband tussen het deel uitmaken van een specifieke loopbaangroep en de wiskundegroei. Causale uitspraken zijn enkel mogelijk voor de vergelijking tussen zittenblijvers en de totale controlegroep. Een eerste suggestie voor vervolgonderzoek is dan ook om het effect van zittenblijven in K3 te onderzoeken in vergelijking met andere schoolloopbanen waarbij loopbaangroepen a priori worden gelijkgesteld en causale uitspraken mogelijk worden gemaakt. Hoewel de technische uitwerking nog in zijn kinderschoenen staat en concrete toepassingen zeldzaam zijn, werden methodes om causale effecten van multiple treatments te onderzoeken aangekaart door onder andere Lechner (2001), Imai en van Dyck (2004) en Dong (2011; 2012).

Een tweede beperking van deze studie is dat enkel de wiskundegroei als uitkomst werd bekeken. Vervolgonderzoek zal de effecten van zittenblijven in K3 op andere criteria onderzoeken, zoals taalvaardigheid, psychosociaal functioneren en de verdere schoolloopbaan van leerlingen.

Verder is het aangewezen om differentieële effecten te onderzoeken. Zoals gesuggereerd in een meta-analyse van Allen et al. (2009) wordt nog te weinig aandacht besteed aan de condities waaronder zittenblijven in K3 wel gunstig kan zijn. De resultaten van deze studie suggereren alvast het bestaan van zulke differentieële effecten. Leerlingen die vergelijkbaar zijn op tal van pre-treatment kenmerken kennen toch een verschillende loopbaan

doorheen het lager onderwijs. Een beperkte groep leerlingen die doorstroomt naar L1, ondanks het feit dat ze een gelijkaardig profiel als dat van zittenblijvers hebben, slaagt erin om de lagere school zonder vertraging succesvol af te werken. Wat is het geheim van deze leerlingen? Wat zijn de (al dan niet geobserveerde) kenmerken van leerlingen die later in hun loopbaan blijven zitten of overstappen naar het buitengewoon onderwijs? Zijn er bepaalde kenmerken van de klas, de praktijk of de leerkrachten die samenhangen met de verschillende loopbanen van leerlingen? Zijn er bepaalde schoolprocessen die samenhangen met de verschillende loopbanen van leerlingen? Verschillende auteurs benadrukken het belang van de rol van de klas, leerkracht en school in het proces van schoolrijp worden en de ontwikkeling doorheen het lager onderwijs. Zoals aangegeven door Morisson et al. (1997), onderschatten we te vaak de interactie-effecten van genetische en omgevingsinvloeden op de ontwikkeling van kinderen. De ene school kan de ontplooiing van de mogelijkheden van een kind belemmeren terwijl hetzelfde kind in een andere school, waar meer geschikte middelen voorzien zijn, wel kan ontplooiën (Bronfenbrenner & Ceci, 1994). Zittenblijven in K3 is in dat opzicht overbodig in scholen die in staat zijn om de kennis en vaardigheden van kinderen te mobiliseren. Andere auteurs benadrukken dat het aanpassen van instructie meer kan baten dan kinderen K3 te laten dubbelen (Leinhardt, 1980; Reynolds, 1992; Tanner & Galis, 1997). Kortom, de rol en het gewicht van klas- en schoolkenmerken in de voorbereiding op en de transitie naar L1 verdient aandacht in vervolgonderzoek.

Met deze studie wordt aangetoond dat een kritische houding bij het lezen van onderzoeksbevindingen in deze context onontbeerlijk is. Quasi-experimentele designs waarin uitkomsten van zittenblijvers met deze van leeftijdsgenoten die overgaan naar het volgende leerjaar en equivalent zijn in zake kenmerken die zittenblijven voorafgaan vergeleken worden, worstelen al te vaak met factoren die na het zittenblijven ontstaan en de uitkomsten beïnvloeden. In deze studie stelden we vast dat ongeveer de helft van de risicoleerlingen die toch overgingen naar L1

later bleef zitten of overgang naar het buitengewoon onderwijs. Het is belangrijk om dit in het achterhoofd te houden bij het interpreteren van de resultaten. Samengevat, het is niet correct om uitsluitend op basis van de bevinding dat *gemiddeld genomen* zittenblijvers beter zouden presteren moesten ze toch overgaan naar L1 te besluiten dat zittenblijven een slechte onderwijspraktijk is.

