



Centraal Planbureau

CPB Achtergronddocument | 17 januari 2015

# Zittenblijven in perspectief

*Behoort bij:*

*CPB Policy Brief 2015/01 'Zittenblijven in het  
primair en voortgezet onderwijs: Een  
inventarisatie van de voor- en nadelen'*

Frank van Erp  
Sander Gerritsen  
Suzanne Heijnen  
Karen van der Wiel



# Inhoud

1	Inleiding—4
2	Effecten van zittenblijven—4
2.1	Inleiding—4
2.2	Generieke kanttekeningen bij de literatuur—7
2.3	Betrouwbare studies over zittenblijven—10
2.4	Conclusie—13
3	Zittenblijven in relatie tot de organisatie van het onderwijs—14
3.1	Samenhang met strenge diploma-eisen—14
4	Effecten van alternatieven voor zittenblijven—16
4.1	Voor- en vroegschoolse educatie—16
4.2	Uniformering onderwijstijd—18
4.3	Gerichte verlenging onderwijstijd—18
4.4	Deelcertificaten—20
	Appendix—21
	Internationale vergelijking zittenblijven—21
	Berekening kosten zittenblijven—24
	Referenties—25

# 1 Inleiding

Dit achtergronddocument gaat nader in op literatuur die genoemd wordt in de CPB Policy Brief *'Zittenblijven in het primair en voortgezet onderwijs: Een inventarisatie van de voor- en nadelen'*. Naast een korte beschrijving van de literatuur behandelen we de betrouwbaarheid van de resultaten in verschillende studies.

Hoofdstuk 2 behandelt de empirische (economische) literatuur over zittenblijven. Welke effecten op leerprestaties en de sociaal-emotionele ontwikkeling van kinderen zijn aangetoond? Hoofdstuk 3 gaat kort in op de verhouding tussen zittenblijven en de organisatie van het onderwijs. Speciale aandacht gaat uit naar harde eindexameneisen en welke effecten hiervan uitgaan. In Hoofdstuk 4 komt kort de literatuur aan bod die te maken heeft met verschillende alternatieven voor zittenblijven. Daarbij wordt het langst stilgestaan bij voor- en vroegschoolse educatie. In de appendix wordt aandacht besteed aan wat we weten en vooral wat we nog niet weten over hoe de zittenblijfpercentages in Nederland zich verhouden tot de percentages in andere landen. Bovendien wordt in de appendix de kostenberekening over zittenblijven uit de Policy Brief expliciet gemaakt in een tabel.

## 2 Effecten van zittenblijven

*Wie de literatuur in vogelvlucht scant kan de indruk krijgen dat zittenblijven geen positieve effecten heeft op de betrokken leerlingen. Veel studies naar de effecten van zittenblijven blijken echter van beperkte econometrische of conceptuele kwaliteit. De groep studies die hier niet of minder onder gebukt gaan, laten een genuanceerder beeld zien.*

### 2.1 Inleiding

De standaardreferentie voor bijna elke rapportage over 'retention policy' is Jimerson (2001).<sup>1</sup> In deze studie zijn de resultaten van eerdere meta-analyses door Jackson (1975), Holmes en Matthews (1984) en Holmes (1989) samengevat en aangevuld met een meta-analyse uitgevoerd op twintig studies uit de periode 1990-1999 (zie Tabel 2.1). In de meeste van deze studies is aandacht voor de effecten op de leerprestaties van leerlingen en de sociaal-emotionele uitkomsten, waarbij binnen elk van deze twee domeinen verschillende uitkomstmaten werden gebruikt.<sup>2</sup>

---

<sup>1</sup> Een recenter overzicht is te vinden in Xia en Kirby (2009).

<sup>2</sup> Zie Jimerson (2001, tabel 2) voor een overzicht van de gemiddelde effecten (in standaarddeviaties) op de verschillende uitkomstmaten in de studies van Holmes en Matthews (1984), Holmes (1989) en Jimerson (2001).

**Tabel 2.1** Meta-studies over zittenblijven

Auteurs	Aantal studies	Periode
Jackson (1975)	30	1911-1973
Holmes en Matthews (1984)	44	1929-1981
Holmes (1989)	63	1925-1989
Jimerson (2001)	20	1990-1999

De voornaamste conclusies uit Jimerson (2001) zijn:

1. dat de meeste studies op lange termijn geen significante verschillen laten zien tussen de prestaties van kinderen die zijn blijven zitten en de prestaties van kinderen die bevorderd zijn.
2. dat in sommige studies hooguit sprake is van een positief effect van zittenblijven op prestaties op de korte termijn, maar deze effecten na twee tot drie jaar zijn weggeëbd.
3. dat, volgens sommige studies, de kinderen die zijn blijven zitten, op lange termijn zelfs slechter af zijn (bijvoorbeeld als gevolg van 'drop out') en
4. dat zittenblijven vaker voorkomt onder kinderen uit de lagere sociaaleconomische milieus, jongens en de jongste leerlingen van de groep (effect geboortemaand).

Los van de vraag of deze conclusies een-op-een vertaald kunnen worden naar de Nederlandse situatie anno 2015, liggen deze conclusies de laatste jaren steeds vaker onder wetenschappelijk vuur.

Het gegeven dat aan het eind van een schooljaar een groep kinderen blijft zitten, terwijl een andere groep leerlingen overgaat naar het volgende leerjaar, impliceert dat de twee groepen naar kenmerken en prestaties verschillen. De eerste groep haalt op grond daarvan het vereiste niveau niet, terwijl de tweede groep met andere kenmerken en prestaties dat wel bereikt. Dit verschil in kenmerken en prestaties tussen de twee groepen bemoeilijkt het onderzoek naar de effecten van zittenblijven als remediërende actie om leerachterstanden te verhelpen.

Om namelijk tot een zuivere conclusie over de effecten van zittenblijven te komen, zouden idealiter leerlingen willekeurig verdeeld moeten worden over een groep die blijft zitten en een groep die overgaat. Op het moment van die verdeling zijn, als gevolg van het willekeurige karakter daarvan, de twee groepen - behoudens het zittenblijven - naar kenmerken en prestaties identiek. Enige tijd later opgetreden verschillen in prestaties moeten dan veroorzaakt zijn door het enige verschil tussen beide groepen: wel/niet zittenblijven. Dit type *random design* onderzoek blijkt in de praktijk moeilijk te realiseren. Als gevolg daarvan hebben onderzoekers in het verleden

1. of zelf vergelijkbare groepen moeten samenstellen,
2. of hun resultaten moeten corrigeren (controleren) voor verschillen tussen de groep zittenblijvers en de groep bevorderden.

Beide benaderingen zijn onderwerp van kritiek.<sup>3</sup> In verschillende studies is betoogd dat de *selection bias* (zie kader) die ontstaan is door een inadequate samenstelling van de groepen of onvolledig controleren voor groepsverschillen, mede aanleiding is geweest voor de conclusies geformuleerd in Jimerson (2001). De statistisch beschouwd betere empirische studies komen juist tot andere of meer genuanceerde conclusies.<sup>4</sup>

In deze notitie bespreken we daarom vooral recente studies naar de effecten van zittenblijven, waarbij streng is geselecteerd ten aanzien van de gevolgde methode om problemen van de *selection bias* uit te sluiten. Ondanks de betere empirische kwaliteit van deze studies hebben ze nog verschillende haken en ogen die een antwoord op de vraag naar de effecten van zittenblijven in Nederland in de weg staan. Deze generiek geldende kanttekeningen betreffen:

1. Het moment van toetsen (*same grade* versus *same age*).
2. De reikwijdte van de conclusies (marginale effecten versus buitenmarginale effecten).
3. De gehanteerde uitkomstmaat om de effecten van zittenblijven aan af te meten en
4. De mate waarin het onderzoek zich gericht heeft op de effecten van zittenblijven *sec*, of op de effecten van een combinatie van maatregelen (zittenblijven versus zittenblijven<sup>+</sup>).

Op elk van deze kanttekeningen gaan we in de volgende paragraaf dieper in. Vervolgens wordt in paragraaf 3.3 uitgebreider ingegaan op de recente studies die een beter empirisch ontwerp hebben om de causale relatie vanuit zittenblijven op een uitkomst vast te stellen. Paragraaf 3.4 rapporteert dan onze conclusies.

### Selection bias

Om de effecten van zittenblijven op bijvoorbeeld leerprestaties te kunnen vaststellen, moeten de leerprestaties van twee groepen kinderen worden vergeleken waarbij de groepen in maar één aspect verschillen: wel/niet zittenblijven. Het vinden van die controlegroep is de grote uitdaging.

Een veel gebruikte methode is om de vergelijkbare groep kinderen te selecteren op basis van vergelijkbare kenmerken met de groep zittenblijvers. Te denken valt dan aan: leerprestaties eerder in de schoolloopbaan, leeftijd, afkomst en SES-categorie. Bijvoorbeeld, bij een in groep 4 zittenblijvende jongen van 7 jaar wonende te Wassenaar kan een even oude jongen uit diezelfde gemeente worden gezocht die niet in groep 4 doubleerde. Door deze matching lijken de twee jongens, behoudens het zittenblijven, in hun achtergrondkenmerken (leeftijd, leefomgeving) op elkaar. Onderzoekers zijn bij die matching echter gebonden aan de informatieset waarover zij beschikken (waarneembare kenmerken).

Echter, de beslissing tot wel/niet zittenblijven wordt ook bepaald door niet voor de onderzoeker waarneembare kenmerken van kind, ouders en/of omgeving (niet-waarneembare factoren). Voorbeelden hiervan zijn: motivatie, eventueel aanwezige leer- of gedragsstoornissen, opleidingsniveau van de ouders, mate waarin het kind geaccepteerd is binnen de groep. Zittenblijven is dus de uitkomst van voor de onderzoeker waarneembare en niet-waarneembare factoren. Een controlegroep is slechts door de onderzoeker samengesteld op basis van de waarneembare determinanten. De twee groepen zijn daarmee niet in alle kenmerken, behoudens zittenblijven, identiek. Dit verschil tussen de twee groepen vertroebelt de meting van het effect van zittenblijven. In het jargon spreken wij van de *selection bias*. Deze zorgt ervoor dat het gemeten effect afwijkt van het echte effect, omdat ook invloeden van niet-waarneembare verschillen tussen de twee groepen in het gemeten effect tot uiting zullen komen.

