

Hoofdafdeling : Institutionele Analyse
Afdeling : Kenniseconomie
Samenstellers : Dinand Webbink, Wolter Hassink
Nummer : 30
Datum : 29 januari 2002

Preventie van onderwijsachterstanden

In dit memorandum staat de vraag centraal hoe de middelen voor preventie van onderwijsachterstanden het best kunnen worden ingezet. Daarbij worden twee deelvragen onderscheiden. In welke onderwijsfase kunnen de middelen het best worden ingezet (de timing van de inzet van middelen) en welke interventie verdient daarbij de voorkeur (waaraan moeten de middelen worden besteed)? Het effect van een bepaalde interventie kan worden vastgesteld met een experiment of met een benadering van een experiment door een toevallige situatie in de werkelijkheid. In Nederland zijn dit soort studies zeldzaam. Daarom wordt in deze notitie een overzicht gegeven van de internationale literatuur. De kosten en effecten van verschillende interventies worden afgezet tegen die van het verkleinen van groepen. In Nederland worden namelijk de middelen voor preventie van achterstanden in het onderwijs vooral hieraan besteed. Het overzicht van de buitenlandse studies laat zien dat vroege interventies het meest opbrengen. De buitenlandse studies maken niet direct duidelijk waaraan de middelen het best kunnen worden besteed. Wel laten deze studies zien dat het verkleinen van groepen een relatief dure interventie is. De resultaten van buitenlandse studies zijn niet zonder meer geldig voor Nederland. Door te experimenteren met interventies in Nederland kan kennis worden verworven over wat werkt bij de bestrijding van onderwijsachterstanden. Het nadeel van experimenteren is echter dat leerlingen ongelijk behandeld kunnen worden.

1	Inleiding	4
2	Onderwijsmiddelen en effecten	5
2.1	Het Onderwijsmiddelen Debat: Does money matter?	5
2.2	Recente ontwikkelingen	8
2.3	Lessen uit het debat	11
3	Effecten van interventies in het voorschoolse traject	12
3.1	Inleiding	12
3.2	Vier modelprogramma's	13
3.3	Grootschalige projecten	15
3.4	Wat weten we over de kosten en baten van voorschoolse interventies?	19
4	Interventies in het schoolse traject	22
4.1	De 'klassieke' onderwijsinputs	23
4.1.1	Verkleining van groepsgrootte	24
4.1.2	Uitbreiding van het aantal lessen	27
4.1.3	Docenten	28
4.1.4	Effecten van ICT	31
4.1.5	Peergroup-effecten op school (spreidingsbeleid)	32
4.2	Veranderingen in de organisatie van het onderwijs	35
4.2.1	Het vergroten van de schoolkeuze	35
4.2.2	Teambeloning	37
4.2.3	Standaarden/ leerdoelen	39
4.2.4	Institutionele aspecten	40
4.3	Specifieke projecten	41
4.4	Conclusies	43
5	Ervaringen in Nederland	44
5.1	Schets van het onderwijsachterstandenbeleid in Nederland	45
5.2	Hoe worden de middelen voor achterblijvers besteed?	47
5.3	Wat weten we over de resultaten?	50
6	Samenvatting en conclusies	53

Literatuur	58
Bijlage 1 Het middelendebat	64
Bijlage 2 Specifieke projecten rond onderwijsuitval	67

Voorwoord

Het onderzoek is begeleid door een klankbordgroep bestaande uit: drs. G. de Boer (Ministerie van Onderwijs, Wetenschappen en Cultuur), mevr. drs. R. Bronneman-Helmers (SCP), mevr. drs. A.M. Bruggert (Ministerie van Onderwijs, Cultuur en Wetenschappen), drs. J.F.P. Hers (Ministerie van Financiën), prof. dr. H. Oosterbeek (Universiteit van Amsterdam), drs. G.H.G. Pelkman (Ministerie van Onderwijs, Cultuur en Wetenschappen) en dr. M. Vermeulen (IVA Tilburg). De onderzoekers danken hen allen hartelijk voor de deskundige begeleiding en constructieve bijdragen.

1 Inleiding

Al vele jaren worden omvangrijke inspanningen ondernomen om onderwijsachterstanden en schooluitval te verminderen. Tot dusver lijken deze inspanningen onvoldoende succesvol gezien het hardnekkige karakter van de problematiek. Om diverse redenen verdient dit beleidsthema de komende jaren nog meer aandacht dan in het verleden. In de eerste plaats zal de ontwikkeling naar een kenniseconomie waarschijnlijk leiden tot een stijging van de opleidingseisen die de arbeidsmarkt stelt. De kennisintensiteit van de productie zal toenemen hetgeen zijn weerslag zal vinden in de vraag naar arbeid. In de Verenigde Staten en in Engeland komt dit tot uiting in het stijgen van 'the labor market rewards to skill' en ook in Nederland zijn de eerste aanwijzingen hiervoor gevonden¹. Leerachterstanden of een niet afgeronde opleiding vergroten de kans op 'mismatch' op de arbeidsmarkt. In de tweede plaats zal als gevolg van de ontgroening en vergrijzing in onze samenleving de vervangingsvraag op de arbeidsmarkt groot zijn. Dit betekent dat een optimale benutting van talent meer dan ooit nodig is. In de derde plaats zal de komende jaren als gevolg van de demografische ontwikkeling de deelname aan het onderwijs toenemen van groepen die in het verleden vaker te maken hadden met achterstanden en uitval (de doelgroepen van het achterstandsbeleid). Een vierde reden voor aandacht heeft te maken met verdelingsvraagstukken. Door de Europese eenwording zullen landen naar verwachting meer op elkaar gaan lijken qua sociale zekerheid en belastingstelsel. Dit beperkt het gewicht van het economisch instrumentarium om de inkomensverdeling te beïnvloeden. Het belang van onderwijsbeleid als instrument ter beïnvloeding van de verdeling van het inkomen neemt daarmee toe.

In de voorliggende studie wordt een onderzoek uitgevoerd naar de kosten en effecten van beleid gericht op preventie van 'achterblijvers in de kenniseconomie'. De centrale vraag is hoe de middelen voor preventie van achterblijvers het best kunnen worden ingezet. Daarbij worden twee aspecten onderscheiden: (1) de *timing* van de inzet van middelen (in welke onderwijsfase), (2) *welke interventies* (waaraan moeten de middelen worden besteed). Hieraan voorafgaand wordt aandacht besteed aan de vraag of de inzet van middelen voor preventie van achterblijvers überhaupt zinvol is. Diverse studies wekken namelijk de indruk dat de omvangrijke uitgaven voor achterblijvers in Nederland weinig hebben opgeleverd².

¹ zie b.v. Card/ Lemieux (2000), voor Nederland zie Leuven en Oosterbeek (2000).

² zie bijvoorbeeld Mulder (1996).

Om deze vragen te kunnen beantwoorden wordt een overzicht gegeven van de economische literatuur. Wat leert de economische literatuur op dit moment over de preventie van achterblijvers? De Nederlandse ervaringen en aanpak van preventie van achterblijvers wordt vervolgens gezien in het licht van deze actuele kennis. Met andere woorden, nagegaan wordt of de besteding van achterstandsmiddelen in Nederland in het licht van de recente kennis efficiënt is of dat winst te behalen is uit een andere allocatie van middelen.

De opbouw van dit rapport is als volgt. In hoofdstuk twee wordt ingegaan op de vraag of het inzetten van middelen voor achterblijvers zinvol is. De vraag wordt besproken als onderdeel van een groter debat over de effecten van hogere onderwijsuitgaven. De lessen uit dit recente debat vormen het kader voor de bespreking van de effecten van verschillende onderwijsinterventies³. In hoofdstuk 3 wordt een overzicht gegeven van de ervaringen met interventies in het voorschoolse traject, hoofdstuk 4 geeft een overzicht van de interventies in het schoolse traject. Hoofdstuk 5 bespreekt de ervaringen in Nederland. In hoofdstuk 6 worden de Nederlandse aanpak afgezet tegen de actuele kennis uit de economische literatuur en worden conclusies getrokken.

2 Onderwijsmiddelen en effecten

Diverse studies suggereren dat de omvangrijke uitgaven voor achterblijvers in Nederland niet geleid hebben tot een vermindering van achterstanden. Hieruit zou de conclusie getrokken kunnen worden dat extra uitgaven voor achterblijvers weinig zinvol zijn. Dit type conclusies zijn in de Amerikaanse onderwijs-economische literatuur in het verleden diverse malen getrokken, waarbij het niet alleen ging om uitgaven voor achterblijvers maar om uitgaven voor onderwijs in den brede. Deze conclusies zijn echter niet onomstreden en de laatste jaren neemt de kritiek toe. Hieronder wordt een overzicht gegeven van dit debat, dat wij het onderwijsmiddelen-debat zullen noemen. De lessen van dit debat zullen we vervolgens gebruiken bij de aanpak van de literatuurstudie naar interventies in het onderwijs.

2.1 Het Onderwijsmiddelen Debat: Does money matter?

Leidt een verhoging van de middelen voor onderwijs tot een verbetering van de resultaten van leerlingen en studenten? Deze vraag staat centraal in het debat dat al enige jaren wordt gevoerd in de Amerikaanse onderwijs-economische literatuur. Aan de ene kant wordt, in een reeks van

³ Onderwijsinterventies zijn gerichte ingrepen/veranderingen in het onderwijs, zoals het verkleinen van groepen, programma's voor voorschoolse educatie of training voor docenten.

invloedrijke studies door Hanushek (1986, 1989, 1996a, 1996b, 1997), de positie ingenomen dat er geen systematische relatie bestaat tussen onderwijsuitgaven en de prestaties van leerlingen. Het onderwijssysteem is kennelijk niet in staat om additionele middelen om te zetten in betere prestaties. Volgens sommige auteurs is dit een belangrijke aanwijzing dat de 'incentive'-structuur in het (publieke) onderwijs niet deugt (b.v. Chubb and Moe, 1990). Aan de andere kant wordt, met name door Card en Krueger (1992, 1996) in enkele al even zeer invloedrijke studies, betoogd dat 'testscores in het onderwijs' slechts één van de opbrengsten zijn van het onderwijsproces. Als gekeken wordt naar een andere opbrengst van het onderwijsproces, namelijk de latere positie op de arbeidsmarkt, vinden zij dat meer middelen voor onderwijs leiden tot betere prestaties op de arbeidsmarkt. Deze schijnbaar tegengestelde bevindingen worden aangeduid als 'de onderwijsuitgaven puzzel' (Burtless, 1996).

Onderwijsmiddelen hebben geen effecten op prestaties van leerlingen

Twee soorten empirisch bewijs liggen ten grondslag aan het eerste deel van de puzzel. Allereerst blijkt de historische toename van de onderwijsuitgaven per leerling niet geleid te hebben tot een verbetering van de leerprestaties. In de Verenigde Staten zijn de reële uitgaven per leerling de afgelopen honderd jaar vrijwel continu gestegen met 3,5 procent per jaar, en ook tussen 1970 en 1990 met 70 procent. Een deel van deze stijging ging naar salarissen van docenten maar een belangrijk deel ging ook naar het intensiveren van de instructie en ondersteuning. De opbrengsten zijn teleurstellend, sinds 1970 zijn de leerprestaties van leerlingen op zijn hoogst constant gebleven dan wel gedaald⁴.

Het tweede soort empirisch bewijs betreft onderzoek naar de relatie tussen enerzijds kenmerken van scholen, leraren, leerlingen en ouders en anderzijds de leerprestaties. Dit type onderzoek start met het befaamde Coleman-report (Coleman, e.a., 1966) waarin geconcludeerd werd dat kenmerken van scholen en leraren slechts een geringe rol spelen bij verschillen in leerprestaties. De sociaal-economische achtergrond van de leerlingen en de samenstelling van de leerlingenpopulatie op scholen zijn, zo werd geconcludeerd, veel belangrijker voor de prestaties van leerlingen. Met andere woorden, scholen spelen slechts een ondergeschikte rol bij leerprestaties en het gezin is het sleutelement voor studiesucces. En, het effect van onderwijsmiddelen op leerprestaties is klein en onzeker.

Na het Coleman-report is veel onderzoek uitgevoerd naar de invloed van scholen, leraren, ouders en medeleerlingen (peers) op de resultaten van leerlingen. Hanushek (1996) heeft de resultaten van 377 schattingen uit 90 studies samengebracht en daarbij gekeken naar het effect van de docent/leerling ratio, de opleiding van de leraar, de ervaring van de leraar, het salaris van de leraar, de uitgaven per leerling, administratieve inputs en faciliteiten. In tabel 2.1 zijn de

⁴ Voor eerdere perioden zijn geen vergelijkbare tijdreeksen beschikbaar.

resultaten van deze exercitie weergegeven. Aangegeven is welk deel van schattingen in statistische zin betekenis heeft (significant is) of niet en wat de richting is van het effect (positief of negatief).

In het overzicht van Hanushek worden bijvoorbeeld 277 schattingen van het effect van de ratio docent-leerling meegenomen. Bij 28 procent van deze studies wordt een statistisch significant effect gevonden waarbij 15 procent met een positief effect en 13 procent met een negatief effect op de leerprestaties. In 52 procent van de studies wordt geen significant effect gevonden, de verdeling over studies met positieve en negatieve effecten is respectievelijk 27 procent en 25 procent.

Tabel 2.1 Verdeling van geschatte effecten van de besteding van middelen aan sleutel inputs van het onderwijs op de leerprestaties, gebaseerd op 377 schattingen uit 90 studies

Middelen	# schattingen	statistisch significant		Niet significant		onbekend
		positief	negatief	positief	negatief	
docent-leerling ratio	277	15%	13%	27%	25%	20%
opleiding docent	171	9	5	33	27	26
ervaring docent	207	29	5	30	24	12
salaris docent	119	20	7	25	20	28
uitgaven per leerling	163	27	7	34	19	13
administratieve inputs	75	12	5	23	28	32
faciliteiten	91	9	5	23	19	44

bron: Hanushek (1996)

Het beeld dat ontstaat uit tabel 2.1 is dat de onderzoeksresultaten over de effecten van 'sleutel-inputs' op de prestaties van leerlingen sterk variëren. Hoewel vaker positieve effecten gevonden worden dan negatieve effecten concludeert Hanushek dat onderzoekers geen systematische relatie gevonden hebben tussen de genoemde 'onderwijsinputs' en de resultaten van leerlingen.⁵

Onderwijsmiddelen zijn van invloed op prestaties op de arbeidsmarkt

Het tweede element uit de onderwijsuitgaven 'puzzle' betreft de bevindingen dat onderwijsmiddelen een meetbare en betekenisvolle invloed hebben op de latere lonen van leerlingen (Card&Krueger (1992, 1996)). Dit effect is samengesteld uit twee delen. Het eerste is dat de financiële opbrengsten van een jaar scholing hoger zijn naarmate dit jaar scholing is gevolgd in een omgeving die beschikte over meer onderwijsmiddelen. Het tweede deel van het

⁵ Deze conclusies van Hanushek bleven niet zonder kritiek. Een kort overzicht van deze kritiek is gegeven in bijlage 1.

effect is dat leerlingen in een 'rijkere' onderwijsomgeving meer (langer) onderwijs volgen. De conclusies van C&K zijn, naast de eigen studie, gebaseerd op een groot aantal en gevarieerde studies. Hun bevindingen sluiten direct aan op resultaten van onderzoek naar loonverschillen tussen 'witte' en 'zwarte' Amerikanen. De loonverschillen namen af voor cohorten die in de jaren zestig de arbeidsmarkt betraden (Welch, 1973). In die periode was ook de ongelijkheid in beschikbare middelen voor onderwijs voor deze groepen sterk gereduceerd.⁶

2.2 Recente ontwikkelingen

Twee recente ontwikkelingen werpen nieuw licht op de discussie. Allereerst is onlangs een replicatie van de klassieke studie van Hanushek uitgevoerd door Krueger (2000) die een belangrijke doorbraak lijkt te betekenen. In de tweede plaats worden de laatste jaren nieuwe methoden toegepast waarmee ernstige tekortkomingen van eerdere studies worden voorkomen.