Vanuit de onderzoeksresultaten kunnen dan ook weinig concrete adviezen geformuleerd worden voor de praktijk. De beslissing over het al dan niet overdoen van de derde kleuterklas moet voor elk kind een weloverwogen, doordachte beslissing zijn. De beslissing moet worden genomen zonder dat men in de toekomst kan kijken om te weten of de leerling op een later tijdstip zal blijven zitten. Kortom, elk advies over al dan niet zittenblijven in de derde kleuterklas is een ander verhaal en erg complex. Het vraagt van de betrokkenen een grondig overleg en het in kaart brengen van verschillende factoren (kindkenmerken, maar ook ondersteuning door de school en buiten de school).

Noten

1. In Nederland bepaalt de basisschool of het kind mag overgaan naar groep 3 (Ministerie van Onderwijs, 2013).
2. Dit percentage is gebaseerd op de referentiesteekproef in het SiBO-project (Maes, Ghesquière, Onghena, & Van Damme, 2002).
3. In Vlaanderen starten kinderen in het eerste leerjaar wanneer ze 6 jaar worden voor 1 januari van dat schooljaar. In Nederland starten kinderen in groep 3 wanneer ze 6 jaar worden voor 1 oktober van dat schooljaar. In Vlaanderen zijn de jongste kinderen dus geboren in november/december. In Nederland zijn dit de kinderen die geboren zijn in augustus/september.
4. De SiBO-databank bevat een representatieve steekproef voor Vlaanderen en een oversampling van GOK-scholen, Gentse scholen en methodescholen. Omwille van hun specifieke methode werden de vijf Steinerscholen in ons onderzoek buiten beschouwing gelaten. Alle andere scholen werden opgenomen in ons onderzoek. Het percentage zittenblijvers ligt

hoger dan het Vlaamse gemiddelde (4%) omwille van deze oversamplings.

5. Het gaat om het aantal klassen en scholen in onze steekproef in schooljaar 2002-2003. Wanneer één administratieve school meerdere vestigingsplaatsen heeft, werd elke vestigingsplaats als een aparte school beschouwd, omdat dit ons inziens meer aansluit bij hoe dat in de praktijk ervaren wordt door de leerlingen, ouders en leerkrachten.
6. De volledige lijst van 'true confounders' kan worden opgevraagd bij de eerste auteur.
7. Imputatie betekent dat een ontbrekend datapunt vervangen wordt door een goede schatting van wat de score zou zijn geweest voor die variabele. Er bestaan zeer veel verschillende methoden om ontbrekende gegevens te imputeren.
8. Een overzicht van de loopbaanposities en het aantal ontbrekende wiskundescores van de zittenblijvers en de controlegroep per schooljaar kunnen opgevraagd worden bij de eerste auteur.
9. Ook de zittenblijversgroep zou opgedeeld kunnen worden in loopbaangroepen. Echter, omwille van een groot aantal ontbrekende wiskundescores voor deze groep opteerden we ervoor dit niet te doen. De subgroepen van de zittenblijversgroep zouden te klein zijn om een correcte vergelijking te kunnen maken.

Literatuur

- Alet, E. (2010). *Is grade repetition a second chance?* Retrieved from http://espe.conference-services.net/resources/321/2017/pdf/ESPE2010_0145_paper.pdf
- Allen, C. S., Chen, Q., Willson, V. L., & Hughes, J. N. (2009). Quality of research design moderates effects of grade retention on achievement: A meta-analytic, multilevel analysis. *Educational Evaluation and Policy Analysis, 31*(4), 480-499. doi:10.3102/0162373709352239
- Bossaert, G., Doumen, S., Buyse, E., & Verschueren, K. (2011). Predicting children's academic achievement after the transition to first grade: A two-year longitudinal study. *Journal of Applied Developmental Psychology, 32*(2), 47-57. doi:10.1016/j.appdev.2010.12.002
- Bronfenbrenner, U., & Ceci, S. J. (1994). Nature-nurture reconceptualized in developmental