<sup>3</sup> Zie Allen et al. (2009) en Diris (2012) voor een kort literatuuroverzicht daaromtrent.

<sup>4</sup> Zie o.a. Allen et al. (2009).

## 2.2 Generieke kanttekeningen bij de literatuur

### Same age of same grade

In de literatuur zijn verschillende manieren beschreven om de effecten van zittenblijven in kaart te brengen. Een Jantje-Pietje-voorbeeld verheldert de verschillende methoden. Laten we aannemen dat Jantje in jaar  $t$  in groep 3 blijft zitten, terwijl Pietje doorgaat naar groep 4. Jantje en Pietje hebben dezelfde leeftijd  $l_t$ .

1. 'Same age': de prestaties van Jantje in groep 3 in jaar  $t+1$ , dat wil zeggen na de interventie 'zittenblijven', worden vergeleken met de prestaties van Pietje in groep 4 in jaar  $t+1$ . Beide kinderen hebben immers dezelfde leeftijd. Maar er is wel een verschil in curriculum dat zij in jaar  $t+1$  gevolgd hebben.
2. 'Same grade': de prestaties van Jantje in groep 4 in jaar  $t+2$ , dat wil zeggen na de interventie 'zittenblijven' plus aansluitend verblijf in groep 4, worden vergeleken met de prestaties van Pietje in groep 4 in jaar  $t+1$ . Voor beide kinderen worden nu de prestaties na het doorlopen van hetzelfde curriculum vergeleken, maar de leeftijd van de kinderen in groep 4 verschilt. Jantje is immers een jaar ouder dan Pietje als groep 4 wordt afgerond. Een leeftijdsverschil dat mede van invloed kan zijn op de prestaties, maar waarvan het de vraag is of dit nu het gevolg is van zittenblijven.

Schwerdt en West (2013, blz. 5-7) laten dit onderscheid wiskundig zien en concluderen dat een effectevaluatie op basis van een *same age*-vergelijking de voorkeur verdient. Het verschil in curriculum is nu net wat de *opportunity costs* zijn van zittenblijven. Dit dient goedgemaakt te worden door het hernieuwd doorlopen van de al onderwezen stof. Hetzelfde standpunt wordt ingenomen door Roderick en Nagaoka (2005, blz. 311-312), en Hong en Raudenbusch (2005, blz. 207). De *same age*-aanpak betekent echter wel dat de toetsen, afgenomen in verschillende leerjaren, onderling vergelijkbaar moeten zijn. In Chicago wordt daarvoor gebruik gemaakt van de *Iowa tests of basic skills* en in Florida van de *Florida Comprehensive Assessment Test*.<sup>5</sup>

Voor de gevonden effecten van zittenblijven lijkt het onderscheid *same age* versus *same grade* niet zonder betekenis. Allen et al. (2009, blz. 12) hebben laten zien dat *same grade*-studies (gemiddeld) een groter negatief effect van zittenblijven laten zien dan de studies die een *same age* strategie hanteren.<sup>6</sup> De literatuur biedt hiervoor geen duidelijke verklaring. Schwerdt en West (2013, blz. 5 e.v.) laten echter wiskundig zien dat een *same grade*-vergelijking van leerprestaties ook een leeftijdseffect omvat, omdat leerlingen van een verschillende leeftijd op eenzelfde onderwijsniveau worden vergeleken. Dit leeftijdseffect zou het effect van doubleren kunnen domineren.

---

<sup>5</sup> Los van de inhoudelijke vergelijkbaarheid van toetsen op verschillende onderwijsmomenten kan ook de motivatie van leerlingen op het toetsmoment verschillen. Zo geldt voor het onderwijs in Chicago dat de toets aan het eind van de third, sixth and eighth grade bepalend is voor de overgang naar de volgende grade, terwijl dat in tussenliggende jaren niet het geval is (zie ook Jacob en Lefgren, 2004, blz. 227).

<sup>6</sup> Allen et al. (2009, blz. 10) hebben een meta-analyse uitgevoerd met - naar hun oordeel - 22 kwalitatief goede studies uit de periode 1990-2007. Van deze studies volgt 68% een 'same grade'-strategie, terwijl maar 23% een 'same age'-vergelijking kende en 9% van de studies beide strategieën hebben gevolgd.

### Marginale effecten versus buiten-marginale effecten

Onderzoek t gebruik maakt van de nieuwste econometrische technieken schat causale effecten van zittenblijven door het verschil in prestaties op tijdstip  $t+s$  van - op tijdstip  $t$  - twee identieke groepen. Die twee groepen zijn, behoudens het zittenblijven, identiek en dat betekent dat zij elkaar ook naar schoolprestaties nauwelijks ontlopen. Empirische analyses maken daarmee dus gebruik van een deelselectie van alle zittenblijvers en alle bevorderden die gevormd wordt door de leerlingen die rond de grenswaarde van wel/niet zittenblijven verkeren. Kinderen die ruim boven of onder de kritieke waarde zitten, worden dus doorgaans in dit type effectstudies buiten beschouwing gelaten. Als zittenblijven een effect heeft, dan moeten de prestaties van de (deel)groep zittenblijvers op  $t+s$  uitgaan boven de prestaties van de (deel)groep bevorderden. Als er geen effect van zittenblijven wordt gevonden, zijn de prestaties van beide (deel)groepen gelijk.

Zittenblijven is echter niet gratis. Schoolgaande kinderen kosten geld en het later op de arbeidsmarkt komen als gevolg van een vertraagde leerweg, betekent misgelopen arbeidsinkomsten. Als uit een empirische analyse waarbij gebruik wordt gemaakt van de resultaten van kinderen rond de grenswaarde, zou blijken dat er geen effect is op de prestaties (baten), dan worden wel deze kosten gemaakt. Het saldo van baten en lasten van zittenblijven is voor deze (deel)groep dan dus negatief. De conclusie lijkt dan ook gerechtvaardigd dat voor deze (deel)groep het zittenblijven beter achterwege kan blijven.

Maar geldt dit ook voor de personen die overduidelijk onder de grenswaarde van overgaan vertoeven? Bijvoorbeeld de leerlingen die in het voorexamenjaar een 3 voor wiskunde hebben. Stel dat we op grond van een nul-effect op de prestaties van de marginale groep en het daaraan verbonden negatieve saldo van baten en lasten zouden concluderen dat we zittenblijven moeten afschaffen, dan bevorderen we ook de leerlingen met een wiskunde-3 automatisch naar het examenjaar. Het is maar zeer de vraag of deze leerlingen, die zich niet in de nabijheid van de grenswaarde bevinden, bij het examen wel de minimale eis van een 5 voor wiskunde zullen halen. Voor zover dat niet het geval is, komen zij op grond van de huidige regelgeving dan niet in aanmerking voor het diploma.

Als doubleren niet langer mag plaatsvinden, zullen zij dus zonder diploma het onderwijs verlaten en eventueel op een lager niveau vervolgonderwijs (bijvoorbeeld mbo in plaats van hbo of wo) genieten. Dit lagere niveau van vervolgonderwijs zal zich vertalen in een lagere loonvoet en wellicht een kortere werkzame periode. Beide factoren komen tot uitdrukking in een lager levensloopinkomen. Het is een open vraag of deze daling in het levensloopinkomen groter of kleiner is dan de daling als gevolg van een jaar later op de arbeidsmarkt verschijnen, maar wel met een hoger opleidingsniveau en de andere kosten van zittenblijven. Het vertalen van de effecten van empirische studies verricht met groepen op de grens van zittenblijven/overgaan (marginale effecten) naar effecten voor de volledige populatie zittenblijvers, dient dan ook met de nodige voorzichtigheid te geschieden.<sup>7</sup>

---

<sup>7</sup> Eigenlijk gaat het hier dus om het aloude bekende 'out-of-sample'-gebruik van de resultaten. Zoals bekend is dat altijd hachelijk.



### Uitkomstmaat t.b.v. effecten van zittenblijven

In veel empirisch onderzoek worden leerprestaties voor rekenen en lezen als uitkomstmaat beschouwd. In veel gevallen omdat hier cijfers voor beschikbaar zijn. Verder zijn deze maatstaven waarschijnlijk ook een redelijke indicator van leerprestaties later in het leven en het daaruit voortkomende inkomen. Het is echter de vraag of leerprestaties daarmee alle effecten van zittenblijven meenemen.

Idealiter zou vanuit een economisch perspectief het effect van zittenblijven worden afgemeten aan de hieruit voortkomende veranderingen in de maatschappelijke kosten en baten.<sup>8</sup> De baten worden gevormd door de veranderingen in het nut over de levensloop van alle mensen. Het begrip nut is ruimer dan (levensloop)inkomen omdat het alle ingrediënten van menselijk welzijn meeweegt. Dus naast consumptie of inkomen, bijvoorbeeld ook vrije tijd, gezondheid en het disnut van stress. Elke andere uitkomstmaat dan nut over de levensloop is dus partieel van aard en herbergt het risico dat conclusies worden getrokken die niet in lijn liggen met het effect op het welzijn van mensen. Zo kunnen kinderen die veel uren moeten besteden aan hun huiswerk, wellicht geen effecten op hun leerprestaties en latere inkomen ondervinden als zij blijven zitten, maar wel meer vrije tijd hebben door een verminderd tijdsbeslag van huiswerk en/of zich prettiger voelen omdat zij niet meer de slechtste cijfers van de groep halen. In de literatuur hebben wij geen goede causale empirische studies gevonden die de effecten van zittenblijven afmeten aan de verandering in het lifetime nut van de kinderen, of aan korte- en langetermijnveranderingen in het sociaal-emotioneel welbevinden hanteren.