Een replicatie van Hanushek's klassieke studie

De eerste recente ontwikkeling betreft de replicatie van de meta-analyse van Hanushek door Krueger. Hij baseert zich daarbij alleen op de studies waarin schattingen worden gemaakt van het effect van de docent-leerling ratio op de leerprestaties, dat zijn 277 schattingen uit 59 studies (zie de studies naar de docent-leerling ratio in tabel 2.1). Het resultaat van Hanushek's 'vote counting' exercitie is dat het aantal schattingen met een positieve en negatieve uitkomst vrijwel gelijk is. Met andere woorden, er bestaat geen relatie tussen docent-leerling ratio en prestaties van leerlingen.

Krueger laat zien dat Hanushek niet alle studies even zwaar laat meewegen. Uit studies die een negatief effect vinden van uitgaven op prestaties worden meer schattingen gebruikt dan uit andere studies. In de meerderheid van de studies worden echter positieve effecten op prestaties gevonden terwijl hieruit slechts een beperkt aantal schattingen wordt gebruikt. Als meerdere schattingen worden gebruikt uit een studie zijn dit meestal schattingen op deelsteekproeven. Vanwege de kleinere steekproef hebben deze schattingen een minder sterke statistische onderbouwing. Uit 9 studies worden 123 schattingen afgeleid, dat is 44,4 % van alle schattingen. Nadere bestudering van de negen zwaarst meewegende studies laat zien dat deze studies niet de meest overtuigende waren. Verschillende keren is een verkeerde specificatie (opzet) van het model gehanteerd. Als alle studies hetzelfde gewicht gegeven wordt blijken onderwijsuitgaven systematisch positief samen te hangen met prestaties van leerlingen. Het positieve effect wordt nog sterker als de 'impact factor' van het tijdschrift waarin de studie is gepubliceerd wordt meegewogen. Met andere woorden, de studies die door Hanushek zijn samengebracht geven

⁶ Voor kritiek op de positie van Card&Krueger, zie bijlage 1.

een positief verband aan tussen onderwijsuitgaven en prestaties van leerlingen⁷. Meer in het algemeen acht Krueger de 'vote counting exercise' niet geschikt: onderzoek is niet democratisch, het gaat om de kwaliteit van de studie. En op dit punt is de laatste jaren belangrijke vooruitgang geboekt. Daarover gaat de tweede ontwikkeling.

Nieuwe methoden om causale effecten te bepalen

Om de causale effecten van de verschillende sleutelinputs op prestaties van leerlingen in het onderwijs vast te stellen zou idealiter een experiment moeten worden uitgevoerd waarbij leerlingen door toeval worden toegewezen aan verschillende situaties. Door een opzet met een experimentele en een controlegroep, waarbij leerlingen door toeval worden toegewezen aan één van beide groepen, is het mogelijk om het effect van een bepaalde input vast te stellen zonder dat dit effect vertekend wordt door (niet waargenomen) verschillen tussen leerlingen. Door Currie (2001) wordt deze experimentele opzet 'the gold standard' van het onderzoek genoemd. Het meest bekende experiment in het onderwijs is het STAR-project in Tennessee naar het effect van klassenverkleining (zie paragraaf 4.1.1). In het merendeel van de studies is echter geen sprake van een experimentele opzet. In dat geval is het lastig om de causale effecten vast te stellen. Zeer vele factoren kunnen namelijk een rol spelen terwijl onderzoekers in de praktijk slechts rekening kunnen houden met een beperkte set van variabelen. Daardoor is het mogelijk dat niet waargenomen factoren de onderzoeksresultaten vertekenen. Anders gezegd, het effect dat in het onderzoek wordt toegeschreven aan een bepaalde onderwijsinterventie kan in werkelijkheid worden veroorzaakt door niet waargenomen factoren die samenhangen met deze interventie. Dit wordt het endogeniteitsprobleem genoemd. Endogeniteitsproblemen komen in het onderwijs vaak voor omdat vrijwel alle inputs van het onderwijs de uitkomst zijn van beslissingen van ouders en scholen. Recent zijn in diverse studies nieuwe methoden toegepast waarmee getracht wordt rekening te houden met de endogeniteit van onderwijsinputs. Het gaat dan om quasi-experimentele benaderingen waarbij gebruik gemaakt wordt van toevallige natuurlijke experimenten. In de onderstaande box is een voorbeeld gegeven uit een recente studie over het effect van groepsgrootte op prestaties van leerlingen.

⁷ In een recente reactie wordt deze conclusie aangevochten (Hanushek, 2000). De argumentatie lijkt echter niet overtuigend omdat de kwestie van de deelsteekproeven onjuist wordt voorgesteld.

Het effect van de groeps grootte op de leerlingprestaties in Israël

Verschillende factoren, die vaak niet door de onderzoeker worden waargenomen, kunnen van invloed zijn op de groeps grootte en tevens bijdragen aan de prestaties van leerlingen. Scholen kunnen er bijvoorbeeld voor kiezen om zwakkere leerlingen of leerlingen met gedragsproblemen in kleinere klassen te plaatsen. Ouders die grote waarde hechten aan de leerprestaties van hun kinderen kunnen een voorkeur hebben voor scholen met gemiddeld kleine groepen. Als bovendien het belang dat ouders hechten aan de prestaties van hun kinderen een onafhankelijk effect heeft op de leerprestaties (bijvoorbeeld door meer hulp bij het huiswerk), maar dit niet wordt waargenomen, ontstaat een vertekend effect van de groeps grootte. Door hiermee geen rekening te houden kunnen de effecten die worden toegeschreven aan de groeps grootte in werkelijkheid veroorzaakt worden door deze niet waargenomen factoren.

In een recente studie wordt rekening gehouden met de endogeniteit van de groeps grootte door gebruik te maken van een (toevallig) element uit de bekostigingssystematiek. In Israël mag bijvoorbeeld een groep maximaal bestaan uit 40 leerlingen, bij 41 leerlingen dienen twee groepen gevormd te worden. De groeps grootte daalt dan abrupt van 40 naar 20,5 leerlingen. Voor deze variatie in de groeps grootte weten we dan dat niet-waargenomen verschillen tussen leerlingen en hun ouders geen rol spelen bij het bepalen van het effect van de groeps grootte. Het causale effect van de verandering in de groeps grootte kan dan met meer zekerheid worden vastgesteld. Als geen rekening wordt gehouden met de endogeniteit wordt in deze studie geen effect gevonden van de groeps grootte op de prestaties van leerlingen. Met de quasi-experimentele benadering wordt daarentegen gevonden dat kinderen in kleinere groepen beter presteren.

In de traditionele studies, zoals in het overzicht van Hanushek, wordt geen rekening gehouden met de endogeniteit van de verschillende inputs van het onderwijs. De effecten van deze inputs kunnen daarmee in meer of mindere mate vertekend worden door niet-waargenomen verschillen tussen leerlingen, ouders en scholen. Aannemelijk is dat dit de oorzaak is van de grote variatie in de uitkomsten van de traditionele studies. Zoals gezegd, deze grote variatie in uitkomsten is het fundament voor de positie van Hanushek in het middelendebat.

2.3 Lessen uit het debat

1. Het voorgaande wijst erop dat het fundament onder Hanushek's 'gloomy vision' op het onderwijssysteem steeds meer afbrokkelt. Meer middelen voor onderwijs lijken wel degelijk te kunnen leiden tot betere leerprestaties. De wijze waarop de middelen besteed worden is daarbij van groot belang.
2. Duidelijk is ook dat het vaststellen van de effecten van additionele onderwijsmiddelen of onderwijsinterventies lastig is. Op dit punt wordt de laatste jaren door experimenten of quasi-experimentele benaderingen belangrijke vooruitgang geboekt⁸. Deze methoden maken het mogelijk om de causale effecten vast te stellen. De resultaten van eerdere studies werden vaak vertekend door niet-waargenomen factoren van ouders, leerlingen of scholen.
3. Het debat maakt ook duidelijk dat de opbrengsten van het onderwijsproces meer omvatten dan de gemeten leerprestaties in het onderwijs. Sterker nog, door alleen te kijken naar gemeten leeropbrengsten kunnen belangrijke opbrengsten buiten beeld blijven, zoals de arbeidsmarktresultaten (hogere lonen, meer kans op werk, minder kans op inactiviteit) of andere sociale opbrengsten in de sfeer van veiligheid en gezondheid .

Aanpak van de literatuurbespreking

De lessen uit het debat vormen het uitgangspunt bij de bespreking van de literatuur over de effecten van interventies in het onderwijs. In de bespreking zal de meeste aandacht worden besteed aan studies die (1) een experimentele of quasi-experimentele aanpak hebben en waarin ook (2) lange termijn effecten worden geanalyseerd. Om inzicht te krijgen in de invloed van de 'timing' van de interventie, worden allereerst interventies in het voorschoolse traject besproken en komen vervolgens interventies in het schoolse traject aan de orde. Per interventie zal de stand van zaken worden weergegeven.

Om welke leerlingen gaat het?

In de voorliggende studie gaat het in de eerste plaats om het effect van interventies op 'students at risk', dat wil zeggen leerlingen die ten gevolge van sociale, economische of culturele factoren onderwijsachterstanden kunnen oplopen. Het gaat derhalve om de doelgroepen van het onderwijsachterstandenbeleid in Nederland. Deze afbakening betekent dat leerlingen die vanwege andere oorzaken onderwijsachterstand kunnen oplopen, zoals fysieke of geestelijke handicaps of leerlingen met ernstige leer- of gedragsproblemen, en die thans vaak verwezen worden naar het speciaal onderwijs, niet tot de doelgroep van deze studie behoren.

⁸ Krueger (2000) stelt in dit verband dat experimenten de enig begaanbare weg zijn om de effecten van onderwijsinterventies vast te stellen en daarmee het onderwijs te kunnen versterken.

Ethische bezwaren tegen experimenten in het onderwijs

Het toepassen van experimenten stuit vaak op ethische bezwaren. Deelnemers aan het experiment worden anders behandeld dan niet-deelnemers. Deze bezwaren dienen te worden afgewogen tegen de voordelen van het experiment. Met het experiment kunnen de effecten van bepaalde interventies met grotere zekerheid worden vastgesteld. Dit kan het onderwijs voor alle volgende generaties versterken. Succesvolle interventies kunnen breder worden toegepast, niet succesvolle interventies kunnen worden gestaakt. Bij de ontwikkeling van geneesmiddelen is het toepassen van experimenten heel gebruikelijk.

3 Effecten van interventies in het voorschoolse traject

3.1 Inleiding

De laatste jaren is sprake van een toenemende belangstelling voor interventies in het voorschoolse traject. Binnen de OECD wordt een programma uitgevoerd dat beoogt het kinderopvang- en vroegschoolse beleid in 12 lidstaten in kaart te brengen (Thematic Review of Early Childhood and Care Policy). De International association for the Evaluation of Educational Achievement (IEA) voert thans een studie uit naar voorschoolse educatie in 17 landen (IEA Preprimary Project). In Nederland worden voor- en vroegschoolse programma's sinds enige jaren ingezet ter bestrijding van onderwijsachterstanden en zijn recent enige experimenten uitgevoerd. De toename van de belangstelling voor deze programma's heeft verschillende redenen. Allereerst is in toenemende mate het besef doorgedrongen dat 'an ounce of prevention is worth a pound of cure'⁹. In de tweede plaats zijn er bevindingen in het onderzoek naar de werking van de hersenen die er op duiden dat de vroege kinderjaren cruciaal zijn voor de ontwikkeling van de hersenen.¹⁰ Een derde (Nederlandse) reden is dat het onderwijsachterstandenbeleid, dat sinds 1985 wordt uitgevoerd, nauwelijks zichtbare resultaten heeft opgeleverd.

De literatuur over de effecten van voorschoolse stimuleringsprogramma's beslaat vele honderden studies, vele reviewstudies en enkele reviews van reviews. Het overgrote deel van de studies heeft betrekking op onderwijsinterventies in de Verenigde Staten, maar ook in vele andere landen zijn studies uitgevoerd. Vanuit verschillende disciplines zijn bijdragen geleverd aan de literatuur. In de psychologische literatuur is veel aandacht besteed aan de mogelijke

⁹ Zie Heckman (2000).

¹⁰ Zie Currie (2001).

schadelijke effecten van kinderopvang voor de emotionele ontwikkeling van het kind. Onderwijswetenschappers richten zich op mogelijke positieve effecten voor de prestaties in het onderwijs. Vanuit de economische wetenschap was er in eerste instantie vooral belangstelling voor de relatie met het arbeidsaanbod van vrouwen. De laatste jaren is er in toenemende mate aandacht voor de effecten van vroege onderwijsinterventies op de opbouw van human capital en de hieraan verbonden kosten en baten. Behalve de effecten van interventies op onderwijsprestaties worden daarbij ook (lange termijn) effecten op de arbeidsmarktpositie (inkomen, kans op werk/ inactiviteit) of op crimineel gedrag, bestudeerd.

Het doel van deze paragraaf is om de internationale inzichten over de kosten en baten van interventies in het voorschoolse traject in kaart te brengen. Dit betekent dat we met name geïnteresseerd zijn in studies waarin de lange termijn effecten van de interventies worden gemeten en waarin gebruik gemaakt wordt van een experimenteel design. Recent zijn enkele uitstekende besprekingen van de literatuur over voorschoolse interventies verschenen die de basis vormen voor het navolgende overzicht.¹¹ In navolging van Currie (2001) bespreken we allereerst de resultaten van vier (model)programma's met relatief veel middelen en goed getraind personeel (paragraaf 3.2). Een beperking van de ervaringen met modelprogramma's is dat we niet weten in hoeverre de resultaten ook geldig zijn bij grootschalige toepassingen. In paragraaf 3.3 gaan we daarom in op de ervaringen met enkele grootschalige programma's. Daarbij komen naast Amerikaanse studies ook enkele Europese studies aan de orde. In paragraaf 3.4 worden de kosten en baten van de programma's, voor zover bekend, op een rijtje gezet.