- perspective: A bioecological model. *Psychological Review*, 101(4), 568-586.
- Brookhart, M. A., Schneeweiss, S., Rothman, K. J., Glynn, R. J., Avorn, J., & Stürmer, T. (2006). Variable selection for propensity score models. *American journal of epidemiology*, 163(12), 1149-1156.
- Byrd, R. S., & Weitzman, M. L. (1994). Predictors of early grade retention among children in the United-States. *Pediatrics* 93(3), 481-487.
- Chen, X., Liu, C., Zhang, L., Shi, Y., & Rozelle, S. (2010). Does taking one step back get you two steps forward? Grade retention and school performance in poor areas in rural China. *International Journal of Educational Development*, 30(6), 544-559.
- Cosden, M., Zimmer, J., & Tuss, P. (1993). The Impact of Age, Sex, and Ethnicity on Kindergarten Entry and Retention Decisions. *Educational Evaluation and Policy Analysis*, 15(2), 209-222.
- Dong, N. (2011). *Using Propensity Score Methods to Approximate Factorial Experimental Designs. SREE Fall 2011*. Retrieved from Society for Research on Educational Effectiveness <http://www.eric.ed.gov/PDFS/ED528839.pdf>
- Dong, N. (2012). *Causal moderation analysis using propensity score methods. SREE Spring 2012*. Retrieved from <http://www.eric.ed.gov/PDFS/ED530452.pdf>
- Dong, Y. Y. (2010). Kept back to get ahead? Kindergarten retention and academic performance. *European Economic Review* 54(2), 219-236. doi:10.1016/j.euroecorev
- Elder, T. E., & Lubotsky, D. H. (2009). Kindergarten Entrance Age and Children's Achievement. *Journal of Human Resources*, 44(3), 641-683.
- Gadeyne, E. (2003). *Longitudinaal onderzoek in het basisonderwijs: Eindejaarsbevraging voor de kleuterleidster schooljaar 2002-2003*. Leuven: Steunpunt LOA.
- Gadeyne, E., Onghena, P., & Ghesquiere, P. (2008). Child and family characteristics associated with nonpromotion in preprimary education. *Exceptional Children*, 74(4), 453-469.
- Goos, M., Van Damme, J., Onghena, P., & Petry, K. (2012). First-grade retention, peer relationships, and academic achievement: An analysis of causal mediation effects. Retrieved from <https://lirias.kuleuven.be/handle/123456789/341956>
- Hauser, R. M., Pager, D. I., & Simmons, S. J. (2001). Race-ethnicity, social background, and grade retention. An analysis of the last thirty years. *CEIC Review*, 19(5), 11-12.
- Hill, J. L., Reiter, J. P., & Zanutto, E. L. (2004). A Comparison of Experimental and Observational Data Analyses. In: A. Gelman & X.L. Meng (Eds.), *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives* (pp. 49-60). Chichester: John Wiley & Sons, Ltd. doi: 10.1002/0470090456.ch5
- Hofer, K. G., Farran, D., Lipsey, M., Aydogan, C., & Bilbrey, C. (2010, February). *Using propensity scores to estimate the effect of early grade retention*. Paper presented at the seventh biennial Conference on Research Innovations in Early Intervention, San Diego, CA. Abstract retrieved from http://www.criei.org/CRIEI2010Presentations/Hofer_CRIEI2010.pdf
- Holmes, C. T. (1989). Grade-level retention effects: A meta-analysis of research studies. In L.A. Shepard & M. L. Smith (Eds.), *Flunking grades: Research and policies on retention* (pp. 16-33). London, United Kingdom: The Falmer Press.
- Hong, G., & Raudenbush, S. W. (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. *Educational Evaluation and Policy Analysis*, 27(3), 205-224. doi:10.3102/01623737027003205
- Hong, G., & Raudenbush, S. W. (2006). Evaluating kindergarten retention policy: A case study of causal inference for multilevel observational data. *Journal of the American Statistical Association*, 101(475), 901-910.
- Hong, G., & Yu, B. (2007). Early-grade retention and children's reading and math learning in elementary years. *Educational Evaluation and Policy Analysis*, 29(4), 239-261.
- Hong, G., & Yu, B. (2008). Effects of kindergarten retention on children's social-emotional development: An application of propensity score method to multivariate, multilevel data. *Developmental Psychology*, 44(2), 407-421. Retrieved from WOS:000254115200010
- Ikedda, M. (2011). *Pisa in focus 6. When students repeat grades or are transferred out of school: What does it mean for education systems?* Retrieved from <http://www.pisa.oecd.org/dataoecd/35/58/48363440.pdf>