Met de focus op maatschappelijke baten gaat het niet alleen om het nut over de levensloop van de zittenblijvers, maar ook om de effecten op het nut van derden. Tot de groep derden kunnen velen worden gerekend. Als voorbeelden twee groepen daarvan: voormalige klasgenoten van zittenblijvers en kinderen van zittenblijvers. Bij volgende generaties kan bijvoorbeeld een verbetering van het nut optreden als zij een hoger onderwijsniveau behalen, omdat hun ouders ook een hoger niveau hebben afgemaakt. Idealiter wordt een dergelijk effect meegenomen in de effecten van zittenblijven.

Empirisch onderzoek dat de effecten van zittenblijven slechts afmeet aan de leerprestaties van de zittenblijver in kwestie, is dus partieel van aard.

### Zittenblijven versus Zittenblijven<sup>+</sup>

De kwalitatief betere studies gebruiken vooral de herzieningen in het onderwijsstelsel in verschillende staten en steden van de VS, met name in Chicago en Florida. Na decennialange promotie op basis van de opvatting dat zittenblijven negatieve effecten had op de sociaal-emotionele en cognitieve ontwikkeling van kinderen (*social promotion*), werd overgestapt op een systeem waarbij nog slechts op basis van 'objectieve' testcores bevordering geschiedde.<sup>9</sup> De gevolgen van deze beleidswijzigingen kunnen nu geanalyseerd worden.

---

<sup>8</sup> De maatschappelijke kosten worden in deze notitie niet uitgewerkt, omdat dit geen betekenis heeft voor de duiding van het empirisch onderzoek. Uitwerking van dit begrip is uiteraard wel relevant voor een maatschappelijke kosten-batenanalyse van zittenblijven, maar die is geen onderwerp van deze notitie.

<sup>9</sup> Er zijn daarop overigens wel uitzonderingen mogelijk. In Chicago zijn de criteria daarvoor niet duidelijk omschreven. In Florida is dat wel het geval (zie Greene en Winters (2009, blz. 136)).

Deze analyses leveren echter geen zuivere meting van het gewijzigde beleid ten aanzien van het zittenblijven. Naast de introductie van zittenblijven op basis van prestaties zijn binnen de stelselherziening ook remediërende programma's gestart die zittenblijven moeten voorkomen (bijvoorbeeld zomerscholen) of catch up-mogelijkheden hebben geïntroduceerd (bijvoorbeeld *academic improvement plans, high performing teacher*) die leerlingen de mogelijkheid moeten bieden de opgelopen vertraging weer in te lopen. Effectevaluaties zullen dan ook in veel gevallen het effect van deze gecombineerde acties omvatten.

## 2.3 Betrouwbare studies over zittenblijven

In deze paragraaf gaan wij in op *state of the art*-betrouwbare studies over zittenblijven.

### Jacob en Lefgren (2004)

Op basis van recente wijzingen in het onderwijssysteem in Chicago schatten Jacob en Lefgren (2004) het causale effect van doubleren en een zomerschool. Zij gebruiken hiervoor een *regression discontinuity design* (zie tekstkader). Leerprestaties worden afgemeten aan gestandaardiseerde testen in juni en augustus. Verder vergelijken zij op basis van *same age*-groepen. De studie wordt verder gekenmerkt door een uitgebreide verantwoording van de gevolgde econometrische strategie.

Op basis van een vergelijking van opeenvolgende juni-scores komen Jacob en Lefgren (2004) tot de conclusie dat de combinatie van zittenblijven en zomerschool gericht op (potentiële) zittenblijvers in de *third grade* (groep 5) een aanzienlijk effect heeft op de lees- en rekenprestaties een jaar later. Voor de potentiële zittenblijvers in de *sixth grade* (groep 8) worden echter nauwelijks effecten bij lees- en rekenprestaties gevonden.

Door een vergelijking van de leerprestaties in augustus en juni van het daarop volgende jaar kunnen de effecten geschoond worden voor de invloed van de zomerschool. Ook hier ontstaat het beeld dat het effect op *third graders* groter is dan bij *sixth graders* waarbij het zittenblijven vooral effectief lijkt voor het verbeteren van de rekenprestaties, terwijl de zomerschool ook kan worden benut voor het verbeteren van leesprestaties.<sup>10</sup>

Ter verklaring van de afwijkende resultaten na twee jaar voor de twee klassen voeren de onderzoekers twee mogelijkheden aan. Ten eerste dat zij *age-by-age* vergelijken door de scores van leeftijdsgenoten die inmiddels in een hogere klas vertoeven, te vergelijken met de scores van de vertraagde leerlingen. Als de toets in *grade 8* zwaarder is dan in *grade 7*, zou onderpresteren van de zittenblijvers hiervoor een verklaring kunnen zijn (Jacob en Lefgren, 2004, blz. 240). Een tweede verklaring voor het verschil zou het moeilijker *classroom management* kunnen zijn. Naarmate de puberteit nadert, wordt lesgeven lastiger en zou de effectiviteit daarvan beperkter kunnen zijn (Jacob en Lefgren, 2004, blz. 240).

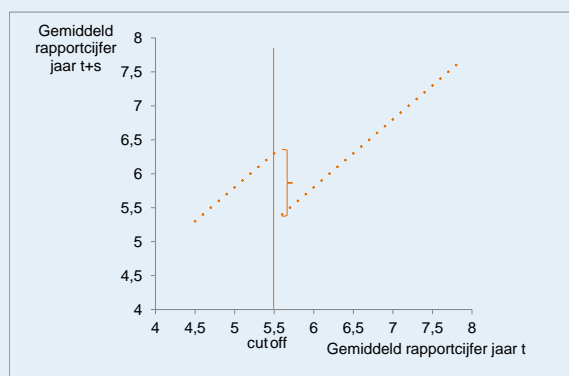
---

<sup>10</sup> Vergelijk daarvoor figuren 4 en 6 uit Jacob en Lefgren (2004).

## Regression discontinuity design

In een zuiver regression discontinuity design (RD) voor zittenblijven is sprake van een duidelijk grens (bijvoorbeeld een 5,5 als gemiddeld rapportcijfer) die bepalend is voor wel/niet doubleren. Alle leerlingen kunnen naar hun gemiddeld rapportcijfer worden gerangschikt. Kinderen met een aanzienlijk lager gemiddelde dan de kritische waarde zijn overduidelijke doubleurs, kinderen met een gemiddeld cijfer vlak onder de kritische grens blijven net zitten, terwijl de kinderen met een gemiddelde daar vlak boven met hakken over de sloot overgaan en de overige kinderen ruimer over zijn. De groepen aan weerszijde van de kritische grens (de 'cut off') zijn onderling het beste vergelijkbaar, terwijl de kinderen aan de staarten van de verdeling de grootste verschillen hebben.

Een of meerdere jaren ( $t+s$ ) later worden de leerprestaties (bijvoorbeeld gemiddeld rapportcijfer) van alle leerlingen opnieuw onderzocht. Ter illustratie bevat de verticale as van onderstaande figuur de leerprestaties op  $t+s$  en staan op de horizontale as de leerlingen naar hun gemiddeld rapportcijfer uit jaar  $t$  gerangschikt. Als zittenblijven een effectieve remediërende actie is, zal rond de 'cut off' uit jaar  $t$  ( $T$ ) het verloop van de leerprestaties in jaar  $t+s$  een sprongetje (discontinuïteit) laten zien. Als zittenblijven tot betere leerprestaties leidt, zal de groep die in het verleden net bleef zitten (dat wil zeggen zich aan de marge van  $T$  bevond) iets beter scoren dan de vrijwel vergelijkbare groep (rechts van de 'cut off') die toen overging. De grootte van deze discontinuïteit beschrijft (zie accolade) dan het positieve effect van zittenblijven, het ontbreken van een discontinuïteit betekent dat zittenblijven geen effect heeft gehad op de leerprestaties.



### Schwerdt en West (2013)

Schwerdt en West (2013) is naar ons oordeel de op dit moment econometrisch beschouwd best uitgevoerde studie naar de effecten van zittenblijven. Schwerdt en West gebruiken een *regression discontinuity design* om na de stelselherziening in Florida de effecten van zittenblijven en de zomerschool bij *same age* groepen te bepalen. Verder zijn de effecten van deze remediërende programma's op een wat langere termijn bepaald. De auteurs komen tot de conclusie dat als gevolg van het zittenblijven en de zomerschool gedurende de eerste drie jaren de toetscore voor lezen met 0,34 standaarddeviatie stijgt en voor rekenen stijgt de toetscore met 0,26 standaarddeviatie.<sup>11</sup>

<sup>11</sup> De bij empirisch onderzoek gebruikte variabelen zijn in verschillende dimensies of op verschillende schalen uitgedrukt. Zo loopt de Cito-score van 500 tot en met 550, terwijl voor de meeste mensen het IQ volgens de WISC-III varieert tussen de 70 en 130 punten. Omwille van vergelijkbaarheid tussen onderzoeken en om econometrische redenen is het wenselijk vergelijkbare grootheden te hanteren. De gebruikelijke methode daarvoor is om de reeksen te 'normaliseren'. Dat wil zeggen om te zetten naar nieuwe reeksen met een gemiddelde van nul en een standaardfout van 1. Wiskundig betekent dit de originele reeks  $y$  eerst verminderen met het gemiddelde van die reeks ( $\mu_y$ ) en dat resultaat te delen door de standaarddeviatie van reeks  $y$  ( $\sigma_y$ ). In een formule  $z = (y - \mu_y) / \sigma_y$  met  $z$  als de nieuwe reeks die in het onderzoek wordt gebruikt. Als vervolgens uit het onderzoek blijkt dat het effect van zittenblijven 0,34 bedraagt, dan betreft dit het effect van zittenblijven op de reeks  $y$  met 0,34. Die reeks beschrijft echter de getransformeerde leerprestaties. Om het effect op de 'echte' leerprestaties te verkrijgen, moet de transformatie worden teruggedraaid. Omdat geldt  $y = \sigma_y z + \mu_y$  is direct te zien dat het effect op de leerprestaties ( $y$ )  $0,34 \sigma_y$  en in gewoon Nederlands 0,34 standaarddeviatie van de te verklaren variabele.