3.2 Vier modelprogramma's

Het doel van voorschoolse interventies is het vergroten van de kans om als volwassenen succesvol te functioneren in de samenleving. Hieruit volgt dat we in de beoordeling van voorschoolse interventies in de eerste plaats geïnteresseerd zijn in de lange termijn effecten. Het functioneren in de samenleving kan gemeten worden in termen van het bereikte onderwijsniveau, de positie op de arbeidsmarkt of deelname aan criminele activiteiten. Bij de bespreking van de middelendiscussie speelde dit aspect, lange-termijn of korte-termijn effecten, ook een belangrijke rol (zie vorige hoofdstuk). Het aantal studies dat daadwerkelijk lange-termijn effecten meet is beperkt¹². De overtuigingskracht van een studie neemt sterk toe als sprake is van een experimenteel design met aselechte toewijzing aan een experimentele en

¹¹ Zie Barnett (1992, 1995), Karoly e.a. (1998) en Currie (2001).

¹² zie Barnett, 1995 voor een overzicht.

controle groep. Een dergelijk design komt echter slechts in een beperkt aantal studies voor. Als vervolgens nog een derde eis wordt opgelegd, namelijk dat de uitval gedurende het onderzoek gering is, resteren slechts vier studies: het Early Training project, het Carolina Abecedarian project, het Perry Preschool project en het Milwaukee project.

Het Early Training werd uitgevoerd in het begin van de jaren zestig in Murfreesboro, Tennessee. Het programma was gericht op 4- en 5-jarigen en bestond uit wekelijkse bezoeken aan het gezin in aanvulling op tien weken deelname aan een deeltijd voorschools traject gedurende twee of drie zomers. Vergeleken met de drie andere programma's is het Early Training project het minst intensief. Als uitkomst werd gevonden een sterke daling van de verwijzing naar het speciaal onderwijs tot 12 jaar (5 % van de experimentele groep versus 29 % in de controle groep) en, hoewel niet significant, werden hogere testcores, minder uitval en het vaker behalen van een high school diploma gevonden.

Het Abecedarian Project werd uitgevoerd in de jaren zeventig en was gericht op kinderen met een relatief ongunstige sociaal-economische achtergrond tussen 4 maanden en 8 jaar. Ook hier was sprake van 'random' toewijzing van kinderen aan het project (en aan de controlegroep). De kinderen werden fulltime opgevangen door het kinderdagverblijf gedurende het hele jaar. Bij de activiteiten met de kinderen lag de nadruk op taalontwikkeling in een rijke leeromgeving. De resultaten werden op de korte en langere termijn, op een leeftijd van 15 en 21 jaar, onderzocht. Op 15 jarige leeftijd bleken de deelnemers t.o.v. de controlegroep beter te presteren in taal en rekenen en minder vaak te doubleren of te worden verwezen naar het speciaal onderwijs. Op 21-jarige leeftijd bleek dat de deelnemers langer onderwijs bleven volgen, vaker een universitaire studie volgden, vaker een baan kregen en later begonnen aan gezinsvorming. Het project liet ook zien dat de effecten sterker waren voor kinderen die op jongere leeftijd deelname aan het project (0-5 jaar) dan voor degenen die later deelnamen.

Waarschijnlijk het meeste bekende en meest uitgebreid besproken programma is het Perry-Preschool programma van de High/Scope Foundation dat in het begin van de jaren zestig is uitgevoerd in Ypsilanti in de Verenigde Staten. Kinderen met een relatief ongunstige sociaal-economische achtergrond en met een IQ lager dan gemiddeld werden 'random' toegewezen aan een intensief één tot tweejarig programma waarbij de ouders wekelijks werden bezocht. De verhouding leerling-docent was 6 op 1 en aan de opleiding van de docent werden hoge eisen gesteld. Het unieke aan het programma is dat de deelnemers langdurig werden gevolgd; er is informatie beschikbaar over de situatie van de deelnemers bij een leeftijd van 27 jaar. Deze informatie is gebruikt om een kosten-baten analyse uit te voeren. Deelnemers aan het programma hadden hogere testcores, volgen meer onderwijs, verdienen hogere salarissen en

zijn minder vaak betrokken bij criminele activiteiten dan degenen in de controlegroep. Gemeten tot en met 27 jaar levert elke dollar besteed aan het programma 5.7 dollar op, projecties naar de toekomst geven een totale verwachte opbrengst van 8.7 dollar. De belangrijkste opbrengst van het programma is de reductie van crimineel gedrag (65 % van de totale opbrengst).

Het Milwaukee-project, uitgevoerd tussen 1968 en 1978, bestond uit voltijds kinderopvang en training van moeders. Het project was gericht op kinderen van 3-6 maanden en liep tot en met 5-jarigen. De resultaten wijken af van de andere projecten in de zin dat kinderen in de experimentele groep hogere IQ-scores hadden rond 14 jaar maar niet beter scoorden in het onderwijs. Ook hier lijkt de IQ-score een ander beeld te geven van de resultaten van de interventie dan andere indicatoren.

In de overzichtsstudie van Barnett (1995) worden nog diverse andere studies besproken waarin lange termijn effecten zijn gemeten. Deze studies kennen echter geen experimenteel design. In het algemeen worden ook hier positieve effecten gevonden (zie ook Karoly, 1998), hoewel de effecten minder vaak significant zijn. De vier modelprogramma's laten zien dat voorschoolse interventies positieve lange termijn effecten kunnen hebben voor kinderen, in termen van betere prestaties in het onderwijs, op de arbeidsmarkt en een afname van de deelname aan criminele activiteiten.

3.3 Grootschalige projecten

De gevonden resultaten van modelprogramma's hoeven niet meteen geldig te zijn bij grootschalige toepassing. Voor modelprogramma's zijn vaak veel middelen beschikbaar. Ook zijn de bedenkers van bepaalde programma's soms direct betrokken bij de implementatie hetgeen een grote inzet en motivatie kan opleveren om het programma tot een succes te maken. Bovendien zijn de steekproeven van de modelprogramma's zonder uitzondering klein. Daarom besteden we hieronder aandacht aan de ervaringen met enkele grootschalige programma's. De studies naar de effecten van deze programma's voldoen echter niet aan 'the gold standard' van het onderzoek: een experimentele opzet met 'random' toewijzing. Allereerst bespreken we de belangrijkste Amerikaanse ervaringen en daarna de Europese.

Amerikaanse ervaringen

Geïnspireerd door het Perry Pre-school project is het programma Head Start gestart in 1965 als een zomerproject voor 561.000 Afro-Amerikaanse kinderen van 4 en 5 jaar. Het doel was het verbeteren van de sociale competentie van jonge kinderen uit arme gezinnen. De interventie bevatte zowel een gezinsgerichte component als een centrum-gerichte component. In de

daaropvolgende jaren veranderde het programma sterk en begin jaren zeventig was sprake van een jaarprogramma met grote verschillen tussen locaties. Vergeleken met het Perry Preschool project zijn voor Head Start projecten veel minder middelen beschikbaar gesteld. Inmiddels hebben meer dan 15 miljoen kinderen het Head Start programma (of diverse varianten hiervan) gevolgd. De evaluatie van de vele projecten binnen het Head Start programma beslaat een omvangrijke literatuur¹³. Een opmerkelijk feit daarbij is dat nooit een evaluatie van een Head Start programma is uitgevoerd met een experimenteel design. Het gevolg daar van is dat het causale effect van het programma een voortdurend terugkerend discussiepunt is bij de evaluaties. Een ander lastig punt is de grote variatie aan Head Start projecten in termen van besteding van middelen. Tot voor kort betrof de evaluatie met name de korte termijn effecten. In het algemeen werden op korte termijn hogere cognitieve skills, verbetering van taalvaardigheid, gezondheid en van sociale skills gemeten. Deze meting is niet gebaseerd op een vergelijking met een controle groep maar met het landelijk gemiddelde. Daarmee kan niet met zekerheid worden gezegd of de gevonden effecten werkelijk veroorzaakt worden door het Head Start programma. De gevonden gunstige effecten bleken na enkele jaren uit te doven. Nadere analyse van het uitdoven van deze effecten leerde dat dit waarschijnlijk te maken had met de kwaliteit van het vervolgonderwijs¹⁴.

Onlangs is voor het eerst een studie uitgevoerd naar de lange termijn effecten van Head Start bij volwassen tot ongeveer 30 jaar (Garces, Thomas & Currie, 2001). In deze studie worden deelnemers aan Head Start vergeleken met niet deelnemers. Met behulp van paneldata kan vrij gedetailleerd rekening worden gehouden met verschillen in sociale achtergrond en de omgeving waarin de kinderen opgroeiden. Voor blanke Amerikanen worden substantiële verbeteringen gevonden voor de onderwijsloopbaan: de kans op voltooiën van de High School (+20 procentpunt), de kans op instromen in een college (+28 procentpunt). Voor zwarte Amerikanen worden geen effecten gevonden op de onderwijsloopbaan maar wel wordt een vermindering van crimineel gedrag gevonden. Het effect op het salaris kon nog niet worden vastgesteld .

In 1967 kende de Amerikaanse federale regering de zogenoemde Title I fondsen toe aan elf 'Child Parent Centers' in Chicago bestemd voor achterstandskinderen. Momenteel ontvangen 24 centra in Chicago en Illinois deze federale fondsen. Anders dan de meeste programma's bestaat het CPC zowel uit een voor- en voerschoolse component als uit een vervolg in het basisonderwijs waardoor de totale interventie kan oplopen tot 6 jaar. Het programma voorziet in halve dagen voerschoolse educatie, halve of hele dagen kleuterschool en hele dagen vervolprogramma's in de eerste drie jaar van het basisonderwijs. Het programma is gericht op

¹³ Volgens Karoly, e.a. (1998) zijn ongeveer 600 citaties te vinden.

¹⁴ zie Currie&Thomas (1998).

ouder en kind en benadrukt de lees- en taalvaardigheden met een kindgerichte benadering. Er zijn gegevens beschikbaar over de effecten van het programma voor kinderen van 13 jaar. De gegevens zijn met diverse statistische technieken geanalyseerd waarbij rekening wordt gehouden met niet waargenomen kenmerken van kinderen in de experimentele en in de controle groep (het project kent geen 'random' toewijzing)¹⁵. Vier jaar na afloop levert het programma winst op voor de prestaties in rekenen en wiskunde en wordt een afname van het zittenblijven gevonden. Recent zijn ook gegevens beschikbaar gekomen voor kinderen aan het eind van de High School. Gevonden is dat het programma de uitval in High School vermindert met 24 % en dat het effect toeneemt naarmate kinderen langer aan het programma hebben deelgenomen. In een andere studie is gekeken naar de uitkomsten ten aanzien van crimineel gedrag en opgedane skills. Een kosten baten analyse op deze uitkomsten laat zien dat elke dollar besteed aan het programma een besparing van de toekomstige kosten van de overheid oplevert van bijna 4 dollar.

Deze studies naar de resultaten van grootschalige Amerikaanse projecten bevestigen de conclusies uit de modelprogramma's dat voorschoolse interventies positieve lange termijn effecten kunnen opleveren voor kinderen.

Europese ervaringen

Hoewel de literatuur over voorschoolse interventies wordt gedomineerd door Amerikaanse programma's zijn de meest ontwikkelde systemen voor voorschoolse kinderopvang te vinden in West-Europese landen¹⁶. De studies naar de effecten van deze programma's komen evenwel in reviews weinig aan de orde. Aannemelijk is dat dit te maken heeft met het design van deze studies. Studies met een experimenteel design waarin deelnemers random worden toegewezen aan de experimentele of aan de controlegroep komen niet voor. Ook worden de lange termijn effecten in Europese studies weinig gemeten. Zoals eerder opgemerkt, bemoeilijkt dit de beoordeling van de resultaten van deze studies omdat de effecten van niet-waargenomen variabelen toegeschreven kunnen worden aan het programma. Ondanks deze methodologische tekortkomingen bespreken we hieronder een aantal Europese ervaringen omdat de context meer lijkt op de Nederlandse.

In Frankrijk neemt momenteel bijna 100 procent van de drie tot vijf-jarigen deel aan de école maternelle. Docenten van deze voorschoolse educatie hebben dezelfde opleiding, ambtelijke status en salariëring als docenten in het primair onderwijs. Tussen 1983 en 1990 zijn de

¹⁵ Reynolds (1998).

¹⁶ Ontleend aan Spence Boocock (1995).

effecten onderzocht. Ieder jaar dat een leerling deelneemt aan de voorschoolse educatie vermindert de kans op schooluitval, vooral voor kinderen met een lagere sociaal-economische achtergrond. Het verschil in uitvalpercentage tussen kinderen met de hoogste sociaal-economische achtergrond en kinderen met de laagste sociaal-economische achtergrond neemt af van 30 % bij één jaar voorschoolse educatie naar 27 % bij twee jaar en naar 24 % bij drie jaar. In het Verenigd Koninkrijk werden de effecten van voorschoolse en kinderopvang programma's onderzocht met een grootschalig survey, de Child Health and Education Study (CHES). Geschat wordt dat 44 % van de Britse drie- en vierjarigen publieke of private kinderopvang bezoeken, en veel kinderen beginnen aan 'public school' onderwijs met vier jaar. De CHES is een longitudinaal onderzoek onder 9000 kinderen geboren in dezelfde week in 1970. Gegevens werden verzameld over deze kinderen op 5- en 10-jarige leeftijd. Het onderzoek laat zien dat deelname aan enig vorm van voorschoolse educatie of kinderopvang bijdraagt aan de cognitieve ontwikkeling en de schoolprestaties van de kinderen. Kinderen met een lagere sociaal-economisch achtergrond profiteerden daarbij licht meer dan andere kinderen. Een vergelijkbare studie werd in 1987 in West-Duitsland uitgevoerd. Ongeveer 65 tot 70 procent van de West-Duitse kinderen tussen drie en zes bezoekt 'kindergartens' gedurende 5 halve dagen per week. Deze voorschoolse educatie wordt geheel gratis aan ouders aangeboden. Voor de evaluatie zijn op 203 basisscholen gegevens verzameld. Het onderzoek laat zien dat basisscholen in districten met een hoge beschikbaarheid van voorschoolse voorzieningen een lager percentage zittenblijvers en minder verwijzing naar speciaal onderwijs hebben. Dit suggereert dat voorschoolse educatie leidt tot een verbetering van de 'school readiness' van kinderen en daarmee de overgang naar het basisonderwijs vergemakkelijkt.

