

Onderwijsbeleid in Nederland

De kwantificering van effecten

CPB Achtergronddocument

6 juni 2011

Roel van Elk, Debby Lanser en Sander van Veldhuizen

Contents

1	Inleiding	7
2	Het belang van menselijk kapitaal	9
3	De opbrengsten van onderwijsbeleid: de methodologie	13
3.1	Deel 1: De kwalitatieve analyse van de effecten van onderwijsbeleid	13
3.2	Deel 2: De kwantitatieve analyse van de effecten van onderwijsbeleid	15
3.2.1	Micro-evidentie revisited: onzekerheid	16
3.2.2	Doorwerking van micro-effecten op langere termijn	17
3.2.3	Cohortmodel	18
3.3	BBP-effecten	20
3.3.1	Arbeidsaanbod	21
3.3.2	Kosten	22
3.3.3	Het totaal beeld	22
3.3.4	Illustratie kwantificering	22
4	Opbrengsten van onderwijsbeleid	25
4.1	Nederland in beeld	27
4.2	Kwantificering beleidsthema's op onderwijsinputs	29
4.2.1	Voor- en vroegschoolse educatie	29
4.2.2	Onderwijstijd	33
4.2.3	Klassengrootte	36
4.2.4	Schooluitval	39
4.2.5	Inspectietoezicht	44
4.2.6	Scholing en begeleiding onderwijspersoneel	46
4.3	Kwantificering institutionele beleidsthema's	51
4.3.1	Centrale examinering	51
4.3.2	Informatievoorziening	54
4.3.3	Accountability	56
4.3.4	Vroege selectie	58
4.3.5	Prestatiebeloning onderwijspersoneel	62
4.3.6	Sociaal leenstelsel	65
4.4	Niet-kwantificeerbare beleidsthema's	67
4.4.1	Salaris onderwijspersoneel	68
4.4.2	Salaris schoolleiders	68
4.4.3	Speciaal onderwijs	69

4.4.4	Excellentie en hoger onderwijs	70
4.4.5	School en omgeving	70
4.4.6	Meer computers en ICT in het onderwijs	71
4.4.7	Leven lang leren	71
4.4.8	Schakelklassen en summer courses	72
4.4.9	Segregatie	73
4.4.10	Concurrentie en vrije schoolkeuze	74
5	Vervolgonderzoek	77
	References	81

Dankwoord

Het onderzoek ten behoeve van het gepresenteerde instrumentarium heeft plaats gevonden op basis van externe financiering door de ministeries van OCW, EL&I en Financiën. Wij danken de leden van de begeleidingscommissie voor hun nuttige commentaar op eerdere versies van dit rapport: Ted Reininga en Cor Katerberg (Ministerie van OCW), Frans Suijker, Broos Brouwers, Piet Donselaar, Herm van Beek en Adiel Jahangir (Ministerie van EL&I) en Matthijs Schreurs, Joost Baeten, Bart Boon en Pieter van Winden (Ministerie van Financiën). Daarnaast heeft de ontwikkeling van het instrumentarium sterk geprofiteerd van de werkzaamheden van CPB-collega's voor Keuzes in Kaart 2011-2015 en de projecten Microevidentie van Onderwijsbeleid en Instrumentarium Productiviteit. In het bijzonder danken wij Marc van der Steeg, Bert Smid, Sander Gerritsen, Patricia Prüfer, Jürgen Antony, Dinand Webbink, Ruud Okker en George Gelauff voor hun waardevolle bijdrage. De gepresenteerde analyse is de verantwoordelijkheid van de auteurs.

1 Inleiding

Onderwijs wordt gezien als een belangrijke determinant van productiviteit. Mede uit het oogpunt van productiviteitsverhoging voert de Nederlandse overheid daarom beleid op deze pijler van de kenniseconomie. De mogelijke beleidsinstrumenten zijn legio en hun reikwijdte en doelstellingen zeer divers. Verhoogt de overheid de lerarensalarissen, zet het in op voor- of vroegschoolse educatie of maakt het een deel van de financiële middelen van een school afhankelijk van haar prestaties? Dergelijke afwegingen vragen om een, bij voorkeur kwantitatieve, analyse van de effecten van onderwijsbeleid. In welke mate verandert de inzet van een specifiek beleidsinstrument het lange termijn productieniveau of de welvaart?

Helaas is dit type analyses nauwelijks voorhanden. Dat heeft verschillende redenen. Allereerst vraagt het om betrouwbare informatie over de effecten van specifieke beleidsinstrumenten op de individuele onderwijsprestaties van leerlingen. Ten tweede vraagt het om inzicht in de wijze waarop de individuele onderwijsprestaties accumuleren tot het menselijk kapitaal van de beroepsbevolking. Ten derde vraagt het om kennis over de wijze waarop menselijk kapitaal bijdraagt aan de productiviteitsverhoging van een land.

Dit rapport stelt deze vragen één voor één centraal. Voor diverse beleidsthema's zetten wij de beschikbare informatie over de effectiviteit van beleidsinstrumenten op een rij en kwantificeren wij, waar mogelijk, de lange termijn effecten. Voor deze kwantificering maken wij gebruik van een instrumentarium dat de micro-empirie van beleidsevaluaties combineert met de macro-economische kennis over de opbrengsten van menselijk kapitaal. De analyses brengen het lange termijn effect op de arbeidsproductiviteit, het arbeidsaanbod en de budgettaire en onderwijskosten in beeld. De thema's zijn zo veel mogelijk afgestemd op de huidige beleidsdiscussie.

Het gepresenteerde instrumentarium vormt een uitbreiding op Cornet et al. (2006), dat de belangrijkste bron vormde voor de kwalitatieve analyse van kennismaatregelen in CPB (2006). De ontwikkeling van dit instrumentarium heeft het mogelijk gemaakt voor het eerst de effecten van onderwijsbeleid ook kwantitatief te analyseren. In Keuzes in Kaart 2011-2015 is de analyse van onderwijsmaatregelen dan ook uitgebreid met een kwantificering van effecten op basis van de uitkomsten van dit rapport. Dit rapport dient derhalve ter achtergrond bij de kwantitatieve analyse van onderwijsmaatregelen in CPB & PBL (2010).

De kwantitatieve analyse van onderwijsbeleid in termen van bbp-effecten is complex en blijft met onzekerheid omgeven. Allereerst verschillen beleidsevaluaties binnen één beleidsthema in hun uitkomsten. Alleen wanneer de studies een consistent beeld laten zien, combineren we deze voor een kwantitatieve effectschatting. Ten tweede zijn veel evaluaties gebaseerd op buitenlands onderzoek, waardoor de uitkomsten veelal niet direct toepasbaar zijn op Nederlands beleid en de Nederlandse beleidsomgeving. Dit vraagt om een vertaling naar de Nederlandse context. Mogelijk voert Nederland meer of minder beleid, of zijn de sociaal-economische

omstandigheden verschillend. Waar mogelijk baseren wij ons op Nederlandse studies, tenzij de kwaliteit onder doet voor de internationale literatuur. In de overige gevallen vertalen wij de buitenlandse situatie zo goed mogelijk naar de Nederlandse context. Ten derde is de exacte vormgeving van een maatregel bepalend voor zijn effect. Onschuldige implementatiefouten kunnen soms grote gevolgen hebben, bijvoorbeeld doordat de prikkels binnen het instrument net anders komen te liggen.

Naast deze onzekerheden betreffende de effectiviteit op het micro-niveau van de beleidsinterventie, bestaat ook onzekerheid in de vertaalslag van deze effectiviteit naar de opbrengsten van beleid op macro-niveau. Zo is allereerst de precieze relatie tussen individuele leerprestaties en het menselijk kapitaal van de beroepsbevolking nog niet volledig ontrafeld. Daarnaast laten de macro-econometrische studies naar de opbrengsten van menselijk kapitaal nog niet altijd een consistent beeld zien. De interpretatie van gevonden resultaten in de literatuur wordt bovendien nog eens bemoeilijkt door de zogenoemde “schooling premium puzzle”. Waarom gaan individuen gegeven de hoge gevonden opbrengsten van onderwijs niet langer naar school?

De onzekerheid binnen de micro-econometrische studies brengen wij in beeld door meerdere varianten te presenteren. Die varianten kennen altijd een voorzichtige insteek om zo de overige onzekerheden ook voldoende recht te doen.

Dit rapport bouwt voort op een tweetal belangrijke overzichtstudies, Cornet et al. (2006) en Webbink et al. (2009). Cornet et al. (2006) zetten voor diverse beleidsinstrumenten de evidentie uit de internationale literatuur op een rij en leiden vervolgens een kwalitatief oordeel af over de effecten ervan. De literatuur die deze studie als uitgangspunt neemt, betreft de (quasi-)experimentele literatuur. In de laatste decennia heeft deze literatuur een enorme vlucht genomen. Nieuwe micro-econometrische technieken maken het mogelijk een causale relatie vast te stellen tussen het onderwijsbeleid en haar effect op leerprestaties. In Webbink et al. (2009) is deze evidentie voor een flink aantal stelselkenmerken nogmaals beschreven en uitgebreid met recente inzichten.

De structuur van dit rapport is als volgt. Hoofdstuk 2 zet de verschillende inzichten uit de internationale literatuur rond het belang van menselijk kapitaal op individuele opbrengsten en productiviteitsverhoging op een rij. Hoofdstuk 3 schetst vervolgens de methodologie die is toegepast in de kwantitatieve analyses. In hoofdstuk 4 illustreren wij deze werkwijze voor een breed scala aan kansrijke beleidsthema's. Hier worden ook een aantal actuele beleidsthema's besproken, die op dit moment vanwege de beperkte beschikbare evidentie nog niet kwantificeerbaar zijn. Hoofdstuk 5 sluit af met mogelijke richtingen voor verder onderzoek.

2 Het belang van menselijk kapitaal

Onderwijs speelt een belangrijke rol in het economisch succes van individuen. Mensen met een hoge opleiding verdienen meer, hebben minder kans om werkloos te worden en participeren meer op de arbeidsmarkt. Een hoog privaat rendement is op zich echter geen reden tot overheidsinterventie. Overheidsinterventie, mits effectief uitgevoerd, is pas legitiem als het maatschappelijk rendement het private rendement overstijgt.

Het private rendement wordt afgemeten aan het extra loon dat een werknemer verdient door het volgen van een jaar extra onderwijs. Betrouwbare schattingen van dit zogenaamde Mincer rendement lopen uiteen van 5% tot 10%.¹ Card (1999) en Heckman et al. (2006) bieden mooie overzichten van de literatuur. De kwaliteit van de schattingen is inmiddels zodanig dat de omvang van dit rendement geschoond is van andere waargenomen en niet-waargenomen individuele karakteristieken. Voor een beschrijving van de achterliggende (quasi-)experimentele techniek verwijzen wij naar het kader “De methodiek van beleidsevaluaties met een (quasi-)experimenteel design” in hoofdstuk 3. Jacobs en Webbink (2006) en CPB (2008) schatten het private rendement specifiek voor Nederland. Tot 1989 neemt het private rendement af. Na 1989 is het rendement oplopend. De geschatte rendementen liggen tussen de 6 en 9%. Deze schattingen zijn echter niet gecorrigeerd voor onobserveerbare kenmerken.

De sociale opbrengsten van onderwijs worden veelal afgemeten aan een relatie tussen de hoeveelheid menselijk kapitaal en het productieniveau van een land. Het wetenschappelijke veld onderscheidt een tweetal benaderingen. De eerste benadering bouwt voort op de neoklassieke aanpak van de groeitheorie, waarin de voorraad menselijk kapitaal van invloed is op het lange termijn productieniveau. Deze relatie wordt geschat met behulp van zogenaamde formele groeiregressies. De tweede benadering baseert zich op de endogene groeitheorie, waarin een eenmalige aanpassing in de voorraad menselijk kapitaal resulteert in een permanente verandering van de economische groei. Sinds de seminal papers van Kormendi en Meguire (1985) en Barro (1991) zijn een breed scala aan relaties geschat, de zogenaamde informele of Barro regressies. In overeenstemming met het merendeel van het wetenschappelijke veld kiezen wij voor schattingen uit de eerste benadering. Het lijkt onwaarschijnlijk dat eenmalige variaties een blijvende impact zullen hebben op economische groei. Deze discussie is echter nog niet beslecht.

Metten en schatten

Empirische analyse vraagt om een geschikte maat voor de voorraad menselijk kapitaal. Deze maat blijkt niet gemakkelijk te geven. Het begrip is abstract en lastig te kwantificeren in een reeks met een enigszins respectabele tijdshorizon of ruimtelijke dimensie. De literatuur hanteert

¹ Bij deze Mincer schattingen van het private rendement wordt rekening gehouden met het misgelopen salaris tijdens een extra studiejaar.

verschillende proxies, zoals scholingsdeelname, of zogenaamde attainment rates, zoals het hoogst behaalde opleidingsniveau of het aantal jaren genoten onderwijs. De laatste maat is de meest gebruikte met als belangrijkste reden de beschikbaarheid van een grote internationale dataset van Barro en Lee (1993), Barro en Lee (2001) en Delafuente en Domenech (2006). Alhoewel veel gebruikt, is de kwaliteit van deze dataset niet onomstreden.

Cohen en Soto (2007) hebben recent een nieuwe verbeterde dataset ontworpen, die rekening houdt met informatie over veranderende onderwijssystemen over de tijd en de leeftijdsopbouw van de beroepsbevolking en haar bijbehorende karakteristieken. Hun dataset met zogenaamde efficiënte onderwijsjaren sluit zodoende beter aan bij de praktijk. Zo is de rol van genoten onderwijs onder ervaren senioren vermoedelijk anders dan voor jonge mensen en heeft een verdubbeling van het aantal jaren onderwijs in een sterk opkomende economie vermoedelijk een andere impact dan voor een gevestigde economie.

Net als de schattingen van het private rendement kampen empirische analyses van het sociale rendement met een endogeniteitsprobleem. Vindt economische vooruitgang nu plaats doordat individuen meer gaan studeren of gaan in een periode van economische voortuitgang individuen meer investeren in onderwijs? Daarnaast kennen macro-economische schattingen vaak dataproblemen. De reeksen over de voorraad menselijk kapitaal zijn kort en van een matige kwaliteit. Het aantal vrijheidsgraden en de nauwkeurigheid van de schattingen wordt hierdoor beperkt. Met de komst van nieuwe econometrische technieken hebben de macroschattingen een sprong voorwaarts gemaakt. Zo kan de schattingsmethode system GMM van Blundell en Bond (1998) een deel van het endogeniteitsprobleem oplossen en zorgen het gebruik van zowel de tijd- als ruimtelijke dimensie in de panelstructuur voor een betere weergave van de heterogeniteit tussen landen.

Tabel 2.1 geeft de geschatte opbrengsten op het lange termijn BBP niveau van een verhoging van het gemiddelde aantal jaren onderwijs van de beroepsbevolking met één jaar. De geschatte sociale opbrengsten voor Nederland liggen zo tussen de 7,5% en 12,5%. Dit geschatte sociale rendement uit macro-studies lijkt het private rendement iets te overstijgen. Dat lijkt te duiden op de aanwezigheid van positieve externaliteiten, waarbij individuen profiteren van de onderwijsinspanningen van anderen. Negatieve externaliteiten, waarbij onderwijs puur dient als signalling device om individuele kwaliteiten te communiceren naar derden, lijken een ondergeschikte rol te spelen. De micro-econometrische studies naar externaliteiten zijn echter niet overtuigend.

Table 2.1 Beknopt literatuuroverzicht: sociaal rendement op onderwijs

Studie	Ondergrens	Bovengrens	Gemiddelde
Delafuente en Domenech (2006)			0,096
Cohen en Soto (2007)	0,096	0,123	0,109
Teulings en Van Rens (2008)			0,076
Coe et al. (2009)	0,097	0,115	0,106

Acemoglu en Angrist (2000) onderzoeken de rol van onderwijs externaliteiten op de individuele opbrengsten. Zo nemen zij aan dat individuen profiteren van het gemiddelde onderwijsniveau in een stad of staat. Hun schattingen bevestigen dit beeld echter niet. Moretti (2004) kiest een gelijke aanpak. Hij vindt dat een stijging van het gemiddeld aantal jaren genoten onderwijs in een stad met 1 jaar, de gemiddelde lonen in de stad doet stijgen met 8% tot 15%. De endogeniteitsproblematiek is hier echter groot. Trekken hoogopgeleiden mensen naar de stad omdat de lonen daar hoog zijn, of zijn de gemiddelde lonen hoog omdat in de stad veel hoogopgeleiden wonen met een goed betaalde baan? Voor een overzicht van deze literatuur verwijzen wij naar Sianesi en Van Reenen (2003).

Rendement op toetsscores

De laatste jaren maken gestandariseerde toetsscores hun opmars in rendementsschattingen. De leerprestaties van individuen of de beroepsbevolking worden daarmee centraal gezet. Deze maat maakt een betere differentiatie tussen landen mogelijk in internationale schattingen. Een jaar onderwijs in Nederland weegt nu eenmaal niet even zwaar als een jaar onderwijs in een ontwikkelingsland. De datareeksen zijn evenwel veelal korter of bevatten data met tijdsspannes groter dan een jaar. Hanushek en Kimko (2001) en Hanushek en Woessmann (2008) laten zien dat indien kwaliteit wordt opgenomen als verklarende variabele in de schattingen het effect van een extra jaar onderwijs insignificant wordt. Helaas schatten zij deze opbrengsten op basis van een informele groeivergelijking. Met andere woorden, een aanpassing in het niveau van de gemiddelde toetsscores heeft een permanent effect op de economische groei in een land. De effecten van toetsscores op de private opbrengsten zijn wel beschikbaar, zie bijvoorbeeld Lazear (2003); Murnane et al. (2000); Mulligan (1999); Rose (2006). Een verhoging van 1 standaarddeviatie toetsscores geeft een privaat rendement van ongeveer 10%. In overeenstemming met deze literatuur, nemen wij aan dat het sociale rendement op 1 standaarddeviatie toetsscores grofweg overeenkomt met het sociale rendement op één jaar extra onderwijs.

Heterogene rendementen

Bovenstaande opbrengsten geven steeds een gemiddeld rendement; soms op een jaar extra onderwijs, soms op toetsscores. Bij voorkeur zouden wij meer weten over het marginale rendement. Hoeveel levert een jaar extra onderwijs vanuit maatschappelijk oogpunt op en voor wie is dat het hoogst? De private opbrengsten lijken af te nemen met een toename van het aantal jaren onderwijs. Een jaar extra opleiding voor een hoog opgeleide lijkt minder effectief dan voor een relatief laag geschoolde leerling. Aan de andere kant kunnen de opportunity kosten voor een hoog opgeleide leerling weleens veel hoger zijn, waardoor het private rendement alsnog lager uitkomt. Sommige macroschattingen naar het maatschappelijk rendement houden rekening met afnemende meeropbrengsten van onderwijs. Teulings en Van Rens (2008) vinden dat landen met

een laag gemiddeld aantal jaren genoten onderwijs in termen van bbp meer profiteren van additionele onderwijsjaren dan landen met een betere uitgangspositie. Aan de andere kant laten studies zoals Vandenbussche et al. (2006) en Aghion et al. (2005) zien dat hoogopgeleide mensen belangrijk zijn voor landen die dicht bij de technologische frontier zitten, zoals Nederland. Sianesi en Van Reenen (2003) zetten de tot dan toe bekende literatuur rond heterogene rendementen op een rij. In onze analyse maken wij gebruik van de beperkte evidentie geen onderscheid tussen de rendementen van verschillende onderwijstypen.

Schooling premium puzzel

Een belangrijke vraag blijft onbeantwoord. Hoe kan het dat wanneer zowel het private als het sociale rendement zoveel hoger liggen dan de risicovrije rente, individuen niet meer investeren in onderwijs en dus langer naar school gaan? En als gevolg daarvan, waarom verplicht de overheid alle individuen niet tot langere onderwijsdeelname om de hoge sociale opbrengsten te kunnen realiseren? Deze puzzel staat in de literatuur bekend als de “Schooling premium puzzel”. Verschillende mogelijke verklaringen doen de ronde. Allereerst berusten de schattingen naar het private rendement op de aanname dat deze volgen uit de directe relatie tussen loonverschillen en aantal jaren genoten onderwijs. Deze Mincer relatie berust op een aantal stevige aannames die in de praktijk niet geheel op zullen gaan. Zo moet onder andere gelden dat

- zowel de arbeidsmarkt als de kapitaalmarkt perfect functioneren,
- de directe kosten van onderwijs verwaarloosbaar zijn,
- alle opbrengsten zijn uit te drukken in monetaire termen,
- de terugverdientijd van onderwijs oneindig is en
- er geen sprake is van externe effecten.

Alhoewel alle relevant, geeft geen van deze factoren een sterke aanwijzing om de opbrengsten substantieel te verlagen, zie Jacobs (2010). Hyperbolic discounting lijkt de beste kandidaat om de puzzel op te lossen. Individuele moeite met het inschatten van private opbrengsten over een lange tijdshorizon en dus met name op de horizon waarop onderwijsopbrengsten zichtbaar worden. Hierdoor hebben zij de neiging toekomstige opbrengsten aanzienlijk lager te waarderen dan op basis van hun korte termijn preferenties mag worden verwacht. Soms is dat zo extreem dat zij lange termijn opbrengsten in het geheel niet meenemen in hun investeringsbeslissing. Onderzoek naar de opbrengsten van onderwijs onder hyperbolic discounting is nog volop in gang.

Wat over blijft is dat gegeven alle kanttekeningen en onzekerheden, een inschatting van het private rendement op onderwijs tussen de 5% en 10% en het sociale rendement op onderwijs tussen de 7,5% en 12,5% redelijk lijkt.

3 De opbrengsten van onderwijsbeleid: de methodologie

De analyse van onderwijsbeleid valt uiteen in twee delen. In het eerste deel kwalificeren wij onderwijsmaatregelen als “kansrijk”, “niet kansrijk”, “neutraal” of “effect onbekend” op basis van empirische evidentie uit wetenschappelijke beleidsevaluaties. In het tweede deel gaan wij, indien mogelijk, over tot kwantificering van deze effecten. Deze kwantitatieve analyse combineert de beschikbare empirie uit beleidsevaluaties over de effectiviteit van verschillende beleidsinstrumenten met de macro-economische kennis over de opbrengsten en kosten van menselijk kapitaal. Paragraaf 3.1 vat eerst kort de werkwijze samen van de eerder gevolgde kwalitatieve analyse in Kansrijk Kennisbeleid ((Cornet et al., 2006)). In deze eerste stap maken wij ook gebruik van Webbink et al. (2009), die een uitgebreid overzicht geven van de relevante empirische evidentie over de effecten van kenmerken van onderwijsstelsels op onderwijsprestaties. Paragraaf 3.2 presenteert de kwantitatieve analyse en paragraaf 3.3 presenteert de uitkomsten in termen van bbp-effecten. Van Veldhuizen (2010) dient als achtergrond memorandum bij deze studie. Dit memorandum behandelt de technische details van het gebruikte instrumentarium en illustreert de werkwijze aan de hand van een tweetal voorbeelden.

3.1 Deel 1: De kwalitatieve analyse van de effecten van onderwijsbeleid

Bij het bepalen van de effectiviteit van een specifiek beleidsinstrument laten wij ons leiden door de wetenschappelijke empirie van beleidsevaluaties. Voor elk beleidsthema is die literatuur in beeld gebracht. Hierbij baseren wij ons enkel op die studies waarvan het design en de data van een zodanige kwaliteit zijn dat met voldoende zekerheid een causale relatie tussen het beleid en de uitkomst kan worden vastgesteld. Dit houdt in dat wij ons voornamelijk concentreren op studies met een (quasi-)experimenteel design. Het kader “De methodiek van beleidsevaluaties met een (quasi-)experimenteel design” licht de vormgeving van dit type studie toe. De gevonden effecten van onderwijsbeleid in dergelijke studies zijn betrouwbaarder aangezien deze niet vertekend zijn door selectie-processen in het onderwijs. Daarnaast gaat onze voorkeur uit naar studies die voor Nederland zijn uitgevoerd. Meestal is op voorhand niet duidelijk of beleidsmaatregelen in verschillende landen dezelfde effecten hebben. De beleidsomgeving kan verschillen en de leerlingen- of lerarenpopulatie heeft mogelijk andere karakteristieken. In de praktijk blijkt dat voor veel maatregelen onvoldoende goede empirische evaluaties voor Nederland beschikbaar zijn. Hierdoor zullen wij ons voornamelijk baseren op buitenlandse empirische studies.

De methodiek van beleidsevaluaties met een (quasi-)experimenteel design

Het vaststellen van causale effecten van onderwijsbeleid is niet eenvoudig omdat allerlei selectieprocessen een rol kunnen spelen. Wanneer we bijvoorbeeld het effect van klassenverkleining willen evalueren, moeten we ons ervan bewust zijn dat keuzes van ouders of scholen kunnen zorgen voor een selectieve toewijzing van leerlingen aan klassen van verschillende groottes. Hierdoor kunnen verschillen in prestaties tussen leerlingen in kleine en grote klassen ook worden veroorzaakt door andere, niet geobserveerde factoren in plaats van door het beleid.

Idealiter wordt beleid geëvalueerd d.m.v. een gecontroleerd experiment, waarbij individuen aselekt worden toegewezen aan een experimentele groep en een controlegroep. In het geval van een dergelijk experimenteel design is geen sprake van verstorende (ongeobserveerde) verschillen tussen de experimentgroep en de controlegroep. In de onderwijspraktijk is aselekte toewijzing echter niet altijd mogelijk. In een quasi-experimenteel design (of natuurlijk experiment) wordt gebruik gemaakt van een aselekte controlegroep die tot stand is gekomen door een "toevallige" situatie in de werkelijkheid. Voorbeelden van natuurlijke experimenten zijn instrumentele variabele analyses en regressie-discontinuïteitsanalyses. Bij een instrumentele variabele analyse wordt gebruik gemaakt van een zogenoemd "instrument" dat zorgt voor exogene variatie in de keuze voor of toewijzing aan de experimentele groep of controlegroep. Denk bijvoorbeeld aan een school met grote klassen die links van een brede rivier ligt en een school met kleine klassen die rechts van deze rivier ligt. Leerlingen die (toevallig) links van de rivier zijn geboren, zullen eerder naar de school met de grote klassen gaan, terwijl de leerlingen aan de andere kant van de rivier eerder naar de school met de kleine klassen zullen gaan. Deze exogene variatie in de keuze voor een grote of kleine klas wordt vervolgens gebruikt bij de evaluatie. Bij een regressie-discontinuïteit wordt gebruik gemaakt van een bepaalde afkapgrens die gehanteerd wordt in regelgeving of instituties. Denk bijvoorbeeld aan een selectietoets die bepaalt of leerlingen in een grote of kleine klas worden geplaatst. Stel dat alle leerlingen die lager dan een 6 scoren in een kleine klas worden geplaatst en alle leerlingen die hoger dan een 6 scoren in een grote klas worden geplaatst. Leerlingen die net onder de 6 scoren, zullen (omdat de school de afkapgrens "toevallig" op een 6 heeft vastgesteld) naar de kleine klas gaan, terwijl vergelijkbare leerlingen die net boven de 6 scoren naar een grote klas zullen gaan.

Een klassiek gecontroleerd experiment vormt de gouden standaard en een natuurlijk experiment de zilveren standaard. De bronzen standaard betreft studies die door middel van een specifiek onderzoeksdesign trachten om niet geobserveerde verschillen tussen individuen of groepen te elimineren, door bijvoorbeeld gebruik te maken van een gematchte controlegroep met voor en nameting (difference-in-differences methode). In het overzicht van de empirische evidentie per beleidsthema maken wij enkel gebruik van de meest geloofwaardige studies. Dit zijn de studies volgens de gouden, zilveren of bronzen standaard.

Kwalificatie van maatregelen

Op basis van de beschikbare empirie kwalificeren wij maatregelen als *kansrijk*, *niet kansrijk*, *neutraal* of *effect onbekend*. Een maatregel is kansrijk als de maatregel daadwerkelijk maatschappelijke baten oplevert en die maatschappelijke baten naar verwachting hoger zijn dan de maatschappelijke kosten verbonden aan het instrument. Een maatregel is niet kansrijk als de maatregel geen maatschappelijke baten oplevert of als de maatschappelijke baten naar verwachting kleiner zijn dan de maatschappelijke kosten verbonden aan het instrument. Een maatregel is neutraal als de maatschappelijke baten en kosten naar verwachting ongeveer aan elkaar gelijk zijn. Dit is bijvoorbeeld het geval als een maatregel budgetneutraal is en geen effect heeft op de kwaliteit van het onderwijs. Als er onvoldoende informatie beschikbaar is voor een oordeel over de efficiency van een maatregel, krijgt deze de kwalificatie effect onbekend. Een kwalificatie effect onbekend betekent dus niet dat de betreffende maatregel niet efficiënt kan zijn. Het betekent alleen dat er op basis van de beschikbare wetenschappelijke evidentie geen uitsluitsel kan worden gegeven over de efficiency van de maatregel.

Voorbeelden van kansrijke beleidsinstrumenten zijn scholing van leraren, prestatiebeloning, voor- en vroegschoolse educatie, intensieve coaching programma's voor risicoleerlingen, prestatiebekostiging, transparantie en toegankelijkheid van informatie over schoolkwaliteit, inspectietoezicht en centrale examens. Voorbeelden van niet kansrijke maatregelen zijn klassenverkleining, meer ict in het onderwijs, en het bevorderen van instroom in bèta studies. De introductie van het sociaal leenstelsel is een voorbeeld van een maatregel met kwalificatie *neutraal*. Deze maatregel zorgt voor een financieringsschuif tussen publieke en private middelen en heeft geen effect op de kwaliteit van het onderwijs.

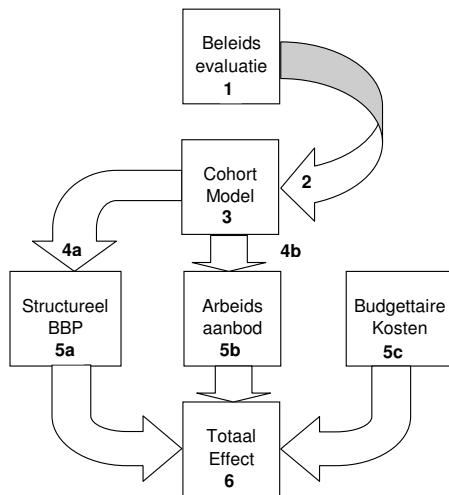
3.2 Deel 2: De kwantitatieve analyse van de effecten van onderwijsbeleid

Maatregelen met een kwalificatie *kansrijk* of *niet-kansrijk* zullen we vervolgens kwantificeren in termen van lange termijn effecten op het bbp, zie figuur 3.1. Hiertoe worden de uitkomsten van de effectiviteitsanalyse op microeconomisch niveau vertaald naar een macroeconomisch effect met behulp van een cohortmodel en een macro rendement van onderwijs. In dit cohortmodel worden individuen gevolgd vanaf hun instroom als leerling in het primair onderwijs tot aan hun pensioen. In het model volgen wij de leerlingen, die uiteindelijk met hogere toetsscores of meer jaren onderwijs in de beroepsbevolking instromen. De veranderende samenstelling van de totale beroepsbevolking resulteert in een verandering van het verwachte lange termijn gemiddelde opleidingsniveau van de beroepsbevolking en dus in een afwijkend lange termijn bbp-niveau, omdat hoger opgeleide werknemers productiever zijn. Ook brengen we de gevolgen voor het arbeidsaanbod en de kosten van een maatregel in beeld. Deze paragraaf beschrijft de verschillende stappen uit de kwantitatieve analyse, zie ook figuur 3.1.