Na vijf jaar zijn de effecten verwaarloosbaar. Er is dus sprake van een wegebbend effect.

Dit paper beschouwen we als de beste studie op dit moment, omdat het een uitgebreide verantwoording van de gevolgde *regression discontinuity* opneemt zoals checks op robuustheid, onderzoek naar een discontinuïteit voor en na de stelselherziening en naar een discontinuïteit rond de grenswaarde van wel/niet zittenblijven.

#### Allen et al. (2009)

Jacob en Lefgren (2004) en Schwerdt en West (2013) zijn niet de enige studies naar de effecten van zittenblijven. Allen et al. (2009) hebben ten behoeve van een meta-analyse 199 studies, verschenen tussen 1990 en 2007, kritisch geëvalueerd. Die evaluatie betrof onder andere de samenstelling van de controlegroep.<sup>12</sup> Slechts 22 van 199 studies voldeden aan alle criteria.

Binnen deze groep van 22 kregen studies met een hoge statistische kwaliteit, waaronder quasi-experimenteel design, een waardering van 3, terwijl studies met een andere vorming van *statistical control* een rating van 1 of 2 kregen.<sup>13</sup> In totaal zes studies van de 22 kregen de hoogste waardering. Van die zes studies richten vier zich op een *same age*-vergelijking, terwijl de andere twee een *same grade*-vergelijking betreffen.<sup>14</sup>

#### Greene en Winters (2007)

In het literatuuroverzicht van Allen et al. (2009) ontbreken Greene en Winters (2007). Zij maken gebruik van de wijzigingen in het onderwijsstelsel van Florida om de effecten van zittenblijven te bepalen.<sup>15</sup> Ook hier gaat het om de effecten één en twee jaar na het zittenblijven, waarbij het - net als bij Jacob en Lefgren (2004) - feitelijk gaat om de gecombineerde effecten van zittenblijven en zomerschool. In tegenstelling tot Jacob en Lefgren (2004) kunnen de effecten van de afzonderlijke interventies niet in deze studie bepaald worden. Voor hun empirische analyse volgen de auteurs zowel een *across-year approach* als een *regression discontinuity approach* waarbij zij gebruik maken van een *same age* vergelijking (Greene en Winters, 2007, blz. 328).

De auteurs komen tot de conclusie dat vooral in het tweede jaar na het zittenblijven sprake is van de grootste vooruitgang in leerprestaties.

---

<sup>12</sup> Zie Allen et al. (2009, blz. 5-6) voor een compleet overzicht van de criteria om toegelaten te worden tot de meta-analyse.

<sup>13</sup> Studies met een zuiver experimenteel design zijn niet gevonden (zie Allen et al., 2009, blz. 5-6).

<sup>14</sup> Zie Allen et al. (2009, blz. 21). Behalve de ook hierboven beschreven studies van Jacob en Lefgren (2004) en Roderick en Nagaoka (2005) kregen ook de 'same age'-studies van Hong en Raudenbusch (2005) en Rust en Wallace (1993) de hoogste waardering voor de samenstelling van de controlegroep. Wij delen die keuze niet. In beide gevallen is geen sprake van een quasi-experimenteel ontwerp, maar worden de controlegroepen vastgesteld op basis van matching. Hong en Raudenbusch (2005) hebben daarbij veel variabelen waarop ze kunnen matchen (207 in totaal, zie blz. 210). Maar het blijft matchen en dat betekent dat er nog steeds niet-observeerbare kenmerken zijn tussen leerlingen die wel/niet zijn blijven zitten en dus is er sprake van een risico op een 'selection bias' bij de resultaten. De dataverzameling in de studie van Rust en Wallace (1993) is in vergelijking met Hong en Raudenbusch (2005) aanzienlijk kleiner. Dit betreft zowel de beschikbare matching-variabelen als het aantal waarnemingen.

<sup>15</sup> Later, Greene en Winters (2009), hebben deze auteurs ook nog onderzoek gedaan naar de effecten van de uitzonderingsregel die bij zittenblijven mogelijk is. Deze studie die eigenlijk niet over zittenblijven sec gaat, maar over de uitzonderingsregel hebben we in dit overzicht buiten beschouwing gelaten.

## 2.4 Conclusie

Ondanks dat in de voorbije decennia een stapel empirische artikelen is verschenen over de effecten van zittenblijven, is de finale conclusie dat er weinig hard bewijs is. Een algemeen gebrek is dat studies zich niet richten op de 'overall effecten': het nut over de levensloop van alle mensen gezamenlijk. In veel onderzoek wordt de toetscore op lezen en/of rekenen als uitkomstmaat gehanteerd. Hoewel dit misschien een indicator is voor het later te verdienen inkomen en daarmee consumptiemogelijkheden, meet deze maatstaf bijvoorbeeld niet het nut van vrije tijd. Zo zal een kind dat de zomerschool volgt om hiaten in zijn kennis bij te spijkeren, hopelijk een positief effect op zijn toetsscores zien, maar tegelijkertijd een disnut ervaren van een misgelopen vakantie. Het effect van die zomerschool op het nut van deze leerling is daarmee ongewis.

Veel onderzoek van voor de eeuwwisseling is niet betrouwbaar genoeg. Men heeft zich vaak gericht op een vergelijking op basis van *same grade* in plaats van *same age* en/of last heeft van een *selection bias* in de resultaten. Recente ontwikkelingen in het onderwijsstelsel in enkele Amerikaanse staten en steden bieden mogelijkheden tot empirisch onderzoek dat aan de *selection bias* ontkomt. Dit empirisch onderzoek beperkt zich echter voorlopig nog tot de effecten van doubleren in de *grades 3, 6 en 8*. Doubleren in het voortgezet onderwijs en op de grens van groepen 2 en 3 in het Nederlandse basisonderwijs is niet onderzocht.<sup>16</sup> Verder wordt in hoofdzaak gekeken naar de effecten op leerprestaties aangaande lezen en rekenen. Slechts in enkele gevallen worden ook andere uitkomstmaten gehanteerd, zoals de kans op doublures later in het onderwijstraject en vroegtijdig schoolverlaten. Effecten op de sociaal-emotionele ontwikkelingen zijn vooralsnog geen onderwerp van betrouwbaar onderzoek. Dit geldt ook voor de effecten op andere mensen dan de zittenblijvers zelf.

De mogelijkheden om conclusies te trekken over de effectiviteit van zittenblijven worden ook beperkt doordat beleidswijzigingen vaak niet alleen betrekking hebben op zittenblijven. Neem bijvoorbeeld de gelijktijdige invoering van een pakket van remediërende acties om zittenblijven te voorkomen, of de gevolgen daarvan te beperken, invoering van een zomerschool, toewijzing van een goed presterende docent in het doublurejaar en extra steunlessen in het doublurejaar.

In de vorige paragraaf zijn een aantal - op het eerste gezicht veelbelovende - recente studies, gericht op een effectmeting van recente wijzigingen in het Amerikaanse onderwijsstelsel, op een rijtje gezet. Uit nadere beschouwing van de gevolgde empirische strategie concluderen wij dat de studies van Jacob en Lefgren (2004) en Schwerdt en West (2013) het meest geloofwaardig zijn. Ze benutten een quasi-experimenteel design en voeren de analyse netjes uit.

---

<sup>16</sup> De Amerikaanse 'grade 3' heeft betrekking op de groep 8 tot 9-jarige kinderen. In de Nederlandse context circa groep 5 van de basisschool.

De rode lijn in het betrouwbare empirisch onderzoek lijkt te zijn:

1. dat zittenblijvers het op lange termijn ten aanzien van leerprestaties niet significant beter of slechter doen dan hun bevorderde groepsgenoten.
2. dat zittenblijvers wel op korte termijn beter presteren.
3. dat de effecten van zittenblijven kunnen verschillen naar de klas waarin zittenblijven plaatsvindt. Zittenblijven in hogere klassen lijkt een grotere kans op vroegtijdig schoolverlaten te kennen (*grade 6* versus *grade 8*) en zittenblijven in *grade 3* (groep 5) lijkt gedurende een langere periode positieve effecten op leerprestaties te geven dan zittenblijven later in de onderwijsloopbaan.
4. dat de gevonden effecten gelden voor de groep kinderen die nabij de grens van wel/niet zittenblijven verkeren.

De vraag is natuurlijk of je op basis hiervan een vertaalslag naar de Nederlandse situatie kan maken. Maar stel dat dit het geval is, dan lijken - toegespitst op de doublure-pieken in de onderbouw van het basisonderwijs en de voorexamenjaren van het voortgezet onderwijs - twee verwachtingen niet onlogisch. Ten eerste lijkt zittenblijven effectief om middelbare scholieren aan een startkwalificatie te helpen zonder daarvoor de exameneisen te verlagen. Er is immers een positief kortetermijneffect van zittenblijven. Ten tweede is, afgaand op het resultaat dat zittenblijven vroeg in de onderwijsloopbaan effectiever is dan later (langere periode van positieve effecten op onderwijsprestaties, verminderde kans op schooluitval), zittenblijven in de groepen 2 en 3 van het basisonderwijs wellicht een effectieve strategie.<sup>17</sup>

## 3 Zittenblijven in relatie tot de organisatie van het onderwijs

*Hoeveel kinderen blijven zitten hangt samen met de organisatie van het onderwijs. Zo is er een verband tussen het percentage zittenblijvers en het feit dat eindexamens extern opgelegd worden. Er bestaat wetenschappelijk bewijs dat die strenge diploma-eisen positief doorwerken op de kwaliteit van het onderwijs. Kanttekening hierbij is wel dat dit bewijs niet keihard te noemen is.*

### 3.1 Samenhang met strenge diploma-eisen

In het voortgezet onderwijs is zittenblijven voor een belangrijk deel gerelateerd aan de harde eindexameneisen die de Nederlandse overheid opstelt. Vanwege deze harde eisen worden voor de eerdere leerjaren ook eisen gesteld aan het overgaan naar het volgende leerjaar. De normen die gelden bij het eindexamen worden dus als het ware naar voren

---

<sup>17</sup> Dat sluit overigens niet uit dat er andere maatregelen zijn die hetzelfde effect kunnen bereiken, maar tegen lagere kosten. Of in andere woorden, efficiëntere maatregelen.

geschoven. Leraren bepalen daarbij zelf hoe deze normen vertaald worden naar de situatie in eerdere jaren.