Zweden beschikt over een uitgebreid systeem van kinderopvang, hetgeen alles te maken heeft met de hoge arbeidsparticipatie van Zweedse moeders met jonge kinderen (86 %). Ongeveer 47 % van de kinderen bezoekt kinderopvang tussen geboorte en 7 jaar. Daarnaast worden veel kinderen door ouders zelf opgevangen gedurende betaald verlof. In een studie worden 128 kinderen geboren in 1975 met een gemiddelde of lagere sociale achtergrond gevolgd van 3 jaar tot 13 jaar. Voor de deelnemers werden belangrijke voordelen gevonden: grotere taalbeschikbaarheid, grotere vasthoudendheid en zelfstandigheid, meer zelfvertrouwen. Deze resultaten werden bevestigd in twee andere studies.

In Ierland is een experiment uitgevoerd gericht op kinderen in een arme wijk in Dublin. Het Dublin-project behelst een tweejarig programma van halve dagen voor 90 driejarigen in 1969. Het doel van het project was het verbeteren van cognitieve ontwikkeling, leervaardigheden, persoonlijkheid en sociale ontwikkeling. De ouders werden direct bij het project betrokken. De cognitieve ontwikkeling, schoolprestaties en ouderbetrokkenheid werd gemeten op 5-jarige, 8-jarige en 16-jarige leeftijd. De controlegroep bestond uit 60 kinderen uit dezelfde wijk.

Op 5-jarige leeftijd scoorden de deelnemende kinderen duidelijk beter op gestandaardiseerde toetsen. Dit gold vooral voor de leerlingen met de grootste achterstanden en voor meisjes. Op 8-jarige leeftijd waren de meeste voordelen evenwel verdwenen. In het voortgezet onderwijs hadden deelnemers aan het programma een grotere kans om in het onderwijs te blijven en twee tot driemaal zoveel kans om examens te doen vereist voor het vervolgonderwijs. Er waren geen verschillen in de situatie op de arbeidsmarkt of crimineel gedrag. De onderzoekers concluderen dat het effect van de interventie beperkt is, waarschijnlijk omdat de interventietijd klein is vergeleken met de buitenschoolse tijd.

Deze Europese studies lijken de Amerikaans bevindingen te bevestigen dat voorschoolse educatie de prestaties in het daaropvolgende onderwijs verbetert. De Europese studies bieden weinig informatie over lange termijn effecten zoals participatie op de arbeidsmarkt, crimineel gedrag of gezondheid.

3.4 Wat weten we over de kosten en baten van voorschoolse interventies?

In voorgaande hebben we gezien dat voorschoolse interventies kunnen leiden tot positieve lange termijn effecten voor kinderen. De cruciale vraag is dan hoe de kosten van de verschillende interventies zich verhouden tot de baten. Op dit punt biedt de literatuur slechts beperkte informatie. Bij drie projecten zijn berekeningen gemaakt over de kosten en baten, dit betreft één modelprogramma (Perry Preschool) en twee grootschalige programma's (CPC en Head Start). Echter, alleen bij het Perry Preschool project is een uitvoerige kosten-baten analyse uitgevoerd. Hieronder worden de verschillende kosten-baten berekeningen besproken waarbij gestart wordt met het modelprogramma. Vooraf moet opgemerkt worden dat kosten-baten analyses veelal gebaseerd zijn op diverse aannamen, bijvoorbeeld over de ontwikkeling van het toekomstig inkomen of de kans op werk/ inactiviteit. Dit betekent dat in de berekening diverse keuzen gemaakt worden. De uitkomsten dienen daarom als indicatief gezien te worden en niet als harde feiten.

Kosten en baten van het Perry Preschool project

In tabel 3.1 zijn de kosten en baten van het Perry Preschool project gegeven. De kosten van deelname aan het project bedroegen 12356 dollar per leerling. Baten zijn er zowel voor de deelnemers (kolom 3) als voor anderen in de samenleving (kolom 4). De totale kosten en baten voor de samenleving, dat is de deelnemers plus de anderen staan in kolom 2. De baten zijn er gedurende de onderwijsloopbaan, bijvoorbeeld minder zittenblijven of verwijzing naar speciaal onderwijs, op de arbeidsmarkt (vaker een baan, minder inactiviteit) en in de vorm van minder criminaliteit. De bovenste helft van tabel 3.1 geeft de baten zoals deze zijn gemeten bij de

deelnemers op 27-jarige leeftijd. De onderste helft geeft projecties naar de toekomst, dat wil zeggen de baten die de komende jaren naar verwachting zullen ontstaan.

Op een leeftijd van 27 jaar levert elke dollar besteed aan het Perry Preschool project een opbrengst van 5,7 dollar. Geprojecteerd naar de toekomst mag een opbrengst worden verwacht van 8,7 dollar. Het merendeel van de baten komt uit de reductie van criminaliteit. Dit betekent dat het merendeel van de baten toevalt aan anderen dan de deelnemers aan het project.

In Karoly e.a. (1998) is berekend dat het Perry Preschool project ook voor de overheid meer oplevert dan het kost. De besparingen door minder uitgaven aan onderwijs, meer belasting inkomsten, minder uitkeringen en minder justitiële kosten wegen op tegen de totale kosten van het project. Deze uitkomst is robuust voor verschillende hoogten van de rente (waarmee toekomstige inkomsten herleid worden tot bedragen in het heden).

<u>Perry Pre-school</u>	gehele samenleving	deelnemers	anderen
kosten	-12356	0	-12356
gemeten baten			
kinderopvang	738	738	0
onderwijs t/m 18 jaar	6872	0	0
vervolgonderwijs	-868	0	-868
volwassen onderwijs	283	0	28
werkgelegenheid	14498	10269	4229
criminaliteit	49044	0	49044
welzijn	219	-2193	2412
totale baten	70876	8814	61972
toekomstige baten			
loon	15833	11215	4618
misdad	21337	0	2133
welzijn	46	-460	506
totale opbrengsten	108002	19569	88433
netto contante waarde	95646	19569	76077

bron: Heckman (2000), p. 30.

De conclusie is dat de maatschappelijke baten van het Perry Preschool programma de kosten ruimschoots overtreffen.

Kosten en baten van grootschalige programma's

Voor twee grootschalige programma's zijn berekeningen uitgevoerd naar kosten en baten. Allereerst zijn bij de Chicago Child Parents Centers de kosten vergeleken met de baten uit verminderd crimineel gedrag. Elke dollar besteed aan het programma blijkt een besparing van de toekomstige kosten van de overheid op te leveren van bijna 4 dollar.

Door Currie (2001) is een (illustratieve) kosten-baten analyse uitgevoerd van Head Start. Een beperking van veel evaluaties van Head Start projecten is dat weinig bekend is over de lange termijn effecten. Daarom is bij de kosten-baten analyse alleen gekeken naar de baten die zich op korte en middellange termijn voordoen. De belangrijkste conclusie is dan dat deze effecten al bijna genoeg zijn om de kosten terug te verdienen¹⁷.)

Conclusies en vragen

Het voorgaande laat zien dat voor bepaalde voorschoolse interventies zeer gunstige resultaten zijn geboekt. Deze gunstige resultaten werden met name gevonden in de betere studies, dat wil zeggen studies waaraan strenge methodologische eisen werden gesteld: (1) experimenteel design, (2) lange termijn effecten gemeten, (3) geringe uitval. Opvallend daarbij is dat de lange termijn effecten vaak een ander beeld geven van de opbrengsten van het programma dan de effecten die gemeten werden op de korte of middellange termijn. Het uitdoven van IQ-scores of het verdwijnen van de winst in termen van onderwijsprestaties kunnen soms ten onrechte suggereren dat een programma niet succesvol is. De schijnbare tegenstelling tussen het op korte of middellange termijn uitdoven van cognitieve effecten versus de gunstige lange termijn effecten heeft mogelijk te maken met niet gemeten sociale vaardigheden. Volgens Heckman (2000) worden de lange termijn voordelen wat betreft de positie op de arbeidsmarkt en vermindering van crimineel gedrag met name veroorzaakt door de verbetering van de sociale vaardigheden door voorschoolse programma's. Sociale vaardigheden kunnen ook gunstige externe effecten voor medeleerlingen opleveren. Bij de evaluatie van het Star-programma, een grootschalig experiment met klasseverkleining in Tennessee (zie paragraaf 4.1.1) werd gevonden dat de prestatie van een leerling sterk toeneemt als meer klasgenoten hebben deelgenomen aan voorschoolse educatie (Krueger(1999)). Gesuggereerd wordt dat dit komt door de sterkere socialisatie van de klas waardoor docenten effectiever les kunnen geven.

Kosten-baten analyses van voorschoolse programma zijn schaars. In de weinige kosten-baten analyses die zijn uitgevoerd worden gunstige tot zeer gunstige resultaten gevonden.

¹⁷ zie Currie, J.(2001), table 3, p. 232.

Ondanks het feit dat voor bepaalde projecten zeer gunstige resultaten zijn geboekt moet worden vastgesteld dat dit vooral geldt voor de zogenoemde modelprogramma's. Dit zijn programma's die werden uitgevoerd in de jaren 60 en 70. We weten niet precies wat de opbrengsten van deze programma's zijn in de huidige context. Er zijn aanwijzingen dat ook grootschalige programma's gunstige (lange termijn) resultaten kunnen opleveren. De studies die hiervoor zijn uitgevoerd zijn echter minder 'hard' dan de studies naar de modelprogramma's. Om meer te leren van de huidige ervaringen is in dat verband door de adviescommissie voor de evaluatie van Head Start onlangs aanbevolen om de evaluatiemogelijkheden te verbeteren. In gevallen dat er te weinig Head Start middelen beschikbaar zijn voor alle deelnemers beveelt de commissie aan om de toewijzing van leerlingen aselekt te laten plaatsvinden. Voor de toekomst wordt daarmee een echte evaluatie van het causale effect van de programma's mogelijk gemaakt.

4 Interventies in het schoolse traject

In veel landen wordt beleid gevoerd gericht op het verminderen van leerachterstanden van 'students at risk' gedurende het schoolse traject. Vaak gaat het dan om het beschikbaar stellen van fondsen voor achterstandsl leerlingen ter financiering van specifieke compensatieprogramma's. De Verenigde Staten kent in dit verband het Title I programma waarin staten op basis van plannen subsidie aanvragen bij de federale overheid. In Australië kunnen besturen voor openbaar of privaat onderwijs subsidie verkrijgen door plannen in te dienen voor het verminderen van leerachterstanden. Ook in verschillende Europese landen, zoals Engeland, Wales en Frankrijk, worden compensatieprogramma's uitgevoerd voor 'students at risk'.

Grootschalige programma's of specifieke interventies

Het doel van deze paragraaf is om een overzicht te bieden van de internationale ervaringen met verschillende interventies voor achterstandsl leerlingen in het schoolse traject, dat is de periode van basisonderwijs tot en met de leerplichtige leeftijd. Het ligt voor de hand om dan vooral te kijken naar de hiervoor genoemde (grootschalige) compensatieprogramma's. Voor het onderhavige project is deze aanpak echter minder geschikt. Het doel is namelijk om in kaart te brengen welke interventies succesvol zijn of niet. Tot dusver is over de causale effecten van grootschalige programma's of delen van programma's weinig bekend. Deze programma's kennen in het algemeen geen experimenteel design; er is er geen controlegroep en leerlingen worden niet door toeval toegewezen aan een controlegroep en een experimentele groep. Voorts worden de beschikbare middelen vaak ingezet in een groot aantal kleinere projecten. Het totale effect van het programma is dan het saldo van de effecten van alle deelprojecten, die ook zelden grondig worden geëvalueerd. Ook het aantal studies waarin getracht wordt de effecten van

grootschalige programma's te evalueren door gebruik te maken van natuurlijke experimenten is nog zeer beperkt.¹⁸

Een ander bezwaar tegen een beperking tot compensatieprogramma's is dat achterblijvers ook profijt kunnen hebben van generiek beleid dat beoogt de kwaliteit van het onderwijs voor alle leerlingen te verbeteren.¹⁹ Er zijn voorbeelden dat generiek beleid in sterkere mate succesvol is bij achterstandsleerlingen. Compensatieprogramma's komen bovendien in het algemeen neer op het beschikbaar stellen van additionele fondsen. Daarmee blijven veranderingen in de organisatie van het onderwijs buiten beeld.

Hieronder wordt daarom een andere aanpak gekozen waarin wordt gekeken naar het effect van specifieke interventies, hetzij gericht op alle leerlingen, hetzij gericht op achterstandsleerlingen. Op hoofdlijnen worden drie soorten interventies onderscheiden:

1. extra middelen voor bepaalde inputs;
2. veranderingen in de organisatie van het onderwijs;
3. specifieke projecten voor achterblijvers.

De eerste groep betreft de 'klassieke input's zoals groeps grootte, lessen, opleiding en ervaring docenten, beloning van docenten en faciliteiten (ICT) (zie tabel 2.1). De tweede groep betreft elementen die te maken hebben met de organisatie van het onderwijs, zoals concurrentie tussen scholen, leerstandaarden, instituties rond het leren zoals de leerplicht. De derde groep interventies betreft specifieke projecten gericht op 'students at risk'.

De beschrijving van de stand van kennis is gebaseerd op studies waarin een experimentele dan wel een quasi-experimentele opzet is gekozen. Dit betekent dat de aandacht voor grootschalige interventieprogramma's beperkt is (omdat de causale effecten niet bekend zijn). Per interventie wordt de stand van de kennis beschreven en wordt ingegaan op het generieke effect (het gemiddelde effect op alle leerlingen) en zo mogelijk op het effect op achterstandsleerlingen.

4.1 De 'klassieke' onderwijsinputs

Allereerst wordt een beschrijving gegeven van de bevindingen rond de klassieke inputs van het onderwijsproces (zie bijvoorbeeld Hanushek(1986)). Tot deze groep van interventies rekenen we ook 'de peer group van de leerlingen' omdat de medeleerling als een input van het

¹⁸ Recent is het effect van de zogenoemde 'equalization law' in Massachusetts in de Verenigde Staten geanalyseerd. Deze wet verschaft additionele middelen voor publieke scholen in districten met weinig onderwijsmiddelen. In de analyse is gebruik gemaakt van een discontinuïteit in de regeling voor de toewijzing van middelen (Guryan, mei 2001). Ook wordt gewerkt aan de evaluatie van het Title I project (Van der Klauw, 2001).

¹⁹ Overigens kan het ook voorkomen dat reguliere leerlingen profiteren van middelen bestemd voor achterblijvers.

onderwijsproces gezien kan worden. Extra middelen voor 'de peergroep' zou dan kunnen neerkomen op het implementeren van spreidingsbeleid.