- Imai, K., & van Dyk, D. A. (2004). Causal inference with general treatment regimes: Generalizing the propensity score. *Journal of the American Statistical Association*, 99(467), 854-866.
- Jimerson, S. R. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *School Psychology Review*, 30(3), 420-437.
- Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, 45(3), 850-867.
- Juchtmans, G., Goos, M., Vandenbroucke, A., & De Fraine, B. (2012). *Zittenblijven in vraag gesteld: Een verkennende studie naar nieuwe praktijken voor Vlaanderen vanuit internationale perspectief*. Leuven: KU Leuven.
- Judkins, D. R., Morganstein, D., Zador, P., Piesse, A., Barrett, B., & Mukhopadhyay, P. (2007). Variable selection and raking in propensity scoring. *Statistics in medicine*, 26(5), 1022-1033
- La Paro, K. M., Rimm-Kaufman, S. E., & Pianta, R. C. (2006). Kindergarten to 1st Grade: Classroom Characteristics and the Stability and Change of Children's Classroom Experiences. *Journal of Research in Childhood Education*, 21(2), 189-202. doi:10.1080/02568540609594588
- Lechner, M. (2001). Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. In: M. Lechner & F. Pfeiffer (Eds.), *Econometric Evaluation of Labour Market Policies* (13th ed., pp. 43-58). Heidelberg: ZEW Economic Studies Physica-Verlag.
- Leinhardt, G. (1980). Transition rooms: Promoting maturation or reducing education? *Journal of Educational Psychology*, 72(1), 55-61. doi:10.1037/0022-0663.72.1.55
- Lorence, J., & Dworkin, A. G. (2006). Elementary grade retention in Texas and reading achievement among racial groups: 1994-2002. *Review of Policy Research*, 23(5), 999-1033.
- Maes, F. (2003a). *Longitudinaal onderzoek in het basisonderwijs: Kleutervragenlijst (schooljaar 2002-2003)* Leuven: Steunpunt LOA.
- Maes, F. (2003b). *Longitudinaal onderzoek in het basisonderwijs: Schoolteamvragenlijst (schooljaar 2002-2003)*. Leuven: Steunpunt LOA.
- Maes, F., Ghesquière, P., Onghena, P., & Van Damme, J. (2002). *Longitudinaal onderzoek in het basisonderwijs: Van doelstellingen tot onderzoeksopzet* (Report No. 1). Leuven, Belgium: Steunpunt LOA.
- Mantzicopoulos, P. Y., & Neuharth-Pritchett, S. (1998). Transitional first-grade referrals: An analysis of school-related factors and children's characteristics. *Journal of Educational Psychology*, 90(1), 122-133.
- Mashburn, A. J., & Pianta, R. C. (2006). Social Relationships and School Readiness. *Early Education & Development*, 17(1), 151-176. doi:doi: 10.1207/s15566935eed1701_7
- Ministerie van de Vlaamse Gemeenschap. (1997). *Decreet basisonderwijs*. Retrieved from <http://www.ond.vlaanderen.be/edulex/database/document/document.asp?docid=12254>
- Ministerie van Onderwijs, C. e. W. (2013). *Wie bepaalt of mijn kind blijft zitten of overgaat naar de volgende groep?* Retrieved from <http://www.rijksoverheid.nl/onderwerpen/basisonderwijs>
- Rasbash, J., Charlton, C., Browne, W. J., Healy, M., & Cameron, B. (2010). MLwiN (Version 2.20) United Kingdom: Centre for Multilevel Modelling, University of Bristol.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Advanced Quantitative Techniques in the Social Sciences, SAGE Publications. Retrieved from <http://books.google.be/books?id=uyCV0CNGDLQC>.
- Reynders, T., Van Heddegem, I., Nicaise, I., & Van Damme, J. (2004). *Longitudinaal onderzoek in het basisonderwijs: Oudervragenlijst schooljaar 2002-2003*. Leuven: Steunpunt LOA.
- Reynolds, A. J. (1992). Comparing measures of parental involvement and their effects on academic achievement. *Early Childhood Research Quarterly*, 7(3), 441-462. doi:10.1016/0885-2006(92)90031-S
- Scheerens, J., & Bosker, R. J. (1997). *The Foundations of Educational Effectiveness*. London: Pergamon Press.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton-Mifflin.
- Smith, M. L., & Shepard, L. A. (1988). Kindergarten readiness and retention: A qualitative study