Externe eindexamens (die niet flexibel genormeerd zijn) lijken gunstig te zijn voor zowel de behaalde leerwinsten in het onderwijs als voor de arbeidsmarkt. Leerlingen uit landen met externe eindexamens die zijn gebaseerd op een vastgesteld curriculum, doen het beter op internationaal vergelijkbare toetsen. De meeste studies in deze literatuur vergelijken scores op internationaal vergelijkbare toetsen als TIMSS en PISA tussen landen en staten met Curriculum Based External Exit Exams (CBEEE's) en zonder CBEEE's. Hierbij kan een *selection bias* niet worden uitgesloten (zie kader in het volgende hoofdstuk), dat wil zeggen dat het mogelijk is dat de reden waarom sommige landen kiezen voor CBEEE's positief samenhangt met de scores. Dit zou dan betekenen dat CBEEE's niet de oorzaak van hoge scores zijn, maar een andere, niet-geobserveerde factor.

Bishop (1997), bijvoorbeeld, vergelijkt scores van 13-jarige leerlingen in de *Third International Mathematics and Science Study* (TIMSS) in 39 landen waaronder Nederland. Het ging hier om de data uit 1994-1995. De onderwijssystemen in deze landen verschillen aanzienlijk. Het vaststellen van externe eindexamens is daar een van. Bishop gebruikt alleen de aanwezigheid van CBEEE's als verklarende variabele. De uitkomst is dat test scores op wiskunde en science hoger zijn in landen met centrale eindexamens. Ook gebruikte Bishop in dit paper gegevens uit het *International Assessment of Educational Progress* (IAEP) uit 1991, met dezelfde resultaten. Hieraan deden vijftien landen mee. Nederland zat niet in deze steekproef.

Bishop vergelijkt vervolgens de uitslagen van de IAEP in verschillende staten in Canada met elkaar, die verschillen in termen van de eindexamens. Voor zover we deze resultaten geloven, laten zij zien dat het effect van CBEEE's waarschijnlijk met prikkels te maken heeft. Schooldirecties nemen betere leraren aan. Leraren doen beter hun best om leerlingen op het gewenste niveau te brengen. Door het 'vooruit schuiven van de normen' en de dreiging van het zittenblijven spijbelen leerlingen minder en kijken zij minder televisie. Ouders lijken ook meer betrokken bij het onderwijs omdat er meer van de examens afhangt.<sup>18</sup>

Als de resultaten op eindexamens goed verifieerbaar zijn voor de buitenwereld, zal er meer waarde aan worden gehecht op de arbeidsmarkt. Dit geeft positieve arbeidsmarktuitkomsten.<sup>19</sup> Piopiunik et al. (2013) vinden bijvoorbeeld aan de hand van GSOEP-data<sup>20</sup> dat in Duitse staten met CBEEE's vooral scholieren die niet aan het hoger onderwijs deelnemen, voordeel hebben van deze transparantie. Hun inkomens zijn hoger en hun kans op werkloosheid kleiner. Het blijft hier echter bij correlaties. Het causaal aantonen van dit verband is erg lastig.

---

<sup>18</sup> Bishop laat verder in een Handbook Chapter in 2006 zien dat ook college attendance lijkt samen te hangen met het vaststellen van externe eindexamens. De inschrijvingen zijn hoger in staten met CBEEE's (alleen New York) of met een *minimum competency exam* (MCE).

<sup>19</sup> Zie Hanushek en Woessman (2010) voor een samenvatting van de literatuur.

<sup>20</sup> German Socio Economic Panel dataset. De auteurs gebruiken de jaargangen 2001-2010.

## 4 Effecten van alternatieven voor zittenblijven

*Over de effectiviteit van voor- en vroegschoolse educatie, als beleidsoptie om verlengd kleuteren te voorkomen, bestaat een aanzienlijke internationale literatuur. Verder weten we voor Nederland dat de uniformering van de onderwijstijd voor kleuters uit achterstandsgroepen effectief is. Een verlenging van onderwijstijd in het vervolg van het basisonderwijs is niet altijd effectief; de intensiteit en uitvoering van de interventie zijn van belang.*

### 4.1 Voor- en vroegschoolse educatie

Voor- en vroegschoolse educatie (VVE) is bedoeld om jonge kinderen met een taal- en ontwikkelingsachterstand voor te bereiden op de basisschool. Voorschoolse educatie wordt aangeboden aan doelgroepkinderen tussen de 2 en 4 jaar op peuterspeelzalen en kinderdagverblijven. Vroegschoolse educatie wordt vervolgens gegeven aan doelgroepkinderen van 4 tot 6 jaar op basisscholen. VVE kan een positief effect hebben op het terugdringen van zittenblijven. Wanneer bij doelgroepkinderen al op jonge leeftijd de achterstand wordt weggewerkt, kan dit in latere klassen zorgen voor minder zittenblijvers.

Er zijn Amerikaanse studies die positieve effecten vinden van VVE op het terugdringen van zittenblijven. Deze studies voldoen een hoge onderzoeksstandaard: er wordt een gecontroleerd experiment (willekeurige toewijzing) uitgevoerd of gebruik gemaakt van een natuurlijk experiment. Er is dus rekening gehouden met de *selection bias*. Een aantal Amerikaanse programma's die wetenschappelijk geëvalueerd zijn, boden een vorm van VVE aan, waarvan de meeste zijn opgericht in het midden van de jaren zestig.

*Head Start* is een programma dat zich sinds 1964 richtte op de sociale en cognitieve ontwikkeling van minderbedeelde of arme kinderen. Studies van Currie en Thomas (1995 en 1999) vergelijken bijna 5000 kinderen uit een landelijke steekproef die aan *Head Start* hebben deelgenomen, met een controlegroep die bestaat uit broers en zussen die niet hebben deelgenomen. Deelname aan *Head Start* verlaagt de kans op zittenblijven voor blanke kinderen met 47% (Currie en Thomas, 1995) en voor kinderen met een Latijns-Amerikaanse achtergrond met 28% (Currie en Thomas, 1999), maar heeft geen effect op Afro-Amerikaanse kinderen. Alhoewel de auteurs dit niet hard kunnen maken, suggereren ze dat het verschil tussen de bevolkingsgroepen kan komen door de heterogeniteit van programma's en in vervolgscholen.

*Pre-K* richtte zich op risicokinderen van 3 en 4 jaar oud op basis van socio-economische status en andere thuis- en buurtomstandigheden. De effectiviteit van het programma is onderzocht door het *Pre-K* programma in New York te evalueren (Irvine et al., 1982). Twee cohorten (1975 en 1976) van in totaal 1348 leerlingen werden tot het einde van de derde klas gevolgd. De controlegroep bestaat uit kinderen die in aanmerking kwamen voor het



programma, maar hier niet aan hebben deelgenomen (kinderen op een wachtlijst). Het percentage van de kinderen dat een keer is blijven zitten of in speciale klassen is geplaatst werd vergeleken. Normale vooruitgang betekende dat ze de derde klas hebben afgemaakt. In vergelijking met de controlegroep maken 33% minder kinderen van de *Pre-K* leerlingen een minder dan normale vooruitgang door.

*The Abecedarian project* is een vorm van VVE waarbij de risicogroep hele dagen voorschools onderwijs werd aangeboden. Het programma bestond uit intensieve activiteiten om voornamelijk taal, maar ook sociale, emotionele en cognitieve ontwikkeling te bevorderen. 112 kinderen, geboren tussen 1972 en 1977, werden in North-Carolina willekeurig in de behandel- en controlegroep geplaatst. Barnett en Masse (2007) vonden, naast cognitieve verschillen, significante verschillen voor zittenblijven. 65% van de kinderen die geen voorschools onderwijs hadden genoten, zijn ooit blijven zitten, terwijl dit 34% is voor de kinderen die wel in de behandelgroep zaten.

Het *Chicago Child/Parent center* programma richtte zich op armoedige gebieden die niet bij *Head Start* of andere programma's aan bod kwamen. 1539 voornamelijk Afro-Amerikaanse kinderen, geboren in 1979 en 1980, werden onderworpen aan halve dagen voorschoolse opvang vanaf het moment dat zij 3 jaar waren. Daarna werden zij 19 jaar gevolgd. Reynolds et al. (2004) kijken naar het effect van deelname aan dit programma. Hieruit blijkt dat ook de effecten op zittenblijven positief zijn; 23% van de kinderen die deze voorschoolse opvang hebben genoten, blijft een keer zitten in de leeftijd 6-15 jaar, terwijl dit 36% is voor de kinderen die hier niet aan deelnamen.

Deze Amerikaanse studies laten zien dat VVE een kansrijke optie is om zittenblijven te verlagen. Er bestaan echter grote verschillen tussen de Amerikaanse en de Nederlandse programma's omtrent het onderwijsachterstandenbeleid. Amerikaanse programma's zijn veelal behoorlijk intensief, goed gestructureerd en soms gericht op behoorlijke achterstandsgroepen. In Nederland omvat VVE maximaal vier dagdelen en bestaat geen specifieke opleiding tot VVE-leid(st)er.<sup>21</sup> Daarnaast zijn er verschillen tussen Amerikaanse en Nederlandse achterstandskinderen. Minder bevoorrechte kinderen in Nederland hebben waarschijnlijk een betere uitgangspositie dan deze kinderen in de VS. Ook bestaan er geen betrouwbare Nederlandse impactevaluaties gericht op de effectiviteit van VVE. Vooralnog laten gegevens uit COOL en pre-COOL<sup>22</sup> geen positieve samenhang van voorschoolse programma's op cognitieve en non-cognitieve effecten zien.<sup>23</sup> Het is dus nog de vraag of deze aanzienlijke positieve effecten van VVE-programma's in de Verenigde Staten naar de Nederlandse situatie vertaald mogen worden.