4.1.1 Verkleining van groepsgrootte

Het effect van het verkleinen van de groepsgrootte op de individuele leerprestaties is een controversieel onderwerp. Het traditionele beeld is dat de effecten ongewis zijn en kunnen afhangen van de range waarop de verkleining betrekking heeft. Hanushek (1986) geeft een overzicht van de studies naar de effecten van verkleining van de groepsgrootte en concludeert dat het aantal schattingen waarin het verwachte negatieve effect gevonden wordt ongeveer even groot is als het aantal schattingen waarin het tegendeel wordt gevonden (zie tabel 2.1). Met andere woorden, de uitkomst is ongewis. Glass en Smith (1979) concludeerden dat er geen effecten voorkomen in de middelrange van de groepsgrootte, ongeveer tussen 20 en 30 leerlingen, en dat er wel effecten gevonden worden beneden deze middelrange. De studies waarop dit traditionele beeld is gebaseerd worden gehinderd door twee problemen:

- de groepsgrootte wordt vaak niet goed gemeten;
- er wordt geen rekening gehouden met het feit dat de groepsgrootte niet door toeval tot stand komt.

Het eerste probleem wordt veroorzaakt doordat onderzoekers vaak alleen beschikken over de leerling-docent ratio voor de hele school. Dit kan betekenen dat de analyse vertekend wordt door meetfouten. Het tweede probleem is de endogeniteit van de groepsgrootte, dat wil zeggen de mogelijkheid dat niet-waargenomen factoren gecorreleerd zijn met de groepsgrootte en tevens van invloed zijn op de leerprestaties (zie paragraaf 2.2).

De laatste jaren zijn verschillende studies uitgevoerd waarin wel rekening wordt gehouden met het 'endogene' karakter van de groepsgrootte. De meest directe manier om hiermee rekening te houden is het uitvoeren van een experiment waarin leerlingen door toeval worden toegewezen aan een experimentele en een controlegroep. Tot dusver is er één grootschalig experiment uitgevoerd, te weten het welbekende Tennessee Star-project waaraan werd deelgenomen door 80 scholen en 11600 leerlingen. Leerlingen en docenten werden door toeval toegewezen aan drie alternatieven: klassen met een omvang van 13-15 leerlingen, klassen van 22-25 leerlingen en klassen met 22-25 leerlingen met een klassenassistent. Het experiment startte in het schooljaar '85-'86 en duurde vier jaar; vanaf 'kindergarten' tot en met het derde jaar basisonderwijs. Inmiddels zijn gegevens beschikbaar over de effecten direct na afloop van het experiment en vijf jaar na afloop van het experiment. Deze gegevens zijn zorgvuldig geanalyseerd door Alan Krueger en anderen. Direct na afloop van het experiment wordt gevonden dat leerlingen die

geplaatst zijn in kleinere klassen beter presteren²⁰. Na één jaar experiment scoren kinderen in kleine klassen gemiddeld 3,5 procentpunt beter op gestandaardiseerde toetsen, ieder extra jaar in een kleine klas levert vervolgens een winst van 0,65 procentpunt. De winst voor leerlingen uit achterstandsgroepen is groter: na het initiële jaar 4,8 procentpunt en ieder extra jaar levert één procentpunt. Uitgedrukt in termen van het verschil tussen 'witte' Amerikanen en Afro-Amerikanen komt dit na het initiële jaar neer op 64 % van het verschil in testscore en na vier jaar op 82 %. Vijf jaar na afloop van het experiment wordt gevonden dat de leerlingen uit kleinere klassen hoger scoren op gestandaardiseerde toetsen en vaker deelnemen aan toelatingsexamens voor hoger vervolgonderwijs (vier procentpunt meer)²¹. Het effect is het grootst voor leerlingen uit achterstandsgroepen (acht procentpunt meer). Het verschil tussen 'witte' en Afro-Amerikanen in het deelnemen aan het examen voor hoger vervolgonderwijs neemt af met 54 procent. Twee uitkomsten van het experiment verdienen extra aandacht. Allereerst valt op dat het grootste effect optreedt na het initiële jaar. In de tweede plaats wordt een groot effect gevonden van de onderrichtservaring van de klas, gemeten als het aandeel kinderen dat 'kindergarten' bezocht (een verandering in het aandeel leerlingen dat 'kindergarten' bezocht met 2 standaarddeviaties geeft 3 procentpunt hogere testcores). Bij beide effecten kan, volgens Krueger, sprake zijn van een socialisatie-effect.

Een andere wijze om rekening te houden met het 'endogene' karakter van de groepsgrootte bestaat uit een quasi-experimentele benadering waarin gebruik gemaakt wordt van toevallige natuurlijke experimenten. Het kan namelijk voorkomen dat de groepsgrootte min of meer toevallig tot stand komt. Het meest bekende, en fraaiste, voorbeeld hiervan is de regel voor de groepsgrootte in Israël (zie box 2.1). Deze regel leidt tot toevallige variatie in de groepsgrootte (discontinuïteiten in de ratio leerling-docent) die door Angrist & Lavy (1999) wordt gebruikt als natuurlijk experiment. Zij vinden het verwachte negatieve effect van groepsgrootte: leerlingen in kleinere klassen presteren beter. Daarbij wordt een sterker effect gevonden bij achterstandsléerlingen. Deze schattingsmethode wordt 'instrumentele variabelen' genoemd. De regel voor de verhouding leerling-docent wordt als instrument gebruikt voor de groepsgrootte. Een vergelijkbare benadering wordt in nog enkele andere studies toegepast. In tabel 4.1 worden de resultaten van de verschillende studies vergeleken. De resultaten van de studies zijn daarbij uitgedrukt in één en dezelfde mate, namelijk de standaardafwijking van het gemiddelde van de scores op gestandaardiseerde toetsen. De effecten geven daarmee als ware aan hoe de verdeling van de scores van alle leerlingen verschuift als gevolg van de verkleining van de groepen.

²⁰ Krueger, A.B. (1999).

²¹ Krueger, A.B. & D.M. Whitmore (2001).

Tabel 4.1 Het causale effect van verkleining van de groeps grootte op de leerprestaties volgens (quasi-)experimentele studies (σ standaardafwijking van gemiddelde van verdeling, n.s. niet significant)

studie	Instrument	Effect		niveau
		gemiddeld	achterblijvers	
Krueger	STAR-experiment	.13 σ	.22 σ	PO
Krueger&Whitmore	STAR-experiment	4 %	8 %	PO
Angrist&Lavy	max groeps grootte	.10-.20 σ	> .20 σ	PO
Akerhielm	gemid. groeps grootte	.15 σ	mixed	grade 8
Hoxby	variatie in natuurlijke populatie (cohorten) + max groeps grootte	n.s.	n.s.	grade 4, 6, 8
Dobbelsteen e.a.	bekostigingsregel	n.s.	sterker effect voor achterblijvers*	PO
Boozer&Rouse	max. groeps grootte speciaal onderwijs	.29 σ	?	grade 8

* bij controle aantal leerlingen met zelfde IQ

In de studies waarin het verwachte positieve effect van een verkleining van de groeps grootte wordt gevonden is de omvang van de effecten vergelijkbaar. De effecten voor achterblijvers zijn daarbij steeds het grootst. In twee studies worden geen significante effecten gevonden. Hoxby gebruikt variatie in de omvang van de natuurlijke populatie in districten als instrument voor de groeps grootte. In kleine districten zal een plotselinge stijging van het aantal jongeren leiden tot grotere groepen. Als tweede instrument gebruikt zij een regel voor het maximum aan leerlingen in een groep, vergelijkbaar met het instrument van Angrist&Lavy. In deze analyse wordt bovendien bekeken in hoeverre de groeps grootte ligt rond de discontinuïteit veroorzaakt door de regel. De gedachte is dat met name rond deze discontinuïteit (bij Angrist en Lavy rond 40 leerlingen) het effect van de groeps grootte niet vertekend wordt door niet waargenomen factoren. Met beide instrumenten worden geen effecten gevonden van de groeps grootte nog voor alle leerlingen noch in districten met veel achterblijvers. Een verschil met de andere studies is dat hier geaggregeerde gegevens worden gebruikt, namelijk het gemiddelde per district. Dobbelsteen e.a. maken gebruik van toevallige variatie veroorzaakt door de systematiek van de bekostiging in het Nederlands basisonderwijs. Ook in deze studie wordt geen effect gevonden. Dit kan te maken hebben met het gekozen instrument. De omvang van de discontinuïteiten zijn namelijk aanmerkelijk kleiner dan die bij Angrist&Lavy hetgeen de toevallige variatie beperkt. Natuurlijk kan ook de Nederlandse context een verklaring zijn voor het niet vinden van het effect. Het is mogelijk dat de Amerikaanse resultaten vooral veroorzaakt worden door socialisatie-effecten en dat deze effecten in de Nederlandse context al zijn ingeboekt door de eerdere start van het basisonderwijs (op vierjarige leeftijd). Bovendien worden in Nederland de middelen voor het achterstandenbeleid voor een belangrijk deel ingezet voor het verkleinen van klassen (zie ook hoofdstuk 5).

Bovenstaande studies laten zien dat de controversie nog niet is uitgewoed. Wel laten de resultaten van het STAR-project onomstotelijk zien dat het verkleinen van de groepen *kan* leiden tot betere resultaten van leerlingen. En, dat dit in sterkere mate geldt voor leerlingen uit achterstandsgroepen. Bij het STAR-project was sprake van een beleidsexperiment. De deelnemers waren zich bewust van het feit dat ze werden geëvalueerd en dat er voordelen te behalen waren bij het gebruik maken van de mogelijkheden die geboden werden door de kleinere groepen²². Hoxby, die gebruik maakt van natuurlijke experimenten, meent dat daarin het verschil in resultaten is gelegen. Bij natuurlijke experimenten zijn deelnemers zich niet bewust van een evaluatie en zijn er bovendien geen voordelen te behalen van betere prestaties. Zij concludeert derhalve dat verkleining van groepen alleen effectief is als de extra middelen gerelateerd worden aan prestaties ('tie resources to performance').

4.1.2 Uitbreiding van het aantal lesuren

Een interventie die sterk lijkt op die van de verkleining van groeps grootte is de uitbreiding van het aantal lesuren. In beide gevallen kan de docent namelijk meer tijd besteden aan de leerling. De meer traditionele literatuur geeft geen heldere uitkomsten over het effect van een vergroting van het aantal lesuren op de prestaties van leerlingen²³. Bijvoorbeeld, Link & Mulligan (1986) vinden geen effect van instructietijd op taal- en wiskundescores. Loeb & Bound (1995) vinden positieve effecten van schooluren gemeten als het aantal schooldagen per jaar. In deze en ook in andere studies wordt geen rekening gehouden met mogelijke versturende factoren die het aantal lesuren kunnen beïnvloeden. Binnen scholen kan er namelijk voor gekozen worden om zwakkere leerlingen meer les te geven. Meer gemotiveerde ouders kunnen een voorkeur hebben voor scholen waarin meer les gegeven wordt.

In een recente studie wordt getracht het causale effect van het aantal schooluren op de leerlingprestaties te bepalen (Lavy, 1999). Als instrument gebruikt hij een bekostigingsregel voor de toekenning van uren aan scholen. Met het instrument wordt de toevallige variatie in lesuren tussen scholen als gevolg van de (discontinuïteiten in de) bekostigingsregel gebruikt als natuurlijk experiment. De analyse laat zien dat lesuren een significant positief effect hebben op de prestaties van leerlingen. Een tentatieve berekening van de kosten van vergroting van de lesuren laat zien dat investeringen in lesuren hogere opbrengsten genereren dan investeringen in het verkleinen van de groeps grootte. Als wordt uitgegaan van hetzelfde effect op de prestaties

²² Volgens Krueger worden zijn bevindingen nauwelijks vertekent door deze zogenoemde Hawthorne-effecten. Hij vindt namelijk ook voor de klassen die te samen de controlegroep vormen dat kinderen in kleinere klassen beter presteren.

²³ In het overzicht van Hanushek komen geen studies voor waarin het effect van het aantal lesuren op leerlingprestaties wordt geanalyseerd.

van leerlingen is voor de uitbreiding van het aantal lessen slechts éénderde van de middelen nodig vergeleken met de besteding bij het verkleinen van de groepsgrootte.

In hoeverre de tijd die wordt doorgebracht op school van invloed is op de prestaties van leerlingen is recent onderzocht.²⁴ Om het effect van de schooltijd op de prestaties van leerlingen te kunnen vaststellen zou men eigenlijk moeten weten hoeveel leerlingen zouden leren als ze niet naar school zouden gaan. Deze ‘counterfactual’ wordt in deze studie ontleend aan de zomervakantie. In het onderzoek worden namelijk testcores van leerlingen geanalyseerd die zijn gemeten aan het eind (tweemaal) en aan het begin van het schooljaar. Daarmee kan vastgesteld worden hoe de testcores veranderen gedurende het schooljaar en daarbuiten. De zomervakantie wordt derhalve gebruikt als natuurlijk experiment.

De studie laat zien dat de tijd die wordt doorgebracht op school meer bijdraagt aan de leerprestaties dan de tijd die wordt doorgebracht buiten school. Dit betekent dat een uitbreiding van de schooltijd tot een verbetering van de leerprestaties leidt.

4.1.3 **Docenten**

In het algemeen wordt aan de docent een sleutelrol toegekend ten aanzien van de kwaliteit van het onderwijs. De resultaten uit Hanusheks klassieke studie zijn in dat opzicht dan ook verbazingwekkend. In zijn overzicht van de literatuur wordt namelijk gevonden dat de ervaring en opleiding van de docent geen invloed hebben op leerlingprestaties (zie tabel 2.1). Maar ook hier kunnen endogeniteits-problemen wederom een belangrijke rol spelen. Als bijvoorbeeld onderzocht wordt wat het effect van scholing van docenten is op de prestaties van leerlingen dient er rekening mee gehouden te worden dat docenten die additionele scholing volgen kunnen verschillen van docenten die geen scholing volgen. Docenten met meer ervaring zijn mogelijk ook de meer gemotiveerde docenten of ook de betere docenten (zij blijven werken in het onderwijs). Een ander element van endogeniteit is de voorkeur van leraren om op bepaalde scholen te werken²⁵.

In de literatuur worden vier kenmerken van docenten onderscheiden: de initiële opleiding, scholing, ervaring en beloning van docenten. Tot dusver is het aantal studies waarin expliciet rekening wordt gehouden met de endogeniteit van kenmerken van leraren schaars. Geen enkele studie kent een experimenteel of quasi-experimenteel design. Wel worden andere technieken gebruikt om te controleren voor niet waargenomen effecten, zoals panel-analyse en matching. Een overzicht van de studies is gegeven in tabel 4.2 (zie volgende bladzijde).

²⁴ Lindahl, M. (2000)

²⁵ Hanushek c.s. (1999) laten bijvoorbeeld zien dat leraren (behalve Afro-amerikaanse leraren) een voorkeur hebben om les te geven aan de beter presterende leerlingen niet afkomstig uit etnische minderheden.