Ter illustratie sluiten we af met een presentatie van de kwantificering van een fictief pakket aan

maatregelen.

Figure 3.1 Weergave stappenplan kwantitatieve analyse onderwijsbeleid



1. In kaart brengen van toepasbare empirische evidentie rond effectiviteit van kennismaatregelen
2. Vertalen van dit effect in een veranderende samenstelling van de leerlingenpopulatie
3. Cohortmodel van leerling tot verplichte uittreding uit de beroepsbevolking
4. Vertaling naar verandering van de beroepsbevolking in termen van
 - a het gemiddeld aantal jaren opleidingsniveau of de gemiddelde leerprestaties
 - b het percentage laag-, midden- en hoogopgeleiden
5. In kaart brengen van de opbrengsten en kosten van een veranderende beroepsbevolking in termen van
 - a een veranderend structureel bbp niveau
 - b een veranderend arbeidsaanbod
 - c de gemaakte budgettaire en directe kosten
6. Totaal effect

3.2.1 Micro-evidentie revisited: onzekerheid

De eerste stap betreft het bundelen van de uitkomsten uit de relevante empirische literatuur tot een effect van een maatregel op toetsscores, aantal jaren genoten onderwijs of slagingskansen. De waardering van de uitkomsten uit de beleidsevaluaties vindt plaats aan de hand van een tweetal validiteitsconcepten. Bij interne validiteit gaat het om de mate waarin de verschillende beleidsevaluaties binnen een beleidsthema een consistent beeld laten zien van het effect van een gelijksoortige beleidsaanpassing. Bij externe validiteit gaat het om de toepasbaarheid van deze studies voor Nederlands beleid en in de Nederlandse beleidscontext. (Heckman en Vytlacil,

2007) zetten de betekenis van beide validiteitsconcepten in een beleidsomgeving op een rij. De onzekerheid omtrent de specifieke waardering van beleidsevaluaties uit de micro-literatuur en de toepasbaarheid van deze uitkomsten voor Nederland wordt in beeld gebracht door één of meerdere onzekerheidsvarianten.

De eerste onzekerheidsvariant betreft de interne validiteit en vormt een afspiegeling van de relevante literatuur. Deze variant presenteert een basiseffect in termen van een toename in toetsscores, het (nominale) aantal jaren gevolgd onderwijs of slagingskansen. De effecten op toetsscores worden hierbij uitgedrukt in standaarddeviaties, om de effecten op verschillende toetsen vergelijkbaar te maken. Het basiseffect kan een gemiddelde zijn van alle studies met een geloofwaardig design, als de studies eenzelfde beeld laten zien. Het basiseffect kan ook het onderzoeksresultaat zijn van een geprefereerde studie, zoals bijvoorbeeld de studie met het beste design. Als de studies tegengestelde resultaten laten zien, is de interne validiteit niet sterk genoeg om een uitspraak te kunnen doen omtrent de effectiviteit van de beleidsmaatregel.

In de overige variant(en) wordt de externe validiteit beschouwd door de beleidscontext in Nederland te vergelijken met die in landen van de betreffende studies. Als bijvoorbeeld op het betreffende thema al veel beleid wordt gevoerd in Nederland, kan dat een reden zijn om te veronderstellen dat gevonden effecten minder sterk zullen zijn vanwege afnemende meeropbrengsten van beleid. Dit komt dan tot uitdrukking in een neerwaartse bijstelling van het basiseffect uit de literatuur.

Op deze manier wordt per beleidsthema een aantal onzekerheidsvarianten gepresenteerd omtrent de effectiviteit van maatregelen op dit terrein. In alle varianten nemen wij aan dat het instrument op effectieve wijze en zonder weeffouten is geïmplementeerd. De genoemde effecten zijn daarnaast partieel. Met andere woorden, wij nemen aan dat de rest van de beleidsomgeving niet verandert. Bovendien gelden de gevonden effecten alleen bij kleine beleidsaanpassingen, dus het gaat om marginale effecten.

3.2.2 Doorwerking van micro-effecten op langere termijn

De empirie van beleidsevaluaties richt zich vaak op de directe effecten van beleidsinterventies, zoals de effecten op toetsscores na 1 of 2 jaar. Zo wordt bijvoorbeeld onderzocht wat de effecten zijn van scholing van leraren op de toetsscores na 1 jaar, maar niet wat de effecten zijn na 4 jaar. Aangezien wij geïnteresseerd zijn in structurele effecten en het aantal studies met een langere tijdshorizon zeer beperkt is, zullen we voor een aantal maatregelen aannames moeten maken over de doorwerking van beleid op de langere termijn. Hierbij maken we een onderscheid tussen maatregelen die aangrijpen op onderwijsinputs (zoals scholing van leraren, kleinere klassen, etc.) en institutionele maatregelen (zoals prestatiebekostiging).

Voor wat betreft de maatregelen die aangrijpen op de onderwijsinputs gaan we uit van afnemende meeropbrengsten van beleid. Hierbij volgen we Krueger (1999), die de effecten van klassenverkleining op onderwijsprestaties over meerdere jaren volgt. Hij vindt dat het

additionele effect van een kleinere klas in het tweede jaar de helft is van het initiële eerstejaars-effect, en dat de effecten in het derde en vierde jaar ieder nog een kwart van het initiële eerstejaars-effect bedragen.

Voor wat betreft de institutionele maatregelen nemen we aan dat deze leiden tot een eenmalige niveau-sprong zonder additionele effecten in volgende jaren. Een institutionele aanpassing in het onderwijsstelsel resulteert op deze manier in een eenmalig, blijvend effect. Deze aanname wordt ondersteund door Chorny en Webbink (2010), die de effecten van accountability beleid in Amsterdam op onderwijsprestaties over meerdere jaren gevolgd hebben.

Daarnaast speelt de vraag in hoeverre beleid gericht op bijvoorbeeld een eenjarige interventie voor de betreffende onderwijsinput een langdurig effect kan hebben. Een voorbeeld: Stel een leraar volgt een cursus in jaar t . De effecten op de prestaties worden afgemeten aan de leerlingen die dat jaar bij hem onderwijs hebben genoten. De jaren daarop zal de leraar, misschien in afnemende mate, echter ook nog profijt hebben van zijn verworven kennis. Bij scholing van leraren nemen wij aan dat zijn leerlingen na 10 jaar geen effect meer zullen ondervinden van de cursus.

3.2.3 Cohortmodel

De uitkomsten van de effectiviteitsanalyse op microeconomisch niveau worden vertaald naar een macroeconomisch effect met behulp van een gestileerd cohortmodel. In dit cohortmodel worden individuen gevolgd van instroom als leerling in het onderwijssysteem tot aan hun pensioen.

Leerling-populatie

De huidige en toekomstige leerlingpopulaties worden bepaald door de huidige inrichting van het Nederlandse onderwijssysteem. Hierbij onderscheidt het model de volgende onderwijsniveaus:

- primair onderwijs (po), bedoeld voor kinderen van 4 tot 12 jaar.
- voorbereidend middelbaar beroepsonderwijs (vmbo). Dit heeft een nominale duur van vier jaar (leeftijd: 12 -16 jaar).
- hoger algemeen voortgezet onderwijs (havo). De nominale duur is vijf jaar (leeftijd: 12 -17 jaar)
- voorbereidend wetenschappelijk onderwijs (vwo). De nominale duur van deze onderwijsvorm is 6 jaar (leeftijd: 12-18 jaar).
- middelbaar beroepsonderwijs (mbo). De nominale duur is vier jaar.
- hoger beroepsonderwijs (hbo). Hbo-opleidingen duren meestal 4 jaar.
- wetenschappelijk onderwijs (wo). De nominale studieduur is meestal 4 jaar.

Dit model laat het speciaal onderwijs buiten beschouwing.

Het cohortmodel kent een aantal belangrijke aannames. Zo wordt in het model een constante bevolkingsgrootte verondersteld. Anders geformuleerd, de instroom in het basisonderwijs is constant verondersteld. Leerlingen ronden het onderwijs af in de nominale studieduur. Er zijn

dus geen zittenblijvers. Aan het eind van het scholingstraject stroomt een deel van de leerlingen de arbeidsmarkt op. De overige leerlingen zullen zich niet meer aanbieden. Een individu dat toetreedt tot de arbeidsmarkt zal blijven werken totdat hij de pensioengerechtigde leeftijd van 65 jaar heeft bereikt. Als een werknemer een bepaalde kwalificatie heeft behaald, dan behoudt hij deze zijn gehele werkzame leven.

Op basis van de tabel “Doorstroom naar herkomst 2007”, zie CBS (2009), zijn de doorstroomkansen tussen de verschillende onderwijsniveaus, en tussen onderwijsniveaus en toetreding tot de arbeidsmarkt bepaald. De doorstroomkans van niveau i naar niveau j is gedefinieerd als het aantal leerlingen dat opleidingsniveau j heeft voltooid gedeeld door het aantal leerlingen dat een diploma op het lagere niveau i heeft behaald. De doorstroomkans van niveau i naar de arbeidsmarkt is de kans dat een leerling geen bekostigd onderwijs meer gaat volgen nadat hij/zij opleidingsniveau i heeft afgerond. De doorstroomkansen zijn te vinden in tabel 3.2 van Van Veldhuizen (2010). We nemen aan dat deze doorstroom- en participatiekansen in steady state zijn. Met andere woorden, zij blijven zonder beleidsinterventies constant over de tijd.

Aangezien wij geïnteresseerd zijn in veranderingen van de beroepsbevolking ten opzichte van het basispad is het beperkte beschrijvende karakter van de beroepsbevolking minder relevant dan voor andere analyses.

Beroepsbevolking

De beroepsbevolking wordt gedifferentieerd naar leeftijd en hoogst genoten opleiding. Afhankelijk van het opleidingsniveau stromen werknemers op een bepaalde leeftijd de beroepsbevolking in. Zo stromen werknemers die het vmbo niet hebben afgemaakt, op hun zestiende in op de arbeidsmarkt (onder de aanname dat ze participeren), terwijl wo-geschoolde werknemers op zijn vroegst op hun 22-ste de beroepsbevolking instromen. Vervolgens werkt iedere werknemer die participeert op de arbeidsmarkt tot de pensioengerechtigde leeftijd van 65 jaar. Aangezien de instroom aan nieuwe leerlingen constant is, definieert de uitstroom vanuit elk opleidingsniveau de bijdrage van het desbetreffende opleidingsniveau aan het gemiddeld aantal jaren scholing van de beroepsbevolking.

Interventie in het onderwijs

Vanuit het oogpunt van het cohortmodel geeft een interventie in het onderwijs een verandering in de doorstroomkansen van het desbetreffende niveau waarop de interventie betrekking heeft. Zo zorgt een kansrijke maatregel in het primair onderwijs er bijvoorbeeld voor dat meer leerlingen naar het havo/vwo in plaats van het vmbo doorstromen. Uiteindelijk stromen dan minder leerlingen met een vmbo-kwalificatie de arbeidsmarkt op. Afhankelijk van de grootte van het effect gevonden in de micro-literatuur, worden de desbetreffende doorstroomkansen tussen de onderwijsniveaus gecalibreerd. Een verandering in de doorstroomkansen leidt tot een

verandering in de leerlingpopulatie ten opzichte van de populatie in het basispad. Dit leidt uiteindelijk tot een andere samenstelling van de beroepsbevolking ten opzicht van het basispad. Het directe gevolg is dat na verloop van tijd het gemiddeld aantal jaren genoten onderwijs van de beroepsbevolking verandert ten opzichte van het gemiddelde in het basispad.

3.3 BBP-effecten

Een hoger opgeleide beroepsbevolking heeft een hogere productiviteit. De micro-uitkomsten over de effectiviteit van specifieke beleidsinstrumenten en de macro-economie worden aan elkaar gelinkt via de bbp-opbrengsten gevonden in groeivergelijkingen met menselijk kapitaal als determinant van productiviteitsgroei. Schattingen van het sociaal rendement van onderwijs wijzen uit dat één jaar nominaal onderwijs een rendement heeft van 10% (zie hoofdstuk 2). Dit wil zeggen dat een toename van het gemiddeld aantal (nominale) jaren gevolgde onderwijs in de beroepsbevolking met 1 jaar leidt tot een verhoging van het bbp met 10%.

Het sociaal rendement op één standaarddeviatie hogere toetsscores is ook ongeveer 10% (Lazear, 2003; Murnane et al., 2000; Mulligan, 1999; Rose, 2006). Een stijging in toetsscores betekent een kwaliteitsverhoging, welke deels zal resulteren in een hogere deelname aan vervolgonderwijs. Leerlingen worden slimmer en blijven daardoor langer in het onderwijs. Murnane et al. (2000) vinden dat ongeveer de helft van de stijging in de arbeidproductiviteit ten gevolge van hogere toetsscores verklaard kan worden doordat leerlingen met hogere toetsscores doorstromen naar hogere vervolgniveaus (het *kwantiteitseffect*). De overige helft wordt verklaard doordat leerlingen productiever worden, omdat zij meer hebben geleerd (het *kwaliteitseffect*). Deze 50-50 verhouding tussen het *kwaliteits*- en het *kwantiteitseffect* houden we ook aan bij het bepalen van de effecten op het arbeidsaanbod of het bepalen van de kosten ten gevolge van extra onderwijs (zie paragraaf 3.3.1 en 3.3.2).

Op deze manier wijzen we een gemiddeld rendement van een extra nominaal jaar genoten onderwijs of een standaarddeviatie hogere toetsscore op het bbp niveau toe. Middels deze procedure wordt het effect van de interventie in het onderwijs uitgedrukt in een effect op het structurele bbp. We merken hierbij op dat we maatregelen beschouwen als marginale aanpassingen op het bestaande beleid en dat er geen algemeen evenwichtseffecten in de analyse worden meegenomen. De aanpassingsperiode voor het onderwijssysteem en de beroepsbevolking na een interventie is maximaal 62,5 jaar.² Na 62,5 jaar zullen geen veranderingen in de leerlingpopulatie noch in de beroepsbevolking worden waargenomen als gevolg van de interventie. Het model is dan wederom in steady state. Het kwantitatieve effect van een beleidsmaatregel wordt nu gegeven door het verschil van het structurele bbp-niveau in

² Een maatregel grijpt op zijn vroegst aan op kinderen in de leeftijd van 2,5 jaar (voorschoolse educatie). Deze leerlingen stromen op 65-jarige leeftijd uit de beroepsbevolking. Wij veronderstellen dat een nieuw ingevoerde maatregel structureel wordt doorgezet.

de nieuwe steady state ten opzichte van het lange termijn bbp-niveau uit het basispad.

3.3.1 Arbeidsaanbod

Naast de zojuist beschreven productiviteitsverbetering heeft onderwijsbeleid ook via een verandering in het arbeidsaanbod effect op het structurele bbp. De mate waarin mensen zich aanbieden op de arbeidsmarkt verschilt per opleidingsniveau. Zo blijken hoger opgeleiden meer te participeren. Een eenvoudige analyse levert het volgende. We volgen de definities van het CBS waarin laag opgeleiden zijn gedefinieerd als werknemers met een po of vmbo diploma, middelbaar opgeleiden als werknemers met een havo/vwo of mbo diploma en hoog opgeleiden als werknemers met een hbo of wo diploma. Uit CBS data van 2009 volgt dat hoger opgeleiden meer participeren dan laag opgeleiden:³

- de kans dat laagopgeleiden participeren is 52%,
- de kans dat middelbaar opgeleiden participeren is 75%,
- de kans dat hoogopgeleiden participeren is 86%.

Deze kansen zijn redelijk constant over de tijd, zie bijvoorbeeld Euwals et al. (2007) en Huizinga en Smid (2004). Het laatste decennium zijn de participatiekansen nauwelijks toegenomen. De uittreedleeftijd is ook over de drie opleidingsniveaus redelijk gelijk.⁴ We nemen aan dat de uittreedleeftijd 65 jaar bedraagt.

Als een interventie beoogt de beroepsbevolking hoger op te leiden, heeft dit twee gevolgen voor het arbeidsaanbod. Hoger opgeleiden gaan langer naar school, waardoor het potentieel aantal werkzame jaren tot aan het pensioen lager ligt dan voor laag opgeleiden. Op de lange termijn staat daartegenover dat hoger opgeleiden meer participeren. Op de korte termijn resulteert eerst een afname van het arbeidsaanbod, waarna op de lange termijn het arbeidsaanbod toeneemt.

Het arbeidsaanbodeffect wordt als volgt berekend. Ten opzichte van het basispad wordt na een interventie gecorrigeerd voor het verschil in aantal individuen dat zich aanbiedt op de arbeidsmarkt. Vervolgens vermenigvuldigen wij dit verschil met het bbp per werknemer uit het basisjaar om het arbeidsaanbodeffect uit te drukken in het percentage bbp van het basisjaar. We benadrukken dat bij deze berekening impliciet wordt aangenomen dat de productietechnologie constant blijft en de waarde van het basisjaar behoudt. Ook de arbeidsparticipatie per opleidingstype blijft onveranderd. Lange termijn evenwichtskoncepten zijn niet geïmplementeerd. Met andere woorden, de arbeidsmarkt ruimt. De effecten van het arbeidsaanbod op het bbp zijn additioneel aan de productiviteitseffecten op het bbp, gebaseerd op de rendementen uit de macro-literatuur.

³ CBS Statline: periode 2001-2009, gegevens beroepsbevolking naar afkomst, leeftijd en volttooide opleiding.

⁴ Zie ook CBS Statline.

3.3.2 Kosten

Een interventie in het onderwijs kent in ons model twee soorten kosten: budgettaire- en onderwijskosten. Onder de budgettaire kosten worden de directe financiële kosten van de maatregel verstaan. Zo zijn voor het verkleinen van de klassengrootte in het basisonderwijs bijvoorbeeld extra leraren en materiële uitgaven nodig. In het geval dat een interventie leidt tot meer scholing, worden ook indirecte kosten gemaakt. Meer mensen gaan langer naar school. Deze secundaire kosten noemen wij onderwijskosten.

3.3.3 Het totaal beeld

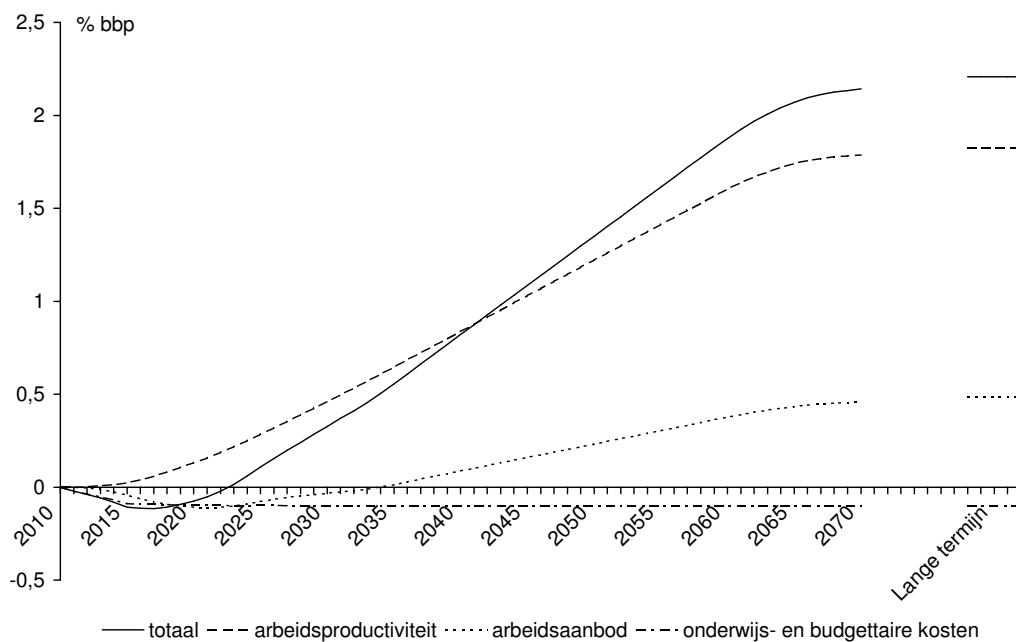
De kosten, het arbeidsaanbod- en arbeidsproductiviteitseffect bepalen uiteindelijk het kwantitatieve effect. In de huidige analyse worden alle posten uitgedrukt als percentage van het huidige BBP en vervolgens gesommeerd tot een totaal beeld. Wij merken uitdrukkelijk op dat wij geen kwantitatieve welvaartsanalyse toepassen. Ten eerste blijven een deel van de maatschappelijke kosten en baten van een maatregel buiten beeld. Voor bepaalde maatregelen kunnen die effecten aanzienlijk zijn. Zo kan voor- en vroegschoolse educatie bijvoorbeeld een positief welvaartseffect genereren door reductie van de criminaliteit van leerlingen uit de doelgroep. Ten tweede vraagt een welvaartsanalyse om een weging van de verschillende effecten over de tijd. Aangezien wij geen evidentie hebben over de discontovoeten van genoemde posten en zij mogelijk onderling variëren, drukken wij het totale effect niet uit in termen van een netto contante waarde.

3.3.4 Illustratie kwantificering

Figuur 3.2 laat de bbp-effecten zien van een fictieve kansrijke maatregel. In eerste instantie zijn de kosten van de maatregelen groter dan hun (financiële) opbrengsten. Deze kosten zijn deels de directe budgettaire kosten van de maatregel. Daarnaast gaan mensen als gevolg van de maatregelen langer naar school, wat leidt tot extra onderwijskosten en een daling van het arbeidsaanbod. Na verloop van tijd neemt het arbeidsaanbod weer toe, omdat hoger opgeleide mensen meer participeren op de arbeidsmarkt. Per saldo resulteert een positief arbeidsaanbodeffect op de lange termijn. Naast dit positieve arbeidsaanbodeffect komt het grootste deel van de opbrengsten voort uit een stijging van de arbeidsproductiviteit ten gevolge van de maatregelen. Het volledige effect is pas na lange tijd bereikt. Op dat moment is de gehele beroepsbevolking door de maatregelen beter geschoold of gelijk geschoold gebleven. Het netto-effect op het bbp is de som van de hogere arbeidsproductiviteit en het hogere arbeidsaanbod, minus de budgettaire- en onderwijskosten.

Er kunnen grote verschillen bestaan tussen verschillende maatregelen wat betreft budgettaire kosten, bbp-effecten en tijdsduur voordat de opbrengsten van de maatregel gerealiseerd zijn. Klassenverkleining is bijvoorbeeld een relatief dure maatregel, terwijl institutionele maatregelen zoals prestatiebekostiging (nauwelijks) extra middelen vereisen. De belangrijkste verklaringen

Figure 3.2 Bbp-effect ten opzichte van het basispad, directe kosten en arbeidsaanbodeffect in % BBP.



liggen in afwijkende doelgroepbereiken en onderlinge verschillen in effectiviteit. Zo kan een maatregel als prestatiebeloning de volledige leerlingpopulatie bereiken, terwijl een maatregel gericht op de aanpak van voortijdig schoolverlaten zich specifiek richt op leerlingen in het vmbo en mbo. En zo kent voor- en vroegschoolse educatie een sterkere invloed op het opleidingsniveau van leerlingen dan een klassenverkleining.

De tijdsduur waarin alle opbrengsten van een maatregel gerealiseerd zijn, hangt samen met de doelgroep van de maatregel. Specifieke inzet in het hoger onderwijs zal eerder opbrengsten genereren, aangezien deze mensen op relatief korte termijn de arbeidsmarkt zullen betreden. Middelen voor voor- en vroegschoolse educatie worden juist ingezet op jonge leerlingen, waardoor het langer duurt voordat de baten van deze maatregel gerealiseerd worden.

4 Opbrengsten van onderwijsbeleid

De allocatie van middelen over beleidsrelevante onderwijsthema's vraagt om inzicht in de maatschappelijke opbrengsten en kosten van verschillende beleidsinstrumenten. Op welk beleidsinstrument kan de overheid het beste inzetten als zij de welvaart wil verhogen? Dit hoofdstuk brengt een deel van de opbrengsten en kosten van onderwijsbeleid in kaart.

Wij onderscheiden 22 beleidsthema's, die centraal staan in de huidige beleidsdiscussie. De thema's zijn verdeeld in twee categorieën. De eerste categorie bestaat uit maatregelen die zich richten op onderwijsinputs, zoals bijvoorbeeld de inzet van meer onderwijssuren, uitbreiding van de voor- en vroegschoolse educatie of scholing voor leraren. Maatregelen in deze categorie vergen veelal een uitbreiding van de huidige financiële middelen. De tweede categorie bestaat uit maatregelen die zich richten op aanpassing van de kenmerken van het huidige onderwijsstelsel, zoals bijvoorbeeld de verplichte deelname van alle leerlingen in het primair onderwijs aan het centraal examen of gedeeltelijke prestatiebekostiging in het primair onderwijs. Maatregelen in deze categorie vergen veelal slechts een geringe inzet van middelen, omdat zij samen gaan met een herinrichting van de prikkels in het huidige systeem.

Binnen een beleidsthema zijn verschillende beleidsinstrumenten voor handen. Zo kan binnen het beleidsthema voor- en vroegschoolse educatie worden gedacht aan de inzet van meer middelen om een grotere doelgroep te bereiken of kan worden ingezet op het aantrekken van hoger geschoolde leidsters. Tabel 4.1 presenteert een overzicht van de beleidsthema's die in dit rapport zijn beschouwd. Elk beleidsthema is op basis van beschikbare empirie gekwalificeerd als kansrijk, niet-kansrijk, neutraal of effect onbekend. Daarnaast staat aangegeven of de maatregel budgettaire gevolgen heeft of een zuiver institutionele doorwerking kent. Wij beschouwen beleidsopties op het niveau van deze thema's en zullen dus in principe geen onderscheid maken tussen verschillende beleidsinstrumenten binnen eenzelfde thema.⁵

Voor beleidsinstrumenten met de kwalificatie kansrijk of niet-kansrijk biedt de internationale literatuur aanknopingspunten om een deel van de opbrengsten en kosten te kwantificeren. Dit hoofdstuk illustreert deze werkwijze per beleidsthema aan de hand van een reeks maatregelen. Hierbij is aangenomen dat de maatregelen plaatsvinden in de huidige beleidsomgeving en een relatief beperkte aanpassing in het huidige beleidspakket behelzen. Per beleidsthema brengen wij een vijftal uitkomsten in beeld:

- het effect op onderwijsprestaties: toetsscores, aantal jaren onderwijs of slagingskansen
- het resulterende arbeidsproductiviteitseffect
- de budgettaire kosten
- de onderwijskosten

⁵ Alleen in het geval er voor meerdere beleidsinstrumenten binnen een thema voldoende empirie beschikbaar is om deze afzonderlijk te beschouwen, maken wij hier een uitzondering op. Een voorbeeld hiervan zijn de beleidsinstrumenten 'intensieve begeleiding' en 'financiële prikkels' binnen het thema 'beperken schooluitval'.

Table 4.1 Overzicht geanalyseerde beleidsthema's met hun kwalificatie

Thema's	Kwalificatie	Budgettaire(B)/Institutioneel(I)
<i>Onderwijsinputs</i>		
Voor- en vroegschoolse educatie	kansrijk	B
Verlenging onderwijstijd	kansrijk	B
Klassenvergroting	kansrijk	B
Beperken schooluitval	kansrijk	B
Uitbreiden inspectietoezicht	kansrijk	B
Scholing en begeleiding onderwijspersoneel	kansrijk	B
Mentoring van startende leraren	kansrijk	B
<i>Stelselkenmerken</i>		
Centrale examinering	kansrijk	I
Informatievoorziening onderwijsprestaties	kansrijk	I
Prestatiebekostiging in het onderwijs	kansrijk	I
Latere selectie	kansrijk	B/I
Prestatiebeloning leraren	kansrijk	B
Sociaal leenstelsel	neutraal	B
<i>Niet-kwantificeerbare thema's</i>		
Verhogen salaris onderwijspersoneel	effect onbekend	B
Verhogen salaris schoolleiders	effect onbekend	B
Meer middelen naar speciaal onderwijs	effect onbekend	B
Bevorderen excellentie in het hoger onderwijs	effect onbekend	B
School en omgeving	effect onbekend	B
Meer ICT in het onderwijs	effect onbekend	B
Leven lang leren	effect onbekend	B
Schakelklassen en summer courses	effect onbekend	B
Terugdringen segregatie in het onderwijs	effect onbekend	B/I
Concurrentie en vrije schoolkeuze	effect onbekend	I

- aanpassingen in het arbeidsaanbod.

Op basis van de kwantitatieve analyses kunnen we de budgettaire kansrijke maatregelen classificeren wat betreft hun (kosten)effectiviteit. Gegeven de verwachte opbrengsten en kosten op de lange termijn, is een beperkte toename in de beschikbare middelen voor prestatiebeloning van leraren effectiever dan bijvoorbeeld de inzet van een gelijk bedrag op mentoring van startende docenten, beperken van schooluitval, inspectietoezicht of scholing van leraren. Voor-en vroegschoolse educatie, verlenging van de onderwijstijd en klassenvergroting zijn het minst kosteneffectief. Deze maatregelen vergen relatief veel middelen om een positief effect te sorteren. Zuiver institutionele maatregelen als prestatiebekostiging en het verplicht stellen van het centraal examen in het po nemen binnen dit overzicht een speciale plaats in. Invoering van deze maatregelen brengt geen kosten met zich mee, waardoor hun kosten-baten verhouding

oneindig is. Onder de randvoorwaarde van een juiste vormgeving is invoering van deze beleidsinstrumenten te allen tijde effectief.

De komende paragrafen lichten deze uitkomsten toe. Paragraaf 4.1 schetst allereerst de contouren van de huidige beleidsomgeving, waarbinnen de maatregelen effect zullen sorteren. Paragraaf 4.2 en paragraaf 4.3 presenteren vervolgens de kwantitatieve effecten van respectievelijk maatregelen die inzetten op onderwijsinputs en maatregelen die zich richten op institutionele kenmerken van het onderwijsstelsel. Paragraaf 4.4 sluit af met een toelichting op de beleidsthema's waarvoor de empirie ontoereikend bleek om te komen tot kwantificering van haar effecten.

4.1 Nederland in beeld

Deze paragraaf geeft een overzicht van de belangrijkste kengetallen in het Nederlandse onderwijssysteem. Hoeveel middelen geeft de overheid uit aan onderwijs? Hoeveel leerlingen nemen deel aan het reguliere onderwijs? En welke kosten worden per leerling gemaakt?

Volgens de OESO definitie van overheidsuitgaven aan onderwijs bedraagt de publieke financiering van het onderwijs in Nederland circa 5% van het bbp.⁶ De uitgaven aan het primair onderwijs bedragen circa 1,75% van het bbp, de uitgaven aan secundair onderwijs circa 2,20% van het bbp en de uitgaven aan tertiair onderwijs circa 1,45% van het bbp.⁷ Figuur 4.1 geeft een overzicht van de ontwikkeling van de Nederlandse publieke uitgaven aan onderwijs als percentage van het bbp over de tijd.