- of teachers' beliefs and practices. *American Educational Research Journal*, 25(3), 307-333. doi:10.3102/00028312025003307
- SPSS Inc. (2010). *IBM SPSS Missing Values 19*. Chicago, IL: IBM. Retrieved from <http://www.uky.edu/ComputingCenter/SSTARS/SPSS/19%20Manuals/IBM%20SPSS%20Missing%20Values%2019.pdf>.
- Tanner, C. K., & Galis, S. A. (1997). Student retention: Why is there a gap between the majority of research findings and school practice? *Psychology in the Schools*, 34(2), 107-114. doi:10.1002/(SICI)1520-6807(199704)34:2<107::AID-PITS4>3.0.CO;2-N
- Van Damme, J. (2004). *Maakt de school het verschil?: effectiviteit van scholen, leraren en klassen in de eerste graad van het middelbaar onderwijs : een exploratie van LOSO-gegevens*. Leuven: Acco.
- Van der Grift, W. (2005). Verlenging en verkorting van de kleuterperiode in het basisonderwijs. *Basisschoolmanagement*, 18(5).
- Verachtert, P. (2003). *Longitudinaal Onderzoek in het Basisonderwijs. Toetsen Schooljaar 2002-2003 (LOA-rapport nr. 14)*. Leuven: Steunpunt LOA.
- Verhaeghe, J.-P. (2004). *Longitudinaal onderzoek in het basisonderwijs: Directievragenlijst (schooljaar 2002-2003)*. Leuven, Belgium: Steunpunt LOA.
- Vlaams Ministerie van Onderwijs en Vorming. (2012). *Vlaamse onderwijsstatistieken en publicaties*. Retrieved from <http://www.ond.vlaanderen.be/onderwijsstatistieken/>
- Wu, W., West, S. G., & Hughes, J. N. (2008). Effect of retention in first grade on children's achievement trajectories over 4 years: A piecewise growth analysis using propensity score matching. *Journal of Educational Psychology* 100(4), 727-740.
- Wu, W., West, S. G., & Hughes, J. N. (2010). Effect of grade retention in first grade on psychosocial outcomes. *Journal of educational psychology*, 102(1), 135-152.

siteit Leuven. **Gudrun Vanlaar** is als doctoranda verbonden aan het Centrum voor Onderwijseffectiviteit en –evaluatie aan de Katholieke Universiteit Leuven. **Mieke Goos** is als postdoctoraal onderzoekster verbonden aan de universiteit van Maastricht. **Bieke De Fraine** is als universitair docent en diensthoofd verbonden aan het Centrum voor Onderwijseffectiviteit en -evaluatie aan de Katholieke Universiteit Leuven. **Jan Van Damme** is als emeritus verbonden aan het Centrum voor Onderwijseffectiviteit en -evaluatie aan de Katholieke Universiteit Leuven.

Correspondentieadres: Machteld Vandecandelaere, Dekenstraat 2, box 3773 – room VHI 2.43, B-3000 Leuven, Tel: + 0032 16 325 744, E-mail: Machteld.Vandecandelaere@ppw.kuleuven.be

Abstract

The effects of repeating kindergarten on mathematics development: A propensity score stratification analysis

In Flanders, 4% repeats kindergarten. By means of a propensity score stratification analysis we investigated the effects of kindergarten retention on mathematics development throughout primary education. Analyzing data from a large-scale longitudinal study, we find that, by the end of the retention year, kindergarten repeaters score significantly lower compared to similar promoted children. At the age of 10, the achievement gap is decreased until less than half. Yet, on average, repeaters would perform better for math until five years after the treatment, were they promoted to first grade instead. Further, the results indicated that the math score five years after the treatment of repeaters and similar children who promoted to first grade but repeated a later grade, do not differ. For mathematics achievement at age 10, the timing of retention doesn't matter.

Auteurs

Machteld Vandecandelaere is als doctoranda verbonden aan het Centrum voor Onderwijseffectiviteit en –evaluatie aan de Katholieke Univer-