Scholing van docenten

In Israël is een programma uitgevoerd, waarbij zittende docenten in het basisonderwijs getraind werden. De training beoogde de instructie-technieken van docenten in taalvaardigheden en rekenen/wiskunde te verbeteren. Een beperkt aantal scholen kreeg geormerkte gelden die waren bestemd voor participatie aan het trainingsprogramma. Angrist en Lavy (1999) hebben via 'matching' leerlingen van 9 scholen uit de interventiegroep vergeleken met leerlingen van 11 scholen uit een controlegroep van scholen met vergelijkbare kenmerken. Deze kinderen worden gevolgd vanaf het vierde leerjaar (in 1994) tot en met het zesde leerjaar (1996). Dit maakt het mogelijk om de ontwikkeling in de leerprestaties voor en na het volgen van de training van docenten te analyseren²⁶. Bij niet-religieuze scholen verbeteren de toetsresultaten. Volgens een voorzichtige schatting is de omvang van dit effect gelijk aan 0,25 standaardafwijking van het gemiddelde van toetsresultaten. Voor religieuze scholen wordt geen effect gevonden, waarschijnlijk omdat het programma bij deze scholen op een kleinere schaal en over een kortere periode werd uitgevoerd. In deze studie is ook de kosteneffectiviteit van het trainingsprogramma vergeleken met die van klassenverkleining. Het resultaat dat wordt bereikt met de klassenverkleining kan ook bereikt worden met de training van docenten. Voor de laatste interventie is echter slechts éénderde van de middelen nodig die eerder gebruikt werden voor de klassenverkleining.

²⁶ Dit heet 'difference in difference' regressie.

Tabel 4.2 Studies naar de effecten van kenmerken van docenten op prestaties van leerlingen

Studie	Doelgroep	Effect*	Methodologie
<i>Scholing</i>			
Angrist en Lavy (1999)	leerjaar 4 en 6, niet religieus, Israël	.25 σ	Matching van scholen
	Religieuze scholen, Israël	Geen effect	Matching van scholen
<i>Initiële opleiding</i>			
Goldhaber en Brewer (1997)	leerjaar 10, VS	Positief effect	panelanalyse
Hanushek e.a. (1998)	Texas leerjaar 4 en 6	Geen invloed	panelanalyse
<i>Ervaring</i>			
Hanushek e.a. (1998)	Texas leerjaar 4 en 6	Toename .10 σ in eerste 2 jaar, daarna geen effect	panelanalyse
Goldhaber en Brewer 1997	leerjaar 10, VS	.07 σ	panelanalyse
<i>Beloning</i>			
Loeb en Page (2000)	50 Staten, high school, VS	Salaris +10% -> dropout -4 %	panelanalyse + instrumentele variabelen
Hanushek (1999)	leerjaar 3-6, VS	Salaris + 10% --> wiskunde + .17 σ lezen + .11 σ	panelanalyse + instrumentele variabelen

* σ standaardafwijking van gemiddelde van verdeling.

De initiële opleiding en ervaring van docenten

In twee recente studies wordt het effect van het opleidingsniveau en de ervaring van de docent op de prestaties van leerlingen geanalyseerd. Goldhaber en Brewer (1997) gebruiken paneldata om te controleren voor niet-waargenomen kenmerken. Met deze techniek kan alleen gecontroleerd worden voor niet-waargenomen kenmerken die constant zijn in de tijd. Gevonden wordt dat het diploma van de docent bijdraagt aan de prestaties van de leerling (2,3 %-effect van een wiskunde-certificaat op prestaties in wiskunde, en 0,82 % effect van BA-degree op wiskundeprestaties). Ook vinden zij een positief effect van de ervaring van de docent op de prestatie. Hanushek e.a. (1998) vinden daarentegen, ook op basis van paneldata, geen effect van de initiële opleiding van docent op de prestaties van leerlingen. Voor de ervaring van de leraren vinden ze alleen een effect aan het begin van de loopbaan. De prestaties van leerlingen in lezen en wiskunde nemen toe gedurende het eerste en tweede jaar dat docenten lesgeven. Voor leraren met een langere anciënniteit is geen toename op onderwijsprestaties waarneembaar. Er zijn geen aanwijzingen dat de inzet van ervaren docenten vooral effectief is voor leerlingen met een lagere sociaal-economische achtergrond.

Het salaris van docenten

Hogere salarissen kunnen de prestaties van leerlingen direct verbeteren, bijvoorbeeld door een hogere inzet van docenten, maar ook indirect via een grotere aantrekkelijkheid van het lerarenberoep voor nieuwe instromers. De directe invloed van het hogere salaris op de leerprestaties is recent onderzocht door Loeb en Page (2000). De analyse is uitgevoerd voor staten in de VS (cohorten '59, '69, '79 en '89). Door gebruik te maken van panelgegevens wordt rekening gehouden met verschillen in niet-financiële opbrengsten en verschillen in lokale arbeidsmarktomsomstandigheden. Gevonden wordt dat een toename van de lerarensalarissen van 10 procent de dropout rate met 3 tot 4 procentpunten vermindert. Het effect van de toename van de lerarensalarissen wordt daarbij niet onmiddellijk gemeten maar 10 jaar later. Opgemerkt wordt dat het hier gaat om een generieke maatregel en dat specifieke salarisverhogingen naar verwachting aanmerkelijk kosten-effectiever zullen zijn.

Hanushek c.s. (1999) analyseren een bestand van alle leerlingen van de publieke scholen in Texas. Gebruik makend van paneldata wordt gevonden dat bij leerlingen die niet van school veranderen er een positief effect is van het salaris op de prestaties in wiskunde en lezen. (10 % meer startsalaries geeft voor wiskunde en taal een winst van respectievelijk $0,17 \sigma$ en $0,11 \sigma$). Niet duidelijk is in hoeverre dit positieve effect daadwerkelijk veroorzaakt wordt door de salaris toename. Het effect wordt namelijk alleen gevonden bij scholen die geen in- of uitstroom van docenten hebben gehad. Dit kan betekenen dat andere kenmerken van deze scholen, die gecorreleerd zijn met de prestaties van leerlingen en met de salarissen, hierbij een rol spelen. Het indirecte effect van de salarisverhoging, namelijk via de aantrekkelijkheid van het lerarenberoep, is geanalyseerd door Figlio (1997). Daarbij wordt een positieve relatie gevonden tussen lerarensalarissen en het opleidingsniveau van nieuwe leraren. Een toename van het gemiddelde loon in het schooldistrict met één procent leidt tot een 1,6 procentpunt hogere kans op het werven van een docent met een opleiding "undergraduate institution". Bovendien leidt één procentpunt loonsverhoging tot een toename van de kans om een docent te werven die een opleiding aan een "college" heeft gevolgd met 0,75 procentpunt.

4.1.4 Effecten van ICT

Al enkele decennia worden computers en ICT ingezet in het onderwijs. De laatste jaren is ICT in veel landen een speerpunt van het onderwijsbeleid geworden. Met de inzet van ICT zijn hoge uitgaven gemoeid. Onderzoek naar de effecten van ICT op leerprestaties is beperkt en geeft veelal geen eenduidige resultaten²⁷. In veel studies is geen sprake van een echte controlegroep. Recent is een studie beschikbaar gekomen die voldoet aan strengere methodologische eisen.

²⁷ Een bekende overzichtsstudie op dit terrein is Kulik&Kulik (1991)

Door Angrist&Lavy (1999) zijn de effecten van ICT op leerprestaties geanalyseerd in het Israëliësch primair en secundair onderwijs. De Israelische staatloterij maakte in 1994 een grootschalige investering in ICT in het onderwijs mogelijk. In juni 1996 konden daardoor 10 procent van de basisscholen en 45 procent van de scholen in het voortgezet onderwijs beschikken over nieuwe computers. Het feit dat niet alle scholen tegelijkertijd voorzien kon worden van ICT-faciliteiten wordt als natuurlijk experiment gebruikt om het effect van ICT op de leerprestaties vast te stellen. De studie laat zien dat de toename van computers in 1994 en 1995 leidde tot een substantiele toename van het gebruik van 'computer aided instruction' (CAI). Het toegenomen gebruik van CAI leidde evenwel niet tot een verbetering van de prestaties van leerlingen. Er werd zelfs een daling van de prestaties in wiskunde gevonden. De uitgaven die gemoeid waren met de inzet van ICT per programmaschool zijn vergelijkbaar met het salaris van vier voltijd docenten. Uiteraard is het de vraag in hoeverre dit resultaat ook zal gelden voor toekomstige ICT-toepassingen. Vooralsnog zijn de resultaten uit de hier besproken studie teleurstellend.

4.1.5 Peergroup-effecten op school (spreidingsbeleid)

Niet alleen de omvang van de groep kan van invloed zijn op de prestaties van leerlingen maar ook de samenstelling van de groep. Leerlingen in een groep kunnen elkaars leerprestaties op tal van manieren beïnvloeden. Dit kan door van elkaar te leren maar ook door het al of niet verstoren van de lessen, door het beïnvloeden van de sfeer in de groep, door de invloed op relatie tussen docent en groep of door middelen die de leerling meebrengt van thuis. Dit geheel van mechanismen noemen we peergroup-effecten²⁸.

In het klassieke Coleman-report en vele latere studies worden sterke effecten gevonden van de peergroup op de prestaties van individuele leerlingen. Deze bevindingen zijn in sommige staten van de Verenigde Staten vertaald in pogingen om de segregatie van leerlingen te doorbreken via spreidingsbeleid. Het grote voordeel van spreidingsbeleid als vorm van achterstandsbeleid is de relatief lage kosten, te weten de kosten die verband houden met het vervoer van leerlingen.²⁹ Cruciaal is derhalve hoe hard de traditionele bevindingen zijn bij meer geavanceerde analysemethoden.

²⁸ Deze effecten beperken zich niet tot de school maar kunnen uiteraard ook daarbuiten voorkomen. Hieronder gaat het om de peergroupeffecten op school.

²⁹ In de Amerikaanse context kan spreidingsbeleid tot meer concurrentie tussen instellingen leiden hetgeen het vaststellen van het 'echte' peergroup-effect lastiger maakt.

Op dezelfde wijze als bij de groepsgrootte spelen bij de analyse van peergroupeffecten endogeniteitsproblemen een rol. Ouders kunnen zich bij de schoolkeuze laten leiden door de samenstelling van de schoolpopulatie of door andere kenmerken van een bepaalde school. Binnen scholen kunnen leerlingen worden toegewezen aan klassen op grond van hun geschiktheid, geslacht, etniciteit of andere factoren. Door Evans, Oates & Schwab (1992) wordt het endogeniteitsprobleem bij de peergroup geïllustreerd in een analyse van tienerzwangerschappen en voortijdige schooluitval. Als geen rekening wordt gehouden met de endogeniteit van de peergroup wordt gevonden dat een peergroup met een groter aandeel achterstandsleerlingen leidt tot meer tienerzwangerschappen en meer uitval. Als rekening wordt gehouden met de endogeniteit, door een simultaan model te schatten, verdwijnt het effect van de peergroup. Het eerder aan de peergroup toegeschreven effect blijkt dan het resultaat te zijn van achtergrondkenmerken van de tiener zelf, zoals verschillen in gezinsstructuur, ouderlijk inkomen en opleiding.

De laatste jaren is een aantal studies uitgevoerd waarbij getracht wordt dit endogeniteitsprobleem op te lossen door een experimentele of quasi-experimentele benadering te kiezen. Een voorbeeld van een echt experiment op dit terrein is Sacerdote (2000) die gebruik maakt van de aselecte toewijzing van eerstejaars studenten aan kamergenoten in het Dartmouth College. Gevonden wordt dat het studieresultaat positief, maar in bescheiden mate, wordt beïnvloed door de resultaten van de kamergenoot en dat de keuze van de specialisatie nauwelijks wordt beïnvloed. Voor het onderhavige project is deze studie echter minder relevant omdat het hier niet gaat om achterblijvers maar om studenten in het hoger onderwijs.

Positieve externaliteiten van de peergroup versus de efficiency van de kennisoverdracht

Lavy (1999) analyseert het effect van de heterogeniteit van de schoolpopulatie op de resultaten in taal en wiskunde in het basisonderwijs in Israël. In zijn analyse onderscheid hij daarbij twee effecten van een grotere heterogeniteit van leerlingen. Allereerst zullen leerlingen bij een grotere heterogeniteit meer van elkaar kunnen leren: sterkere peergroup-effecten. Met name achterstandsleerlingen zullen daarvan kunnen profiteren. Tegenover deze positieve effecten staat dat de grotere heterogeniteit van de leerlingen in de klas het voor een docent lastiger maakt om het onderwijs af te stemmen op het niveau van de leerling. Meer heterogeniteit vermindert de efficiency van de kennisoverdracht. In hoeverre spreidingsbeleid gunstig is voor achterstandsleerlingen hangt derhalve af van het saldo van deze twee effecten: het externaliteitseffect versus het efficiency effect. In de studie van Lavy is de heterogeniteit in de klas gemeten als variatie in de sociaal economische status van de ouders (SES). Als instrument voor de peergroup kiest hij de heterogeniteit van de schoolpopulatie van 7 jaar daarvoor. De

gedachte daarbij is dat recente beslissingen van ouders en scholen, die de analyse kunnen verstoren, hiermee worden geëlimineerd. De analyses wijzen uit dat een grotere heterogeniteit van de klas leidt tot slechtere leerresultaten in taal en wiskunde. Het verminderen van de efficiency van het leerproces is kennelijk sterker dan de winst door peergroupeffecten. De gevonden effecten zijn relatief groot. Om een vergelijkbaar effect te bereiken als bij een reductie van de heterogeniteit in een klas met 10 procent is een reductie van de groepsgrootte met 15 procent nodig of een uitbreiding van het aantal lessen per leerling met 6 procent. Een studie waarin dezelfde elementen (externaliteiten versus efficiency) naar voren komen is Figlio & Page (2000). Zij analyseren het effect van het indelen van leerlingen naar groepen met een vergelijkbare geschiktheid (tracking). In hun analyse worden verschillende manieren gebruikt om rekening te houden met de selectiviteitseffecten³⁰. Het resultaat van de analyse is dat 'tracking' niet schadelijk is voor de zwakkere leerlingen. In een analyse waarin de schoolkeuze gebruikt wordt als instrument wordt zelfs gevonden dat 'tracking' vooral de prestaties van zwakkere leerlingen verbetert.

Deze resultaten staan lijnrecht tegenover de traditionele studies waarin geen rekening werd gehouden met de endogeniteitsproblemen. In het algemeen werd hier gevonden dat 'tracking' gunstig is voor de goede leerlingen en ongunstig is voor de zwakkere leerlingen (zie bijvoorbeeld Summers & Wolfe, 1977).