---

<sup>21</sup> Driessen (2012) beschrijft wel uitgebreid de huidige Nederlandse praktijk.

<sup>22</sup> COOL – cohort onderzoek onderwijsloopbanen. Onder deze paraplu wordt gedetailleerde informatie over leerlingen verzameld.

<sup>23</sup> Zie Karssen et al. (2013)

## 4.2 Uniformering onderwijstijd

Een alternatieve optie voor verlengd kleuteren is een uniformering van de onderwijstijd van kinderen in groep 1 en 2. Nu kan er tussen kinderen die regulier groep 3 instromen, een verschil van maximaal 40 weken onderwijstijd optreden. Door augustus- en septemberkinderen, en wellicht ook oktober-, november- en decemberkinderen, al direct na de zomervakantie te laten starten met groep 1 wordt dit verschil iets minder groot. Dit kan vooral voor achterstandskinderen - allochtone kinderen en kinderen met laagopgeleide ouders - positieve effecten opleveren. Deze onderwijsverlenging zou daarom het best op deze groep leerlingen gericht kunnen worden.

Een betrouwbare studie van Leuven en coauteurs (2010) maakt gebruik van de optredende verschillen in onderwijstijd in groep 1 en 2 (ook veroorzaakt door de zomervakantie) om het effect van een maand onderwijstijd op testcores te laten zien. Hiervoor zijn de enquêtegegevens van de PRIMA-cohorten van 1995-2003 gebruikt. Zij tonen aan dat van iedere maand extra onderwijstijd in groep 1 en 2 de scores op toetsen van achterstandsleerlingen met 5-6% van een standaarddeviatie toenemen. De leerprestaties van andere kinderen werden niet door de extra onderwijstijd beïnvloed.

## 4.3 Gerichte verlenging onderwijstijd

Als alternatief voor zittenblijven in het primair en in het voortgezet onderwijs is een gerichte verlenging van de onderwijstijd een voor de hand liggende optie. In het primair onderwijs hebben we het dan bijvoorbeeld over een intensivering van het vroegschoolse-educatieprogramma. Dit betekent individuele of kleinschalige scholing van kinderen met een vroege achterstand door specifiek opgeleide leerkrachten. Voorbeelden van gerichte verlenging van de onderwijstijd in het voortgezet onderwijs zijn bijspijkertrajecten, weekendscholen en zomerscholen. Hierbij krijgt de leerling individueel of in een kleine groep extra onderwijs op zijn of haar probleemvak.

Uit een meta-analyse van betrouwbare Amerikaanse studies door Kidron en Lindsay (2014) komt naar voren dat de opzet van een verlenging van de onderwijstijd cruciaal is. Opvallend bij deze analyse was dat er in eerste instantie 7000 studies zijn geselecteerd, maar dat uiteindelijk maar 30 studies betrouwbaar genoeg werden geacht om te worden opgenomen. Er bevond zich onder deze 30 studies maar een studie onder *high school* scholieren. Positieve effecten op reken- en taalvaardigheid werden het meest gevonden als de leraren gekwalificeerd zijn, het onderwijs traditioneel en klassikaal is, en het om basisschoolleerlingen (en geen *middle school*-leerlingen) ging. Voor motivatie en sociaal-emotionele ontwikkeling leken juist niet-traditionele methodes die niet op school plaatsvinden, een groter effect te sorteren. Voor de basisschoolperiode bleek al uit een eerdere internationale literatuurstudie van Driessen en coauteurs (2010) dat de vormgeving van een verlenging van de onderwijstijd erg belangrijk is voor de effectiviteit. Driessen et al. concluderen dat de kwaliteit van docenten en de aansluiting op de standaardleerstof erg belangrijk is.

Over zomerscholen bestaat ook behoorlijk wat evidentie in de internationale literatuur. We nemen hier twee studies op die aansluiten bij de artikelen die in Hoofdstuk 3 zijn opgenomen over de effectiviteit van zittenblijven. Omdat zomerscholen bij de stelselherzieningen in Chicago en Florida in de VS onderdeel zijn geworden van het zittenblijf-regime, of beter geformuleerd een herkansing bieden om zittenblijven te voorkomen, schatten veel studies het gecombineerde effect van doubleren en het volgen van de zomerschool.

Matsudaira (2008) gebruikt een *regression discontinuity design* en vindt een gemiddeld effect van het meedoen aan een zomerschool van 0,12 standaarddeviatie voor zowel lees- als schrijfprestaties. Dit is ruim de helft van het effect dat gevonden is voor zittenblijven. De heterogeniteit in effecten betreft vooral *grade 3* (groep 5) versus *grades 5* en *6* (groep 8). Mogelijk reageren deze groepen verschillend. Voor de hoogste groepen komt Matsudaira tot 0,24 standaarddeviatie voor zowel wiskunde als lezen. Het is opvallend, en in afwijking van eerder onderzoek, dat de effecten van de zomerschool op beide domeinen vergelijkbaar zijn, en ook dat het effect sterker is voor oudere kinderen. Door het benutten van de twee verschillende momenten bij de bevorderingsbeslissing (juni-resultaten; augustus-resultaten) kunnen Jacob en Lefgren (2004) ook het effect van de zomerschool isoleren. Net als voor de leerprestaties komen zij tot de conclusie dat de effecten van een zomerschool vroeg in het leertraject (*grade 3*) het grootst zijn.

Voor Nederland bestaan maar weinig betrouwbare onderzoeken naar de effectiviteit van (gerichte) onderwijstijdverlenging. De evaluatie door Mulder et al. (2012) van schakelklassen tijdens de basisschoolperiode, die werden mogelijk gemaakt door een subsidie voor kinderen met een taalachterstand, is een interessante studie. De auteurs laten positieve effecten op leerprestaties zien, die op de langere termijn weer wat lijken uit te doven, net als bij zittenblijven. Deze studie bevat echter alleen beschrijvende statistieken, geen hard bewijs. Wel bestaat er Nederlandse evidentie voor de (in)effectiviteit van een bepaald bijlesprogramma en een specifieke weekendschool. Meijer en Van Klaveren (2011) gebruikten een gerandomiseerd experiment voor een evaluatie van een Nederlands bijlesprogramma voor 8-12 jarigen. Van Klaveren et al (2011) rapporteren ook over een experiment voor een betrouwbare evaluatie van een weekendschool voor groep 7. Beide programma's vonden nauwelijks effecten op leerprestaties. Het feit dat beide interventies niet erg intensief waren, zou dit kunnen verklaren. In feite is er behoefte aan veel meer van dit soort studies om te begrijpen onder welke omstandigheden verlengde lestijd welk effect sorteert.

Voor de middelbareschoolperiode kennen wij geen (internationale) evidentie van de effectiviteit van een gerichte verlenging van de onderwijstijd. Mogelijk vinden deze interventies momenteel niet plaats, of alleen op de private markt van huiswerk cursussen of eindexamentrainingen.

## 4.4 Deelcertificaten

Het gangbaarder maken van deelcertificaten in het onderwijs (bijvoorbeeld vwo-diploma met havo natuurkunde) is een manier om de harde eindexameneisen per schoolniveau gedeeltelijk naar beneden bij te stellen. Er zullen waarschijnlijk minder mensen blijven zitten als er op deelgebieden lagere drempels worden gehanteerd. Dit uiteraard met de kanttekening dat het gedragseffect – dat leerlingen minder hard hun best doen als de eisen minder hoog zijn – niet overheerst. Alhoewel het voor de hand ligt dat deelcertificaten zittenblijven in het VO zullen terugdringen, is hier geen wetenschappelijk bewijs voor. Er is simpelweg in de praktijk nauwelijks gevarieerd met deze mogelijkheid.

Waar we wel behoorlijk veel vanaf weten is hoe ontvangers van certificaten het in de VS doen op de arbeidsmarkt. Het gaat hierbij niet zozeer om deelcertificaten, maar om personen die buiten het reguliere onderwijs om toch een bewijs van voldoende kennis weten te vergaren via een General Educational Development (GED) certificaat. Cameron en Heckman (1993) laten zien dat personen met een GED weliswaar slagen voor hun 'eindexamen' maar in hun verdere leven statistisch toch veel meer overeenkomsten tonen met *high school dropouts*. Ze participeren minder op de arbeidsmarkt en verdienen minder dan personen die slaagden in het reguliere onderwijs. Een deel van de verklaring is selectie: personen die kiezen voor een GED zijn gemiddeld genomen minder getalenteerd dan reguliere diploma-ontvangers. Een ander deel van de verklaring zit in het effect van naar school gaan: een kind leert tijdens de lessen allerlei niet-cognitieve vaardigheden die van waarde zijn op de arbeidsmarkt (zie bijvoorbeeld Heckman en Rubinstein, 2001).

Uit het onderzoek naar GEDs in de VS kan ook worden afgeleid dat werkgevers een dergelijk certificaat minder waard vinden dan een regulier diploma. Nederlandse werkgevers kunnen op vergelijkbare manier reageren door individuen met deelcertificaten minder kansen te bieden. Deelcertificaten maken immers transparant dat een leerling moeite had met een bepaald vak. Dat een persoon is blijven zitten op school, is over het algemeen minder transparant.

# Appendix

## Internationale vergelijking zittenblijven

### Primair onderwijs

In een vergelijking tussen landen op basis van PISA-gegevens lijkt het erop dat Nederland erg hoog zit met het doublure-percentage in het PO (zie Figuur A.1). Van alle OESO-landen blijven alleen in Portugal en Luxemburg meer kinderen zitten op de basisschool. Dit is echter deels een vertekend beeld, zie het kader hierover.