In het schooljaar 2008/2009 namen circa 3,8 miljoen leerlingen deel aan het onderwijs gefinancierd door de ministeries van OC&W en LNV.⁸ Tabel 4.2 presenteert de leerlingaantallen per sector. Gegeven de publieke uitgaven per sector en de verdeling van de leerlingen over de verschillende sectoren vat tabel 4.3 de gemiddelde onderwijsuitgaven per leerling samen. Een uitgebreid overzicht van de kwantitatieve gegevens rond het Nederlandse onderwijsbeleid is te

⁶ De overheidsuitgaven aan regulier onderwijs bestaan volgens de OESO bepaling uit de overheidsuitgaven aan onderwijsinstellingen en de overheidsuitgaven aan huishoudens, bedrijven en non-profit instellingen. De studietoelagen die aan huishoudens worden verstrekt, worden in de overheidsuitgaven meegenomen. De overheidsontvangsten worden niet meegerekend. De overheidsuitgaven aan onderwijsinstellingen zijn de uitgaven van de Rijksoverheid, provincies en gemeenten aan onderwijsinstellingen. Het gaat dan om de lumpsum financiering, de apparaatskosten van het Ministerie van OCW en uitgaven voor contractonderzoek. De overheid geeft huishoudens subsidie voor onderwijs in de vorm van de tegemoetkoming in de schoolkosten en studiefinanciering. Hiermee worden leerlingen/studenten of ouders (gedeeltelijk) gecompenseerd voor de uitgaven aan les- en collegegelden, boeken en leermiddelen en de kosten van openbaar vervoer en levensonderhoud. De overheid geeft ook subsidies en fiscale regelingen aan bedrijven. Dit zijn subsidies voor het verzorgen van leerlingenvervoer en subsidies en fiscale regelingen om bedrijven te stimuleren om leerwerkplekken en stageplaatsen beschikbaar te stellen. Met de belastingkorting of subsidie krijgen bedrijven een deel van hun kosten voor het begeleiden van duale leerlingen en stagiairs vergoed.

⁷ secundair onderwijs: voortgezet- en middelbaar beroepsonderwijs
tertiair onderwijs: hoger beroeps- en wetenschappelijk onderwijs

⁸ Bron: CBS Statline, table 'Historie onderwijs; leerlingen en studenten, vanaf 1950'

Figure 4.1 Totale publieke uitgaven onderwijs in percentage van bbp volgens de OESO definitie. Bron: CBS Statline

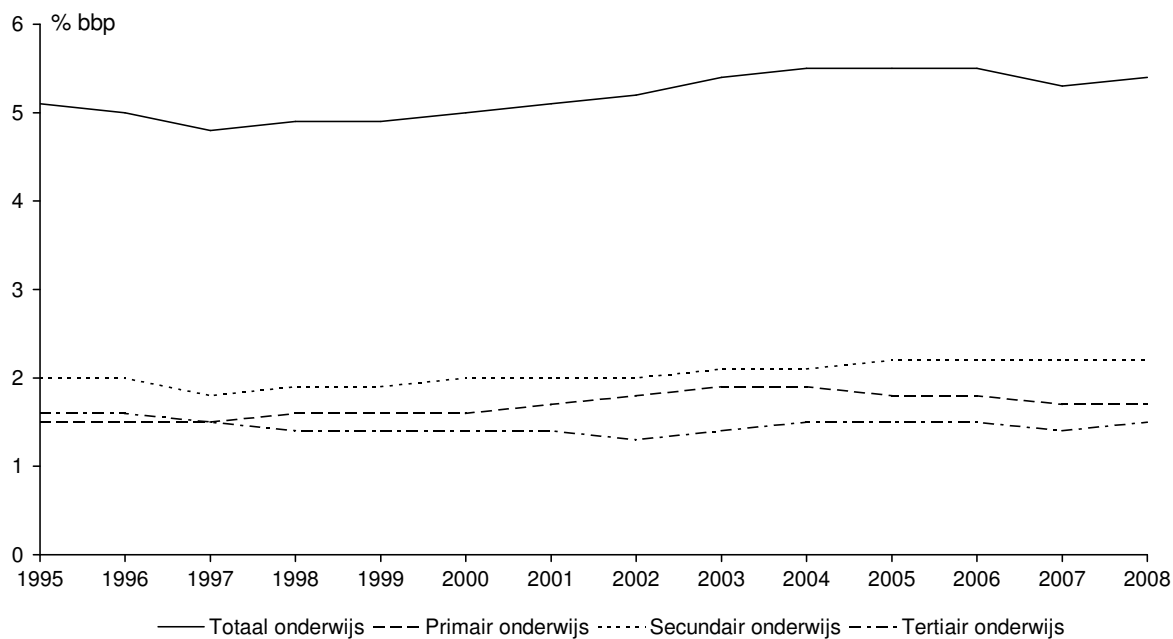


Table 4.2 Leerlingaantallen

	leerlingaantallen (in duizendtallen)
primair onderwijs	1 553
voorbereidend middelbaar beroepsonderwijs	445
havo/vwo	491
middelbaar beroepsonderwijs	484
hoger beroepsonderwijs	383
wetenschappelijk onderwijs	219

Table 4.3 Onderwijsuitgaven per deelnemer per jaar (in lopende prijzen), euro's

	ocw uitgaven	ocw uitgaven inclusief gemeentelijke bijdragen
primair onderwijs	4600	6600
voortgezet onderwijs	7100	7800
middelbaar beroepsonderwijs	6400	6700
hoger beroepsonderwijs	5800	-
wetenschappelijk onderwijs	5900	-

vinden in OC&W (2008).

4.2 Kwantificering beleidsthema's op onderwijsinputs

4.2.1 Voor- en vroegschoolse educatie

Beleidsintensivering gericht op voor- en vroegschoolse educatie (vve) heeft tot doel de vroege ontwikkeling van kinderen te bevorderen, hun mogelijkheden voor vervolgonderwijs te vergroten, en daarmee uiteindelijk hun latere maatschappelijke positie en kansen op de arbeidsmarkt te versterken. Een lage of niet afgemaakte opleiding vergroot de kans op inactiviteit, maatschappelijke uitsluiting en zelfs op crimineel gedrag en kan invloed hebben op de latere gezondheid. De doelgroep bestaat uit kinderen met een achterstand in de leeftijdsgroep van 2,5 tot 6 jaar. De Werkgroep Brede Heroverwegingen rapporteert dat de doelgroep voor voorschoolse educatie bestaat uit zo'n 40 duizend leerlingen, welke kan worden uitgebreid naar 60 duizend.⁹ Het huidige doelgroepbereik in de voorschoolse educatie is 80%. De doelgroep in de vroegschoolse educatie bestaat uit ongeveer 47 duizend leerlingen. Van deze leerlingen wordt nu 57% bereikt. Voor de kabinetsperiode van Balkenende IV was de beleidsopgave dat 70% van de doelgroepkinderen moest worden bereikt met minimaal drie dagdelen vve per week. Per augustus 2010 is in het wetsvoorstel *Ontwikkelingskansen door Kwaliteit en Educatie* vastgelegd dat vve naar minimaal vier dagdelen per week wordt uitgebreid. Aan de huidige vve-regeling wordt circa 250 miljoen euro besteed.

Nieuwe beleidsvoorstellen richten zich met name op vergroting van het bereik of uitbreiding van de doelgroep. Deelname aan vve is in de huidige vormgeving niet verplicht. Daarnaast richten nieuwe voorstellen zich op de vormgeving van het programma, bijvoorbeeld door de uitbreiding naar meer dagdelen of een kwaliteitsverbetering, bijvoorbeeld door te investeren in hoogopgeleide leid(st)ers.

Empirie

Cornet et al. (2006) concluderen dat beleidsintensivering in voor- en vroegschoolse educatie kansrijk is. De hoge potentiële maatschappelijke baten overstijgen de kosten. Het aantal jaren genoten onderwijs en de prestaties op de arbeidsmarkt onder de deelnemers nemen toe en de deelname aan criminele activiteiten en inactiviteit nemen af, zie Currie (2001) en Heckman (2000). Op basis van een gemiddelde over de uitkomsten van evaluaties van een vijftal programma's in de VS (het Early Training project, het Abecedarian project, het Perry-Preschool programma van de High/Scope Foundation, het Milwaukee-project en Head Start) vinden wij een toename in het aantal jaren genoten onderwijs van 0,6 jaar. Het effect van deelname op toetsscores verdwijnt na verloop van tijd, waar vve programma's vrij consistent een blijvend

⁹ Zie Rapport Brede Heroverwegingen, 5, Kindregelingen: Het kind van de regeling, 2010

effect hebben op het aantal jaren onderwijs, zie bijvoorbeeld Currie en Thomas (1995) en Schweinhart et al. (2005). Dat is opmerkelijk, aangenomen dat juist beter presterende leerlingen veelal langer in het onderwijssysteem blijven. Volgens Heckman et al. (2006) is een mogelijke verklaring dat door deelname aan vve de non-cognitieve vaardigheden van de deelnemers toenemen, die vervolgens een blijvend effect genereren op het behaalde onderwijsniveau.

De gemiddelde toename in jaren onderwijs is met onzekerheid omgeven. De projecten verschillen o.a. in hun vormgeving, intensiteit, beoogde doelgroep en omvang. Zo is het Head Start programma recenter en veel omvangrijker dan de anderen. In de VS nemen op dit moment meer dan 900.000 kinderen deel aan dit programma, waarmee ongeveer 7 miljard dollar gemoeid is. Het Milwaukee-project was veruit het intensiefste programma. Uitgevoerd tussen 1968 en 1978, bestond dit programma uit voltijdse kinderopvang en training van moeders. Het Early Training project daarentegen bood wekelijkse bezoeken door een coach gedurende het jaar en een tienweekse summerschool van enkele dagdelen. Intensieve programma's van minimaal vier dagdelen per week met goed opgeleide begeleiders lijken betere resultaten te geven dan minder intensieve programma's. Ook zijn er aanwijzingen dat deelname aan een vve-programma op jongere leeftijd effectiever is, zie Campbell et al. (2002).

In Nederland zijn de opbrengsten van voor- en vroegschoolse educatie in termen van genoten aantal jaren onderwijs vermoedelijk beperkter dan in de VS. De Nederlandse risicogroepen zijn minder zwaar dan in de VS. Bovendien kunnen kinderen in Nederland al vanaf 4 jaar naar de basisschool, waardoor de meerwaarde van vve-programma's ten opzichte van de bestaande praktijk minder kan zijn dan in de VS. In de VS gaan kinderen pas vanaf hun zesde levensjaar naar het reguliere onderwijs. Voor Nederland is geen overtuigend bewijs beschikbaar over de effectiviteit van een vve-programma. Veen et al. (2000) evalueren de Nederlandse vve-programma's Piramide en Kaleidoscoop. Het betreft geen (natuurlijk) experiment, maar de auteurs maken gebruik van "matching-technieken". Zij vinden geen eenduidige resultaten van deelname op toetsscores. Op sommige toetsen lijkt deelname de prestaties te verhogen, terwijl voor andere toetsen het effect nul is. Ook vinden zij geen effect van deelname op de non-cognitieve vaardigheden van de kinderen.

Samenvattend werken wij voor Nederland met een onzekerheidsvariant, waarin deelname in een vve programma resulteert in een toename van het aantal jaren onderwijs van 0,3 jaar. Indien de ingezette middelen een grotere reikwijdte hebben dan de huidige risicogroep, dan neemt de effectiviteit van het vve-programma af. Dit resulteert in een tweede onzekerheidsvariant voor Nederland, waarin het aantal jaren genoten onderwijs ten gevolge van een vve-programma toeneemt met 0,15 jaar.

Table 4.4 Onzekerheidsvarianten vve

	Literatuurvariant	NL variant I	NL variant II
Effect op aantal jaren onderwijs	+ 0,60 jaar	+ 0,30 jaar	+ 0,15 jaar
Doelgroep	risicoleerlingen	risicoleerlingen	afnemend risicoprofiel leerlingen

Illustratie

Wij beschouwen een maatregel gericht op uitbreiding en verbetering van het huidige bereik van de voor- en vroegschoolse educatie. De inzet van middelen bedraagt 225 mln euro. Van dit bedrag kunnen jaarlijks zo'n 75.000 kinderen deelnemen aan voor- en vroegschoolse educatie met een jaarlijkse nieuwe instroom van zo'n 20.000 kinderen. De kosten per deelnemer bedragen gemiddeld per jaar zo'n 3.000 euro. Een toelichting is te vinden in het kader "VVE: gemiddelde kosten per leerling per jaar".

Vve: gemiddelde kosten per leerling per jaar

Het wetsvoorstel *Ontwikkelingskansen door Kwaliteit en Educatie* (OKE) vormt de basis voor de berekening van de kosten per kind per jaar in voor- en vroegschoolse educatie. OKE legt vast dat per 1 augustus 2010 de voorschoolse educatie wordt uitgebreid naar minimaal vier dagdelen per week (ofwel minimaal 10 uur per week). Deze wet verplicht gemeenten om op een groep van 16 kinderen twee leid(st)ers aan te stellen en gebruik te maken van brede programma's voor voorschoolse educatie. Vroegschoolse educatie bestrijkt ook vier dagdelen per week met één leid(st)er per groep van 18-24 kinderen. De jaarlijkse kosten van een leid(st)er bedragen 90 dzd euro. De jaarlijkse kosten van vve komen als volgt tot stand.

Voorschoolse educatie

Groepen van 16 leerlingen met 2 docenten voor vier dagdelen per week. De (standaard)kosten voor een docent per jaar, inclusief overhead, is 90 dzd euro. De kosten per leerling per jaar worden gegeven door

$$\text{kosten per leerling voorschoolse educatie per jaar} = \frac{2 \times 90 \text{ dzd} \times 4/10}{16} = 4500 \text{ euro.} \quad (4.1)$$

Vroegschoolse educatie

Groepen van 18-24 leerlingen met één docent voor vier dagdelen per week. Wederom zijn de kosten van een docent 90 dzd euro per jaar. De gemiddelde groepsgrootte in vroegschoolse educatie is 21, hetgeen impliceert dat de kosten per leerling per jaar gelijk zijn aan

$$\text{kosten per leerling vroegschoolse educatie per jaar} = \frac{90 \text{ dzd} \times 4/10}{21} = 1700 \text{ euro.} \quad (4.2)$$

Gemiddelde kosten vve

Voorschoolse educatie beslaat 1,5 cohort en vroegschoolse educatie beslaat 2 cohorten. Derhalve zijn de gemiddelde kosten van een volledig (3,5-jarig) vve programma per leerling per jaar

$$\text{kosten per leerling vve per jaar} = \frac{4500 \times 1,5 + 1700 \times 2}{3,5} \approx 3000 \text{ euro.} \quad (4.3)$$

Het bedrag van 225 mln euro vormt tevens een bovengrens voor de inzet van middelen op verbetering van het bereik van de huidige vve en de doelgroep uitbreiding zoals beoogd in de

Heroverwegingswerkgroep Productiviteit Onderwijs. Deze berekening is terug te vinden in het kader “VVE: doelgroep uitbreiding en vergroting van het bereik”. Additionele bestedingen voor voor- en vroegschoolse educatie zijn mogelijk, maar zullen naar alle waarschijnlijkheid resulteren in een lagere effectiviteit per ingezette euro. Voor de huidige populatie aan risicoleerlingen is het bedrag van 225 mln toereikend. Andere doelgroepen hebben een lager risicoprofiel, waardoor vve vermoedelijk minder effectief zal zijn.

VVE: doelgroep uitbreiding en vergroting van het bereik

Hoeveel leerlingen kunnen nog worden bereikt met de huidige vve-programma's en hoeveel middelen zijn hiermee gemoeid? Uit het rapport "Productiviteit Onderwijs" van de brede heroverwegingen komt naar voren dat de doelgroep van voorschoolse leerlingen zich uitbreidt van 40 dzd naar 60 dzd. Het huidige bereik in de voorschoolse educatie bedraagt 80%. In totaal kunnen dus nog zo'n 28 dzd kinderen in de vve worden geplaatst. Per cohort betreft dat 19 dzd kinderen. Het voorschoolse traject duurt 1,5 jaar. Onder de aanname dat ieder kind het volledige vve traject doorloopt, het vve traject 3,5 jaar duurt, van 2,5 tot 6 jaar en de kosten per jaar per kind 3000 euro bedragen, resulteert dit in een mogelijke inzet van 200 mln euro. Dit bedrag vullen wij aan met de middelen die naar aanleiding van OKE moeten worden ingezet om het aantal dagdelen te verruimen. Landelijk bedraagt de gemiddelde intensiteit van de vve programma's 3,4 dagdelen per week. In de grote steden krijgen de bereikte kinderen al de beoogde vier dagdelen per week, zie Werkgroep Heroverwegingen rapport 5: Het kind van de regeling. De verruiming van het aantal dagdelen in de voorschoolse educatie kost ongeveer 25 mln euro.^a

^a Onder de aanname dat momenteel circa 32 dzd kinderen in de voorschoolse educatie deelnemen, zie Werkgroep Heroverwegingen rapport 5: Het kind van de regeling, geldt

$$\text{kosten aantal dagdelen naar vier} = \frac{0,6}{3,4} \times 4500 \times 32 \text{ dzd} \approx 25 \text{ mln euro.} \quad (4.4)$$

Voor onze analyse is een verruiming van het aantal dagdelen in de vroegschoolse educatie niet zinvol, omdat deze leerlingen het effect in onze analyse niet meekrijgen.

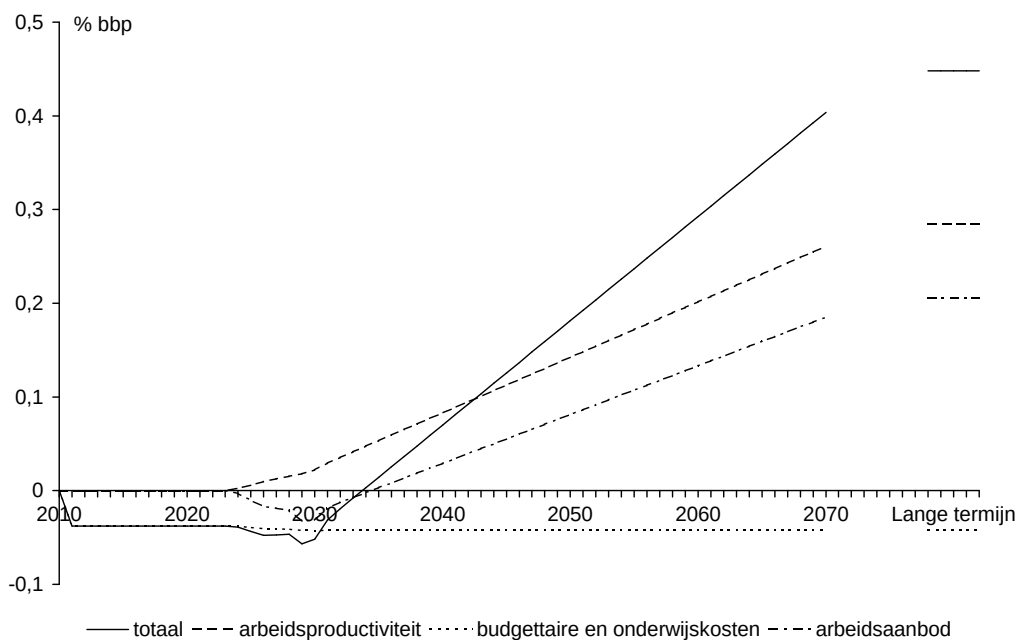
Uitkomsten

Figuur 4.2 presenteert de uitkomsten van de maatregel op basis van de eerste Nederlandse onzekerheidsvariant. De inzet van extra middelen vindt plaats vanaf 2011 en blijft constant (in prijzen van 2010). De eerste nieuwe leerlingen stromen na circa 13 jaar in in de beroepsbevolking. Vanaf 2024 begint het gemiddeld opleidingsniveau van de beroepsbevolking, en daarmee hun arbeidsproductiviteit te stijgen. Na ongeveer 63 jaar heeft iedere werknemer uit de beoogde doelgroep de gelegenheid gehad deel te nemen aan voor- en vroegschoolse educatie. De samenstelling van de beroepsbevolking verandert niet langer. De groep leerlingen, die vve heeft genoten, is aanzienlijk en bedraagt op de lange termijn 21,3% van de beroepsbevolking. 10,9% punt van deze groep wordt bereikt met onze maatregel. Merk op dat hierbij is verondersteld dat de doelgroep van voorschoolse leerlingen over de tijd gelijk blijft. De vraag is of dit blijft gelden als de ouders van leerlingen in deze doelgroep hoger zijn opgeleid.

Alhoewel de maatregel zich slechts richt op een deel van de beroepsbevolking resulteert een aanzienlijke toename van de arbeidsproductiviteit ten gevolge van de grote effectiviteit van de

maatregel. Op de lange termijn heeft 10,9% van de beroepsbevolking 0,3 jaar meer scholing. Dit leidt tot een stijging van het bbp met ongeveer 0,3% vanwege de hogere arbeidsproductiviteit van deze groep. Daarnaast resulteert de maatregel op de lange termijn in een stijging van het arbeidsaanbod, omdat hoger opgeleiden meer participeren op de arbeidsmarkt. Het karakter van de maatregel brengt met zich mee dat de opbrengsten zich relatief laat manifesteren en er in eerste instantie over een lange periode alleen kosten worden gemaakt. De structurele inzet van 225 mln euro voor uitbreiding van de huidige vve-regeling leidt per saldo op de lange termijn tot een stijging van het bbp met 0,5%.

Figure 4.2 Uitwerking Nederlandse variant voor de vve



4.2.2 Onderwijstijd

Een verlenging van de onderwijstijd kan bijdragen aan de onderwijsprestaties van leerlingen. In de Nederlandse beleidscontext richten nieuwe beleidsvoorstellen binnen dit thema zich op het uitbreiden van het aantal uren reken- en taalonderwijs, kop-voet-en schakelklassen en 'summer schools'. Deze voorstellen kunnen zeer van elkaar verschillen wat betreft omvang en doelstelling. Bij extra uren reken-en taalonderwijs kan het gaan om een additionele inzet van één of enkele uren in de week op bestaande leerdoelen binnen het reguliere schooljaar. Schakelklassen zijn specifiek bedoeld voor leerlingen met onderwijsachterstanden in de overgang tussen verschillende typen onderwijs. Bij 'summer schools' gaat het om intensievere cursussen gedurende een korte periode buiten het reguliere schooljaar. Van de laatste twee beleidsinterventies ontbreekt voldoende empirische evidentie (zie paragraaf 4.5.8). Deze krijgen derhalve de kwalificatie 'effect onbekend' en zullen niet kwantitatief worden geanalyseerd.

Empirie

De beschikbare empirie over verlenging van de onderwijstijd laat over het algemeen positieve effecten zien op toetsscores. De literatuur is vrij divers: een aantal studies richt zich op extra onderwijstijd aan het begin van de schoolcarrière (Leuven et al., 2010; Berlinski et al., 2009; Cascio en Lewis, 2006), andere studies analyseren de effecten van extra lesdagen of extra uren rekenen en taal op latere leeftijden (Loeb en Bound, 1996; Link en Mulligan, 1986; Lavy, 1998). Op basis van het gemiddelde van deze studies resulteert een verlenging van de onderwijstijd met één maand tot een toename van de toetsscores met 0,04 standaarddeviatie. Wanneer we studies die zich richten op extra onderwijstijd aan het begin van de schoolcarrière (wat raakt aan voorschoolse educatie) buiten beschouwing zouden laten, resulteert een gemiddeld effect van 0,06 standaarddeviatie. Dit wordt met name veroorzaakt door de relatief sterke effecten in Lavy (1998).

Deze toename is met onzekerheid omgeven. Zo vinden de studies bijvoorbeeld verschillende resultaten voor specifieke doelgroepen. Leuven et al. (2010) vinden dat een maand extra onderwijstijd leidt tot een verbetering van de toetsscores op rekenen en taal (op 6-jarige leeftijd) met respectievelijk 0,05 en 0,06 standaarddeviatie voor achterstandsleerlingen. Voor niet-achterstandsleerlingen worden geen effecten gevonden. Cascio en Lewis (2006) analyseren effecten van verschillen in genoten onderwijstijd tussen tieners in de VS en richten zich hierbij vooral op effecten voor etnische minderheden. Zij vinden dat extra onderwijstijd de toetsscores voor 'hispanics' en 'blacks' significant doet toenemen. Link en Mulligan (1986) analyseren de effecten van extra onderwijsuren per week voor rekenen en taal in het basisonderwijs in de VS. Alleen op rekentoetsen in sixth grade en taaltoetsen in third and sixth grade voor 'hispanics' worden significante effecten gevonden. De auteurs vinden ook enige aanwijzing voor afnemende meeropbrengsten van onderwijstijd.¹⁰ Lavy (1998) vindt relatief grote effecten van extra lessen in de week op toetsscores in het primair onderwijs in Israël. Hij laat zien dat 3,8 extra lessen per week leidt tot een toename in de toetsscores op rekenen en taal met 0,15 standaarddeviatie.

Aangezien we binnen dit thema beschikken over een kwalitatief goede Nederlandse studie (Leuven et al., 2010), gebruiken we deze voor twee Nederlandse onzekerheidsvarianten. De eerste Nederlandse variant geeft het effect van een maand extra onderwijstijd op de totale populatie (+ 0,0125 standaarddeviatie), de tweede geeft het effect op de populatie van achterstandsleerlingen. Als de middelen voor extra onderwijstijd gericht worden ingezet op deze groep leerlingen, kan een maand extra onderwijstijd leiden tot een toename van de toetsscores voor deze groep met 0,05 standaarddeviatie. Tabel 4.5 presenteert de onzekerheidsvarianten.

¹⁰ Zij vinden een negatieve coëfficiënt voor een tweede orde term van onderwijstijd.

Table 4.5 Onzekerheidsvarianten onderwijstijd verlenging met één maand

	Literatuurvariant	NL variant I	NL variant II
Effect op toetsscores	+ 0,04 SD	+ 0,0125 SD	+ 0,05 SD
Doelgroep	totale populatie	totale populatie	achterstandsleerlingen

Illustratie

De maatregel die wordt doorgerekend beoogt achterstandsleerlingen in het primair onderwijs per jaar 40 uur extra taal- en rekenonderwijs te geven. De bekostiging van de extra lessen per leerling wordt verondersteld proportioneel te zijn aan de jaarlijkse uitgaven per leerling in het primair onderwijs. Vermoedelijk is dit een kleine overschatting, daar in de jaarlijkse uitgaven per leerling ook de materiële kosten zijn opgenomen. De materiële kosten zijn ongeveer 15% van de lumpsum financiering in het primair onderwijs, en dus worden de kosten met circa 15% overschat. Veertig uur extra onderwijs betekent een toename van de onderwijstijd op jaarbasis met 4,25%. Derhalve stijgen ook de jaarlijkse kosten per achterstandsleerling per jaar met 4,25%. Om de volledige populatie van achterstandsleerlingen in het basisonderwijs te bereiken is 96 miljoen euro per jaar nodig. Een toelichting is te vinden in het kader “De kosten van 40 uur extra taal- en rekenonderwijs voor achterstandsleerlingen in het basisonderwijs”.

De kosten van 40 uur extra taal- en rekenonderwijs voor achterstandsleerlingen in het basisonderwijs

Om achterstanden in het primair onderwijs tegen te gaan bestaat in Nederland de gewichtenregeling. Dit beleid kent op basis van etnische achtergrond en de hoogst genoten opleiding van de ouders gewichten aan leerlingen toe. Achterstandsleerlingen krijgen een gewicht groter dan 1. De Referentieraming 2009 van OCW gaat uit van 121 500 leerlingen met gewicht 1.3 en 83 900 leerlingen met gewicht 2.2 in 2009 verdeeld over de 8 groepen van de basisschool. Per cohort zijn dat circa 25 duizend leerlingen.

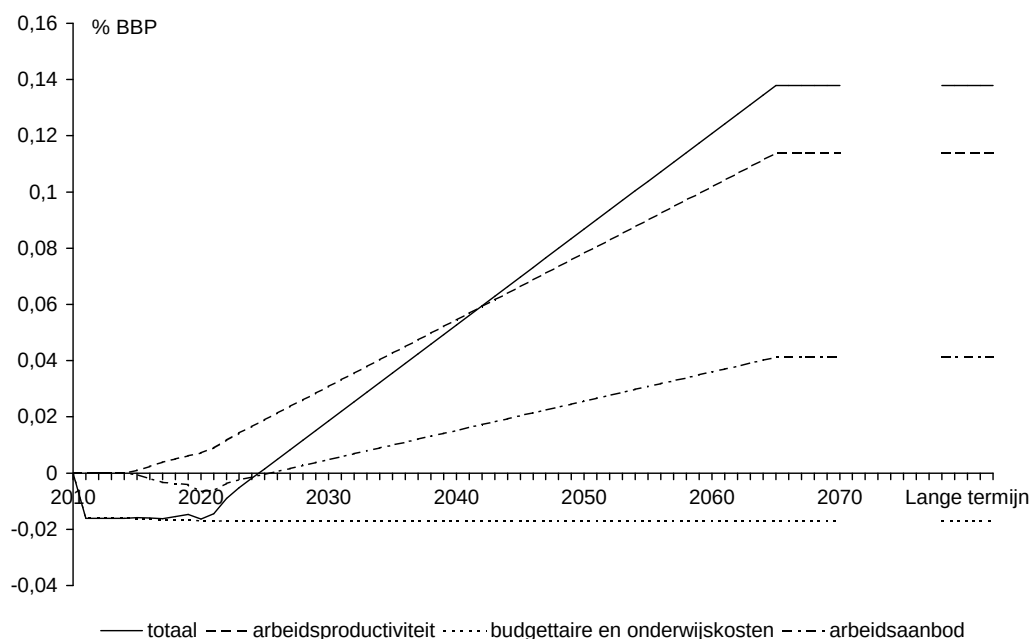
De wettelijke onderwijstijd over acht schooljaren basisonderwijs is minimaal 7520 uur. Alle leerlingen krijgen dan minimaal 940 uur per jaar onderwijs (het zogenoemde ‘Hoorns model’). Veertig uur extra taal- en rekenonderwijs betekent een toename van 4.3% van het aantal uren onderwijs voor achterstandsleerlingen.

De jaarlijkse uitgaven per leerling in het primair onderwijs bedragen 6600 euro, zie OC&W (2008). Dit bedrag is inclusief de kosten voor huisvesting welke voor rekening van de gemeente komen. Scholen ontvangen voor leerlingen met gewicht 1.3 30%, en voor leerlingen met gewicht 2.2 120%, meer bekostiging per leerling. Een stijging van 4,3% in de jaarlijkse kosten voor 205 duizend gewichtleerlingen in het basisonderwijs, resulteert in

$$(121\,500 \times 1.3 + 83\,900 \times 2.2)6600 \times 4,25\% = 96 \text{ mln euro}$$

Extra taal- en rekenonderwijs in het basisonderwijs sorteert niet meteen een effect. Als de maatregel in 2011 aangrijpt, zijn de eerste effecten meetbaar in 2015 wanneer de eerste leerlingen die extra taal- en rekenonderwijs hebben genoten de arbeidsmarkt instromen. Op de lange termijn heeft 13% van de beroepsbevolking baat bij deze interventie. Hierbij is verondersteld dat het aantal achterstandsleerlingen en de bevolkingsgrootte gelijk blijft.

Figure 4.3 Uitwerking Nederlandse variant voor extra taal- en rekenonderwijs voor achterstandsleerlingen in het basisonderwijs



De structurele inzet van 96 miljoen euro voor extra taal- en rekenonderwijs leidt op de lange termijn tot een stijging van het bbp met 0,1%. Hiervan is het overgrote deel toe te schrijven aan de stijging van de arbeidsproductiviteit.

4.2.3 Klassengrootte

Klassenverkleining heeft tot doel de kwaliteit van het onderwijs te verhogen, waardoor leerprestaties kunnen verbeteren. Vaak wordt gedacht dat het verkleinen van klassen bijdraagt aan de prestaties van leerlingen omdat een leraar meer tijd beschikbaar heeft per leerling. Dit effect is echter niet vanzelfsprekend, omdat verschillende mechanismen zich kunnen voordoen in een grote of kleine klas. Zo heeft een leerling in een grotere klas een grotere kans een medeleerling te treffen waarvan hij/zij veel kan leren. Ook is het mogelijk dat de leraar zijn gedrag niet verandert in een kleinere klas of zelfs minder effectief gaat werken in een kleinere klas door te kiezen voor een andere verhouding tussen individuele en klassikale aanpakken. Daarnaast leidt het verkleinen van klassen tot een grotere vraag naar leraren waardoor er meer onervaren docenten voor de klas komen hetgeen de kwaliteit negatief kan beïnvloeden. In hoeverre deze mogelijke mechanismen daadwerkelijk van belang zijn is niet bekend. Het betekent echter wel dat het effect van het verkleinen of vergroten van klassen op voorhand niet duidelijk is.