Recent is ook door Hoxby (2000) een studie uitgevoerd naar de invloed van de peergroup op de prestaties van leerlingen. Haar benadering is vergelijkbaar met die in de studie naar het effect van de groepsgrootte. Ook hier wordt gebruik gemaakt van toevallige variatie tussen cohorten uit opeenvolgende jaren. In dit geval gaat het niet om de omvang van de cohorten maar om variatie in de samenstelling naar geslacht en etniciteit. Dit wordt gebruikt als instrument voor de endogeniteit van de peergroup. De gedachte is dat de toevallige variatie in de samenstelling van cohorten onafhankelijk is van beslissingen van ouders en scholen. In de analyse is gekeken naar de samenstelling van een cohort naar geslacht en etniciteit. Gevonden wordt dat de prestaties van de peergroep een positief effect hebben op individuele prestaties. Een (exogene) toename van de prestaties van de peergroup met 1 punt op de leesscore leidt tot een toename van de individuele score tussen 0.15 en 0.40 punten. De peergroup effecten zijn sterker binnen etniciteitsgroepen dan tussen etniciteitsgroepen. In groepen waarin het aandeel meisjes groter is worden betere prestaties gehaald in wiskunde³¹.

³⁰ Onder andere door te instrumenteren met variabelen die concurrentieverschillen tussen staten meten. De veronderstelling is dat de behoefte aan tracking groter zal zijn naarmate er meer concurrentie is in een staat.

³¹ Andere recente studies zijn Rivkin (2000) en Feinstein & Symons (1999). De resultaten zijn niet eenduidig.

Stand van zaken

De huidige kennis over peergropeffecten geeft geen eenduidig beeld. In hoeverre achterblijvers kunnen profiteren van spreidingsbeleid is dan ook niet duidelijk. Wel lijkt er steeds meer bewijs te ontstaan dat de traditionele sterke peergropeffecten in werkelijkheid vaak verband houden met de sociaal-economische achtergrond van de leerling. De peergroup hangt namelijk samen met beslissingen van ouders en scholen. Een belangrijke kanttekening bij het bovenstaande onderzoek is dat steeds gekeken is naar leerprestaties of prestaties op de arbeidsmarkt. Zaken als sociale integratie of ontwikkeling van normen en waarden in relatie tot de peergroup zijn niet onderzocht.

4.2 Veranderingen in de organisatie van het onderwijs

Als tweede groep interventies wordt gekeken naar het effect van veranderingen in de organisatie van het onderwijs. Dit kan neerkomen op het vergroten van de schoolkeuzemogelijkheden, het versterken van de concurrentie tussen scholen, het introduceren van standaarden voor prestaties van leerlingen of aanpassingen van algemene instituties zoals de leerplicht.

4.2.1 Het vergroten van de schoolkeuze

Vouchers voor achterstandsleerlingen

In de Verenigde Staten zijn de mogelijkheden om een school te kiezen in het primair en voortgezet onderwijs beperkt. Het schooldistrict waarbinnen men woont bepaalt het aanbod van publieke scholen. Deze scholen zijn in het algemeen kwalitatief nauwelijks van elkaar te onderscheiden. Private scholen vormen het alternatief, voor achterstandsleerlingen zijn deze evenwel vaak niet toegankelijk vanwege de hoge kosten. In de jaren negentig zijn enkele experimenten uitgevoerd waarbij achterstandsleerlingen 'vouchers' kregen die hen de mogelijkheid boden om onderwijs te volgen in private scholen.

Allereerst werd in 1995 het 'Milwaukee Parental School Program' uitgevoerd. Met een voucher van ongeveer 3200 \$ konden achterstandsleerlingen onderwijs volgen in 12 private scholen. De leerlingenpopulatie op deze private scholen mocht maximaal voor 49 procent bestaan uit 'voucherleerlingen'. Scholen mochten niet selecteren onder de 'voucherleerlingen'. In een evaluatie van het project is de procedure bij de selectie gebruikt als natuurlijk experiment. De 'voucherleerlingen' die niet konden worden geplaatst op de private scholen vormden de controlegroep die werd vergeleken met de leerlingen die wel konden worden geplaatst³².

³² Rouse, C.E. (2000)

Gevonden werd dat het volgen van onderwijs op private scholen voor achterstandsleerlingen leidde tot een verbetering van de wiskundeprestaties met 0,05 - 0,12 standaardafwijking van het gemiddelde van de toetsresultaten. De effecten van de leesprestaties verbeterden niet.

De meest overtuigende resultaten tot dusver komen uit drie recente experimenten met vouchers in New York, Dayton Ohio en Washington.³³ In alle drie experimenten bestond de doelgroep uit leerlingen uit lagere inkomensgroepen op publieke scholen. Bij twee experimenten ging het vooral om leerlingen in het basisonderwijs, aan het derde experiment konden ook leerlingen uit het voorgezet onderwijs deelnemen. De vouchers die deelname aan privaat onderwijs mogelijk maakten werden door loting toegewezen. Voorafgaand aan het experiment werden de leerlingen getest. De tests werden herhaald 1 jaar en 2 jaar na de start van het experiment. Bij alle drie experimenten werd gevonden dat leerlingen met een Afro Amerikaanse achtergrond profiteerden van de switch van een publieke naar een private school. Gemiddeld verbeterden hun testcores na 1 jaar met 0.17 standaardafwijking van het gemiddelde en na 2 jaar met 0.33 standaardafwijking van het gemiddelde. Na 2 jaar is ongeveer een derde van het verschil in prestaties tussen blanke en zwarte Amerikanen teniet gedaan. Als deze trend zou doorzetten in de daaropvolgende jaren zou het gehele verschil tussen blanke en zwarte Amerikanen in korte tijd kunnen verdwijnen.

Ook vergeleken met andere interventies is sprake van een relatief grote verbetering van de prestaties. Zo leidde de reductie van de groepsgrootte met 7 leerlingen in het STAR-project na 4 jaar tot een verbetering van de testcores van Afro Amerikanen met 0.22 standaardafwijking.³⁴

Meer concurrentie

Door het vergroten van de mogelijkheden voor schoolkeuze van leerlingen kan de concurrentie tussen scholen worden aangewakkerd. Het feit dat leerlingen (en ouders) met 'hun voeten kunnen stemmen' kan scholen stimuleren de kwaliteit van het onderwijs te verbeteren. Een ruimere schoolkeuze kan tevens een betere 'match' opleveren tussen leerling en school. Een nadeel van een vrije schoolkeuze is dat de segregatie van leerlingen over scholen kan toenemen. Recent is het effect van concurrentie op prestaties van leerlingen geanalyseerd in metropolitane gebieden in de Verenigde Staten (Hoxby, 2000). De keuzemogelijkheden bestaan uit schooldistricten in deze gebieden (binnen schooldistricten valt namelijk niet veel te kiezen zoals eerder is opgemerkt). Door binnen het metropolitane gebied te verhuizen naar een bepaald district kan men gebruik maken van het publieke onderwijs in dit district. Nagegaan is of de leerprestaties in gebieden met meer districten (keuzemogelijkheden) beter zijn dan in andere

³³ Peterson, Howell, Wolf and Campbell (2001)

³⁴ In het STAR-project zaten leerlingen gemiddeld 2,3 jaar in kleinere klassen.

gebieden. Het aantal districten in een bepaald gebied kan echter afhankelijk zijn van eerdere verschillen in leerprestaties (districten met kwalitatief matig onderwijs kunnen verdwijnen door te geringe vraag naar onderwijs). Daarom is in de analyse bekeken in hoeverre het aantal te kiezen districten samenhangt met natuurlijke barrières zoals grote rivieren. De variatie in keuzemogelijkheden als gevolg van deze natuurlijke barrières, de exogene variatie in de keuzemogelijkheden, is vervolgens gerelateerd aan de verschillen in leerprestaties. Gevonden wordt dat een toename van de keuzemogelijkheden de prestaties van scholen vergroot. Tussen districten met de minste en de meeste keuze bedraagt het verschil in prestatie 0,25 tot 0,5 standaardafwijking van de leerprestatie. Dit effect wordt gevonden onafhankelijke van de sociaal-economische positie van de leerling (gemeten als ouderlijk inkomen en etniciteit). Deze studies laten zien dat het bevorderen van de schoolkeuze, en de hieruit voortvloeiende concurrentie tussen scholen, gunstig zijn voor de prestaties van leerlingen. De toepasbaarheid in de Nederlandse situatie lijkt echter beperkt omdat de mogelijkheden voor schoolkeuze in Nederland van oudsher (sinds de pacificatie van 1917) veel groter zijn dan in de Verenigde Staten. In feite kent het Nederlands basis en voortgezet onderwijs al een soort vouchersysteem. Het bevorderen van concurrentie kan daarentegen wel worden toegepast in de BVE-sector waar de afgelopen jaren door de ROC-vorming regionale monopolies zijn ontstaan.

4.2.2 Teambeloning

Concurrentie tussen scholen kan worden aangemoedigd door goed presterende scholen te belonen met een extra budget, dat bijvoorbeeld benut kan worden voor een prestatietoeslag op het salaris voor alle docenten binnen de school. Dit kan de prestatieprikkels voor docententeams versterken. Recent zijn hiermee ervaringen opgedaan in Israël en de Verenigde Staten. In Israël is een programma toegepast waarbij de best presterende scholen een bonus konden verwerven. De rangschikking van scholen is gebaseerd op de toegevoegde waarde van de prestaties van leerlingen. Hierbij wordt de vooruitgang van leerprestaties bepaald ten opzichte van de prestatie die verwacht mag worden vanwege de sociaal-economische achtergrond van de leerlingen. De leerprestaties worden aan de hand van vijf indicatoren vastgesteld. Vier hiervan hebben betrekking op de prestaties van de leerlingen in het examenjaar: het gemiddelde niveau van de examenvakken, het aantal leerlingen in de algemene richting, de gemiddelde eindexamencijfers en het slaagpercentage. De vijfde indicator, het percentage voortijdig schoolverlaters wordt voor alle leerlingen van de school bepaald. De beste 33% van de scholen krijgen een bonus, de hoogte hiervan hangt af van de relatieve positie van een school. In 1996 varieerde de bonus per school van \$13.000 tot \$105.000. Circa 25% hiervan wordt gebruikt voor verbetering van de arbeidsomstandigheden. Voor het overige wordt de bonus in de vorm van groepsbeloning onder alle docenten van de school verdeeld, als een vast percentage van hun

bruto salaris. De bonus per docent varieerde van \$200 tot \$715 op een jaarinkomen van gemiddeld \$30.000. Heeft prestatiebeloning de prestaties van leerlingen verbeterd? Het programma is over de periode 1995 - 1997 zorgvuldig geëvalueerd.³⁵ 62 middelgrote scholen zijn toegelaten tot het programma. Het selectie criterium voor de toelating was dat scholen de enige van hun soort (religieus versus niet religieus) zijn binnen de gemeente. Voor de evaluatie van het programma wordt gebruik gemaakt van dit selectie criterium. Scholen die de enige in hun soort zijn binnen een gemeente (de experimentele groep) worden vergeleken met scholen waarvan er twee van hun soort zijn binnen de gemeente (de controle groep). Verwacht mag worden dat de kenmerken van de controlegroep niet sterk zullen verschillen van de experimentele groep.

Voor vier van de vijf indicatoren werd een toename ten opzichte van de controlegroep vastgesteld. Het gemiddelde examencijfer steeg met 1,75 punten (op een schaal van 1 op 100). De schooluitval van groep 9 naar groep 10 daalde gemiddeld van 3,7 procent naar 3,2 procent. Voor de overgang van groep 10 naar 11 daalde de schooluitval van 6,2 procent naar 4,8 procent. Alleen het gemiddelde slaagpercentage vertoonde geen verbetering. Opmerkelijk is dat het programma bij alle vijf indicatoren vooral leidde tot een verbetering van de leerprestaties van leerlingen met een lage sociaal-economische achtergrond. Het slaagpercentage van deze leerlingen nam toe terwijl dit bij de andere licht afnam. De resultaten waren aanleiding om het programma in Israël op nationale schaal in te voeren.

Ook in een schooldistrict in Dallas (VS) is prestatiebeloning in het basisonderwijs (BO) en het voortgezet onderwijs (VO) ingevoerd. In hoofdlijnen lijkt het sterk op het Israëlische programma. Het belangrijkste verschil is dat de bonussen vast zijn, onafhankelijk van de positie van een school binnen een rangordening of van het bruto salaris van een docent. De 50% best-presterende scholen krijgen een bonus: \$2000 voor het schoolactiviteiten fonds, \$1000 voor elke docent en \$500 voor elke medewerker van het onderwijsondersteunende personeel. De rangordening van de scholen is gebaseerd op een aantal indicatoren. Eén hiervan is de gemiddelde toename van de individuele leerprestaties gedurende het schooljaar. De leerprestaties van leerlingen worden aan het begin en aan het einde van een schooljaar gemeten. Hiervoor worden meerdere toetsen gebruikt om te vermijden dat leerlingen alleen voor een specifieke toets worden opgeleid. Van elke leerling wordt de toename van de leerprestaties bepaald, waarbij de uitkomst wordt gecorrigeerd voor de sociaal-economische kenmerken van de leerling. Andere prestatie-indicatoren van scholen in het BO zijn de aanwezigheid van de leerlingen en het percentage zittenblijvers. Bij scholen in het VO is hieraan nog het percentage

³⁵ Lavy, V. (1999), "Evaluating the Effect of Teachers' Performance Incentives on Pupils' Achievements," mimeo, Hebrew University, Jerusalem.

voortijdig schoolverlaters toegevoegd. Wat zijn de effecten van prestatiebeloning op de prestaties van de leerlingen? Ladd heeft het slaagpercentage voor lezen en wiskunde van de scholen (eerste klas VO) in Dallas vergeleken met die van scholen in vijf andere grote steden in Texas.³⁶ Over de periode 1990 - 1995 nam in Dallas het slaagpercentage voor lezen met 10 tot 15 procentpunten toe ten opzichte van de andere schooldistricten. Voor wiskunde steeg het slaagpercentage 14 tot 17 procentpunten. Niet alle groepen leerlingen gingen beter presteren. Vooral blanke en Spaanstalige leerlingen profiteerden, maar bij de Afro-Amerikaanse leerlingen was er geen toename van de leerprestaties. Voor risicoleerlingen zijn de uitkomsten wisselend. Voor de overige indicatoren wordt gevonden dat prestatiebeloning niet leidde tot een verandering van aanwezigheid van leerlingen op de scholen. Wel nam in Dallas het percentage voortijdig schoolverlaters af ten opzichte van de andere districten.

Teambeloning lijkt voor achterblijvers tot gunstige prestaties te leiden. De ervaring in Israël leert dat hun slaagpercentage toeneemt. In Dallas nemen die leerprestaties niet zonder meer toe. Wel neemt in beide programma's het percentage voortijdig schoolverlaters af.