Deze verschillen zijn het resultaat van een samenspel van verschillende factoren in heden en verleden.<sup>24</sup> Zo verschillen landen in hun onderwijsstelsel, bijvoorbeeld qua leerplichtleeftijd, groepsgrootte, opleidingsniveau van docenten, aangeboden curriculum en regelgeving omtrent zittenblijven. Bijvoorbeeld in Noorwegen en IJsland is zittenblijven verboden en start de leerplicht pas bij 6 jaar.<sup>25</sup> In sommige landen bevindt de gemiddelde groepsgrootte in het basisonderwijs zich rond de 18 leerlingen, terwijl dat in andere landen - zoals Nederland - rond de 22 ligt.<sup>26</sup> En als laatste voorbeeld, in sommige landen is sprake van een jaarklassensysteem waarin beperkte ruimte is voor differentiatie naar mogelijkheden van individuele kinderen (bijvoorbeeld Nederland), terwijl in andere landen het onderwijs explicieter is afgestemd op het niveau van het kind (bijvoorbeeld Finland).<sup>27</sup>

Ook verschillen landen in de invulling van het voorschoolse traject. Zo is in sommige landen sprake van hogeropgeleide docenten in de tweede fase van dit traject.<sup>28</sup> In IJsland is er bijvoorbeeld een nationaal curriculum opgesteld voor de opvang in deze periode.<sup>29</sup>

---

<sup>24</sup> Goos et al. (2013) betogen dat cultuur ook een belangrijke rol speelt.

<sup>25</sup> Zie OECD (2014, tabel C1.1a, blz. 312).

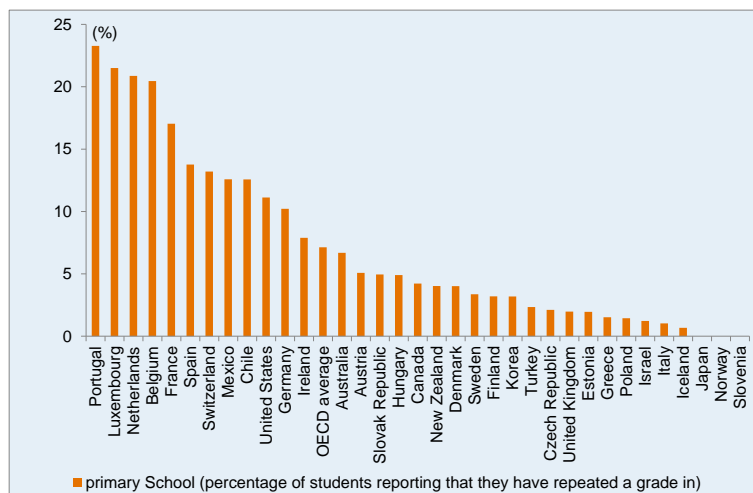
<sup>26</sup> Zie OECD (2014, chart D2.1, blz. 442).

<sup>27</sup> OECD (2013a, blz. 6).

<sup>28</sup> Zie Eurydyce (2009, blz. 110 e.v.).

<sup>29</sup> Zie met name artikel 2 van de Preschool Act (no. 90, 12 June 2008) van The Ministry of Education, Science and Culture in IJsland.

**Figuur A.1 Percentage PISA-scholieren (15-jarigen) die aangeven ooit te zijn blijven zitten in het primair onderwijs (Bron: OESO 2013, primary school)**



### Meetproblemen bij internationale vergelijkingen

De internationaal vergelijkende gegevens over zittenblijven bevatten veel problemen. Zo zijn de gegevens niet ontleend aan administratieve data, maar aan een enquête onder 15/16-jarigen. Van enquêtes is bekend dat er een grotere kans is op inconsistenties tussen de verschillende antwoorden van eenzelfde respondent. Verder vereist een internationale vergelijking een uniform definitiekader. Bij de rapportage over zittenblijven in de Pisa-enquête is twijfel daarover gerechtvaardigd. Zo wordt in de Nederlandse versie van de Pisa-vragenlijst aan 15/16-jarigen gevraagd of zij ooit op de basisschool zijn blijven zitten, dat wil zeggen of zij een jaar gedoubleerd hebben in de opleidingsniveaus ISCED-0 en/of ISCED-1. In veel landen beperkt het primair onderwijs zich tot het ISCED-1 niveau. Hierdoor komt in die landen het verlengd kleuteren niet in de cijfers van zittenblijven op de basisschool tot uitdrukking. Om internationaal vergelijkend te zijn, zou Nederlandse jongeren dan ook gevraagd moeten worden naar zittenblijven in de groepen 3 tot en met 8. Om de effecten van zittenblijven op bijvoorbeeld leerprestaties te kunnen vaststellen, moeten de leerprestaties van twee groepen kinderen worden vergeleken waarbij de groepen in maar één aspect verschillen: wel/niet zittenblijven. Het vinden van die controlegroep is de grote uitdaging.

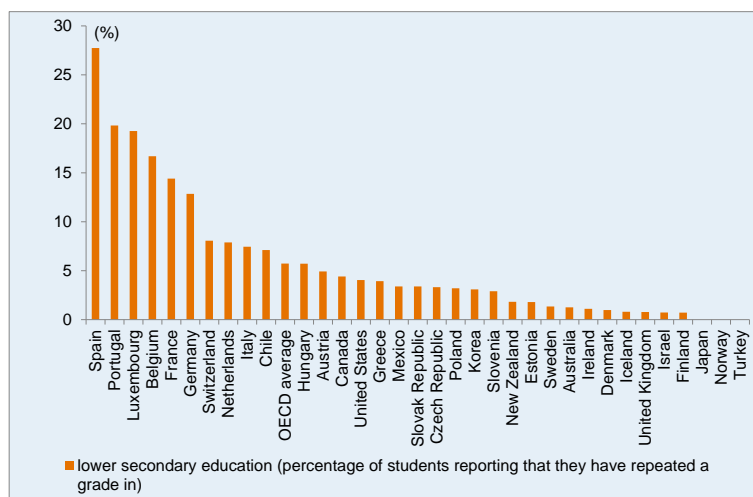
### Voortgezet onderwijs

Voor de internationale vergelijking van zittenblijven in het VO is minder informatie beschikbaar dan voor die vergelijking in het PO. De PISA-gegevens zijn immers gebaseerd op enquêtes onder 15-jarigen. Over de situatie van oudere middelbare scholieren in andere landen weten we niets. Juist in de bovenbouw blijven in Nederland relatief veel kinderen zitten. Nederland scoort in de internationale vergelijking van zittenblijven in de onderbouw van het vo (voortgezet onderwijs) gemiddeld (zie Figuur A.2). Het kan dus zijn dat dit beeld voor het gehele vo te rooskleurig is, gezien onze hoge zittenblijpercentages in de bovenbouw.

Ook hier hangen zittenblijpercentages mede samen met het onderwijsstelsel. Zo zien sommige landen af van zittenblijven en bepalen aan het eind van de onderbouw van het vo het bereikte niveau van een leerling. Dit niveau is dan richtinggevend voor

vervolgonderwijs.<sup>30</sup> In andere landen, zoals Nederland, is de eindtoets bedoeld om vast te stellen of de leerling over de minimaal vereiste kennis beschikt. Ook is er internationaal sprake van een verschillend criterium voor beëindiging van de leerplicht. In veel landen is leeftijd (variërend van 16 tot 18 jaar) het enige criterium.<sup>31</sup> In bijvoorbeeld Nederland wordt additioneel een bereikt kennisniveau (startkwalificatie) vereist. Verschillende landen bieden de mogelijkheid verschillende vakken in een verschillend tempo te doorlopen door slechts eenmaal per twee jaar het bereikte niveau ten opzichte van de norm vast te stellen.<sup>32</sup>

**Figuur A.2 Percentage PISA-scholieren (15-jarigen) die aangeven ooit te zijn blijven zitten in de onderbouw voortgezet onderwijs (Bron: OESO 2013, lower secondary education)**



<sup>30</sup> Zie OECD (2013b, laatste kolom table 4.A1.3) voor een overzicht.

<sup>31</sup> Zie OECD (2014, tabel C1.1a, blz. 312).

<sup>32</sup> Zie Eurydice (2011, hoofdstuk 3) voor een vollediger overzicht van wanneer en de wijze waarop in verschillende landen tot doubleren wordt besloten.

## Berekening kosten zittenblijven

De kosten van zittenblijven worden in paragraaf 4 van de CPB Policy Brief 'Zittenblijven in het primair en voortgezet onderwijs: Een inventarisatie van de voor- en nadelen' op een eenvoudige wijze berekend. Het voordeel van deze eenvoud is de transparantie.

Tabel A3 geeft een schematisch overzicht van de berekening. De jaarlijkse onderwijsuitgaven zouden met ongeveer 500 miljoen euro dalen als alle zittenblijvers hetzelfde einddiploma zonder vertraging zouden behalen. In deze berekening worden kosten voor een eventueel alternatief voor zittenblijven niet meegenomen. De uitgaven dalen dus alleen per saldo wanneer het alternatief goedkoper is dan 500 miljoen euro.

Naast deze uitgavendaling zou afschaffen van zittenblijven leiden tot een inkomstenstijging voor de overheid met een bedrag in de orde van 900 miljoen euro per jaar. Minder zittenblijven betekent in de regel een langer verblijf op de arbeidsmarkt, waardoor de belasting- en premie-inkomsten stijgen. De gestileerde berekening abstraheert van verschillen in inkomensprofielen over de levensloop en veronderstelt dat een vmbo'er succesvol doorstudeert aan het mbo, een havist een hbo-opleiding voltooit en een vwo'er zijn mastertitel behaalt. Bovendien wordt verondersteld dat zittenblijvers hetzelfde inkomen bereiken als de gemiddelde persoon met dat opleidingsniveau. Op deze aannames valt in de praktijk uiteraard af te dingen, waardoor het berekende bedrag voor de indirecte kosten slechts als indicatief mag worden beschouwd.