Nieuwe beleidsvoorstellen kunnen bestaan uit extra middelen om de klassen in specifieke onderwijsniveaus te verkleinen. Ook is het mogelijk dat beleidsvoorstellen gericht op het korten

van de lump-sum indirect zullen leiden tot grotere klassen.

Empirie

In Cornet et al. (2006) wordt klassenverkleining op basis van de beschikbare empirie als 'niet kansrijk' gekwalificeerd. Dit houdt in dat de maatschappelijke baten van klassenverkleining naar verwachting niet opwegen tegen de maatschappelijke kosten.

De afgelopen jaren zijn verschillende studies uitgevoerd naar de effecten van klassengrootte op basis van een (quasi-)experimenteel onderzoeksdesign. Een deel van de studies vindt overwegend kleine effecten (Krueger (1999); Angrist en Lavy (1999); Urquiola (2006); Browning en Heinesen (2007); Boozer en Rouse (2001)), een ander deel vindt geen effect van veranderingen in klassengrootte (Dobbelsteen et al. (2002); Hoxby (2000b); Leuven et al. (2008)). Wanneer we het gemiddelde van deze studies nemen leidt een verkleining van de klas met één leerling tot een toename in de toetsscores van 0,012 standaarddeviatie.

De meest bekende van de studies, en de enige met een experimenteel design waarbij leerlingen random werden toegewezen aan grote en kleine klassen, is Krueger (1999). Hij vindt dat kleinere klassen leiden tot beter onderwijsprestaties. De effectgroottes liggen tussen de 0,013 en 0,025 standaarddeviatie. Angrist en Lavy (1999) vinden iets kleinere effecten. De grootste effecten worden gevonden in Urquiola (2006) (0,019 tot 0,035 standaarddeviatie). Hij vindt vooral grote effecten in klassen vanaf ongeveer 30 leerlingen. Dit suggereert dat effecten van klassengrootte niet lineair zijn.

De genoemde studies richten zich op het primair onderwijs of de onderbouw van het voortgezet onderwijs. De resultaten wijzen niet op aanzienlijke verschillen in effectgrootte tussen beide onderwijstypen. Voor zover ons bekend zijn er zijn nog geen studies die de effecten van klassenverkleining in latere fasen van het onderwijs analyseren.

In Nederland zijn de effecten van klassenverkleining vermoedelijk minder sterk dan het gemiddelde effect uit de literatuur. Nederland kent de gewichtenregeling, die additionele middelen beschikbaar stelt voor scholen met veel achterstandsleerlingen. Deze middelen worden veelal ingezet voor het verkleinen van klassen. Daarnaast zijn de klassen in de onderbouw van het basisonderwijs sinds het eind van de jaren negentig substantieel verkleind. De gemiddelde klassengrootte in het Nederlandse primair onderwijs is lager dan de gemiddelde klassengrootte in de meeste empirische studies.¹¹ De literatuur ondersteunt ons vermoeden. Een quasi-experimentele studie van Dobbelsteen et al. (2002) vindt geen effecten van kleinere klassen op de leerprestaties in de groepen 4, 6 en 8 van het Nederlands primair onderwijs. Op basis van bovenstaande overwegingen werken wij voor Nederland met een onzekerheidsvariant, waarin een klassenverkleining met één leerling resulteert in een toename van de toetsscore met

¹¹ De gemiddelde klassengrootte in het Nederlandse primair onderwijs bedraagt 23,9 leerlingen OECD (2009). In de studies van Angrist en Lavy (1999), Urquiola (2006) en Boozer en Rouse (2001) zijn de gemiddelde klassengrootten respectievelijk 32, 30 en 25 leerlingen.

0,004 standaarddeviatie, zie tabel 4.6. De Nederlandse effecten van Dobbelsteen et al. (2002) worden niet volledig gevolgd, omdat het design van de buitenlandse studies betrouwbaarder is.

Table 4.6 Onzekerheidsvarianten Klassenverkleining (met één leerling)

	Literatuurvariant	NL variant
Effect op toetsscores	+ 0,012 SD	+ 0,004 SD
Doelgroep	PO, VO	PO, VO

Illustratie

We beschouwen een interventie waarin in het primair onderwijs de klassen met één leerling worden verkleind. Gegeven de huidige leerlingpopulaties in het primair onderwijs, is met deze interventie een bedrag van 450 miljoen euro gemoeid. De additionele kosten per leerling bedragen 290 euro. Voor een berekening van deze bedragen verwijzen we naar het kader “Klassenverkleining met één leerling in het primair onderwijs”.

Klassenverkleining met één leerling in het primair onderwijs

In dit tekstkader laten we zien hoe de kosten van klassenverkleining met één leerling in het primair onderwijs worden berekend. Folger en Parker (1990) schatten dat de additionele kosten van klassenverkleining proportioneel zijn aan de totale jaarlijkse uitgaven per leerling. Anders gezegd, de kosten van klassenverkleining met één leerling volgen uit vermenigvuldiging van de procentuele omvang van de klassenverkleining met de totale kosten per leerling per jaar.

De gemiddelde klassengrootte in het primair onderwijs is 23,9, zie OECD (2009) en onderstaande tabel. Een verkleining van de gemiddelde klassengrootte naar 22,9 impliceert dat er 4,4% meer fte's nodig zijn om dit voor elkaar te krijgen. De totale uitgaven per leerling per jaar zijn 6600 euro, en dus zijn de additionele uitgaven per leerling per jaar voor klassenverkleining 290 euro (4,4% van 6600). Met ruim 1.5 miljoen leerlingen in het primair onderwijs komt dit neer op 450 mln euro per jaar.

Uitgaven per leerling, leerlingaantallen en gemiddelde klassengroottes

	Uitgaven per leerling	aantal leerlingen	Gemiddelde klassengrootte
primair onderwijs	6 600 euro ^a	1 552 900	23.9

^a De uitgaven per leerling zijn opgebouwd uit de uitgaven van OC&W per leerling plus de gemeentelijke bijdrage per leerling. In het primair onderwijs komt de huisvesting van scholen voor rekening van de gemeenten.

Uitkomsten

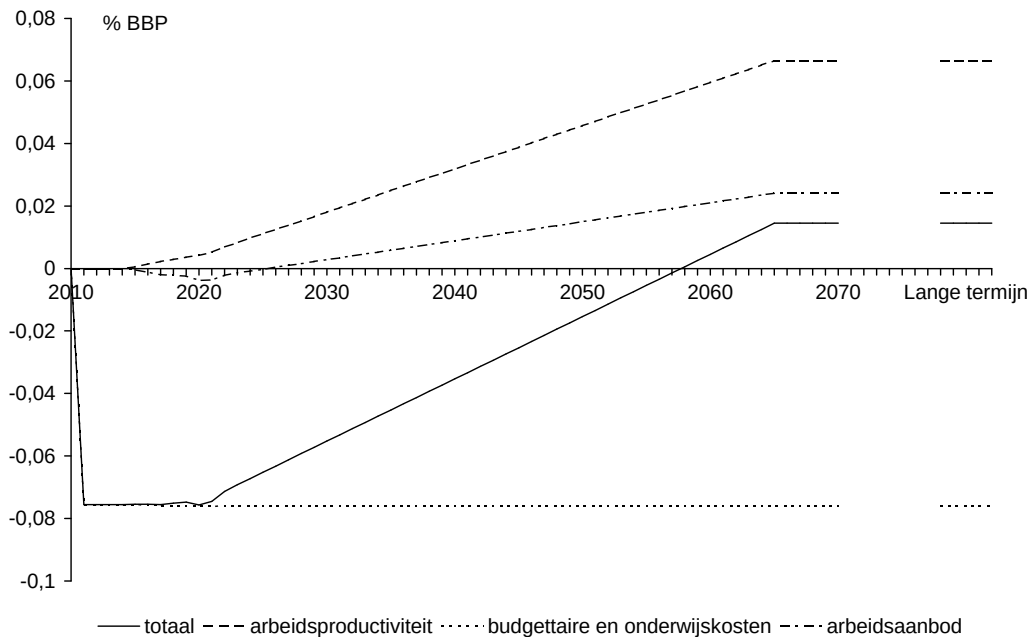
Figuur 4.4 vat de effecten op het bbp samen als gevolg van een brede invoering van klassenverkleining met één leerling in het primair onderwijs. Vanaf 2011 bedraagt de structurele inzet van middelen 450 miljoen euro.

Klassenverkleining vereist veel middelen en leidt slechts tot een geringe verbetering van de onderwijsprestaties. Op de lange termijn stijgt de arbeidsproductiviteit niet veel, terwijl de

structurele kosten hoog zijn. Op de lange termijn leidt deze interventie tot een stijging van het bbp met 0,0%.

De empirische literatuur laat zien dat andere interventies kosten-effectiever zijn dan klassenverkleining, zoals bijvoorbeeld scholing van leraren (Angrist en Lavy, 2001). Daarnaast heeft het verkleinen van klassen als nadeel dat de vraag naar leraren toeneemt. In een periode waarin tekorten aan leraren worden verwacht, neemt de druk op de arbeidsmarkt voor docenten toe. In deze analyse wordt hiermee geen rekening gehouden.

Figure 4.4 Uitwerking Nederlandse variant voor klassenverkleining in het primair onderwijs



4.2.4 Schooluitval

Beleidsintensivering gericht op vermindering van schooluitval zonder startkwalificatie heeft tot doel het bereikte onderwijsniveau van jongeren te verhogen, en daarmee hun latere maatschappelijke positie en kansen op de arbeidsmarkt te versterken. De kansen op de arbeidsmarkt en de sociale integratie verbeteren naarmate men een hoger opleidingsniveau heeft bereikt. Het voorkomen van voortijdig schoolverlaten kan het beroep op sociale uitkeringen verminderen en de kans op sociale uitsluiting en deelname aan criminele activiteiten verkleinen. Het percentage voortijdig schoolverlaters (VSV-ers) bedroeg in Nederland in 2007 12%. Hiermee staat Nederland binnen de EU op de 7e plaats (Lissabon Monitor 2009).

Het Nederlands beleid voor het terugdringen van voortijdig schoolverlaten kent verschillende elementen. Zo is er de kwalificatieplicht die inhoudt dat leerlingen zonder startkwalificatie tot hun achttiende jaar verplicht onderwijs moeten volgen. Daarnaast zijn er ook veel middelen in de bekostiging ingebouwd om achterstanden te bestrijden (zoals het leerwegondersteunend

onderwijs en de leerplusarrangementen in het voortgezet onderwijs). In het schooljaar 2006-2007 is gestart met het zogenoemde convenantenbeleid. Dit beleid belooft de verantwoordelijke regio's als zij het aantal voortijdig schoolverlaters weten te verminderen. In 2010 was voor dit beleid in totaal 81 miljoen euro beschikbaar.¹²

Nieuwe beleidsvoorstellen kunnen zich richten op intensieve begeleiding van jongeren in het vmbo of mbo, of financiële prikkels voor risicojongeren. De financiële prikkels kunnen worden gericht op behaalde studieprestaties, of op inzet en deelname aan studietaken.

Empirie rondom intensieve begeleiding

Cornet et al. (2006) concluderen dat het verminderen van voortijdig schoolverlaten een kansrijke beleidsoptie is. Hierbij worden programma's met intensieve coaching als veelbelovende beleidsrichting aangemerkt. Webbink et al. (2009) geven een overzicht van empirische evaluaties van deze programma's.

Heckman (2000) concludeert dat langdurige en intensieve programma's met coaches invloed kunnen hebben op studieprestaties en latere arbeidsmarktpositie. Hij stelt bovendien dat preventieve interventies gericht op jongeren die nog op school zitten veel effectiever zijn dan curatieve interventies gericht op jongeren die al zijn uitgevallen. Voorbeelden van effectieve programma's in de VS zijn 'Big Brothers/Big Sisters' en het 'Quantum Opportunity Program'. In deze programma's krijgen risicoleerlingen een mentor toegewezen die zich met name richt op de sociale ontwikkeling van de jongere. Bij 'Big Brothers/Big Sisters' worden positieve effecten gevonden op toetsscores, waarbij met name meisjes beter presteerden (Grossman en Tierney, 1998). Deze verbetering in leerprestaties is op basis van de beschikbare gegevens niet uit te drukken in termen van standaarddeviaties. Bij het 'Quantum Opportunity Program' wordt coaching gecombineerd met financiële prikkels voor de leerlingen. Er worden positieve effecten gevonden op schoolprestaties (Hahn et al., 1994; Taggart, 1998; Maxfield et al., 2003), waarbij de slagingskansen op high school met ongeveer 50% toeneemt.

Lavy en Schlosser (2005) evalueren een programma met intensieve begeleiding, waarbij onderpresterende leerlingen in kleine groepen extra onderwijstijd krijgen in de laatste klas voor het examen. Zij vinden dat dit programma de slagingskansen met 22% doet toenemen.

In de onzekerheidsvarianten baseren wij ons op deze laatste studie, aangezien dit de enige studie is die kwantificering van de effecten van intensieve begeleiding mogelijk maakt. De sterkere gevonden effecten bij het Quantum Opportunity Program (waar ook financiële prikkels een rol spelen) ondersteunen dit resultaat. Voor de Nederlandse onzekerheidsvariant wegen wij mee dat er in Nederland al relatief veel beleid wordt gevoerd om voortijdig schoolverlaten tegen te gaan. Zo kennen wij een kwalificatieplicht tot 18 jaar, en worden relatief veel middelen

¹² Zie www.rijksbegroting.nl, tabblad: Home > 2010 > Voorbereiding > Begroting > VIII Onderwijs, Cultuur en Wetenschap > Vaststelling begroting Ministerie van Onderwijs, Cultuur en Wetenschap (VIII) voor het jaar 2010

ingezet voor zowel curatieve als preventieve bestrijding van voortijdig schoolverlaten.

Table 4.7 Onzekerheidsvarianten VSV (intensieve begeleidingsprogramma's)

	Literatuurvariant	NL variant
Effect op slagingskans	+ 22%	+ 15%
Doelgroep	vmbo, mbo	vmbo, mbo

De onzekerheid rond de effecten van intensieve programma's is groot. Er is veel variatie mogelijk in de precieze invulling van een programma. De intensiteit en duur, de kwaliteit van de coaches en de mate waarin maatwerk geboden kan worden, zijn in hoge mate bepalend voor het uiteindelijke resultaat.

Illustratie intensieve begeleiding coaches

We beschouwen een maatregel die zich richt op intensieve begeleiding van risicoleerlingen in het vmbo en mbo door coaches. De inzet van middelen bedraagt 50 miljoen euro. Van dit bedrag kunnen jaarlijks 20 duizend risicoleerlingen een uur per week worden begeleid door een speciaal daarvoor aangestelde coach. Het beschikbare budget per risicoleerling is 2 500 euro per jaar. In het kader "Voortijdig schoolverlaten: kosten en bereik van coaches" worden deze cijfers toegelicht. Achtergrondinformatie over het huidige beleid en aantallen voortijdig schoolverlaters is te vinden in het kader "Voortijdig schoolverlaters".

Voortijdig schoolverlaten: kosten en bereik van coaches

Per vsv-er is een jaarlijks budget beschikbaar van 2 500 euro.^a Het gemiddelde van de programma's uit de literatuur is 2425 euro per leerling per jaar.^b Uitgaande van 2500 euro per leerling per jaar, kunnen met 50 miljoen euro 20 duizend risicoleerlingen in het vmbo en mbo worden bereikt.

Omdat de definitie van een risicoleerling niet eenduidig is, is het exacte aantal risicoleerlingen in het vmbo en mbo niet bekend. Een bovengrens van het aantal risicoleerlingen in deze sectoren is de hele groep. Een ondergrens wordt gegeven door het aantal geregistreerde voortijdig schoolverlaters.

^a Bron: Brief aan de Tweede Kamer van de Minister van onderwijs, cultuur en wetenschap van 18 februari 2010, met als onderwerp: Voortijdig schoolverlaten: cijfers en beleid

^b De kosten van het Quantum Opportunity Program bedragen onder de aanname dat 1 dollar gelijk is aan 1 euro, 3750 euro per leerling per jaar. De kosten van de interventie beschouwd in Lavy en Schlosser (2005) zijn 1100 euro. Het gemiddelde van beide programma's is dus 2425 euro.

Uitkomsten intensieve begeleiding coaches

Figuur 4.5 presenteert de uitkomsten van de maatregel. De extra structurele inzet van 50 miljoen euro aan intensieve coaching voor voortijdig schoolverlaters vindt plaats vanaf 2011 en blijft constant. Op de lange termijn leidt deze interventie tot een stijging van het bbp met 0,2%. De interventie sorteert meteen effect na de start in 2011. De groep risicoleerlingen die intensieve coaching heeft gehad bedraagt op de lange termijn 10,7% van de beroepsbevolking. Dit geldt

Voortijdig schoolverlaters

Voortijdig schoolverlaters worden gedefinieerd als leerlingen die het (bekostigd) onderwijs verlaten zonder dat zij een startkwalificatie hebben behaald. Een leerling heeft een startkwalificatie met ten minste een afgeronde havo- of vwo-opleiding, of een basisberoepsopleiding (mbo niveau 2).

De doelstelling van het in 2006 gestarte beleid "Aanval op uitval", is het aantal vsv-ers terug te dringen tot 35 duizend in 2012. Gezamenlijk met alle betrokkenen (scholen, gemeenten, RMC's, zorginstellingen, de werkgevers, CWI's, politie, justitie, ouders en leerlingen) probeert men dit doel te realiseren. OCW heeft in alle 39 RMC-regio's convenanten afgesloten met contactgemeenten en onderwijsinstellingen in het voortgezet onderwijs en het middelbaar beroepsonderwijs. Als extra stimulans om de convenantsafspraken te halen stelt OCW aanvullende subsidie beschikbaar voor onderwijsprogramma's. Binnen het beschikbare budget voor vsv-beleid is het normbedrag vastgesteld op 2500 euro per verminderde vsv-er. Sinds de start van het vsv-beleid is het aantal vsv-ers op landelijk niveau gedaald, zie onderstaande tabel.^a

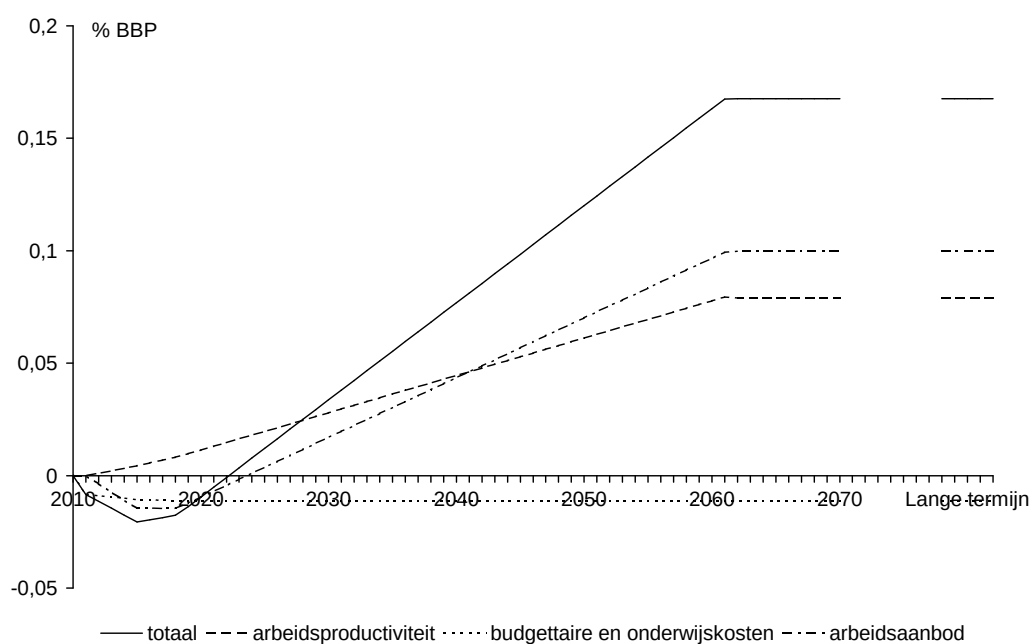
Landelijk aantal vsv-ers

	2002	04/05	05/06	06/07	07/08	08/09	09/10	10/11
aantal vsv-ers	71 000	58 900	52 700	50 900	46 800	42 600		35 000

^a Bron: cijfers.minocw.nl

onder de aanname dat er per jaar minstens 20 duizend voortijdig schoolverlaters in het vmbo en mbo zijn, én onder de aanname dat er een perfecte targetting van risicoleerlingen is.

Figure 4.5 Uitwerking Nederlandse variant voor intensieve begeleiding van risicoleerlingen door coaches



Empirie rondom financiële prikkels

In Cornet et al. (2006) wordt het gebruik van financiële prikkels voor leerlingen als kansrijke beleidsoptie aangemerkt om voortijdig schoolverlaten tegen te gaan. Webbink et al. (2009) geven een overzicht van empirische evaluaties van de betreffende literatuur. In het algemeen blijkt hieruit dat financiële prikkels een positief effect hebben op de studieprestaties (Angrist en Lavy, 2009; Dearden et al., 2009; Bettinger, 2008; Kremer et al., 2009). Angrist en Lavy (2009) vinden in een experiment met random toewijzing dat een bonus van 1500 dollar voor het slagen op een zogenoemd ‘matriculation exam’ (het noodzakelijke examen voor toegang tot het meeste post-secundair onderwijs) de slagingskansen met gemiddeld 16% doen toenemen. Opvallend is dat zij grote effecten vinden voor meisjes en geen effecten voor jongens. Dearden et al. (2009) maken gebruik van ‘matching technieken’ om de effecten van het Education Maintenance Allowance (EMA) programma in Engeland te evalueren, waarbij families met lage inkomens een vergoeding kregen voor zowel deelname aan het onderwijs na de leerplichtige leeftijd als voor het behalen van onderwijsprestaties. Zij vinden dat de kans op het afronden van twee jaar post-leerplichtig onderwijs met 12% toeneemt. Hierbij moet worden opgemerkt dat een deel van de effecten mogelijk verklaard kan worden door bestaande liquiditeitsrestricties onder deelnemers. Bettinger (2008) vindt in een experiment met random toewijzing dat financiële bonussen voor leerlingen in het primair onderwijs de toetsscores met 0,15 standaarddeviatie doet toenemen. De resultaten suggereren echter ook dat deze effecten in volgende jaren niet standhouden wanneer de leerlingen niet langer in aanmerking komen voor een bonus. Kremer et al. (2009) analyseren de effecten van beloningen voor meisjes in een gerandomiseerd experiment. Zij vinden dat een financiële beloning resulteert in een toename van de toetsscores met 0,12 tot 0,19 standaarddeviatie.

Op basis van deze studies valt niet met zekerheid te zeggen wat de minimaal benodigde grootte van de financiële prikkel is om effect te sorteren en wat de effecten zijn van verschillen in grootte van de prikkel.

In de onzekerheidsvarianten gaan wij uit van de experimentele studie van Angrist en Lavy (2009), die effecten vinden in termen van slagingskansen. Voor de Nederlandse onzekerheidsvariant wegen wij mee dat er in Nederland al relatief veel beleid is gericht op voortijdig schoolverlaten.

Table 4.8 Onzekerheidsvarianten VSV (financiële prikkels)

	Literatuurvariant	NL variant
Effect op slagingskansen	+ 16%	+ 10%
Doelgroep	VMBO, MBO	VMBO, MBO

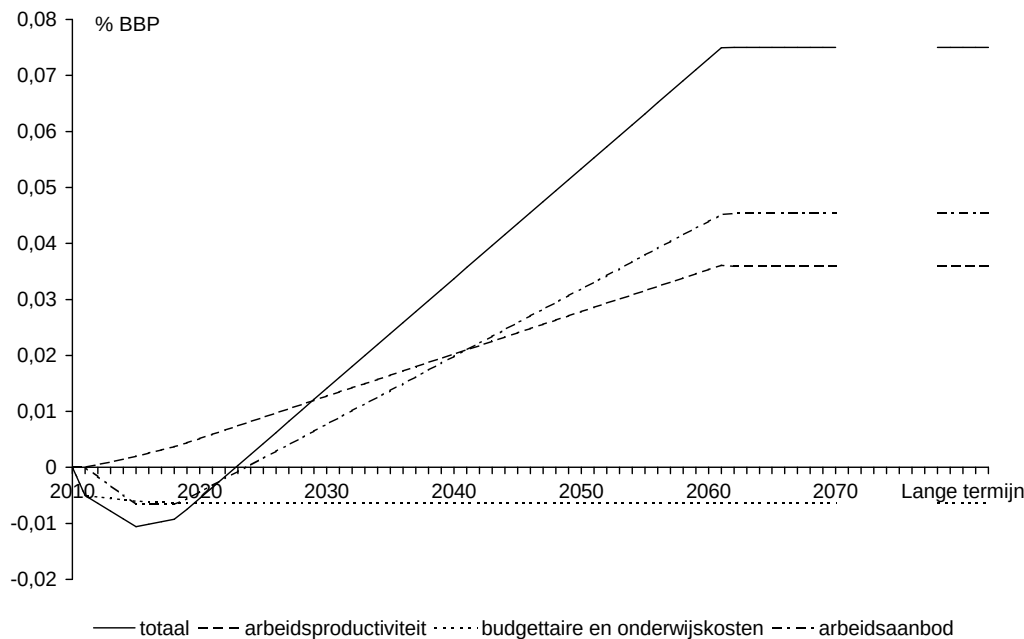
Illustratie financiële prikkels

We beschouwen een maatregel waarbij aan de 20 duizend risicoleerlingen in het vmbo en mbo een bonus van 1500 euro wordt toegekend wanneer zij slagen voor hun opleiding. De inzet van middelen bedraagt dan maximaal 30 miljoen euro.

Uitkomsten financiële prikkels

Figuur 4.6 presenteert de uitkomsten van de maatregel. Op de lange termijn leidt deze interventie tot een stijging van het bbp met 0,1%.

Figure 4.6 Uitwerking Nederlandse variant voor financiële prikkels voor risicoleerlingen



4.2.5 Inspectietoezicht

Veel Europese landen kennen een overheidsorgaan met de taak om scholen te inspecteren. De Nederlandse Inspectie van het Onderwijs voert inspecties uit op scholen in het primair en voortgezet onderwijs. Deze inspecties zijn bedoeld om te controleren of de scholen voldoen aan de wettelijke bepalingen in het onderwijs en om de kwaliteit van het onderwijs op de scholen te verbeteren. De inspecteurs evalueren de kwaliteit van het onderwijs en schrijven hierover een publiek rapport met aanbevelingen voor de school ter verbetering. De kwaliteit van het onderwijs wordt mogelijk verbeterd door de verantwoording die scholen moeten afleggen en door de aanbevelingen die volgen uit het inspectietoezicht.

Beleidsintensiveringen in het inspectietoezicht hebben tot doel deze mechanismen te versterken.

Empirie

Een overzicht van empirische evaluaties van de effecten van inspectietoezicht wordt gepresenteerd in Webbink et al. (2009). Deze studies richten zich op het effect van een belangrijk instrument van dit toezicht, namelijk de schoolbezoeken. Ladd (2010) vergelijkt het accountability systeem in de VS met het inspectiesysteem in Nederland en Nieuw-Zeeland en vindt dat prestaties in deze laatste twee (inspectie)landen hoger zijn dan in de VS. Het blijft echter de vraag in hoeverre deze verschillen het gevolg zijn van het bestaan van de Onderwijsinspectie. Een aantal studies in Engeland (Cullingford et al., 1999; Wilcox en Gray, 1996; Shaw et al., 2003) vindt negatieve effecten van inspectietoezicht, maar deze studies zijn mogelijk vertekend door selectie-effecten (inspecteurs die een voorkeur hebben voor het bezoeken van zwakkere scholen). De internationale empirie waarbij gebruik wordt gemaakt van een (quasi-)experimenteel onderzoeksdesign is dus nog erg dun.

Luginbuhl et al. (2009) onderzoeken de invloed van bezoeken van de onderwijsinspecteurs in het Nederlandse primair onderwijs. Zij maken hierbij gebruik van een aselechte steekproef van scholen die door de Inspectie worden bezocht in het kader van het jaarlijkse onderwijsverslag. De bezoeken hebben positieve effecten op de onderwijsprestaties. Een inspectiebezoek doet de toetsscores met 0,01 tot 0,04 standaarddeviaties toenemen. Voor de Nederlandse onzekerheidsvariant gaan we uit van het gemiddelde effect uit deze studie.

Table 4.9 Onzekerheidsvarianten Inspectietoezicht

	Literatuurvariant	NL variant
Effect op toetsscores	onbekend	+ 0,025 SD
Doelgroep	primair onderwijs	primair onderwijs

Illustratie doorrekening

We beschouwen een interventie waarbij 25 miljoen euro extra wordt vrijgemaakt voor de inspectie van het primair onderwijs. De huidige middelen besteed aan de Inspectie van het Onderwijs zijn 59 miljoen euro.¹³ Aangezien bij deze bekostiging gemiddeld één op de vier scholen een bezoek van de schoolinspectie krijgt, nemen we aan dat met het staande beleid reeds een kwart van de doelgroep is bereikt. Dit betekent dat per 10 miljoen euro aan middelen een effect van $0.025/(4 \cdot 5,9) = 0,001$ standaarddeviatie op toetsscores wordt gerealiseerd.

De effectiviteit van het inspectiebezoek hangt van de frequentie van het bezoek af. Bij een te hoge frequentie zijn er afnemende meeropbrengsten en zullen scholen hun beleid en prestaties niet meer verbeteren. Bij een te lage bezoeksfrequentie zal er geen effect zijn op de schoolprestaties. We nemen aan dat schoolinspectiebezoeken effectief blijven als ze twee maal

¹³ Bron:

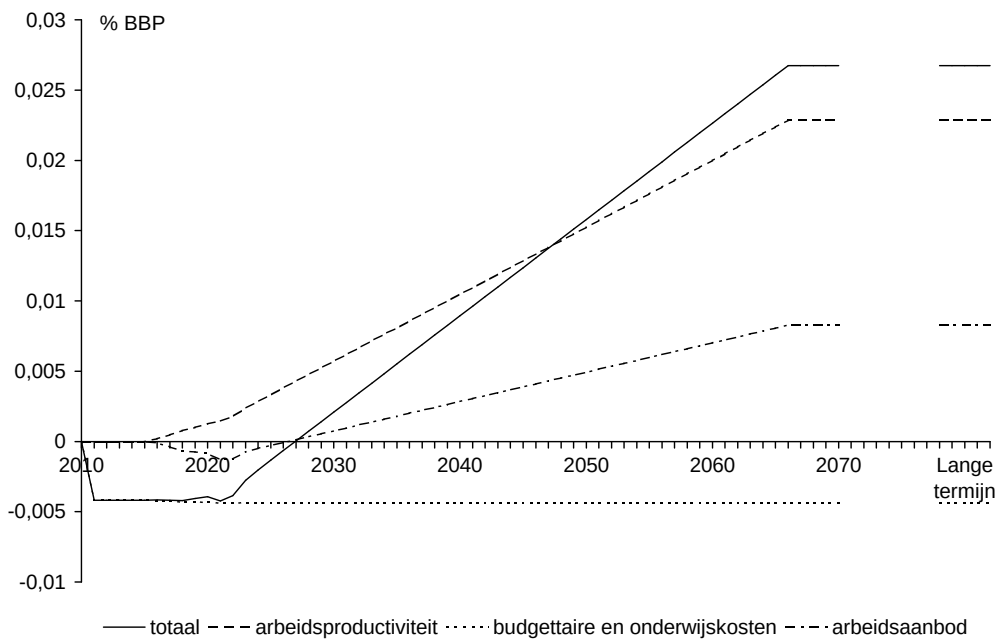
http://rijksbegroting.minfin.nl/2010/voorbereiding/begroting,kst132824b_24.html

zo vaak plaatsvinden. Derhalve kan er op de onderwijsinspectie maximaal 60 miljoen geïntensiveerd worden.

Uitkomsten

Figuur 4.7 presenteert de uitkomsten van de intensivering op de onderwijsinspectie. Ondanks dat het aantal inspecties met ingang van 2011 omhoog gaat, duurt het vier tot vijf jaar voordat de arbeidsproductiviteit begint te stijgen. Leerlingen doorlopen het primair onderwijs en vervolgens een vervolgopleiding, omdat zij leerplichtig zijn tot en met 16 jarige leeftijd. Op de lange termijn leidt deze interventie tot een stijging van 0,0% van het bbp.

Figure 4.7 Uitwerking Nederlandse variant voor uitbreiding Inspectie van het Onderwijs



4.2.6 Scholing en begeleiding onderwijspersoneel

Leraren spelen een sleutelrol in het onderwijs. Empirisch onderzoek laat zien dat de kwaliteit van leraren van groot belang is voor de onderwijsprestaties (Rivkin et al., 2005). Beleid gericht op scholing van leraren heeft als doel de kwaliteit van leraren te verbeteren en daarmee de prestaties van leerlingen te bevorderen. Ook laat de empirie zien dat een gebrek aan ervaring met name in de eerste twee jaar van de beroepsloopbaan van invloed kan zijn op de leerprestaties van leerlingen (Hanushek et al., 1998; Kane et al. (2008)). Na twee jaar in het onderwijs blijkt ervaring nog maar weinig verschil te maken. Beleidsvoorstellen die zich richten op mentoring of coaching van startende docenten kunnen bijdragen aan de effectiviteit van docenten in deze eerste jaren.

Het huidige beleid in Nederland is vooral gericht op scholing van leraren. Naast de reguliere

scholingsbudgetten voor training van onderwijspersoneel is in 2008 de zogenoemde 'lerarenbeurs' geïntroduceerd. Deze biedt leraren de mogelijkheid om, zonder inmenging van de school, een beurs te krijgen voor het volgen van een opleiding gericht op het verkrijgen van een hogere of bredere kwalificatie (een andere en/of hogere onderwijsbevoegdheid) of op het verwerven van specifieke bekwaamheden.

Empirie rondom scholing van onderwijspersoneel

Cornet et al. (2006) concluderen dat scholing van leraren een doeltreffend instrument is om de prestaties van leraren te verbeteren. Een overzicht van de literatuur wordt gegeven in Webbink et al. (2009). De empirische studies wijzen op positieve effecten van scholing van leraren; gemiddeld wordt gevonden dat scholing leidt tot een verbetering van de toetsscores met 0,07 standaarddeviatie. Deze gemiddelde toename in de toetsscores is met onzekerheid omgeven. Het resultaat wordt sterk bepaald door een economische studie van Angrist en Lavy (2001), die zich richten op de effecten van training in instructietechnieken op de prestaties van leerlingen in het basisonderwijs. Op niet-religieuze scholen nemen de toetsscores met 0,25 standaarddeviatie toe. Voor religieuze scholen worden echter geen statistisch significante effecten gevonden. Jacob en Lefgren (2004b) proberen het effect van scholing te identificeren door gebruik te maken van een regeling in Chicago die slecht presterende scholen tijdelijk extra middelen toekende, die onder meer bestemd waren voor scholing van leraren. Het hoofdresultaat van de studie is dat investeringen in ontwikkeling van personeel de onderwijsprestaties van leerlingen niet hebben verbeterd. De auteurs merken op dat dit mogelijk komt door de relatief beperkte omvang van de investering.

Het is de vraag in hoeverre beleid de scholingsdeelname van leraren kan vergroten. Van der Steeg et al. (2010) analyseren de effecten van de introductie van de lerarenbeurs in Nederland op de (hoger) onderwijsdeelname door docenten. Van elke 10 beurzen die verstrekt worden voor een hogeronderwijsopleiding blijkt er gemiddeld 1 beurs te worden gebruikt voor een opleiding die anders niet gevolgd zou zijn; 1 op de 10 beurzen wordt helemaal niet gebruikt, en 8 van de 10 beurzen worden gebruikt voor opleidingen die anders ook gevolgd zouden zijn. Er vindt verdringing plaats van zowel reguliere scholingsbudgetten van scholen als van eigen bijdragen van docenten. De additionaliteit in scholingsdeelname neemt toe met de lengte van de opleiding. Voor hogeronderwijsopleidingen van meer dan een jaar is circa 2 op de 10 toegewezen beurzen additioneel, terwijl er bij opleidingen tot en met 1 jaar geen aanwijzingen zijn voor significante effecten op scholingsdeelname. Verder wordt opgemerkt dat de additionaliteit verder vergroot zou kunnen worden door een andere vormgeving. Zo werd de lerarenbeurs in de huidige vorm ook ter beschikking gesteld voor docenten die al met een opleiding waren begonnen, wat de additionaliteit aan scholingsdeelname niet ten goede komt. Informatie over de mogelijkheden om scholingsdeelname door beleid te beïnvloeden kan ook worden ontleend aan een studie van

Messer en Wolter (2009), die de effecten van scholingsvouchers voor volwassenen op de scholingsdeelname in Zwitserland onderzoeken. Hierbij werden vouchers (met verschillende waardes) random verdeeld over volwassenen van alle opleidingsniveaus in de leeftijd van 20 tot 60 jaar. De auteurs vinden voor werkenden een procentuele stijging van de scholingsdeelname van ongeveer 15 procent. De additionaliteit bleek hier ook af te hangen van de waarde van de vouchers: het verstrekken van de grootste vouchers deed de scholingsdeelname met 26 procent toenemen.

De onzekerheidsvarianten in tabel 4.10 presenteren de effecten van scholing van leraren op de toetsscores van leerlingen. Specifiek in de Nederlandse beleidscontext heeft de studie van Van der Steeg et al. (2010) laten zien dat extra middelen voor scholing niet direct hoeven te leiden tot additionele scholingsdeelname. De additionaliteit bij de opleidingen met een langere opleidingsduur was 20%, waarbij wordt opgemerkt dat een hogere additionaliteit te behalen valt door een betere vormgeving. Op basis hiervan gaan wij er vanuit dat extra middelen voor scholing van docenten zullen leiden tot een additionaliteit in scholingsdeelname van 30%. Deze additionaliteit komt ongeveer overeen met het grootste gevonden effect in de studie van Messer en Wolter (2009). Op deze manier komen we tot een onzekerheidsvariant voor Nederland waarin extra middelen voor scholing van leraren leiden tot een toename in de toetsscores met 0,02 standaarddeviatie voor de bereikte leerlingen. Dit effect volgt uit vermenigvuldiging van het gemiddelde effect uit de literatuur met de additionaliteit van 30%.

Table 4.10 Onzekerheidsvarianten scholing van leraren

	Literatuurvariant	NL variant
Effect op toetsscores	+ 0,07 SD	+ 0,02 SD
Doelgroep	PO, VO, MBO	PO, VO, MBO

Illustratie scholing van onderwijspersoneel

Wij beschouwen een uitbreiding van de huidige lerarenbeurs, met het doel alle leraren in het primair-, voortgezet en middelbaar beroepsonderwijs de mogelijkheid te bieden een tweejarige bachelor- of masteropleiding te volgen. We nemen aan dat leraren die meedoen met de huidige regeling ook kunnen deelnemen aan deze regeling. De jaarlijkse kosten per leraar zijn 1.920 euro. De inzet van middelen bedraagt 430 miljoen euro, waarmee alle leerlingen in het primair-, voortgezet- en middelbaar beroepsonderwijs worden bereikt. Een toelichting is te vinden in het kader “Alle leraren in po, vo en mbo een twee-jarige master”.

Het is vanzelfsprekend dat 430 miljoen euro een bovengrens vormt van de in te zetten middelen. Additionele middelen kunnen worden gebruikt om het bereik te vergroten naar het hoger onderwijs. Echter, de empirie is niet toereikend om een effect voor deze doelgroep te identificeren.

Alle leraren in po, vo en mbo een tweejarige master

De kosten van scholing voor onderwijspersoneel bestaan enerzijds uit cursus- of collegegeld, aanschaffen van studiemiddelen en reiskosten, en anderzijds uit vergoedingen voor de werkgever om vervanging tijdens studieverlof te bekostigen. Het referentiekader waaruit de kosten worden bepaald is de lerarenbeurs. De gemiddelde uitgekeerde lerarenbeurs, voor bachelor- en master trajecten, is 9.600 euro per docent.^a Dit bedrag is inclusief de kosten van het vervangen van de docent.

Wij nemen aan dat de kennis en vaardigheden opgedaan in deze opleiding in 10 jaar lineair wordt afgeschreven. Met andere woorden, een leraar is na afronding van zijn scholing gemiddeld 5 jaar lang productiever. Als we aannemen dat een leraar om deze reden eens in de vijf jaar geschoold mag worden, dan zijn de jaarlijkse kosten van deze interventie 1.920 euro per docent.

Voor het aantal bereikte leerlingen per docent wordt uitgegaan van de verhoudingen gegeven in onderstaande tabel. Het maximum aan te besteden middelen, in totaal 430 miljoen euro, is ook opgenomen in deze tabel.

Scholing onderwijspersoneel: bereik

	aantal leerlingen ^b	leraar/leerling ratio ^b	maximaal aantal middelen
primair onderwijs	1 552 900	1 : 13	225 mln euro
voortgezet onderwijs	900 100	1 : 12	149 mln euro
middelbaar beroepsonderwijs	484 100	1 : 17	56 mln euro

^b Bron: OC&W (2008)

^a Volgt uit eigen berekeningen met de data uit Van der Steeg et al. (2010)

Uitkomsten scholing van onderwijspersoneel

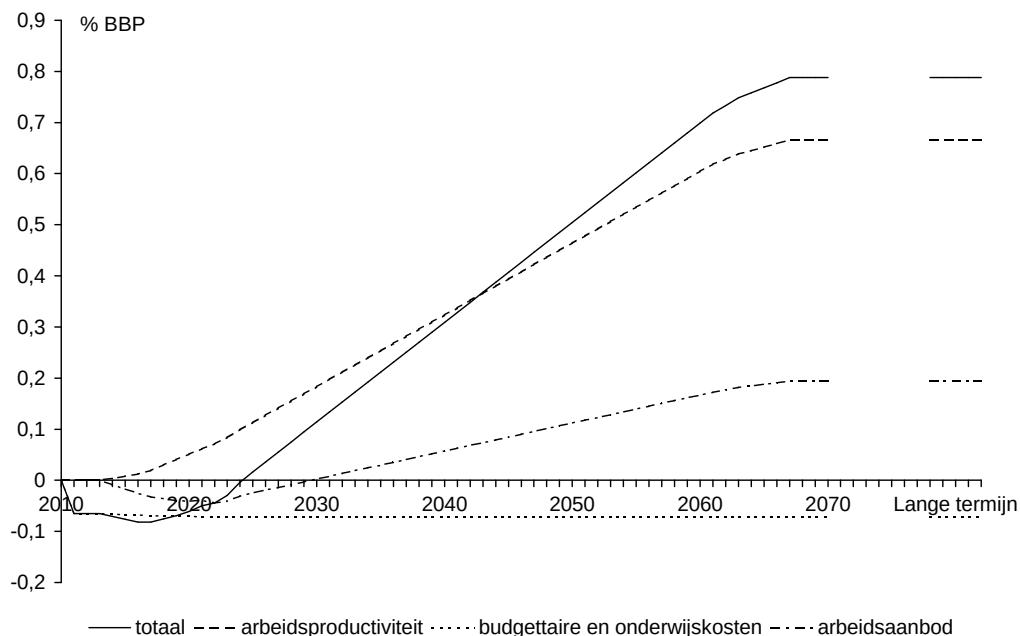
Figuur 4.8 laat de berekende effecten zien van de maatregel waarbij het maximum aan middelen (430 miljoen euro) structureel wordt ingezet vanaf 2011. De duur van de opleidingen is 2 jaar, zodat vanaf 2013 de arbeidsproductiviteit van de beroepsbevolking begint te stijgen. Op de lange termijn leidt de structurele inzet van 430 miljoen euro voor scholing van leraren tot een stijging van het bbp-niveau met 0,8%.

In deze illustratie is aangenomen dat het lerarencorps in het po, vo en mbo niet van omvang verandert en dat iedere docent een beurs krijgt.

Empirie rondom mentoring van startende docenten

Twee recente studies vinden positieve effecten van mentoring van startende docenten. Rockoff (2008) onderzoekt de effecten van mentoring van beginnende docenten (met minder dan één jaar ervaring) in het basisonderwijs in New York op leerprestaties van leerlingen en de kans op vertrek uit het onderwijs. Hij vindt dat extra beschikbare uren mentoring voor leraren leidt tot significant betere leerprestaties op taal en rekenen. Ook blijkt dat mentoring een positief effect heeft op de kans dat de docent het schooljaar afmaakt. Mentoring kan hiermee ook het behoud van startende leraren in het onderwijs positief beïnvloeden. Bressoux et al. (2009) onderzoeken de effecten van training van startende leraren in Frankrijk en vinden ook positieve effecten op de

Figure 4.8 Uitwerking Nederlandse variant voor scholing van onderwijspersoneel in primair-, voortgezet- en middelbaar beroepsonderwijs



onderwijsprestaties van leerlingen.

Gemiddeld genomen doet mentoring van startende docenten de toetsscores met 0,14 standaarddeviatie toenemen. Hierbij wordt uitgegaan van programma waarin docenten ongeveer 35 uur mentoring krijgen. Vanwege het nog beperkte aantal studies binnen dit thema, zijn de resultaten uiteraard met de nodige onzekerheid omgeven. Wij zien geen reden voor de Nederlandse onzekerheidsvariant af te wijken van het gemiddelde van de literatuur.

Table 4.11 Onzekerheidsvarianten Mentoring startende docenten

	Literatuurvariant	NL variant
Effect op toetsscores	+ 0,14 SD	+ 0,14 SD
Doelgroep	PO, VO, MBO	PO, VO, MBO

Illustratie mentoring van startende docenten

We beschouwen een fictieve interventie waarin startende docenten in het primair, voortgezet- en middelbaar beroepsonderwijs twee jaar lang worden begeleid door een speciaal daarvoor aangestelde mentor. Voor de huidige instroom van startende docenten in deze sectoren, bedraagt de inzet van middelen 73,8 miljoen euro. Dit bedrag is tevens een bovengrens voor de inzet van middelen, waarbij wordt aangenomen dat de instroom van beginnende docenten gelijk blijft. De inzet van additionele middelen krijgt 'effect onbekend'. Een toelichting is te vinden in het tekstkader "Tweejarige begeleiding van startende docenten".

Tweejarige begeleiding van startende docenten

We nemen aan dat de interventie is ingericht zoals in de interventie uit Rockoff (2008). De mentor-leraar ratio is in deze studie gelijk aan 1 : 15. Het programma in Rockoff (2008) duurt twee jaar, waarin beginnende docenten 35 uur scholing krijgen.

Het aantal startende docenten in het primair-, voortgezet- en middelbaar beroepsonderwijs zijn gegeven in onderstaande tabel. De leraar-leerling ratio's en het aantal leerlingen dat wordt bereikt middels deze interventie zijn ook gegeven.

Startende docenten	Aantal startende leraren	leraar-leerling ratio	aantal te bereiken leerlingen
primair onderwijs	5.800	1:15	87 dzd
voorgezet onderwijs	4.700	1:13	61 dzd
middelbaar beroepsonderwijs	1.800	1:19	34 dzd

Maximaal kunnen 12 300 leraren per jaar worden bereikt. Hiervoor zijn 820 mentoren nodig. De jaarlijkse kosten van een mentor zijn 90 duizend euro. Dit impliceert een maximale inzet van middelen van 73,8 miljoen euro.

Uitkomsten mentoring van startende docenten

Figuur 4.9 presenteert de uitkomsten van de fictieve maatregel. De inzet van middelen vindt plaats vanaf 2011 en blijft constant (in prijzen van 2010). Na twee jaar resulteren de eerste effecten op de arbeidsproductiviteit van de beroepsbevolking. Dit is vanwege de tweejarige duur van het mentoring-programma. Op de lange termijn heeft, bij gelijke instroom van startende docenten en een constante bevolkingsgrootte, 17% van de bevolking onderwijs genoten van een beginnende docent die ondersteuning heeft gekregen van een mentor. De structurele inzet van 73,8 miljoen euro leidt dan tot een stijging van het bbp met 0,6%.

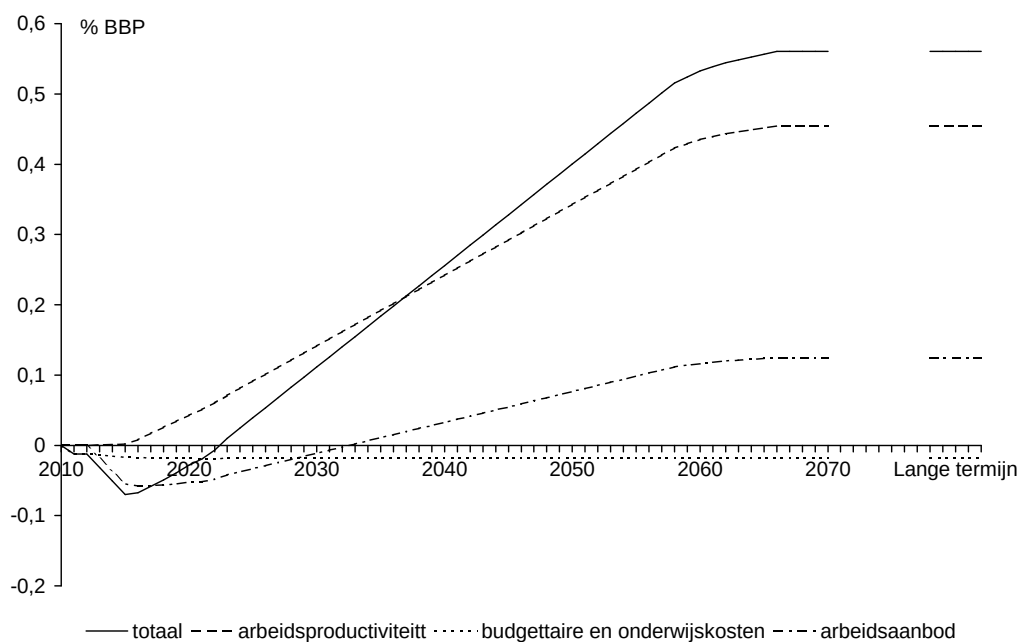
4.3 Kwantificering institutionele beleidsthema's

4.3.1 Centrale examinering

Centrale examens voorzien in een objectieve meting van prestaties en kunnen zorgen voor transparantie over en verantwoording van onderwijsprestaties. Dit geeft het schoolpersoneel prikkels om de prestaties van leerlingen te verbeteren. Zonder de juiste prikkels zouden leraren mogelijk een voorkeur kunnen hebben voor de meest vertrouwde lesmethoden in plaats van de meest veelbelovende. Centrale examens geven ook prikkels voor leerlingen om te leren. Wanneer absolute normen ontbreken en er alleen sprake is van relatieve prestatiemeting kunnen leerlingen baat hebben bij een gezamenlijk afspraak om niet te studeren.

Nederland kent centrale examinering in het primair onderwijs in de vorm van de CITO-eindtoets. Hoewel het overgrote deel van de basisscholen deelneemt aan deze toets, is deelname niet verplicht. Ongeveer 15% van alle scholen doet op dit moment niet mee

Figure 4.9 Uitwerking Nederlandse onzekerheidsvariant voor het begeleiden van startende docenten in het primair, voorgezet- en middelbaar beroepsonderwijs



(Stroucken et al., 2008). Ook het voortgezet onderwijs kent centrale examens. Het eindresultaat in het examenjaar wordt voor de helft bepaald door het centraal examen en voor de helft door schoolexamens. In het mbo zullen centrale examens voor taal en rekenen geleidelijk worden ingevoerd vanaf 2012. Nieuwe beleidsvoorstellen kunnen zich bijvoorbeeld richten op het verplicht stellen van deelname aan de CITO-eindtoets in het po, het afschaffen van de schoolexamens in het vo, of aanpassingen bij de invoering van centrale examinering in het mbo.

Empirie

In Webbink et al. (2009) wordt een overzicht gegeven van de empirische literatuur op het gebied van centrale examens. In het algemeen worden positieve effecten gevonden van centrale examens op de onderwijsprestaties van leerlingen. De empirische studies bevestigen daarmee de theoretische verwachtingen. Het onderzoek betreft vooral studies waarin de prestaties van leerlingen in landen (of regio's) met een centraal examen worden vergeleken met de prestaties van leerlingen in landen (regio's) zonder een centraal examen (Bishop, 1997; Woessmann, 2003; Fuchs en Woessmann, 2007; Jürges et al., 2005).

De studie van Jürges et al. (2005) heeft het sterkste onderzoeksdesign. Zij maken gebruik van verschillen in regionale schoolwetgeving in Duitsland en vinden in de door hun geprefereerde specificatie dat centrale examens de toetsscores met 0,13 standaarddeviatie doet toenemen. Dit resultaat is in overeenstemming met de uitkomsten uit de overige studies. Woessmann (2003) en Fuchs en Woessmann (2007) vinden effecten tussen de 0,10 en 0,20

standaarddeviatie. Bishop (1997) vindt zelfs nog wat grotere effecten.

Voor Nederland kwalificeren wij de invoering (uitbreiding) van centrale examens in het vo en het mbo als ‘effect onbekend’. In het vo wordt momenteel de helft van het resultaat in het examenjaar bepaald door een centraal examen en de helft door de schoolexamens. Een uitbreiding van het gewicht van het centraal examen in het eindresultaat zou, gezien het voorstaande, tot betere onderwijsprestaties kunnen leiden. Aan de andere kant zorgt de huidige methodiek voor meerdere toetsmomenten, waardoor een groter deel van de stof getoetst kan worden. Bovendien worden mogelijkheden tot manipulatie van de schoolexamens door scholen ingeperkt door de verplichting de schoolexamens ook door een externe docent (van een andere school) na te laten kijken en door controle van de inspectie op verschillen in resultaten tussen de schoolexamens en het centraal examen. Het is onbekend of een uitbreiding van het gewicht van een centraal examen in het eindresultaat, of het afschaffen van de schoolexamens, een verbetering betekent ten opzichte van deze bestaande praktijk.

Binnen het mbo bestaan veel verschillende beroepsopleidingen, waardoor de ontwikkeling en invoering van een centrale toets die aansluit bij de specifieke leerdoelen van deze opleidingen moeilijk is.

In het po is een uitbreiding van de centrale examinering mogelijk door bijvoorbeeld de CITO eindtoets verplicht te stellen, waardoor ook de resterende 15% van de leerlingen aan een centraal examen onderworpen zal worden. Tabel 4.12 presenteert de onzekerheidsvarianten.

Table 4.12 Onzekerheidsvarianten Centrale examens

	Literatuurvariant	NL variant
Effect op toetsscores	+ 0,13 SD	+ 0,13 SD
Doelgroep	primair- en voortgezet onderwijs	1/6 van het primair onderwijs

Illustratie doorrekening

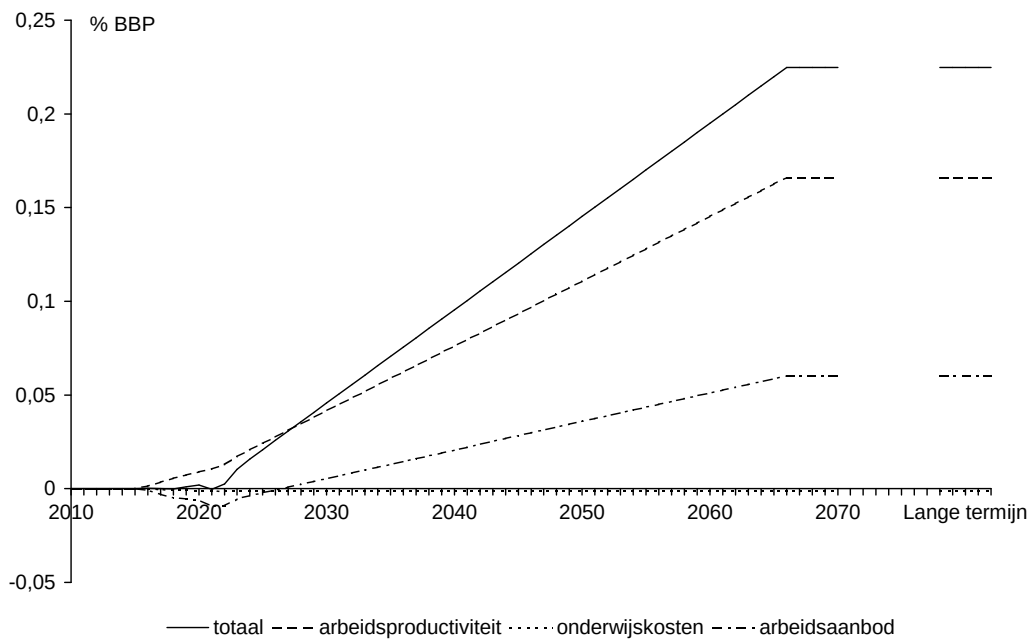
De interventie betreft ‘het verplicht stellen van deelname aan de CITO-toets voor alle leerlingen in het primair onderwijs’. Dit betekent dat de resterende 15% van de leerlingenpopulatie in het primair onderwijs zal worden onderworpen aan de CITO-toets. Het verplichten van de eindtoets in het primair onderwijs kost, op verwaarloosbare uitvoeringskosten na, niets. De toets zelf is in principe al ontwikkeld.

Figuur 4.10 geeft de uitkomsten weer van het verplicht stellen van de CITO-toets in het basisonderwijs vanaf 2011. Deze interventie leidt tot een stijging van het bbp met 0,2% op de lange termijn.

Wij nemen aan dat de eerste effecten waarneembaar zijn drie jaar na verplichte deelname aan de CITO-toetsen. Wij achten het onwaarschijnlijk dat in het jaar van invoering de leerlingen direct beter gaan presteren. In de praktijk vindt de voorbereiding op de CITO-toets plaats in de

laatste twee jaar van het basisonderwijs. Deze aanname heeft geen effect op het lange termijn resultaat.

Figure 4.10 Uitwerking Nederlandse variant voor verplichten CITO-toets in basisonderwijs



4.3.2 Informatievoorziening

De vrije schoolkeuze in Nederland kan mogelijk leiden tot meer ongelijkheid wanneer gezinnen met een lagere sociaaleconomische achtergrond meer kiezen voor scholen met een lagere kwaliteit. Het verschaffen van toegankelijke, eenvoudige informatie over de kwaliteit van scholen kan het keuzegedrag van de ouders beïnvloeden. Doel hiervan is dat leerlingen met een lagere sociaaleconomische achtergrond op kwalitatief betere scholen terecht komen, wat een positief effect kan hebben op hun onderwijsprestaties. Het Nederlandse beleid besteedt aandacht aan de mogelijke gevolgen van de vrije schoolkeuze voor achterstandsleerlingen. Zo is er recent beleid in de gemeente Nijmegen om via een centraal meldpunt te komen tot een andere spreiding van leerlingen over scholen.

Mogelijke nieuwe beleidsvoorstellen kunnen zich richten op transparante informatievoorziening aan gezinnen met een lage sociaal economische achtergrond over de kwaliteit van de scholen in hun omgeving.

Empirie

De recente studie van Hastings en Weinstein (2008) laat zien dat ouders met een lagere sociaal-economische achtergrond vaker scholen kiezen met een lagere kwaliteit. De auteurs onderzoeken vervolgens de invloed van eenvoudig toegankelijke informatie over prestaties van

scholen in de buurt op het keuzegedrag van ouders met een lager of midden inkomen. De kinderen van de ouders die informatie hebben ontvangen, blijken gemiddeld uit te komen op scholen waar de toetsscores gemiddeld 0,05 standaarddeviatie hoger liggen. Als gevolg hiervan blijken de onderwijsprestaties van deze leerlingen te verbeteren met 0,02 standaarddeviatie. In dit experiment uit de VS sorteerden overinschrijvingen op bepaalde scholen het effect dat niet elke leerling kon worden toegelaten tot de school van zijn of haar eerste keus. Als alle leerlingen zouden zijn toegelaten tot de school van hun voorkeur, zouden de leerlingen van ouders die informatie hebben gehad terecht zijn gekomen op scholen waarvan de gemiddelde toetsscore 0,10 standaarddeviatie hoger lag.

Voor de Nederlandse onzekerheidsvariant nemen we aan dat er geen sprake zal zijn van overinschrijving. Wanneer je voor een betere school kiest, word je in principe ook toegelaten. In dat geval zal informatievoorziening aan gezinnen met een lagere sociaaleconomische achtergrond de toetsscores voor de leerlingen uit deze groep met 0,04 standaarddeviatie kunnen verbeteren.

Table 4.13 Onzekerheidsvarianten Informatievoorziening kwaliteit scholen

	Literatuurvariant	NL variant
Effect op toetsscores	+ 0,02 SD	+ 0,04 SD
Doelgroep	Leerlingen met lage sociaaleconomische achtergrond (po, vo)	Leerlingen met lage sociaaleconomische achtergrond (po, vo)

Illustratie

Wij beschouwen een interventie waarbij aan ouders van leerlingen in het primair onderwijs met een lage sociaaleconomische achtergrond eenvoudige, toegankelijke informatie over de kwaliteit van scholen wordt verstrekt. Dit kan bijvoorbeeld een samenvatting van de onderwijsprestaties van scholen in de buurt zijn.

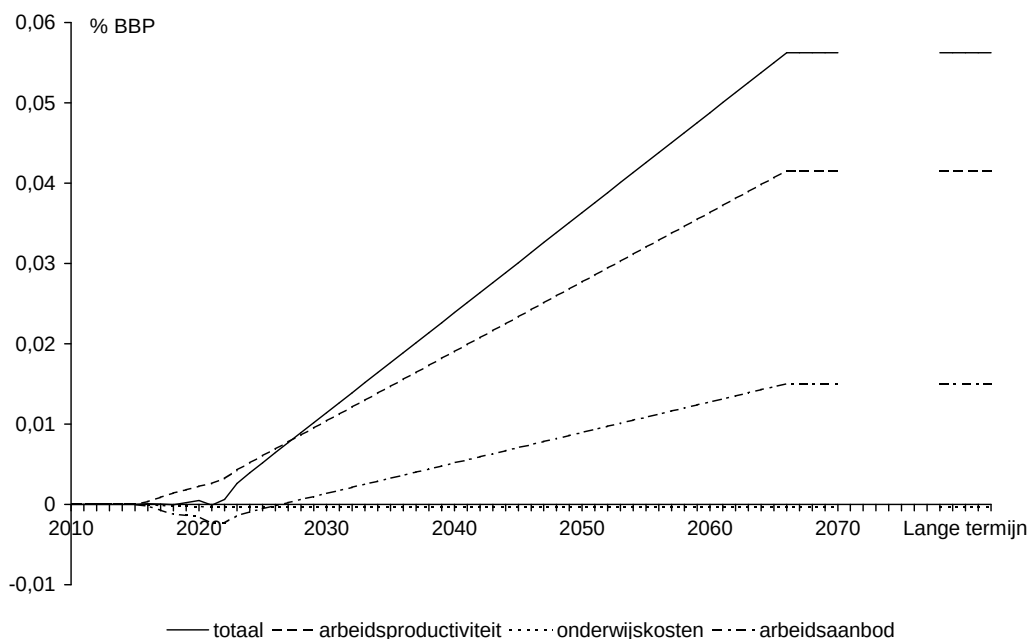
Het aantal leerlingen afkomstig uit huishoudens met een lage sociaaleconomische status is moeilijk vast te stellen. Het meten van sociaaleconomische status is geen eenvoudige opgave omdat de definitie niet eenduidig is. Een ondergrens van het aantal leerlingen uit huishoudens met een lage sociaaleconomische status wordt gegeven door het aantal gewichtenleerlingen.¹⁴ Het aantal leerlingen die tot de doelgroep behoren waarmee wij rekenen, is 25% van het totaal aantal leerlingen in het primair onderwijs.