**Tabel A.3 Toelichting berekening kosten zittenblijven**

	Vmbo	Havo	Vwo	Totaal
Zittenblijvers (x 1000) (a)*	41	23	9	73
Directe onderwijskosten (in mrd, (a) x 6500 euro**)	0,3	0,1	0,1	0,5
waarvan toe te rekenen aan PO	0,1	0,0	0,0	0,2
waarvan toe te rekenen aan VO	0,2	0,1	0,0	0,3
Gemiddeld jaarinkomen (x 1000) (b)***	25	36	50	
Totaal inkomensverlies (in mrd, (a) x (b)) (c)	1,0	0,8	0,5	2,3
Indirecte kosten voor de overheid: totaal aan gedeerde belastingen/premies (in mrd, (c) x 40%)	4,0	0,3	0,2	0,9
Saldo overheidsrekening (in mrd)	0,7	0,5	0,2	1,4
waarvan toe te rekenen aan PO	0,3	0,1	0,0	0,5
waarvan toe te rekenen aan VO	0,4	0,4	0,2	0,9

\* Eigen berekening op basis van referentieramingen OCW. Hier worden alle leerlingen die in de zomer van 2012 een diploma ontvingen, met een jaar of meer vertraging meegenomen. Het aantal jaar vertraging is weliswaar een onderschatting, maar hiertegenover staat de vertraging die niet is opgelopen door zittenblijven maar door 'opstroom' naar een hoger niveau.

\*\* Cijfer afkomstig uit Stamos van het CAOP.

\*\*\* Bron: Mooij et al. (2012).



# Referenties

Allen, C., Q. Chen, V. Willson en J. Hughes, 2009, Quality of research design moderates effects of grade retention on achievement: A meta-analytic, multi-level analysis, *Educational Evaluation and Policy Analysis*, vol. 1: 480-99.

Barnett, W.S. en L.N. Masse, 2007, Comparative benefit-cost analysis of the Abecedarian program and its policy implications. *Economics of Education Review*, vol. 26(1): 113-25.

Bishop, J., 1997, The effect of national standards and curriculum-based examinations on achievement, *American Economic Review*, vol. 87(2): 260-64.

Bishop, J., 2006, Drinking from the fountain of knowledge: Student incentive to study and learn -- Externalities, information problems, and peer pressure, in: E. Hanushek en F. Welch (eds) *Handbook of the Economics of Education*, Amsterdam: North Holland.

Cameron, S. en J. Heckman, 1993, The Nonequivalence of High School Equivalents, *Journal of Labor Economics*, vol. 11: 1-47.

Currie, J. en D. Thomas, 1995, Does Head Start make a difference? *American Economic Review*, vol. 85(3): 341-64.

Currie, J. en D. Thomas, 1999, Does Head Start help Hispanic children? *Journal of Public Economics*, vol. 72: 235-62.

Diris, R., 2012, *Estimating the Impact of Grade Retention: A Heterogeneous Analysis*, Working Paper Maastricht University.

Driessen, G., A. Klaassen en F. Smit, 2010, *Variatie in Schooltijden*, ITS, Radboud Universiteit Nijmegen.

Driessen, G., 2012, *Voor- en vroegschoolse Educatie: Een onderzoek naar de uiteenlopende wijzen waarop in gemeenten vorm wordt gegeven aan VVE*, ITS, Radboud Universiteit Nijmegen.

Driessen, G., B. Leest, L. Mulder, T. Paas en T. Verrijt, 2014, *Zittenblijven in het Nederlandse basisonderwijs: een probleem?*, ITS, Radboud Universiteit Nijmegen.

Eurydice, 2011, Grade retention during compulsory education in Europe: regulations and statistics, Tech. Rapp.

Goos, M., B. Schreier, H. Knipprath, B. De Fraine, J. van Damme en U. Trautwein, 2013, How can cross-country differences in the practice of grade retention be explained? A closer look at national educational policy factors, *Comparative Education Review*, vol. 57(1): 54-84.

- Greene, J. en M. Winters, 2007, Revisiting grade retention: an evaluation of Florida's test-based promotion policy, *Education Finance and Policy*, vol. 2: 319-40.
- Greene, J. en M. Winters, 2009, The effects of exemptions to Florida's test-based promotion policy: Who is retained? who benefits academically?, *Economics of Education Review*, vol. 28: 135-42.
- Hanushek, E. en L. Woessman, 2010, The Economics of International Differences in Educational Achievement, in: Eric A. Hanushek, S. Machin en L. Woessman (eds) *Handbook of the Economics of Education*, Amsterdam: North Holland.
- Heckman, J. en Y. Rubinstein, 2001, The Importance of Noncognitive Skills: Lessons from the GED Testing Program, *American Economic Review*, vol. 91: 145-49.
- Holmes, C.T. en K.M. Matthews, 1984, The effects of nonpromotion on elementary and junior high school pupils: A meta-analysis, *Review of Educational Research*, vol. 54: 225-36.
- Holmes, C., 1989, Grade-level retention effects: a meta-analysis of research studies, in: L. Shepard en M. Smith (eds), *Flunking grades: research and policies on retention*, pp. 16-33, The Falmer Press.
- Hong, G. en S. Raudenbush, 2005, Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics, *Educational Evaluation and Policy Analysis*, vol. 27(3): 205-44.
- Irvine, D.J., M.D. Horan, D.L. Flint, S.E. Kukuk en T.L. Hick, 1982, Evidence supporting comprehensive early childhood education for disadvantaged children, *Annals of the American Academy of Political and Social Science*,
- Jackson, G., 1975, The research evidence on the effects of grade retention, *Review of Educational Research*, vol. 45(4): 613-35.
- Jacob, B. en L. Lefgren, 2004, Remedial education and student achievement: a regression-discontinuity analysis, *The Review of Economics and Statistics*, vol. 86(1): 226-44.
- Jacob, B. en L. Lefgren, 2009, The Effect of Grade Retention on High School Completion, *American Economic Journal: Applied Economics* 2009, vol. 1(3): 33-58.
- Jimerson, S.R., 2001, Meta-analysis of grade retention research: Implications for practice in the 21st century, *School Psychology Review*, vol. 30(3): 420-37.
- Karssen, M., I. van der Veen, A. Veen, M. van Daalen en J. Roeleveld, 2013, *Effecten van deelname aan en kwaliteit van voor- en vroegschoolse educatie op de ontwikkeling van kinderen*, Kohnstamm Instituut, Amsterdam.
- Kidron, Y. en T. Lindsay, 2014, *The effects of increased learning time on student academic and nonacademic outcomes: Findings from a meta-analytic review* (REL 2014-015), Washington,

DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance

Klaveren, C. van, H. Terwijn en E. Meyer, 2011, The Weekend School Effect on Perceived Cognitive and Social Competences: Evidence from a Randomized Controlled Experiment, TIER Working Paper Series 11/01.

Leuven, E., M. Lindahl, H. Oosterbeek en D. Webbink, 2010, Expanding schooling opportunities for 4-year-olds, *Economics of Education Review*, vol. 29: 319-28.

Matsudaira, J., 2008, Mandatory summer school and student achievement, *Journal of Econometrics*, vol. 142: 829-50.

Mooij, M. de, M. Geerdinck, L. Oostrom en C. van Weert (2012), Studeren loont. Inkomens van afgestudeerden in het mbo, hbo en wetenschappelijk onderwijs, *Socialeconomische trends*, CBS

Meijer, E. en C. van Klaveren, 2011, Evaluation of an Extended Day Program in the Netherlands: A Randomized Field Experiment, Tier Working Paper Series 11/02.

Ministry of Education , Science and Culture, 1998, The educational system in Iceland.

Mulder, L., I. van der Veen, M. Derriks en D. Elshof, 2012, *De schakelleerlingen verder gevolgd: Het tweede vervolgonderzoek bij leerlingen die in 2006/07 of 2007/08 in een schakelklas hebben gezeten*, ITS en Kohnstamm instituut, Nijmegen en Amsterdam.

OECD, 2013a, *Education policy outlook: Finland*, Paris.

OECD, 2013b, *Synergies for better learning: an international perspective on evaluation and assessment education policy outlook*, Paris.

OECD, 2014, *Education at a Glance*, Paris.

Piopiunik, M., G. Schwerdt en L. Woessmann, 2013, Central school exit exams and labor-market outcomes, *European Journal of Political Economy*, vol. 31: 93-108.

Reynolds, A.J., 2004, *Success in early intervention: The Chicago Child-Parent Centers*. Lincoln:University of Nebraska Press.

Roderick, M. en J. Nagaoka, 2005, Retention under Chicago's high-stakes testing program: Helpful, harmful, or harmless?, *Educational Evaluation and Policy Analysis*, vol. 27: 309-40.

Roeleveld, J. en I. van der Veen, 2007, Kleuterbouwverlenging in Nederland: omvang, kenmerken en effecten, *Pedagogische Studiën*, vol, 84(6): 448-62.

Rust, J. en K. Wallace, 1993, Effects of grade level retention for four years, *Journal of Instructional Psychology*, vol. 20: 152-66,  
<http://researchroom.wikispaces.com/effects+of+grade+level+retention+93>.

Schwerdt, G. en M. West, 2013, The effects of test-based retention on student outcomes over time: Regression discontinuity evidence from Florida, Discussion Paper 7314, IZA.

Xia, N. en S. Kirby, 2009, Retaining students in grade: a literature review of the effects of retention on students' academic and nonacademic outcomes, Tech. Rapp., Rand Education.

Zambrana, I.M., F. Pons, P. Eadie en E. Ystrom, 2014, Trajectories of language delay from age 3 to 5: persistence, recovery and late onset, *International Journal of Language & Communication Disorders*, vol. 49(3): 304-16.



Dit is een uitgave van:

Centraal Planbureau  
Van Stolkweg 14  
Postbus 80510 | 2508 GM Den Haag  
T (070) 3383 380

[info@cpb.nl](mailto:info@cpb.nl) | [www.cpb.nl](http://www.cpb.nl)

Januari 2015