¹⁴ In het basisonderwijs krijgen leerlingen met een potentiële onderwijsachterstand op basis van bepaalde criteria een gewicht. Deze gewichten bepalen hoeveel geld een basisschool krijgt voor het wegwerken van onderwijsachterstanden. De gewichtenregeling kent geldbedragen toe aan leerlingen op basis van het opleidingsniveau van de ouders. Sinds het schooljaar 2009-2010 is de regeling ook gekoppeld aan het postcodegebied waarbinnen een school is gevestigd. Het aantal gewichtenleerlingen onderschat dus de doelgroep. Zie ook OC&W (2008).

De verbetering van de informatievoorziening in de vorm van een pamflet in de brievenbus, of het opzetten van een website, heeft in deze analyse verwaarloosbare kosten.

Figuur 4.11 presenteert de uitkomsten van de fictieve interventie. Het invoeren van deze institutionele maatregel in 2011, resulteert in 2015 in een eerste effect. De leerlingen die als gevolg van de interventie een ‘beter’ presterende basisschool uitzoeken, komen op zijn vroegst pas na 4 jaar op de arbeidsmarkt. Vanaf 2015 begint het gemiddeld opleidingsniveau van de beroepsbevolking te stijgen, en daarmee ook de arbeidsproductiviteit. Op de lange termijn heeft 25% van de beroepsbevolking profijt gehad van deze interventie, welke leidt tot een stijging van het bbp met 0,1%.

Figure 4.11 Uitwerking Nederlandse variant voor invoering een informatiesysteem voor het primair onderwijs



4.3.3 Accountability

In de VS wordt sinds enkele jaren gebruik gemaakt van zogenoemde ‘school accountability’ systemen. Dit zijn systemen waarin de prestaties van scholen worden beoordeeld volgens van tevoren vastgestelde standaarden en waarin consequenties worden verbonden aan het niet halen van bepaalde normen. Dit kan inhouden dat een deel van de financiering van scholen afhankelijk wordt gemaakt van de prestaties. Het doel van een accountability systeem is dat scholen geprikkeld worden om de kwaliteit van het onderwijs te verbeteren.

Het Nederlands beleid heeft in toenemende mate belangstelling voor ‘resultaatgericht onderwijs’. Nieuwe beleidsvoorstellen richten zich bijvoorbeeld op verschillende vormen van prestatiebekostiging, waarbij een deel van de bekostiging van scholen afhankelijk wordt gemaakt van hun prestaties.

Empirie

Webbink et al. (2009) presenteren een samenvatting van de internationale evidentie over de invloed van accountability systemen op onderwijsprestaties. Dit betreft vooral studies uit de VS. Hieruit blijkt dat de invoering van een accountability systeem een sterk positief effect heeft op de toetsscores van leerlingen (Hanushek en Raymond, 2004; Jacob, 2005; Dee en Jacob, 2009). Dit gaat om accountability systemen met relatief sterke prikkels waarbij beloningen en sancties gekoppeld worden aan de jaarlijkse onderwijsprestaties. Zo kunnen zwakke prestaties leiden tot een verlaging van de bekostiging en kunnen zwak presterende scholen onder toezicht worden gesteld of gedwongen worden te reorganiseren. Wanneer we het gemiddelde van deze studies nemen, doet accountability de toetsscores met 0,25 standaarddeviatie toenemen. De variatie in gevonden effectgroottes is gering: alle studies vinden effecten tussen de 0,20 en 0,30 standaarddeviatie. Wel zijn verschillen waarneembaar tussen verschillende groepen leerlingen. Zo concluderen Hanushek en Raymond (2004) bijvoorbeeld dat 'witte' studenten meer baat hebben bij dit systeem dan 'zwarte' studenten en, derhalve, het accountability beleid de 'black-white gap' vergroot.

Er zijn ook studies die suggereren dat school accountability systemen strategisch gedrag in de hand kunnen werken. Zo zijn er aanwijzingen dat de invoering van accountability beleid kan leiden tot *teaching-to-the-test* en tot meer verwijzingen van leerlingen naar trajecten die niet getest werden, zoals het speciaal onderwijs, zie Jacob (2005). Ook worden aanwijzingen gevonden voor fraude door docenten en het strategisch inzetten van schorsingen om zwakke leerlingen buiten de tests te houden, zie Jacob en Levitt (2003) en Figlio (2006).

Voor Nederland maken we gebruik van een studie van Chorny en Webbink (2010), die de effecten van accountability beleid in Amsterdam onderzoeken. Hierbij werd onder andere de toekenning van gemeentelijke middelen voor achterstanden (de zogenoemde GOA-middelen) afhankelijk gemaakt van de prestaties van de scholen op de CITO eindtoets. Dit is de enige Nederlandse studie die zich richt op de effecten van accountability beleid. De auteurs vinden dat het accountability beleid heeft geleid tot een toename in de scores op de CITO eindtoets met 0,25 standaarddeviatie. Om te corrigeren voor mogelijke effecten van *teaching-to-the-test* worden ook de resultaten op de zogenoemde PRIMA toetsen onderzocht (vergelijkbare toetsen die worden afgenomen in de groepen 2, 4, 6). Deze toetsen maakten geen onderdeel uit van de beoordeling van de scholen, waardoor de resultaten minder kwetsbaar zijn voor *teaching-to-the-test*. Ook op deze toetsen worden positieve effecten gevonden van de invoering van het beleid: de scores op de PRIMA toetsen nam gemiddeld met 0,15 standaarddeviatie toe. Bij deze resultaten dient wel te worden opgemerkt dat de CITO-scores in Amsterdam voor de invoering van het beleid behoorlijk onder het landelijk gemiddelde lagen, waardoor het wellicht relatief eenvoudiger was om grote effecten te realiseren. Bovendien maakte de prestatiebekostiging onderdeel uit van een groter gemeentelijk beleidsplan in Amsterdam ter bevordering van de onderwijsprestaties in het primair onderwijs.

Op basis van bovenstaande nemen wij in de Nederlandse onzekerheidsvariant aan dat de invoering van accountability beleid de toetsscores met 0,05 standaarddeviatie doet toenemen. We betrachten hier enige voorzichtigheid om recht te doen aan de moeilijkheden die gepaard kunnen gaan met de invoering van accountability beleid, zoals verschillende vormen van strategisch gedrag van scholen. De optimale of minimaal benodigde mate van prestatiebekostiging is niet bekend. In Amsterdam bedroegen de GOA-middelen 3 procent van de totale financiering. De resultaten van accountability beleid zijn sterk afhankelijk van de juiste randvoorwaarden, zoals bijvoorbeeld het bestaan van objectieve standaarden voor de onderwijsprestaties. Om deze reden kwalificeren wij de invoering van accountability beleid in het mbo, waar nog geen centrale toetsing bestaat, als 'effect onbekend'. Tabel 4.14 presenteert de onzekerheidsvarianten.

Table 4.14 Onzekerheidsvarianten Accountability

	Literatuurvariant	NL variant
Effect op toetsscores	+ 0,25 SD	+ 0,05 SD
Doelgroep	primair- en voortgezet onderwijs	primair- en voortgezet onderwijs

Illustratie

We beschouwen de maatregel waarbij 3 procent van de financiering van scholen in basis- en voortgezet onderwijs afhankelijk wordt van hun prestaties. De kosten van dit beleid bestaan louter uit uitvoeringskosten. De interventie is immers een verschuiving van middelen. De uitvoeringskosten zijn, ten opzichte van de opbrengsten, verwaarloosbaar klein. In beide onderwijstypen wordt de hele leerlingpopulatie bereikt.

Uitkomsten

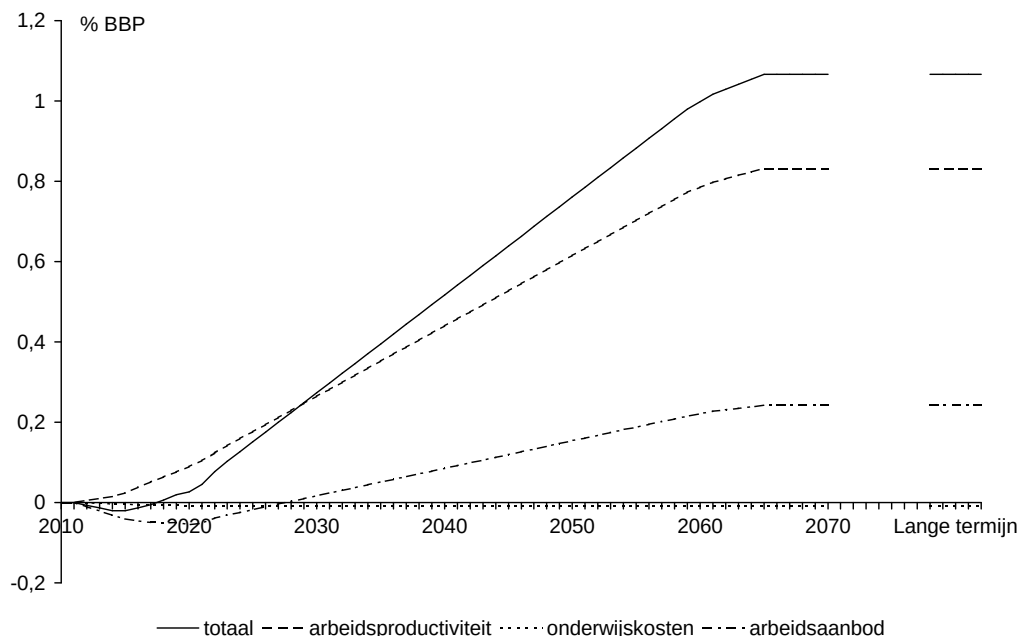
Figuur 4.12 presenteert de resultaten van deze maatregel. Ondanks het kleine effect op toetsscores per leerling, is het geaggregeerde effect enorm doordat de hele populatie leerlingen wordt bereikt. Op de lange termijn leidt dit beleid tot een stijging van het bbp met 1,1%.

4.3.4 Vroege selectie

Het Nederlandse onderwijs kent, vergeleken met andere landen, een vroege selectie van leerlingen. Direct na het verlaten van het basisonderwijs, op een leeftijd van ongeveer 12 jaar, stromen de leerlingen door in verschillende typen vervolgonderwijs. Deze vroege selectie heeft tot gevolg dat leerlingen verschillende niveaus van onderwijs krijgen aangeboden en in aanraking komen met verschillende medeleerlingen.

Op voorhand zijn de effecten van vroege selectie niet direct duidelijk. Er kan sprake zijn van een tradeoff tussen efficiëntie en ongelijkheid. Enerzijds kan vroege selectie leiden tot meer efficiëntie vanwege een grotere homogeniteit van de klassen. Anderzijds kan vroege selectie

Figure 4.12 Uitwerking Nederlandse variant voor invoering van accountability systeem in het primair- en voortgezet onderwijs



nadelig zijn voor relatief zwakkere leerlingen. Dit kan veroorzaakt worden door bijvoorbeeld peer group effecten of verschillen in kwaliteit van leraren, wanneer de betere leraren een voorkeur hebben voor het lesgeven aan betere klassen. Vroege selectie van leerlingen brengt meer onzekerheid met zich mee over de werkelijke mogelijkheden van een kind, waardoor een groter risico bestaat dat de leerling in een verkeerd type onderwijs geselecteerd wordt.

Volgens de OECD (2007) vormt de vroege selectie van leerlingen een belangrijke barrière voor de groei van de deelname aan het hoger onderwijs in Nederland. De Onderwijsraad (2007) stelt dat leerlingen vaak op het verkeerde spoor zitten in het voortgezet onderwijs en dat tweejarige gecombineerde brugklassen nodig zijn. Nieuwe beleidsvoorstellen binnen dit thema richten zich veelal op het verhogen van de leeftijd van selectie.

Empirie

Webbink et al. (2009) presenteren een overzicht van de internationale empirische literatuur die kijkt naar de gevolgen van een vroege selectie van leerlingen. Er bestaat geen eenduidig beeld van de effecten van vroege selectie op de kwaliteit van het onderwijs. Wel lijkt het empirisch onderzoek te wijzen op vergroting van de ongelijkheid in onderwijsprestaties (Hanushek en Woessmann, 2006) en versterking van de invloed van sociaaleconomische achtergrond van leerlingen op hun onderwijsuitkomsten (Brunello en Checchi, 2007; Schütz et al., 2008; Ammermüller, 2005).

Van Elk et al. (2009) onderzoeken de gevolgen van het tijdstip van selectie in het Nederlands

onderwijs voor de deelname aan (en het afronden van) het hoger onderwijs. Dit onderzoek richt zich op leerlingen die het basisonderwijs verlaten met een mavo (vmbo-tl) advies en vergelijkt leerlingen die in het voortgezet onderwijs direct instromen in een categorale klas met leerlingen die instromen in een gecombineerde brugklas. Deze laatste groep stelt de keuze voor het type vervolgonderwijs met één tot twee jaar uit. Het blijkt dat deze latere selectie de deelname aan het hoger onderwijs met 4%-punt doet toenemen, van 35 tot 39%. Dit komt overeen met een toename in de deelname met 11%.

We gebruiken deze studie als uitgangspunt voor de Nederlandse onzekerheidsvariant. Aangezien de internationale literatuur nog geen eenduidig beeld laat zien van de effecten van vroege selectie op onderwijsprestaties, zullen we deze alleen gebruiken voor de kwantificering van beleidsvoorstellen die in lijn zijn met de specifieke onderzoeksvraag in de Nederlandse studie (het uitstellen van het tijdstip van selectie door gecombineerde vmbo-havo brugklassen voor vmbo-tl leerlingen).

Table 4.15 Onzekerheidsvarianten Latere selectie

	Literatuurvariant	NL variant
Effect op doorstroom hoger onderwijs	onbekend	+ 11%
Doelgroep	VO	Leerlingen met VMBO-tl advies

Illustratie doorrekening

De maatregel betreft het “stimuleren van gecombineerde vmbo-havo klassen door middel van extra geld voor gecombineerde klassen”. Het verschil in de bekostiging voor een vmbo-leerling in een gecombineerde vmbo-havo brugklas van een scholengemeenschap ten opzichte van een leerling in een brugklas op een categorale vmbo is 7,5%, zie Onderwijsraad (2010). De doelgroep, leerlingen met een vmbo-tl advies die toch naar een categorale vmbo gaan, bestaat jaarlijks uit circa 5 500 leerlingen. De structurele inzet van middelen om deze groep leerlingen te bereiken is 3 miljoen euro. Hoe deze cijfers tot stand zijn gekomen, is terug te vinden in het kader “Leerlingen met een vmbo-tl advies”.

Uitkomsten

Figuur 4.13 vat de uitkomsten van deze interventie samen. Na vijf jaar blijkt dat meer leerlingen hoger onderwijs gaan volgen. Nadat de eerste generatie van deze interventie het hoger onderwijs heeft afgerond, beginnen de arbeidsproductiviteit en het arbeidsaanbod wederom te stijgen. Op de lange termijn leidt deze maatregel tot een stijging van het bbp met 0,0%.

Onder de aanname dat de groep vmbo leerlingen met een vmbo-tl advies gelijk blijft, en dat de bevolkingsgrootte gelijk blijft, heeft op de lange termijn 2,9% van de beroepsbevolking deze interventie ondergaan.

Leerlingen met een vmbo-tl advies

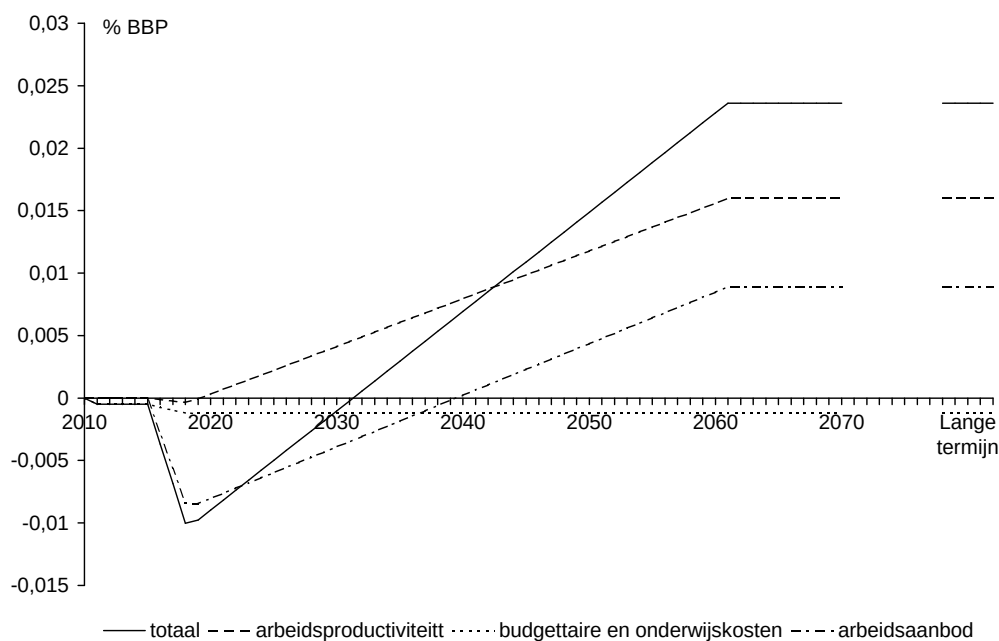
Scholen of scholengemeenschappen met gecombineerde vmbo-havo brugklassen krijgen per brugklasleerling tussen de 5% en 10% meer middelen, zie Onderwijsraad (2010). Het verschil in bekostiging hangt af van de grootte van de school. Wij rekenen met een opslag van 7.5% op de jaarlijkse bekostiging voor leerlingen met een vmbo-tl advies die als gevolg van de interventie naar een gecombineerde vmbo-havo brugklas gaan.

Van de leerlingen in groep acht van het basisonderwijs krijgt circa 58% een vmbo advies. Ongeveer 46% van deze leerlingen krijgt een vmbo-tl advies.^a Een ondergrens van het aantal leerlingen met een vmbo-tl advies die er voor kiezen naar een categorale vmbo te gaan is 11%.^b Derhalve bestaat de doelgroep uit circa 5 500 vmbo-leerlingen, hetgeen een intensivering vraagt van 3 miljoen euro per jaar.

^a Bron: Statline tabel 'studievoortgang brugklas 2003-2007 advies obv score eindtoets'

^b Bron: Statline tabel 'studievoortgang brugklas 2003-2007 advies obv score eindtoets'

Figure 4.13 Uitwerking Nederlands scenario voor meer gecombineerde brugklassen voor vmbo-havo



4.3.5 Prestatiebeloning onderwijspersoneel

Prestatiebeloning voor leraren kan leiden tot betere onderwijsprestaties van leerlingen via betere prikkels. De beloning kan gekoppeld zijn aan de individuele prestaties (individuele prestatiebeloning) of aan de prestaties van een groep leraren, zoals een afdeling of school (teambeloning). De prestatiebeloning zorgt er in het laatste geval voor dat managers en werknemers dezelfde doelen nastreven.

Prestatiebeloning in het onderwijs is een controversieel onderwerp. Problemen rond het meten van de 'output' van scholen of individuele docenten maken het gebruik van prestatiebeloning lastiger dan in sommige andere sectoren. In het onderwijs worden prestaties van leerlingen immers niet alleen bepaald door de bijdrage van de leraar of school, maar ook door de inputs van het gezin en klasgenoten.

Ook kunnen onderwijsprestaties worden gemanipuleerd door bijvoorbeeld zwakke leerlingen buiten de testen te houden (Ladd en Walsh, 2002; Jacob, 2002; Figlio en Getzler, 2002) of door teaching-to-the-test (Jacob, 2002).

Naast problemen met het meten van onderwijsprestaties zijn er ook in algemene zin risico's verbonden aan het gebruik van prestatiebeloning. Zo zou individuele prestatiebeloning ertoe kunnen leiden dat collega's minder goed gaan samenwerken (Lazear en Rosen, 1981). Teambeloning kan dit probleem van non-coöperatief gedrag verhelpen, maar daar staat tegenover dat een beloning gekoppeld aan een groepsprestatie kan leiden tot liftersgedrag (Lazear, 1999). Er bestaat bovendien enige evidentie dat prestatiebeloning nadelig kan uitpakken wanneer werknemers intrinsiek gemotiveerd zijn (Kreps, 1997; Frey en Jegen, 2001).

De laatste jaren zijn veel nieuwe beleidsprogramma's met prestatieprikkels in het onderwijs gestart, met name in de VS en Engeland. Ook in Nederland is in toenemende mate sprake van belangstelling voor resultaatgericht onderwijs. Beleidsvoorstellen richten zich veelal op hogere salarissen of een jaarlijkse bonus voor de best presterende leraren. Versterking van de functiemix rekenen wij niet tot dit thema. De "functiemix" betreft de verdeling van leraren (in voltijdbanen, fte's) over de verschillende salarisschalen. Versterking van de functiemix biedt schoolbesturen extra geld om leraren promotie te kunnen geven naar hogere leraarsfuncties, met bijbehorende beloning. Deze maatregel kent een eenmalige prikkel en geen doorlopende prikkel zoals de prestatiebeloning beoogt.

Empirie

Cornet et al. (2006) stelt dat prestatiebeloning een doeltreffend instrument is om de prestaties van leraren te verbeteren. Webbink et al. (2009) presenteren een uitgebreid literatuuroverzicht. Hieruit blijkt dat de empirische evaluaties van beleidsprogramma's met prestatiebeloning wijzen op positieve effecten van deze programma's op de prestaties van leerlingen.

De studies die de effecten van prestatiebeloning op de toetsscores (uitgedrukt in standaarddeviaties) van leerlingen onderzoeken, geven een gemiddeld effect van 0,13

standaarddeviatie. Deze literatuur behelst zowel evaluaties van individuele beloningsprogramma's (Lavy, 2003; Atkinson et al., 2009; Eberts et al., 2002; Muralidharan en Sundararaman, 2009) als van teambeloningsprogramma's (Lavy, 2002; Ladd, 1999; Glewwe et al., 2003). De resultaten van de verschillende studies zijn goed vergelijkbaar en geven geen aanleiding om aan te nemen dat de effecten van individuele- en teambeloningsprogramma's erg van elkaar afwijken. We zullen dan ook geen onderscheid maken tussen beide typen prestatiebeloning.

De studies richten zich allemaal op beleidsprogramma's waarbij een bonus in het vooruitzicht wordt gesteld voor de best presterende leraren of scholen. De programma's verschillen weliswaar enigszins in de hoogte van de te behalen bonus en in het percentage leraren (of scholen) dat in aanmerking komt voor de bonus, maar bij alle programma's is sprake van een substantiële financiële prikkel. Gemiddeld stellen de programma's een bonus van ongeveer 10% van het jaarsalaris voor de 20% beste leraren in het vooruitzicht.

Voor de Nederlandse onzekerheidsvariant houden we rekening met de risico's die verbonden zijn aan het invoeren van prestatiebeloning, zoals manipulatie van onderwijsprestaties door bijvoorbeeld zwakkere leerlingen uit te sluiten van deelname aan de toetsen of door teaching-to-the-test. Op basis hiervan halveren we het gemiddelde effect uit de literatuur, wat leidt tot een toename van de toetsscores met 0,07 standaarddeviatie. Invoering van prestatiebeloning kan alleen werken als onderwijsprestaties van leerlingen objectief meetbaar zijn. Wij gaan ervan uit dat het beleid invulling zal geven aan deze belangrijke randvoorwaarde.

De literatuur richt zich met name op interventies in het po en het vo en niet op andere onderwijsniveaus. Vanwege het ontbreken van empirie en het gegeven dat het moeilijker is aan de randvoorwaarde voor objectieve toetsing te voldoen, krijgt prestatiebeloning in de onderwijsniveaus mbo, hbo en wo de kwalificatie 'effect onbekend'.

Table 4.16 Onzekerheidsvarianten prestatiebeloning onderwijspersoneel

	Literatuurvariant	NL variant
Effect op toetsscores	+ 0,13 SD	+ 0,07 SD
Doelgroep	primair en voortgezet onderwijs	primair en voortgezet onderwijs

Illustratie

We beschouwen een maatregel gericht op de invoering van prestatiebeloning in het primair en voortgezet onderwijs. De inzet van middelen bedraagt 180 miljoen euro. Van dit bedrag kunnen de 20% best presterende docenten in het primair en voortgezet onderwijs een bonus van 10% van

hun jaarsalaris krijgen. Deze berekening is terug te vinden in het kader “Inzet voor prestatiebeloning”. De bonus regeling is in overeenstemming met de regelingen die empirisch geëvalueerd zijn. Met deze omvang heeft de interventie gevolgen voor alle leerlingen in het basis- en voortgezet onderwijs. Additionele middelen aan prestatiebeloning kunnen worden besteed aan zowel uitbreiding van het percentage dat een bonus ontvangt, of door een verhoging van de bonus. Beide zijn vermoedelijk minder effectief, want de prikkels zullen afnemen.¹⁵

Inzet voor prestatiebeloning

In onderstaande tabel zijn per onderwijstype het aantal fte's en de bijbehorende beste 20% daarvan opgesomd. Het gemiddelde jaarsalaris in 2010 per fte in lokaal onderwijs is 60 duizend euro en 71 duizend euro bij universiteiten en hbo-instellingen.^a Deze cijfers zijn afgeleid uit de nationale rekeningen en vervolgens in het kader van CEP en MEV geëxtrapoleerd. De laatste kolom uit onderstaande tabel bevat de toe te wijzen middelen per onderwijsniveau.

Aantallen fte's in het onderwijs met bovengrens in te zetten middelen

	Aantal fte	20% best presterenden	10% bonus voor 20% best presterenden
po ^a	87 200	17 440	105 mln euro
vo ^a	64 300	12 860	77 mln euro
mbo ^a	23 400	4 680	28 mln euro
hbo ^a	14 900	2 980	21 mln euro
wo ^b	11 500	2 300	16 mln euro
totaal	201 300	40 940	247 mln euro

^a Bron: OC&W (2008)

^b Bron: VSNU website (www.vsnunl.nl), tabblab: feiten & cijfers

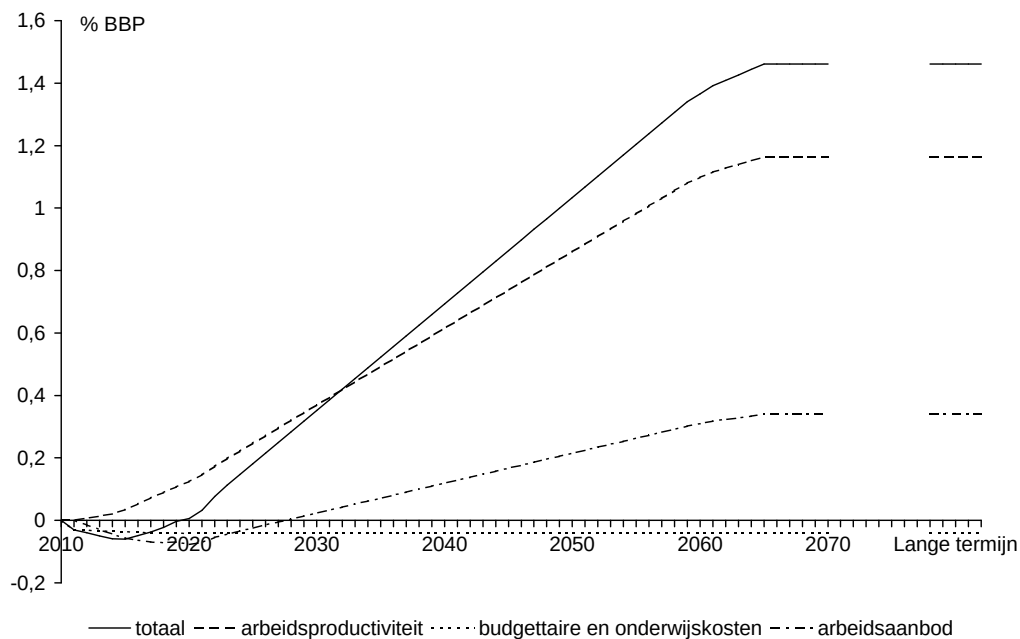
^a Voor de berekening van de individuele bonus is de overhead niet meegenomen in het salaris.

Uitkomsten

In figuur 4.14 zijn zowel het totale effect, als het effect op de arbeidsproductiviteit en de budgettaire en onderwijskosten van deze fictieve maatregel gepresenteerd. Alhoewel de effecten voor individuele leerlingen relatief klein zijn, is het effect op de lange termijn aanzienlijk. De reden hiervoor is dat de maatregel gemiddeld het gehele lerarencorps beter doet presteren. Het gevolg is dat de gehele leerlingenpopulatie in het primair- en voortgezet onderwijs beter gaat presteren tegen geringe kosten. De structurele inzet van 180 mln euro voor prestatiebeloning in het primair- en voortgezet onderwijs leidt op de lange termijn tot een stijging van het bbp-niveau met 1,5 %.

¹⁵ In CPB & PBL (2010) hanteerden het CPB een bovengrens van 250 miljoen euro te besteden aan prestatiebeloning van leraren in het primair en voortgezet onderwijs. De onzekerheid over afnemende meeropbrengsten werden boven dit bedrag te groot verondersteld.

Figure 4.14 Uitwerking Nederlands scenario voor prestatiebeloning onderwijspersoneel in het primair- en voortgezet onderwijs



4.3.6 Sociaal leenstelsel

De introductie van een sociaal leenstelsel is een specifiek beleidsvoorstel waarbij de overheid de kosten van collegegeld met een lening voorschiet. Ex-studenten betalen na afloop van hun studie deze studielening terug. Deze terugbetalingen zijn gekoppeld aan het verdiende inkomen.

Deelname aan een sociaal leenstelsel is vrijwillig en er is een maximum leenbedrag. Het doel van de introductie van het sociaal leenstelsel is het bevorderen van de efficiëntie van de inzet van publieke middelen in het hoger onderwijs. In het huidige stelsel van studiefinanciering is het verstrekken van subsidies het belangrijkste middel om toegankelijkheid van het hoger onderwijs te waarborgen: er is een basisbeurs voor alle studenten en een aanvullende beurs voor studenten met minder draagkrachtige ouders. Door de invoering van een sociaal leenstelsel wordt de private bijdrage aan het hoger onderwijs vergroot. Problemen rond de toegankelijkheid van het hoger onderwijs zijn in essentie terug te voeren op problemen bij het lenen voor een studie of het niet kunnen verzekeren van de risico's van studeren. Een sociaal leenstelsel corrigeert dit falen van de kapitaal- en verzekeringsmarkt door het aanbieden van een lening en een verzekering (de aflossing is afhankelijk van de hoogte van het inkomen). Introductie van het sociaal leenstelsel maakt het dus mogelijk het profijtbeginsel verder toe te passen zonder de toegankelijkheid van het systeem te beperken.

Empirie

Cornet et al. (2006) geven een overzicht van de beschikbare literatuur op dit thema. Het merendeel van deze literatuur richt zich op de effecten van het verhogen van de private bijdrage op de deelname aan het hoger onderwijs. De empirische literatuur suggereert dat de daling in deelname aan het hoger onderwijs ten gevolge van deze hogere private bijdrage niet groot zal zijn. De meeste van deze studies hebben betrekking op prijsveranderingen in de Verenigde Staten. In enkele studies is gevonden dat een verlaging (verhoging) van de collegegelden met 1000 dollar leidt tot 3 tot 4%-punt meer (minder) deelname aan hoger onderwijs (Dynarski (2003), zie overzicht in Kane (2003)). Dynarski (2005) vindt bij bedragen variërend tussen de 1000 en 3000 dollar per jaar effecten op de slagingskansen van 3%-punt.

Een belangrijke kanttekening bij bovenstaande schattingen voor de VS is dat het in deze studies draait om de effecten van een verandering in de prijs van studeren zonder de mogelijkheid om de extra benodigde middelen te lenen. De gevonden prijselasticiteit is dus een samengesteld effect van zowel de werkelijke prijsgevoeligheid als mogelijke financiële beperkingen. Belly and Lochner (2007) vinden bewijs voor het bestaan van zogenoemde “credit constraints”. Bij de introductie van een sociaal leenstelsel in Nederland zal de uitbreiding van private bijdragen voor hoger onderwijs gepaard gaan met een uitbreiding van de mogelijkheden om te lenen. Aannemelijk is dat een dergelijke uitbreiding de effecten op de deelname zullen verminderen. Empirische onderbouwing hiervoor wordt gegeven door onderzoek uit het Verenigd Koninkrijk, Canada en Australië. In het Verenigd Koninkrijk en Canada hebben een verhoging van de collegegelden in combinatie met een verhoging van de leenmogelijkheden niet geleid tot een daling in de deelname aan het hoger onderwijs (OESO, 2009; Barr (2004)). In Australië heeft de introductie van een sociaal leenstelsel in de vorm van het zogenoemde Higher Education Contribution Scheme (HECS) in 1989 niet geleid tot een afname in de deelname aan het hoger onderwijs (zie voor een overzicht van evaluaties CPB, 2001).

Bij de uiteindelijke effecten van de introductie van een sociaal leenstelsel speelt mogelijk ook leenaversie een rol. Dit houdt in dat aan het hebben van een schuld ook “psychologische kosten” verbonden kunnen zijn, waardoor studenten minder bereid zullen zijn een lening aan te gaan. Een recente studie in de Verenigde Staten (Field, 2009) geeft aanwijzingen voor leenaversie bij studenten. De kennis over leenaversie bij de deelname aan hoger onderwijs in Nederland is echter nog beperkt.

Nederlandse studies naar prijsgevoeligheid van de deelname aan hoger onderwijs vinden dat de totale instroom in het hoger onderwijs nauwelijks verandert bij prijsverhogingen. Ook de ontwikkelingen in de jaren negentig duiden op een geringe prijsgevoeligheid. In deze periode zijn de private bijdragen aan hoger onderwijs substantieel toegenomen: de collegegelden zijn verhoogd en, waarschijnlijk nog belangrijker, het recht op studiefinanciering is beperkt van 6 naar 4 jaar. Ondanks deze prijsverhogingen is de deelname aan hoger onderwijs in deze jaren gestegen. Belot et al. (2007) vinden bovendien positieve effecten op toetsscores en

slagingskansen voor de propedeuse. Dit resultaat wordt ondersteund door een studie van Garibaldi et al. (2007) die ook positieve effecten van een hogere private bijdrage op studieprestaties vinden.

In de onzekerheidsvarianten baseren we ons allereerst op de “harde” empirische studies uit de VS. Uitgaande van een interventie waarbij in Nederland de basisbeurs wordt afgeschaft en vervangen door een sociaal leenstelsel, zullen de private bijdragen met ongeveer 3000 euro en 1200 euro per jaar toenemen voor respectievelijk uit- en thuiswonenden studenten. Op basis van de empirische studies in de VS zal hierdoor de kans op afronding van het hoger onderwijs met ongeveer 3%-punt afnemen. Voorts houden we rekening met het feit dat de Amerikaanse studies enkel betrekking hebben op verhoging van de private bijdrage, terwijl in het sociaal leenstelsel de verhoging van de private bijdrage gepaard gaat met het aanbieden van een lening. Hierdoor zullen effecten bij introductie van een sociaal leenstelsel minder sterk zijn. Alles overziend komen we voor de literatuurvariant uit op een effect van -1,5%-punt op het afronden van het hoger onderwijs.

Voor de Nederlandse onzekerheidsvariant wegen de specifieke evidentie uit de Nederlandse studie en internationale ervaringen met invoering van een sociaal leenstelsel zwaarder mee. In deze studies worden geen effecten op deelname gevonden. Bovendien zijn er ook positieve effecten gevonden van verhoging van de private bijdrage op studieprestaties. Bovenstaande overwegend gaan wij ervan uit dat de invoering van een sociaal leenstelsel in Nederland geen effect zal hebben op de kans op afronding van het hoger onderwijs. Aangezien de introductie van een sociaal leenstelsel geen effect heeft op onderwijsprestaties, en er slechts sprake is van een schuif tussen publieke en private middelen, kwalificeren wij dit beleidsvoorstel vanuit maatschappelijk oogpunt als “neutraal”.¹⁶

Table 4.17 Onzekerheidsvarianten sociaal leenstelsel

	Literatuurvariant	NL variant
Effect op afronden hoger onderwijs	- 1,5%-punt	0,0 %-punt
Doelgroep	HBO, WO	HBO, WO

4.4 Niet-kwantificeerbare beleidsthema's

Op een aantal thema's worden beleidsvoorstellen gedaan, die moeilijk kwantitatief zijn te analyseren. Dit is het geval bij voorstellen waarbij de wetenschappelijke empirische literatuur van beleidsevaluaties op dit moment onvoldoende basis geeft voor een oordeel over de effectiviteit van het beleid. Deze beleidsthema's krijgen de kwalificatie '*effect onbekend*'.

¹⁶ Deze kwalificatie wijkt af van Cornet et al. (2006), waarin de publieke baten als restsom zijn genomen.

Dit hoofdstuk bespreekt centrale thema's in de beleidsdiscussie met onvoldoende empirische evidentie over hun effectiviteit.

4.4.1 Salaris onderwijspersoneel

Het salaris van leraren is een potentiële determinant van de kwaliteit van leraren. Hogere salarissen voor leraren kunnen leiden tot betere onderwijsprestaties via een hogere productiviteit van leraren (efficiency wages) of doordat betere leraren worden aangetrokken in het onderwijs. In Nederland zouden hogere salarissen een oplossing kunnen bieden voor een dreigend lerarentekort. Beleidsvoorstellen binnen dit thema richten zich op een generieke salarisverhoging voor leraren of schoolleiders.

Webbink et al. (2009) bespreken de empirische literatuur. In Hanushek en Rivkin (2006) wordt een overzicht gegeven van studies naar de effecten van het salaris van leraren op onderwijsprestaties van leerlingen. De meer overtuigende studies vinden in de overgrote meerderheid (82%) van in totaal 17 studies geen significant effect. De auteurs merken bovendien op dat het moeilijk is het effect van salaris op prestaties precies te identificeren. Zo is het mogelijk dat variaties in nominale salarissen voor een deel slechts verschillen reflecteren in de compensatie voor kosten van levensonderhoud, de aantrekkelijkheid van bepaalde scholen of werkomstandigheden. Een ander punt is het feit dat salarissen effect hebben op de beroepskeuze, de keuze van de eerste baan en mobiliteit tussen banen. Aangezien vrijwel alle studies het effect van salaris op de effectiviteit van de huidige verzameling leraren analyseert, worden effecten op de verzameling (en verdeling) van leraren niet meegenomen.¹⁷

Gezien bovenstaande moeilijkheden bij het empirisch vaststellen van effecten van salarissen en het feit dat de huidige studies geen eenduidig beeld laten zien, kwalificeren wij beleidsvoorstellen waarbij het salaris voor onderwijspersoneel generiek verhoogd wordt als 'effect onbekend'. Hierbij is het van belang te benadrukken dat wij onze analyse richten op de marginale effecten van een generieke salarisverhoging op onderwijsprestaties. Hogere salarissen kunnen mogelijk ook bijdragen aan het oplossen van een dreigend lerarentekort. Hogere lonen voor onderwijspersoneel kunnen leiden tot een hoger arbeidsaanbod in het onderwijs. Er zal mogelijk meer nieuw personeel worden aangetrokken en een groter deel van het huidige lerarencorps zou behouden kunnen blijven voor het onderwijs. Dergelijke algemeen evenwichtsanalyses maken geen onderdeel uit van onze analyse.

4.4.2 Salaris schoolleiders

Het generiek verhogen van de salarissen van schoolleiders zou kunnen bijdragen aan de kwaliteit van het onderwijs. Er is één studie, Lavy (2008) die vindt dat een salarisverhoging voor

¹⁷ Een uitzondering is Hanushek et al. (2005) die een steekproef gebruiken van leraren die van regio veranderen om het verband tussen salaris en kwaliteit van lesgeven te identificeren. Hierbij vinden zij niet dat een hoger salaris effectievere leraren aantrekt.

schoolleiders de toetsscores met 7% doet toenemen. Hoewel de studie een generieke verhoging van salarissen betrof die niet gekoppeld was aan prestaties, worden prikkels om goed te presteren toch genoemd als mogelijke verklaring van het gevonden effect. Schoolleiders zouden bang kunnen zijn bij slechte prestaties hun salaris alsnog te verliezen. Bovendien zou het openbaar maken van deze salarisverhoging meer druk hebben kunnen leggen op de prestaties van de school. Lavy (2008) geeft aan dat de resultaten zijn gevonden op een relatief kleine steekproef en dat de empirie nog erg dun is. Wij volgen zijn advies eerst meer evidentie te vinden alvorens te zware conclusies te verbinden aan de gevonden resultaten. Vooralsnog kwalificeren wij deze beleidsoptie als 'effect onbekend'.

4.4.3 Speciaal onderwijs

Vrijwel alle landen kennen naast het reguliere onderwijs vormen van speciaal onderwijs voor leerlingen met een bepaalde handicap. In het Nederlandse basisonderwijs wordt onderscheid gemaakt tussen drie soorten zorgleerlingen: leerlingen op het speciaal onderwijs, zogenaamde 'rugzakleerlingen' op het reguliere onderwijs en leerlingen op het speciaal basisonderwijs¹⁸. Het speciaal onderwijs op het po en vo is ongeveer vier keer zo duur als het reguliere basis- of voortgezet onderwijs. In Nederland is de laatste jaren sprake van een toename van leerlingen in de relatief dure vormen van speciaal onderwijs, te weten in het speciaal onderwijs en rugzakleerlingen. Nieuwe beleidsvoorstellen kunnen zich richten op de indicatiestelling voor toelating tot het speciaal onderwijs, de mate waarin leerlingen met het speciaal onderwijs in het reguliere onderwijs worden opgevangen en de wijze van financiering in het speciaal onderwijs. Zo zijn er bijvoorbeeld voorstellen de huidige rugzaksystematiek te vervangen door budgetfinanciering.

Webbink et al. (2009) bespreken de beschikbare literatuur op het gebied van het speciaal onderwijs. Zij concluderen dat de empirische kennis omtrent de effecten van het speciaal onderwijs zeer beperkt is. De literatuur geeft alleen enige aanwijzingen dat de vraag naar speciaal onderwijs gevoelig is voor de financieringssystematiek (Cullin, 1999; Sack, 1998; Monk, 1990).

Onderzoek naar de effectiviteit van beleid voor het onderwijs aan zorgleerlingen wordt bemoeilijkt door verschillen tussen stelsels van het speciaal onderwijs en het regulier onderwijs enerzijds en moeilijk waarneembare verschillen in capaciteiten tussen zorgleerlingen en gewone leerlingen anderzijds. Daarnaast is er een schaarse beschikbaarheid van cijfers over de deelname aan en kosten van het onderwijs aan zorgleerlingen. Ook blijkt dat de stelsels voor het speciaal onderwijs internationaal zo sterk verschillen dat de gegevens moeilijk internationaal vergelijkbaar zijn. Er zijn dan ook geen internationaal vergelijkende studies naar de effectiviteit van speciaal onderwijs beschikbaar. Dit thema krijgt zodoende de kwalificatie "effect

¹⁸ Het speciaal basisonderwijs is voor leerlingen met minder ernstige stoornissen dan het speciaal onderwijs.

onbekend”.

4.4.4 Excellentie en hoger onderwijs

Excellentie in het onderwijs kan de welvaart bevorderen voor zowel het individu als de samenleving als geheel. Hanushek en Woessmann (2007) vinden sterke aanwijzingen dat het aandeel mensen in de top van de verdeling van kennis en vaardigheden, gemeten als het aandeel leerlingen dat behoort tot de besten van de wereld, belangrijk is voor de productiviteit van landen en voor economische groei. Er zijn ook aanwijzingen dat een toename van het aandeel hoger opgeleiden bijdraagt aan de productiviteit van anderen (Iranzo en Peri, 2006). Kennisspillovers spelen hierbij een belangrijke rol. Zo is het mogelijk dat hoger opgeleiden de productiviteit van lager opgeleiden verhogen, bijvoorbeeld doordat zij snel leren met nieuwe technologie om te gaan en dit overdragen aan lager opgeleiden. De hoge niveaus van kennis en vaardigheden blijken vooral belangrijk te zijn in hoogproductieve landen, zoals Nederland. Naarmate landen dichterbij de zogenoemde ‘technology frontier’ opereren is de mogelijkheid tot imiteren van andere landen beperkter en wordt innovatie belangrijker voor de groei van de productiviteit (Vandenbussche et al., 2006; Aghion et al., 2005). Een overzicht van deze literatuur wordt gegeven in Minne et al. (2007).

Een vergelijking van de prestaties van Nederlandse leerlingen met leerlingen uit andere rijke landen laat zien dat Nederland goed scoort aan de linkerkant van de verdeling van kennis en vaardigheden, maar minder goed scoort aan de rechterkant van de verdeling (Minne et al., 2007). Er lijkt dus ruimte voor verbetering van excellente leerlingen in het Nederlandse onderwijssysteem. Over het bevorderen van excellentie is echter weinig bekend. De empirische literatuur levert weinig inzicht in welk beleid het meest effectief is. Toekomstig onderzoek zou zich kunnen richten op het effect van recent geïntroduceerde programma’s voor excellente leerlingen in het voortgezet en hoger onderwijs.

Ook op het gebied van het hoger onderwijs is nauwelijks empirische kennis over kansrijke interventies om de kwaliteit te verbeteren. Beleidsvoorstellen waarbij extra middelen ter beschikking worden gesteld ter verbetering van de kwaliteit van het hoger onderwijs, kunnen daarom momenteel niet kwantitatief worden geanalyseerd.

4.4.5 School en omgeving

De laatste jaren worden scholen steeds meer verbonden met andere actoren of instellingen uit de omgeving van de school. Een voorbeeld hiervan is de zogenoemde ‘Brede school’, waarin de school bijvoorbeeld samenwerkt met culturele en sportinstellingen en de leerlingen of kinderen uit de buurt na de reguliere schooltijd activiteiten aanbiedt. Onderzoek naar deze activiteiten suggereert dat leerlingen op deze scholen beter presteren. Zo vinden Sheldon (2007) en Sheldon en Epstein (2005) bijvoorbeeld dat de aanwezigheid van leerlingen hoger is op scholen die programma’s ontwikkelen waarbij school, familie en gemeenschap betrokken zijn en dat

specifieke betrokkenheid van de familie of de omgeving een positief effect heeft op wiskunde toetsscores van leerlingen. Deze studies kennen geen experimentele opzet. Het is daarom niet duidelijk of de 'brede school' daadwerkelijk leidt tot positieve uitkomsten omdat de uitkomsten vertekend kunnen zijn door de selectie van ouders en scholen. Een opvallende ontwikkeling in de omgeving van het Nederlandse publieke onderwijs is de groei van huiswerkbegeleiding door private aanbieders. Op dit gebied bestaan ook geen experimentele studies naar de effecten op de prestaties van leerlingen.

Vanwege het ontbreken van (kwalitatief goede) empirie krijgen beleidsvoorstellen binnen dit thema momenteel de kwalificatie 'effect onbekend'.

4.4.6 Meer computers en ICT in het onderwijs

Door de technologische ontwikkeling speelt ICT een steeds grotere rol in onze samenleving en dus ook op scholen. Extra inzet van meer computers en ICT in het onderwijs kan de arbeidsproductiviteit van de docent vergroten. Hierbij kan bijvoorbeeld gedacht worden aan digitaal leermateriaal, video- of webconferencing (om leerlingen lessen te laten volgen bij een bevoegde leraar los van tijd en plaats) en ingeblikte lessen (zoals opgenomen lessen van leraren of educatieve games). Inzet van ICT zou ook een oplossing kunnen bieden voor het oplossen van het lerarentekort of kunnen bijdragen aan de kwaliteit van het onderwijs doordat met digitaal leermateriaal mogelijk beter kan worden ingespeeld op verschillen tussen leerlingen.

Cornet et al. (2006) kwalificeren meer inzet van computers en ICT in het onderwijs als een 'niet kansrijke' beleids optie. Zij concluderen dat de beschikbare empirische evidentie van goede kwaliteit vooralsnog niet wijst op positieve effecten van ICT gebruik op de kwaliteit van het onderwijs. Vier van de vijf wetenschappelijke studies vinden geen verbetering in leerprestaties en soms zelfs een verslechtering van de leerprestaties (Angrist en Lavy, 2002; Goolsbee en Guryan, 2002; Rouse en Krueger, 2004; Leuven et al., 2007). In een vijfde studie wordt wel een verbetering van leerprestaties gevonden na ICT investeringen in scholen (Machin et al., 2006).

Bij deze weinig overtuigende resultaten dient wel de kanttekening gemaakt te worden dat dit te maken kan hebben met het leertraject. Het is mogelijk dat scholen en docenten meer tijd nodig hebben voordat ICT op een effectieve wijze gebruikt kan worden. Gezien het mogelijke bestaan van deze leereffecten en de positieve gevonden resultaten in de recente studie van (Machin et al., 2006) kwalificeren wij de beleidsvoorstellen gericht op meer inzet van ICT in het onderwijs op dit moment als 'effect onbekend'.

4.4.7 Leven lang leren

Het uitbreiden van de subsidies voor scholing van werknemers ('een leven lang leren') heeft tot doel het up-to-date houden van kennis en vaardigheden gedurende de beroepsloopbaan.

Werknemers kunnen langer blijven werken en komen bij baanverlies sneller weer aan het werk.

Cornet et al. (2006) kwalificeren de kansrijkheid van de beleids optie als onbekend. De eerste

reden is de onzekerheid of additionele middelen daadwerkelijk leiden tot additionele deelname aan scholing. Zo bleek de scholingsaftrek voor werknemers van 40 jaar en ouder in Nederland slechts te leiden tot uitstel van deelname aan scholing en niet tot een stijging van de deelname (Leuven en Oosterbeek, 2004). Daarentegen is in een recente Nederlandse studie gevonden dat de fiscale scholingsaftrek voor werknemers de deelname aan scholing doet toenemen (Leuven en Oosterbeek, 2006). Een tweede reden is dat het rendement op de inzet van middelen voor onderwijs en scholing in het algemeen het hoogst blijkt te zijn in de eerste levensfasen Heckman (2000). Het versterken van het initieel onderwijs lijkt daarom kosteneffectiever dan scholing van werknemers.

Beleidsvoorstellen waarbij ingezet wordt op een 'leven lang leren', krijgen de kwalificatie "effect onbekend".

4.4.8 Schakelklassen en summer courses

Schakelklassen en summer courses zijn specifieke beleidsinstrumenten met als doel het bestrijden van leerachterstanden en het tegengaan van voortijdig schoolverlaten. Vanaf het schooljaar 2006-2007 heeft het ministerie van OCW aan gemeenten met gewichtenleerlingen middelen ter beschikking gesteld voor de invoering van schakelklassen. Schakelklassen zijn bestemd voor leerlingen in het primair onderwijs die te kampen hebben met achterstanden. Het doel is om die leerlingen gedurende een schooljaar zodanig bij te spijkeren dat zij daarna in staat zijn om het onderwijs op hun eigen niveau te kunnen vervolgen.

Nieuwe beleidsvoorstellen richten zich bijvoorbeeld op landelijke invoering van verschillende varianten van schakelklassen. Door het ontbreken van empirie kunnen deze voorstellen momenteel niet kwantitatief geanalyseerd worden. De grote aandacht die dit thema momenteel heeft in het Nederlands beleid maakt het tot een interessant en belangrijk onderwerp voor toekomstig onderzoek.

Ook over de effecten van summer courses is de empirische evidentie zeer beperkt. Een studie van Jacob en Lefgren (2004a) onderzoekt de effecten van een 'remedial summer program', waarbij laag presterende leerlingen in het primair onderwijs een gestructureerd zesweeks programma kregen in kleine klassen met speciaal geselecteerde docenten. De auteurs vinden positieve resultaten op de toetsscores op taal en rekenen. Dit suggereert dat summer courses mogelijk kunnen bijdragen aan de onderwijsprestaties, maar de empirie is nog erg dun. Bovendien zijn de gevonden effecten op toetsscores op basis van de beschikbare informatie niet uit te drukken in standaarddeviaties, wat kwantificering van effecten volgens onze methodiek onmogelijk maakt.

Wel lijken summer courses mogelijk effectief. Op dit moment kunnen wij beleidsvoorstellen in deze richting echter nog niet kwantificeren.

4.4.9 Segregatie

Segregatie in het onderwijs, bijvoorbeeld naar sociaaleconomische achtergrond of naar etniciteit (witte en zwarte scholen) kan van invloed zijn op de prestaties van leerlingen. In het Nederlands beleid bestaat grote zorg dat de concentratie van leerlingen met een lagere sociaaleconomische achtergrond of niet-Nederlandse etniciteit leidt tot zwakke onderwijsuitkomsten. Deze zorg komt voort uit de veronderstelling dat de samenstelling van de school of klas (peers) van invloed kan zijn op de prestaties van een leerling.

Het vaststellen van peereffecten is methodologisch gezien lastig vanwege allerlei selectieprocessen van ouders (leerlingen) en scholen. Daarnaast is de vraag hoe de invloed van de peers verloopt: wie beïnvloedt wie? In de recente onderwijsliteratuur zijn peereffecten onderzocht door te zoeken naar situaties waarin sprake is van toevallige variatie in de klassensamenstelling (natuurlijke experimenten). Zo is bijvoorbeeld gebruik gemaakt van variatie door veranderingen tussen cohorten van de populatie (Hoxby, 2000c; Ammermüller en Pischke, 2006; Lavy en Schlosser, 2007), variatie door de plotselinge instroom van immigranten (Gould et al., 2004), variatie door random toewijzing van kamergenoten (Sacerdote, 2000), variatie door vroege selectie (tracking) van leerlingen (Lefgren (2004)) en variatie door spreiding van leerlingen (busing) (Angrist en Lang, 2004). Driessen (2007) geeft een compleet overzicht van deze recente literatuur. Hoewel de resultaten niet eenduidig zijn, lijkt de belangrijkste conclusie dat peereffecten positief samenhangen met leerprestaties (betere peers leiden tot betere prestaties) maar dat de effecten klein zijn. Veranderingen in klassensamenstelling kunnen daarmee leiden tot een verbetering of verslechtering van prestaties van bepaalde groepen. Voor het beleid levert dit een dilemma. Herschikking van leerlingen kan winst opleveren voor bepaalde groepen leerlingen maar daar staan slechtere prestaties van andere leerlingen tegenover.

Alleen als peereffecten niet-lineair zouden zijn is het wel mogelijk om via herschikking van leerlingen de totale resultaten te verbeteren zonder dat bepaalde groepen slechter gaan presteren. Het aantonen van niet-lineaire peereffecten is echter gecompliceerd en vereist veel gegevens. Twee recente studies vinden bewijs voor het bestaan van bepaalde nonlineariteiten. Ding en Lehrer (2007) vinden dat goede leerlingen meer dan matige leerlingen profiteren van goede medeleerlingen. Hoxby en Weingarth-Salyer (2005) vinden dat leerlingen het sterkst beïnvloed worden door leerlingen die in de beginsituatie sterk op hen lijken. Ook vinden zij dat beperkte heterogeniteit in een klas niet van invloed is op de prestaties maar dat veel heterogeniteit een negatief effect heeft op de prestaties van alle leerlingen in de klas.

De directe implicaties van deze recente studies voor het Nederlands beleid zijn niet duidelijk. Vooralsnog kwalificeren wij beleidsvoorstellen die segregatie in het onderwijs willen tegenaan, met 'effect onbekend'.

4.4.10 Concurrentie en vrije schoolkeuze

Het vergroten van de vrije schoolkeuze in het onderwijs kan de prestaties in het onderwijs mogelijk verbeteren. Dit vergroot enerzijds de kans op een goede match tussen leerling en school, en draagt anderzijds bij aan meer concurrentie tussen scholen. Concurrentie kan scholen stimuleren de kwaliteit van het onderwijs te verbeteren. Meer vrije schoolkeuze zou ook kunnen leiden tot meer segregatie en mogelijk tot meer ongelijkheid, bijvoorbeeld wanneer leerlingen met een lagere sociaaleconomische achtergrond vaker kiezen voor scholen van een lagere kwaliteit.

Het Nederlandse onderwijsstelsel kenmerkt zich vanuit internationaal perspectief als een stelsel met vrije schoolkeuze. Nieuwe beleidsvoorstellen kunnen zich richten op het introduceren van restricties in de vrije schoolkeuze om segregatie en ongelijkheid in het onderwijs tegen te gaan. Een voorbeeld van dergelijk beleid is het postcodebeleid in Amsterdam. Beleidsvoorstellen gericht op beïnvloeding van de concurrentie tussen scholen kunnen zich bijvoorbeeld richten op het verminderen (vergroten) van toetredingsbarrières voor nieuwe scholen, het versoepelen (versterken) van de fusietoets, of het veranderen van de regelgeving omtrent minimum schaalgrootte.

Webbink et al. (2009) geven een overzicht van de empirische literatuur over concurrentie in het onderwijs. In het algemeen worden kleine positieve effecten gevonden van concurrentie op de prestaties van leerlingen (Belfield en Levin, 2003; Hoxby, 2000a; Gibbons et al., 2008; Noailly et al., 2009; Holmes et al., 2003; Hsieh en Urquiola, 2003).

In deze studies is de Herfindahl-index een veelgebruikte maat om de mate van concurrentie aan te duiden. Gemiddeld vinden deze studies dat een toename van de concurrentie met 1 standaarddeviatie op de Herfindahl-index leidt tot een toename met ongeveer 0,10 standaarddeviatie in de toetsscores. Voor Nederland, waar in grote mate sprake is van vrije schoolkeuze, vinden Noailly et al. (2009) effecten tussen de 0,05 en 0,10 standaarddeviatie. In het algemeen kwalificeren wij bevordering van concurrentie tussen scholen dan ook als een kansrijke beleidsoptie.

Een recente studie van De Haan et al. (2010) laat echter zien dat een afname in het aantal scholen (en dus afname in de concurrentie) in het Nederlands basisonderwijs in de jaren negentig een positief effect heeft gehad op de toetsscores op de CITO eindtoets. De auteurs noemen als mogelijke verklaring dat positieve effecten van schaalgrootte de negatieve effecten van minder concurrentie domineren. Hierdoor zijn de effecten van beleidsvoorstellen die zowel concurrentie als schaalgrootte beïnvloeden, moeilijk te bepalen. Dergelijke maatregelen, zoals een aanscherping van de regels wat betreft minimum schaalgrootte, krijgen dan ook 'effect onbekend'.

Overige beleidsvoorstellen om de concurrentie te bevorderen, krijgen weliswaar de kwalificatie 'kansrijk', maar zijn niet kwantitatief te analyseren. Hiervoor is het immers essentieel het precieze effect van de maatregel op de concurrentie, uitgedrukt in termen van de

Herfindahl-index, vast te kunnen stellen. Aangezien de benodigde kennis hierover ontbreekt, zullen effecten van beleidvoorstellen op het gebied van concurrentie niet gekwantificeerd worden.

5 Vervolgonderzoek

Het huidige instrumentarium combineert de beschikbare micro-empirie uit beleidsevaluaties over de effectiviteit van verschillende beleidsinstrumenten met de macro-economische kennis over de opbrengsten van menselijk kapitaal om beleidseffecten te kwantificeren. De wetenschappelijke literatuur van beleidsevaluaties op het gebied van onderwijs heeft de afgelopen decennia belangrijke vorderingen geboekt. Studies die gebruik maken van moderne micro-econometrische technieken en een (quasi-)experimentele opzet vormen de wetenschappelijke standaard en bieden de mogelijkheid om causale effecten van specifieke beleidsinterventies op een kwantitatieve effectmaat (zoals toetsscores) te identificeren. De macro-literatuur biedt inzicht in de rendementen van onderwijs op basis van groeivergelijkingen met menselijk kapitaal als determinant van de productiviteit.

De kwantitatieve analyse van onderwijsbeleid in termen van bbp-effecten is complex en blijft met onzekerheid omgeven. Met de ontwikkeling van dit instrumentarium zijn voor het eerst de effecten van onderwijsbeleid in Nederland kwantitatief geanalyseerd. Dit heeft ook een nieuwe reeks vragen opgeroepen. Deze variëren van wetenschappelijke vragen over bijvoorbeeld de hoogte van de rendementen op onderwijs, tot meer modelmatige vragen over de mate waarin thans gemaakte keuzes of vereenvoudigingen in het instrumentarium noodzakelijk en/of gewenst zijn. Beantwoording van deze vragen is niet eenvoudig, zeker gezien het feit dat antwoorden niet altijd in de wetenschappelijke literatuur voorhanden zijn. Gericht vervolgonderzoek kan bijdragen aan een verdere ontwikkeling van de mogelijkheden om onderwijsbeleid in Nederland kwantitatief te analyseren.

Vervolgonderzoek ter verbetering en uitbreiding van het huidige instrumentarium kent drie pijlers.

Beleidsevaluaties

De eerste pijler betreft onderzoek op het micro-niveau van de beleidsevaluaties. Hoewel de afgelopen jaren belangrijke vorderingen zijn gemaakt op het gebied van beleidsevaluaties in het onderwijs, zijn hier nog stappen te maken. Zo is het in specifieke sectoren als het mbo en het hoger onderwijs bijvoorbeeld moeilijk om beleid te evalueren, vanwege de grote differentiatie aan opleidingen en het ontbreken van centrale normering van toetsen. Voor een aantal belangrijke thema's is op dit moment onvoldoende empirie beschikbaar om beleidseffecten te kwantificeren. Voorbeelden hiervan zijn de effectiviteit van kop-voet-en schakelklassen en summer courses, de effectiviteit van excellentie programma's of een verlenging van de studieduur in het hoger onderwijs. Evaluaties van beleid, bijvoorbeeld via het uitvoeren van experimenten, kunnen bijdragen aan inzet van het instrumentarium op een breder spectrum van beleidsthema's.

Ook op beleidsthema's waar reeds buitenlandse empirie beschikbaar is, zijn aanvullende

Nederlandse beleidsevaluaties zeer waardevol. De vertaalslag van deze empirie naar de Nederlandse situatie is niet triviaal. Allereerst zijn de sociaal-economische omstandigheden in het buitenland vaak anders dan in Nederland. Ten tweede verschillen landen in de mate waarin zij al beleid ingevoerd hebben. Dit is belangrijk, omdat de effectiviteit van een bepaalde beleidsoptie kan afnemen naarmate meer middelen worden ingezet. In het huidige instrumentarium wegen we deze twee factoren mee en stellen effecten uit de literatuur waar nodig bij. Meer Nederlandse empirie zal bijdragen aan een meer betrouwbare inschatting van effecten in Nederland.

Maatschappelijke baten en kosten van onderwijs

De tweede pijler betreft onderzoek naar de maatschappelijke opbrengsten en kosten van onderwijs, waaronder de rendementen op onderwijs. De huidige literatuur geeft een redelijk consistent beeld van het rendement op een jaar gevolgd onderwijs. Desalniettemin bieden de gevonden private en sociale rendementen veel ruimte voor discussie. Zo is het onduidelijk waarom het rendement van onderwijs zoveel hoger is dan het rendement op sparen. Waarom studeren individuen niet langer als het rendement op het volgen van een jaar onderwijs daadwerkelijk zo hoog is? De literatuur kent verschillende verklaringen, die slechts in beperkte mate ondersteund worden door de empirie. Deze vraagstelling staat bekend als de “schooling premium puzzle”.

Ook kan nog meer inzicht worden verkregen in de relatie tussen private en sociale rendementen van onderwijs. Het sociale rendement bevindt zich in dezelfde orde van grootte als het private rendement, mogelijk met een kleine opslag. Welke drivers bepalen dit lange termijn rendement?

Daarnaast is nog weinig bekend over de “heterogene rendementen” op onderwijs. Welke variatie kennen de verschillende opbrengsten van onderwijs over verschillende opleidingstypen en welk effect heeft een beter opgeleide bevolking op de rendementen?

Tot slot is de empirie over rendementen op toetscores nog erg dun. Meer inzicht in de relatie tussen de rendementen op kwaliteit en kwantiteit van het onderwijs kunnen bijdragen aan een verbetering van de vertaling van beleid naar productiviteitseffecten.

Ander onderzoek binnen deze pijler kan zich richten op de relatie tussen opleidingsniveau en werkgelegenheid. Hoger opgeleiden participeren meer dan laagopgeleiden en zijn relatief minder lang werkloos. Een hoger opgeleide beroepsbevolking zal zodoende op de lange termijn in een groter arbeidsaanbod resulteren. Momenteel maken we gebruik van cijfers van het CBS voor het bepalen van de arbeidsaanbodeffecten. Hierbij wordt het arbeidsaanbod gepresenteerd voor laag-, midden-, en hoogopgeleiden. Kunnen we deze effecten verder specificeren naar opleidingstype? Wat weten wij over de ontwikkeling van het arbeidsaanbod over de tijd? Welke algemeen evenwichtsmechanismen kunnen optreden waardoor de toename bij een stijgend arbeidsaanbod wellicht stagneert? Meer inzicht in deze vragen kan bijdragen aan een

nauwkeuriger raming van het arbeidsaanbod.

Onderhoud en uitbreiding instrumentarium

De derde pijler richt zich op het onderhoud en mogelijke uitbreiding van het huidige instrumentarium. Verder onderzoek op dit terrein kan de werking van het instrumentarium verbeteren. Zo nemen we nu bijvoorbeeld aan dat elke maatregel een vast effect sorteert, onafhankelijk van ander beleid dat gevoerd wordt. Is er sprake van synergie tussen verschillende beleidsopties of werken verschillende maatregelen elkaar juist tegen? Ook is het mogelijk dat inzet op meerdere beleidsinstrumenten tegelijk minder effectief is vanwege afnemende meeropbrengsten van beleid.

Daarnaast is de vraag hoe een beleidsinstrument op de langere termijn doorwerkt. De empirie van beleidsevaluaties richt zich vaak op de directe effecten van beleidsinterventies, zoals de effecten op toetsscores na 1 of 2 jaar. Er is echter nog weinig bekend over de doorwerking van beleid op langere termijn. In het huidige instrumentarium worden aannames gemaakt op basis van de beperkte beschikbare empirie. Hoeveel groter is het effect wanneer je 4 jaar les krijgt van een geschoolde docent dan wanneer je 1 jaar les krijgt van deze docent? En hoe lang werkt een eenmalige scholing van een docent door op de kwaliteit van zijn onderwijs? Dit zijn enerzijds empirische vragen. Anderzijds is een aanpassing van het instrumentarium nodig om dergelijke effecten mee te kunnen nemen in de analyse.

Om de praktische inzet van het instrumentarium ook bij nieuwe maatregelen te waarborgen kan het nodig zijn het cohortmodel aan te passen. Zo lijkt het voor de analyse van schakelklassen bijvoorbeeld nodig om de vaste koppeling tussen opleidingsniveau en het aantal nominale onderwijsjaren op enige wijze te flexibiliseren. Ook kan het van belang zijn het cohortmodel uit te breiden met bijvoorbeeld zittenblijvers of post-initieel onderwijs te modelleren. Een uitbreiding van het instrumentarium met modellering van werkloosheid kan de analyse van arbeidsmarkteffecten van onderwijsbeleid verbeteren.

Tot slot nemen wij in het huidige instrumentarium aan dat elke maatregel slechts een relatief beperkte aanpassing op het huidige beleidspakket behelst. Het zou interessant kunnen zijn te onderzoeken in welke mate onze resultaten robuust zijn voor grotere aanpassingen in het beleid.

Voor alle pijlers geldt dat het van groot belang is de wetenschappelijke literatuur nauwgezet te volgen om het instrumentarium up-to-date te houden. Nieuwe beleidsevaluaties kunnen mogelijk leiden tot aanpassing van de huidige inzichten omtrent de effectiviteit van een beleidsthema of kwantificering op nieuwe terreinen mogelijk maken. Hernieuwde inzichten rond bijvoorbeeld de “schooling premium puzzle”; de rendementen op onderwijs; een betere duiding van de effecten van bepaalde beleidsthema’s en meer inzicht in de doorwerking van beleidsmaatregelen in de tijd en op elkaar kan leiden tot verbetering van het instrumentarium.

References

Acemoglu, D. en J. Angrist, 2000, How large are the social returns to education?, working paper 7444, NBER.

Aghion, P., L. Boustan, C. Hoxby en J. Vandenbussche, 2005, Exploiting states' mistakes to identify the causal impact of higher education on growth, Harvard university/cepr/nber.

Ammermüller, A., 2005, Educational opportunities and the role of institutions, ZEW Discussion Paper 05-44, Centre for European Economic Studies, Mannheim.

Ammermüller, A. en J.S. Pischke, 2006, Peer effects in European primary schools, evidence from PIRLS, NBER Working Paper 12180.

Angrist, J. en V. Lavy, 2009, The effects of high stakes high school achievement awards: Evidence from a randomized trial, *American Economic Review*, vol. 99, nr. 4, pag. 1384–1414.

Angrist, J.D. en K. Lang, 2004, Does school integration generate peer effects? Evidence from Boston's Meteo Program, *American Economic Review*, vol. 94, nr. 5, pag. 1613–1634.

Angrist, J.D. en V. Lavy, 1999, Using Maimonides' rule to estimate the effect of class size on scholastic achievement, *Quarterly Journal of Economics*, vol. 114, pag. 533–575.

Angrist, J.D. en V. Lavy, 2001, Does teacher training affect pupil learning? Evidence from matched comparisons in Jerusalem public schools, *Journal of Labor Economics*, vol. 19, nr. 2, pag. 343–369.

Angrist, J.D. en V. Lavy, 2002, New evidence on classroom computers and pupil learning, *The Economic Journal*, vol. 112, pag. 735–765.

Atkinson, A., S. Burgess, B. Croxson, P. Gregg, C. Propper, H. Slater en D. Wilson, 2009, Evaluating the impact of performance-related pay for teachers in England, *Labour Economics*, vol. 16, pag. 251–261.

Barr, N., 2004, Higher education funding, oxford review of economic policy, *Oxford Review of economic policy*, vol. 20, nr. 2, pag. 264–283.

- Barro, R., 1991, Economic growth in a cross section of countries, *Quarterly Journal of Economics*, vol. 106, pag. 407–443.
- Barro, R. en J. Lee, 1993, International comparison of educational attainment, *Journal of Monetary Economics*, vol. 32, pag. 363–394.
- Barro, R. en J. Lee, 2001, International data on educational attainment updates and implications, *Oxford Economic Papers*, vol. 53, pag. 541–563.
- Belfield, C.R. en H.M. Levin, 2003, The effects of competition on educational outcomes: A review for the United States, Occasional Paper 35, National Center for the Study of Privatization in Education, Columbia University, New York, NY.
- Belot, M., E. Canton en D. Webbink, 2007, Does reducing student support affect scholastic performance, *Empirical Economics*, vol. 32, pag. 261–275.
- Berlinski, S., S. Galiani en P. Gertler, 2009, The effect of pre-primary education on primary school performance, *Journal of Public Economics*, vol. 93, pag. 219–234.
- Bettinger, E.P., 2008, Paying to learn: The effect of financial incentives on elementary school test scores, in *CESifo / PEPG Conference on Economic incentives: Do they work in education? Insights and findings from behavioral research*.
- Bishop, J.H., 1997, The effect of national standards and curriculum-based examinations on achievement, *American Economic Review*, vol. 87, pag. 260–264.
- Blundell, R. en Bond, 1998, Initial conditions and moment restrictions in dynamic panel data models, *Journal of econometrics*, vol. 87, pag. 115–143.
- Boozer, M. en C. Rouse, 2001, Intrascchool variation in class size: Patterns and implications, *Journal of Urban Economics*, vol. 50, pag. 163–189.
- Bressoux, P., F. Kramarz en C. Prost, 2009, Teachers' training, class size and students' outcomes: Learning from administrative forecasting mistakes, *The Economic Journal*, vol. 119, pag. 540–561.
- Browning, M. en E. Heinesen, 2007, Class size, teacher hours and educational attainment, *Scandinavian Journal of Economics*, vol. 109, pag. 415–438.

- Brunello, G. en D. Checchi, 2007, Does tracking affect equality of opportunity? New international evidence, *Economic Policy*, vol. 52, pag. 781–861.
- Campbell, F.A., C.T. Ramey, E. Pungello, J. Sparling en S. Miller-Johnson, 2002, Early childhood education: Young adult outcomes from the Abecedarian project, *Applied Development Science*, vol. 6, nr. 1, pag. 42–57.
- Card, D., 1999, *Handbook of Labour Economics*, hfdst. The causal effect of education on earnings, Elsevier.
- Cascio, E.U. en E. Lewis, 2006, Schooling and the armed forces qualifying test: Evidence from school-entry laws, *The Journal of Human Resources*, vol. 41, nr. 2, pag. 294–318.
- Chorny, V. en D. Webbink, 2010, The effect of accountability policies in primary education in Amsterdam, CPB Discussion Paper 144.
- Coe, Helpman en Hoffmaister, 2009, International r&d spillovers and institutions, *European Economic Review*, vol. 53, pag. 723–741.
- Cohen en Soto, 2007, Growth and human capital: good data, good results, *Journal of Economic Growth*, vol. 12, pag. 51–76.
- Cornet, M., F. Huizinga, B. Minne en D. Webbink, 2006, Kansrijk kennisbeleid, CPB Document 124.
- CPB, 2006, Keuzes in kaart 2008-2011, Tech. rapp., CPB.
- CPB, 2008, Centraal economisch plan, Tech. rapp., CPB.
- CPB & PBL, 2010, Keuzes in kaart 2011-2015, CPB Bijzondere Publicatie 85.
- Cullin, J.B., 1999, The impact of fiscal incentives on student disability rates, NBER Working Paper 7173.
- Cullingford, C.I., S. Daniels en J. Brown, 1999, The effects of Ofsted inspection on school performance, *School Leadership and Management*, vol. 19, nr. 4, pag. 323–526.

- Currie, J., 2001, Early childhood education programs, *Journal of PEconomic Perspectives*, vol. 15, nr. 2, pag. 213–238.
- Currie, J. en D. Thomas, 1995, Does Head Start make a difference?, *American Economic Review*, vol. 85, nr. 3, pag. 341–364.
- De Haan, M., E. Leuven en H. Oosterbeek, 2010, The effect of school choice on pupil test scores: Evidence from dutch reform data, work in progress, wordt een discussion paper.
- Dearden, L., C. Emmerson en C. Meghir, 2009, Conditional cash transfers and school dropout rates, *Journal of Human Resources*, vol. 44, nr. 4, pag. 827–857.
- Dee, T.S. en B.A. Jacob, 2009, The impact of No Child Left Behind on student achievement, NBER Working Paper 15531.
- Delafuente, A. en R. Domenech, 2006, Human capital in growth regressions: how much difference does data quality make?, *Journal of the European Economic Association*, vol. 4, pag. 1–36.
- Ding, W. en S.F. Lehrer, 2007, Do peers affect student achievement in China's secondary schools, *Review of Economics and Statistics*, vol. 89, nr. 2, pag. 300–312.
- Dobbelsteen, S., J. Levin en H. Oosterbeek, 2002, The causal effect of class size on scholastic achievement: Distinguishing the pure class size effect from the effect of changes in class composition, *Oxford Bulletin of Economics and Statistics*, vol. 64, pag. 17–38.
- Driessen, G., 2007, Peer group effecten op onderwijsprestaties. Een internationaal review van effecten, verklaringen en theoretische en methodologische aspecten, Its, Radboud Universiteit Nijmegen.
- Dynarski, S., 2003, Does aid matter? measuring the effect of student aid on college attendance and completion, *American Economic Review*, vol. 93.
- Dynarski, S., 2005, Building the stock of collega educated labor, working paper 11604, NBER.
- Eberts, R., K. Hollenbeck en J. Stone, 2002, Teacher performance incentives and student outcomes, *Journal of Human Resources*, vol. 37, pag. 913–927.

Euwals, R., M. Knoef en D. van Vuuren, 2007, The trend in female labour force participation; what can be expected for the future, Discussion Paper 93, CPB.

Field, E., 2009, Educational debt burden and career choice: Evidence from a financial aid experiment at nyu law school, *American Economic Journal: Applied Economics*, vol. 1, pag. 1–21.

Figlio, D., 2006, Testing, crime, and punishment, *Journal of Public Economics*, vol. 90, pag. 837–851.

Figlio, N.D. en L.S. Getzler, 2002, Accountability, ability and disability: Gaming the system, NBER Working Paper 9307.

Folger, J. en J. Parker, 1990, The cost-effectiveness of adding aides or reducing class size, mimeo, Vanderbilt University.

Frey, B. en R. Jegen, 2001, Motivation crowding theory, *Journal of Economic Surveys*, vol. 15, nr. 5, pag. 589–611.

Fuchs, T. en L. Woessmann, 2007, What accounts for international differences in student performance? A re-examination using PISA data, *Empirical Economics*, vol. 32, pag. 433–464.

Garibaldi, P., F. Giavazzi, A. Ichino en E. Rettore, 2007, College cost and time to complete a degree: Evidence from tuition discontinuities, working paper 12863, NBER.

Gibbons, S., S. Machin en O. Silva, 2008, Choice, competition and pupil achievement, *Journal of the European Economic Association*, vol. 6, nr. 4, pag. 912–947.

Glewwe, P., N. Ilias en M. Kremer, 2003, Teacher incentives, NBER Working Paper 9671.

Goolsbee, A. en J. Guryan, 2002, The impact of internet subsidies in public schools, NBER Working Paper, nr. 9090.

Gould, E.D., V. Lavy en M.D. Paserman, 2004, Does immigration affect the long-term outcomes of natives, quasi-experimental evidence, NBER Working Paper 10844.

Grossman, J.B. en J.P. Tierney, 1998, Does mentoring work? An impact study of the Big Brothers Big Sisters program, *Evaluation Review*, vol. 22, nr. 3, pag. 402–425.

- Hahn, A., T. Leavitt en P. Aaron, 1994, Evaluation of the Quantum Opportunities Program (QOP). Did the program work? A report on the post secondary and cost-effectiveness of the QOP program (1989-1993), Brandeis university heller graduate school center for human resources, Waltham, MA:.
- Hanushek, E., J. Kain, D. O'Brien en S. Rivkin, 2005, The market for teacher quality, NBER Working Paper 11154.
- Hanushek, E. en D. Kimko, 2001, Schooling, labour force quality and the growth of nations, *American Economic Review*, vol. 90, pag. 1184–1208.
- Hanushek, E. en S.G. Rivkin, 2006, Teacher quality, in E. Hanushek en F. Welch, red., *Handbook of the Economics of Education*, vol. 2, hfdst. 18, pag. 1051–1078, Elsevier.
- Hanushek, E. en L. Woessmann, 2006, Does education tracking affect performance and inequality? Differences in differences evidence across countries, *The Economic Journal*, vol. 116, pag. C63–C76.
- Hanushek, E. en L. Woessmann, 2007, The role of school improvement in economic development, NBER Working Paper 12832.
- Hanushek, E. en L. Woessmann, 2008, The role of cognitive skills in economic development, *Journal of Economic Literature*, vol. 46, pag. 607–668.
- Hanushek, E.A. en M.E. Raymond, 2004, The effect of school accountability systems on the level and distribution of student achievement, *Journal of the European Economic Association*, vol. 2, nr. 2-3, pag. 406–415.
- Hastings, J. en J.M. Weinstein, 2008, Information, school choide and academic achievements: Evidence from two experiments, *Quarterly Journal of Economics*, vol. 123, nr. 4, pag. 1373–1414.
- Heckman, J.J., 2000, Policies to foster human capital, *Research in Economics*, vol. 53, nr. 1, pag. 3–56.
- Heckman, J.J., J. Stixrud en S. Uruza, 2006, The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior, NBER Working Paper 12006.

- Heckman, J.J. en E.J. Vytlacil, 2007, *Handbook of Econometrics, Volume 6*, hfdst. 70
Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and
Econometric Policy Evaluation, Elsevier.
- Holmes, G.M., J. DeSimone en N.G. Rupp, 2003, Does school choice increase school quality?,
NBER Working Paper 9683.
- Hoxby, C.M., 2000a, Does competition among public schools benefit students and taxpayers,
American Economic Review, vol. 90, nr. 5, pag. 1209–1238.
- Hoxby, C.M., 2000b, The effects of class size on student achievement: New evidence from
population variation, *Quarterly Journal of Economics*, vol. 115, pag. 1239–1285.
- Hoxby, C.M., 2000c, Peer effects in the classroom: Learning from gender and race variation,
NBER Working Paper 7867.
- Hoxby, C.M. en G. Weingarth-Salyer, 2005, Taking race out of the equation: school
reassignment and the structure of peer effects, mimeo, Stanford University.
- Hsieh, C.T. en M. Urquiola, 2003, When schools compete, how do they compete? An assesment
of Chile's nationwide school voucher programme, NBER Working Paper 10008.
- Huizinga, F. en B. Smid, 2004, Vier vergezichten op nederland; productie, arbeid en
sectorstructuur in vier scenario's tot 2040, Discussion Paper 55, CPB.
- Iranzo, S. en G. Peri, 2006, Schooling externalities, technology and productivity: Theory and
evidence from U.S. states, NBER Working Paper 12440.
- Jacob, B. en L. Lefgren, 2004a, Remedial education and student achievement: A
regression-discontinuity analysis, *Review of Economics and Statistics*, vol. LXXXVI, nr. 1, pag.
226–244.
- Jacob, B. en S. Levitt, 2003, Rotten apples: An investigation of the prevalence and predictors of
teacher cheating, *Quarterly Journal of Economics*, vol. 118, nr. 3, pag. 843–877.
- Jacob, B.A., 2002, Accountability, incentives and behavior: The impact of high-stakes testing in
Chicago public schools, NBER Working Paper 8968.

Jacob, B.A., 2005, Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools, *Journal of Public Economics*, vol. 89, pag. 761–796.

Jacob, B.A. en L. Lefgren, 2004b, The impact of teacher training on student achievement: quasi-experimental evidence from school reform efforts in Chicago, *Journal of Human Resources*, vol. 39, nr. 1, pag. 50–79.

Jacobs, B., 2010, Consequenties van rendementsberekeningen voor onderwijsbeleid, Tech. rapp., Erasmus universiteit.

Jacobs, B. en H. Webbink, 2006, Rendement onderwijs blijft stijgen, *Economische Statistische Berichten*.

Jürges, H., K. Schneider en F. Büchel, 2005, The effect of central exit examinations on student achievement: Quasi-experimental evidence from TIMMS Germany, *Journal of the European Economic Association*, vol. 3, pag. 1134–1155.

Kane, T., J.E. Rockoff en O. Staiger, 2008, What does certification tell us about teacher effectiveness? Evidence from New York City, *Economics of Education Review*, vol. 27, nr. 6, pag. 615–631.

Kane, T., 2003, A quasi-experimental estimate of the impact of financial aid on college-going, working paper 9703, NBER.

Kormendi, R. en P. Meguire, 1985, Macroeconomic determinants of growth: Cross-country evidence, *Journal of Monetary Economics*, vol. 16, pag. 141–163.

Kremer, M., E. Miguel en R. Thornton, 2009, Incentives to learn, *Review of Economics and Statistics*, vol. 91, nr. 3, pag. 437–456.

Kreps, D., 1997, Intrinsic motivation and extrinsic incentives, *American Economic Review*, vol. 87, nr. 2, pag. 359–364.

Krueger, A.B., 1999, Experimental estimates of education production functions, *Quarterly Journal of Economics*, vol. 114, pag. 497–532.

Ladd, H.F., 1999, The Dallas school accountability and incentive program: An evaluation of its impacts on student outcomes, *Economics of Education Review*, vol. 18, pag. 1–16.

- Ladd, H.F., 2010, Education inspectorate systems in New Zealand and the Netherlands, *Education Finance and Policy*, vol. 5, nr. 3, pag. 378–392.
- Ladd, H.F. en R.P. Walsh, 2002, Implementing value-added measures of school effectiveness: Getting the incentives right, *Economics of Education Review*, vol. 21, pag. 1–17.
- Lavy, V., 1998, Using quasi-experimental designs to evaluate the effect of school hours and class size on student achievement, mimeo, Jerusalem: Hebrew University of Jerusalem, Department of Economics.
- Lavy, V., 2002, Evaluating the effect of teachers' group performance incentives on pupil achievement, *Journal of Political Economy*, vol. 110, nr. 6, pag. 1286–1317.
- Lavy, V., 2003, Paying for performance: The effect of teachers' financial incentives on students' scholastic outcomes, CEPR Discussion Paper 3862.
- Lavy, V. en A. Schlosser, 2005, Targeted remedial education for underperforming teenagers: Costs and benefits, *Journal of Labor Economics*, vol. 23, nr. 4, pag. 839–874.
- Lavy, V. en A. Schlosser, 2007, Mechanisms and impacts of gender peer effects at school, NBER Working Paper 13292.
- Lavy, V., 2008, Does raising the principal's wage improve the school's outcomes? Quasi-experimental evidence from an unusual policy experiment in Israel, *Scandinavian Journal of Economics*, vol. 110, nr. 4, pag. 639–662.
- Lazear, E.P., 1999, Personnel economics: Past lessons and future directions, presidential address to the Society of Labor Economics, *Journal of Labor Economics*, vol. 17, nr. 2, pag. 199–236.
- Lazear, E.P. en S. Rosen, 1981, Rank-order tournaments as optimum labor contracts, *Journal of Political Economy*, vol. 89, nr. 5, pag. 841–864.
- Lazear, E., 2003, Teacher incentives, *Swedish Economic Policy Review*, vol. 10, pag. 179–214.
- Lefgren, L., 2004, Educational peer effects and the Chicago public schools, *Journal of Urban Economics*, vol. 56, nr. 2, pag. 169–191.
- Leuven, E., M. Lindahl, H. Oosterbeek en D. Webbink, 2007, The effect of extra funding for

- disadvantaged pupils on achievement, *The Review of Economics and Statistics*, vol. 89, nr. 4, pag. 721–736.
- Leuven, E., M. Lindahl, H. Oosterbeek en D. Webbink, 2010, Expanding schooling opportunities for 4-year-olds, *Economics of Education Review*, vol. 29, pag. 319–328.
- Leuven, E. en H. Oosterbeek, 2004, Evaluating the effect of tax deductions on training, *Journal of Labor Economics*, vol. 22, nr. 1, pag. 461–488.
- Leuven, E. en H. Oosterbeek, 2006, The effectiveness of human capital policies for disadvantaged groups in the Netherlands, in P. Peterson en L. Woessmann, red., *Schools and the Equal Opportunity Problem*, MIT Press.
- Leuven, E., H. Oosterbeek en M. Roenning, 2008, Quasi-experimental estimates of the effect of class size on achievement in Norway, *Scandinavian Journal of Economics*, vol. 110, nr. 4, pag. 663–693.
- Link, C.R. en J.G. Mulligan, 1986, The merits of a longer school day, *Economics of Education Review*, vol. 5, nr. 4, pag. 373–381.
- Loeb, S. en J. Bound, 1996, The effect of measured school inputs on academic achievement: Evidence from the 1920s, 1930s and 1940s birth cohorts, *The Review of Economics and Statistics*, vol. 78, nr. 4, pag. 653–664.
- Luginbuhl, R., D. Webbink en I.F. De Wolf, 2009, Do school inspections improve primary school performance?, *Educational Evaluation and Policy Analysis*, vol. 31, nr. 3, pag. 221–237.
- Machin, S., S. McNally en O. Silva, 2006, New technology in schools: Is there a payoff?, mimeo, London School of Economics.
- Maxfield, M., A. Schirm en N. Rodriguez-Planas, 2003, The Quantum Opportunities Program demonstration: Implementation and short-term impacts, Tech. rapp., Mathematica Policy Research, Washington DC.
- Messer, D. en S. Wolter, 2009, Money matters: evidence from a large-scale randomized field experiment with vouchers for adult training, IZA Discussion Paper 4017.

- Minne, B., M. Rensman, B. Vroomen en D. Webbink, 2007, Excellence for productivity, CPB Bijzondere Publicatie 69.
- Monk, D.H., 1990, *Educational finance: An econometric approach*, McGraw-Hill.
- Moretti, E., 2004, Estimating the external return to higher education, *Journal of Econometrics*, vol. 121, pag. 175–212.
- Mulligan, C., 1999, Galton versus the human capital approach to inheritance, *Journal of Political Economy*, vol. 107, pag. 184–S224.
- Muralidharan, K. en V. Sundararaman, 2009, Teacher performance pay: Experimental evidence from India, NBER Working Paper 15323.
- Murnane, R., J. Willett, Y. Duhaldeborde en J.H. Tyler, 2000, How important are the cognitive skills of teenagers in predicting subsequent earnings?, *Journal of Policy Analysis and Management*, vol. 19, pag. 547–568.
- Noailly, J., S. Vujic en A. Aouragh, 2009, The effects of competition on the quality of primary schools in the Netherlands, CPB Discussion Paper 120.
- OC&W, 2008, *Kerncijfers 2004-2008*, Ministerie van Onderwijs, Cultuur en Wetenschap.
- OECD, 2007, Thematic review of tertiary education - The Netherlands, country note.
- OECD, 2009, *Education at Glance 2009*, OECD, Paris.
- Onderwijsraad, 2007, Doorstroom en talentontwikkeling : onderwijs voor 12- 18-jarigen, No. 20070256/883, Onderwijsraad, Den Haag.
- Onderwijsraad, 2010, Advies Vroeg of laat, uitgebracht aan de Minister en aan de Staatssecretaris van Onderwijs, Cultuur en Wetenschap, Uitgave van de onderwijsraad, nr. 20100040/969, Den Haag.
- Rivkin, S., E. Hanushek en J. Kain, 2005, Teachers, schools and academic achievement, *Econometrica*, vol. 73, nr. 2, pag. 417–458.

- Rockoff, J.E., 2008, Does mentoring reduce turnover and improve skills of new employees? Evidence from teachers in New York City, NBER Working Paper 13868.
- Rose, H., 2006, Do gains in test scores explain labor market outcomes?, *Economics of Education Review*, vol. 25, pag. 430–446.
- Rouse, C.E. en A.B. Krueger, 2004, Putting computerized instruction to the test: a randomized evaluation of a scientifically based reading program, *Economics of Education Review*, vol. 23, nr. 4, pag. 323–338.
- Sacerdote, B., 2000, Peer effects with random assignment: Results for Dartmouth roommates, *Quarterly Journal of Economics*, vol. 116, nr. 2, pag. 681–704.
- Sack, J.L., 1998, Spec. ed. designation varies widely across country, *Education Week*, vol. 17.
- Schütz, G., H.W. Ursprung en L. Woessman, 2008, Education policy and equality of opportunity, *Kyklos*, vol. 61, nr. 2, pag. 279–308.
- Schweinhart, L.J., J. Montie, Z. Xiang, W.S. Barnett, C.R. Belfield en M. Nores, 2005, *Lifetime effects: The HighScope Perry Preschool study through age 40*, Monographs of the HighScope Educational Research Foundation, 14, HighScope Press, Ypsilanti, MI.
- Shaw, I., D.P. Newton, M. Aitkin en R. Darnell, 2003, Do Ofsted inspections of secondary schools make a difference to GCSE results?, *British Educational Research Journal*, vol. 29, nr. 1, pag. 63–75.
- Sheldon, S., 2007, Improving student attendance with school, family and community partnerships, *Journal of Educational Research*, vol. 100, nr. 5, pag. 267–275.
- Sheldon, S.B. en J.L. Epstein, 2005, Involvement counts: Family and community partnerships and math achievement, *Journal of Educational Research*, vol. 98, nr. 4, pag. 196–206.
- Sianesi, B. en J. van Reenen, 2003, The returns to education: macroeconomics, *Journal of economic surveys*, vol. 17, pag. 157–200.
- Steeg, M. van der, R. van Elk en D. Webbink, 2010, Het effect van de lerarenbeurs op scholingsdeelname docenten, CPB Document 205.

- Stroucken, L., D. Takkenberg en A. Béguin, 2008, Citotoets en de overgang van basisonderwijs naar voortgezet onderwijs, *Sociaaleconomische trends: Statistisch kwartaalblad over arbeidsmarkt, sociale zekerheid en inkomen*, vol. 2, pag. 7–16.
- Taggart, R., 1998, The quantum opportunities program, in D. Elliot, red., *Blueprints for violence protection: book four*, Center for the Study and Prevention of Violence, Boulder, CO.
- Teulings, C. en T. van Rens, 2008, Education, growth and income inequality, *Review of Economics and Statistics*, vol. 90, pag. 89–104.
- Urquiola, M., 2006, Identifying class size effects in developing countries: Evidence from rural Bolivia, *Review of Economics and Statistics*, vol. 88, pag. 171–176.
- Van Elk, R., M. Van der Steeg en D. Webbink, 2009, The effect of early tracking on participation in higher education, CPB Document 182.
- Van Veldhuizen, S., 2010, Een instrument voor de kwalitatieve analyse van onderwijsbeleid, CPB Memorandum 3/2010/01.
- Vandenbussche, J., P. Aghion en C. Meghir, 2006, Growth, distance to frontier and composition of human capital, *Journal of Economic Growth*, vol. 11, nr. 2, pag. 97–127.
- Veen, A., J. Roeleveld en P. Leseman, 2000, Evaluatie van Kaleidoscoop en Piramide. Eindrapportage, Amsterdam: SCO-KohnstammInstituut.
- Webbink, D., I. de Wolf, R. L. Woessmann, van Elk, B. Minne en M. van der Steeg, 2009, Wat is bekend over de effecten van kenmerken van onderwijsstelsels? Een literatuurstudie, CPB Document 187.
- Wilcox, B. en J. Gray, 1996, *Inspecting Schools: Holding Schools to Account and Helping Schools to Improve*, Open University Press, Buckingham.
- Woessmann, L., 2003, Central exams as the “currency” of school systems: International evidence on the complementarity of school autonomy and central exams, *Journal for Institutional Comparisons*, vol. 1, pag. 46–56.