4.2.3 Standaarden/ leerdoelen

Leerstandaarden kunnen gezien worden als een prikkel voor de inspanning van leerlingen en docenten. Aan het behalen van de standaard of het doel is namelijk in het algemeen een beloning verbonden (b.v. doorstromen naar een volgende groep). De laatste jaren is in toenemende mate belangstelling voor het effect van standaarden of leerdoelen in het onderwijs. De aandacht richt zich daarbij op de eisen die individuele docenten stellen aan hun leerlingen (hoe hoog leggen ze de lat) maar ook op het effect van nationale standaarden. Vooral in de theoretische literatuur is aandacht besteed aan standaarden. In het algemeen wordt daarin gevonden dat hoge standaarden de inspanning van de gemotiveerde leerlingen bevorderen en die van anderen juist verminderen als de standaard buiten bereik dreigt te komen. Hoge standaarden zouden in dat geval een negatief effect kunnen hebben op achterstandsleerlingen. Recent zijn ook enkele empirische studies verschenen. Figlio & Lucas (2000) analyseren het effect van leerstandaarden van docenten op de prestaties van leerlingen in het basisonderwijs. De leerstandaard van de docent wordt bepaald door de beoordeling door de docent te vergelijken met de score op een gestandaardiseerde toets. Daarbij wordt gebruikt gemaakt van panelgegevens van leerlingen waardoor de mogelijk verstoringen als gevolg van niet geobserveerde kenmerken wordt tegengegaan. De analyse laat zien dat veeleisende docenten (hoge standaarden) een groot positief effect hebben op de leerprestaties. Dit effect geldt met

³⁶ Ladd, H.F. (1999), "The Dallas School Accountability and Incentive Program: an Evaluation of its Impacts on Student Outcomes," *Economics of Education Review*, (18) pp 1-16.

name voor de goede leerlingen. Echter ook zwakke leerlingen profiteren en dit geldt sterker naarmate zij meer goede medeleerlingen hebben.

Betts & Grogger (2000) analyseren het effect van standaarden op testcores, voltooiën van high school, instroom in 'college' en startsalariissen. De standaard wordt ontleend aan de beoordeling door de school in vergelijking met de score op een landelijke gestandaardiseerde toets.

Gevonden wordt dat testcores stijgen door hogere standaarden, dit geldt meer voor goede leerlingen dan voor zwakke leerlingen. Er worden geen effecten gevonden op het bereikte onderwijsniveau, en voor leerlingen uit minderheden daalt het bereikte onderwijsniveau. Ook wordt enige bewijs gevonden voor een verbetering van de startsalariissen, het gaat hier echter om een kleine steekproef. De verbetering van de testcores en de hogere uitval van leerlingen uit minderheden heeft mogelijk te maken met de verslechtering van hun relatieve positie, hun achterstand neemt toe door de hogere standaarden.

Deze uitkomsten laten zien dat hogere standaarden resulteren in 'winners' en 'losers'. Dit is in overeenstemming met de theoretische literatuur. Anders gezegd, hogere standaarden leiden niet tot Pareto verbeteringen in onderwijsuitkomsten.

Een gerelateerd onderwerp is het effect van nationale standaarden, zoals een centraal schriftelijk eindexamen, op leerprestaties. Bishop (1997) vergelijkt de leerprestaties in verschillende Amerikaanse staten en in verschillende landen waarbij het al of niet hebben van een nationale standaard als verklarende variabele wordt gebruikt. Gevonden wordt dat nationale standaarden de leerprestaties verbeteren. Bishop schrijft dit toe aan het verschuiven van de aandacht van relatieve prestaties naar absolute prestaties door nationale standaarden.

4.2.4 Institutionele aspecten

In diverse studies wordt de laatste jaren aandacht besteed aan de impact van institutionele aspecten/hervormingen van het onderwijssysteem, zoals een verlenging van de leerplicht. De belangstelling hiervoor is ontstaan vanuit een indirecte weg, namelijk via het onderzoek naar het causale effect van onderwijs op de lonen. Onderwijshervormingen kunnen namelijk leiden tot toevallige variatie in onderwijsdeelname en kunnen daardoor als instrument gebruikt worden om de bijdrage van onderwijs aan de lonen te bepalen. De onderwijsdeelname veroorzaakt door dit soort hervormingen hangt namelijk niet samen met verschillen in geschiktheid, talent of motivatie. Inmiddels zijn dit soort studies voor verschillende landen uitgevoerd. Angrist & Krueger (1991) gebruiken de vaste startdatum van schooljaren in combinatie met de leerplicht in de Verenigde Staten als instrument voor verschillen in onderwijsdeelname. Naarmate men eerder in het jaar is geboren wordt gemiddeld minder onderwijs gevolgd. Harmon & Walker (1995) gebruiken de verlenging van de leerplicht die zich tweemaal heeft voorgedaan in het Verenigd Koninkrijk. Meghir & Palme (1999) maken gebruik van de verlenging van de

leerplicht in Zweden die gefaseerd naar gemeenten is ingevoerd. Zoals gezegd, deze studies beogen allereerst het causale effect van onderwijs op het loon vast te stellen. Op dat punt wordt gevonden dat de financiële opbrengsten van een jaar extra onderwijs als gevolg van dit soort institutionele hervormingen eerder hoger ligt dan lager voor het geval de opbrengst bepaald wordt zonder rekening te houden met versturende factoren.³⁷

Voor het onderhavige project is interessant dat deze hervormingen met name effect hebben op de onderwijsdeelname van achterblijvers en dat de financiële opbrengst van de extra deelname aan onderwijs relatief hoog is. In Zweden werd zelfs een soort 'knock on' effect gevonden in de zin dat achterblijvers nog langer onderwijs gingen volgen dan verplicht opgelegd door de hervorming. In alle studies werd tevens gevonden dat de financiële opbrengst van een extra jaar onderwijs hoger is voor achterblijvers dan voor anderen. Het extra onderwijs vertaalt zich dus in een hoger loon. Op individueel niveau geeft een kosten-baten analyse, gezien het hoge private financiële rendement, derhalve een voordelig resultaat.

Het algemene beeld is dat een verlenging van de leerplicht positief uitpakt voor achterblijvers. In de hiervoor besproken studies gaat het echter steeds om hervormingen die al geruime tijd geleden zijn doorgevoerd. Het is de vraag in hoeverre op dit moment nog ruimte bestaat voor bijvoorbeeld een verlenging van de leerplicht en of de beweegredenen die indertijd speelden op dit moment nog relevant zijn gezien de sterke toename van de onderwijsdeelname over alle linies. Als er al ruimte zou zijn voor een uitbreiding van de leerplicht ligt die ruimte waarschijnlijk eerder aan het begin van het traject.

4.3 Specifieke projecten

De derde groep van interventies betreft specifieke projecten gericht op 'students at risk'. Op hoofdlijnen kunnen daarbij twee typen interventies worden onderscheiden: (1) integrale schoolhervormingen, (2) specifieke projecten gericht op preventie van uitval.

Integrale schoolhervormingen

In de Verenigde Staten worden thans op veel scholen projecten uitgevoerd die worden aangeduid als 'whole school reforms'. De achterliggende gedachte is dat het veranderen van één element uit het onderwijsproces onvoldoende is en dat alleen met een integrale aanpassing van het onderwijsproces resultaten kunnen worden geboekt. Opvallende voorbeelden in dit verband

³⁷ Een verklaring voor de hogere opbrengsten van een jaar extra onderwijs is dat onderwijshervormingen met name de deelname beïnvloeden van groepen die relatief weinig deelnemen aan onderwijs. Het rendement op onderwijs voor deze groepen ligt vaak hoger dan het gemiddelde rendement over alle werkenden.

zijn het Accelerated Schools project³⁸, het Edison project, het Modern Red Schoolhouse, Roots&Wings en Succes for All. Al deze projecten zijn gebaseerd op een soort basisfilosofie over effectief onderwijs. Sommige projecten worden op grote schaal toegepast. Zo werd Succes for All in 1998 toegepast op 474 scholen met 250.000 leerlingen en wordt het Edison project thans toegepast op 113 scholen met 57.000 leerlingen. Hoewel deze projecten bemoedigende resultaten suggereren zijn tot dusver geen grondige evaluaties uitgevoerd waarbij daadwerkelijk gekozen is voor een experimenteel design. Daardoor is niet duidelijk wat de bijdrage is van de programma's aan de verbetering van de leerresultaten.

Projecten rond onderwijsuitval

Verschillende programma's rond onderwijsuitval zijn de laatste jaren op experimentele wijze geëvalueerd. Het betreft zowel programma's gericht op preventie van onderwijsuitval als programma's gericht op het alsnog behalen van certificaten door leerlingen die het onderwijs voortijdig hebben verlaten. Een overzicht van de inhoud en resultaten van deze programma's is gegeven in bijlage 2. Bij sommige specifieke projecten gericht op het voorkomen van voortijdig schoolverlaten zijn gunstige effecten gevonden. Deze effecten lijken echter bescheiden en gaan in sommige gevallen gepaard met hoge kosten. Heckman (2000) concludeert ten aanzien van de programma's voor preventie van uitval: 'sustained interventions targeted at adolescents still enrolled in school can positively impact learning and subsequent employment and earnings. (...) Impacts are modest, but positive. (...) These studies also suggest that interventions for dropouts are much less successful.'

³⁸ Voor een uitvoerige beschrijving zie www.acceleratedschools.net

4.4 Conclusies

In het voorgaande zijn de kosten en effecten van diverse interventie in het schoolse traject besproken. Een overzicht is gegeven in tabel 4.3. In de tweede en derde kolom zijn de effecten op de leerprestaties weergegeven van respectievelijk alle leerlingen en de achterblijvers (+ is positief effect, = is geen effect, - is negatief effect, +/- wisselend effect, =/- geen of negatief effect, ? is onbekend effect). In de vierde kolom is een kwalitatieve typering gegeven van de kosten-effectiviteit van de interventie *vergeleken met de kosten-effectiviteit van het verkleinen van de groepsgrootte*. Deze kolom geeft aan of een gulden besteed aan een bepaalde interventie meer of minder leerwinst voor de leerling oplevert dan een gulden besteed aan het verkleinen van de groepsgrootte. De effecten worden afgezet tegen het effect van groepsgrootte omdat, zo zal in het volgende hoofdstuk naar voren komen, deze interventie het meest wordt toegepast in Nederland. De kwalitatieve typering zijn voor zover mogelijk gebaseerd op bevindingen uit de literatuur. Zo zijn de typering bij de interventies 'lesuren', 'scholing docent', en 'teambeloning' gebaseerd op kosten-effectiviteit analyses. De positieve typering voor de verlenging van de leerplicht is gebaseerd op het hoge financiële rendement van het onderwijs dat additioneel wordt gevolgd als gevolg van deze interventie (dit is in feite een kosten-baten analyse op individueel niveau). Voor de andere interventies waren geen directe kosten-effectiviteit analyses beschikbaar. De positieve typering voor vouchers en concurrentie zijn gebaseerd op de relatief grote effecten en de relatief lage kosten.³⁹ De negatieve typering voor ICT en specifieke projecten rond uitval zijn gebaseerd op de hoge kosten en matige/bescheiden effecten.

³⁹ De kosten van de voucherexperimenten varieerden tussen 1200 \$ en 1700 \$ per leerling per jaar, terwijl de kosten van het STAR-programma ongeveer 3500 \$ per leerling per jaar bedroegen (Krueger, 2000, p. 25). Het STAR-programma had bovendien een langere duur. De kosten van het vergroten van concurrentie zijn niet bekend. In de Verenigde Staten zou dit kunnen neerkomen op het vergroten van de mogelijkheden om onderwijs te volgen buiten het eigen schooldistrict. De kosten lijken in dat geval relatief laag.

Tabel 4.3 Schatting van de kosten en effecten van interventies voor preventies van onderwijsachterstanden volgens buitenlandse studies

Interventie	Effecten		kosten-effectiviteit
	algemeen	achterblijvers	t.o.v. verkleinen groepsgrootte
verkleining groepsgrootte	+	+	=
meer lesuren / schooltijd	+	+	>
scholing docent	+	+	>
initiële opleiding docent	?	?	?
ervaring docent	?	?	?
beloning docent	?	?	?
ICT	= / -	= / -	<
peergroup (spreidingsbeleid)	?	?	?
vouchers	?	+	>
vergroting concurrentie	+	?	>
teambeloning	+	+	>
standaarden	+ / -	+ / -	?
verlenging leerplicht	+	+	>
specifieke projecten rond uitval	?	+	<

+ positief effect, = geen effect, - negatief effect, + / - wisselend effect, =/- geen of negatief effect, ? is onbekend effect; > kosten-effectiviteit gunstiger dan bij verkleinen van groepsgrootte, >> veel gunstiger, < ongunstiger.

Het overzicht laat zien dat in de literatuur voor zes soorten interventies in het schoolse traject een gunstiger kosten-effectiviteit is gevonden dan voor het verkleinen van de klassen. Dit zijn:

- (1) de scholing van docenten
- (2) vergroting van het aantal lesuren/ schooltijd
- (3) voucherprogramma's
- (4) vergroting van concurrentie
- (5) teambeloning
- (6) verlenging van de leerplicht.

5 Ervaringen in Nederland

In deze paragraaf worden de ervaringen met het bestrijden van achterstanden in het onderwijs in Nederland besproken. Allereerst wordt een korte schets gegeven van het gevoerde beleid tot dusver (paragraaf 5.1). Vervolgens wordt ingegaan op de besteding van middelen (paragraaf 5.2). De aandacht richt zich daarbij wederom op twee elementen: (1) de *timing* van de inzet van middelen, (2) de verdeling van de middelen over de verschillende interventies (waaraan worden

de middelen besteed). In paragraaf 5.3 komen de resultaten van het beleid aan de orde: wat weten we over de opbrengst van de inzet van achterstandsmiddelen.

5.1 Schets van het onderwijsachterstandenbeleid in Nederland

In Nederland wordt sinds de jaren zeventig beleid gevoerd ter bestrijding van onderwijsachterstanden. De geschiedenis en achtergronden van het gevoerde beleid zijn elders uitvoerig beschreven⁴⁰. Hier volgt een korte weergave van de hoofdlijnen grotendeels ontleend aan eerdere beschrijvingen.

Onderwijsstimuleringsbeleid

Het talentenproject van Van Heek e.a. (1968) liet zien dat kinderen uit een arbeidersmilieu gemiddeld veel slechter presteerden in taal en rekenen dan kinderen uit hogere milieus. Om hier verandering in te brengen kregen scholen met veel arbeiderskinderen extra personeel (Beleidsplan voor het onderwijs aan groepen in achterstandssituaties, OC&W, 1974). De toekenning van extra personeel vond plaats aan de hand van de zogeheten beroepenfactor, een getal dat is afgeleid van het aandeel arbeiderskinderen per school. Dit geheel van maatregelen ter verbetering van de prestaties van de schoolloopbanen van de arbeiderskinderen werd aangeduid met de term 'onderwijsstimuleringsbeleid' (OSB). Naast het OSB werd een beleid ontwikkeld gericht op kinderen van buitenlandse werknemers. Het ging daarbij met name om kinderen van Turkse en Marokkaanse werknemers die in de jaren zestig naar Nederland waren gehaald.

Onderwijsvoorrangsbeleid

In het midden van de jaren tachtig werden de faciliteiten voor minderheden en het onderwijs aan leerlingen uit achterstandsgroepen in één kader ondergebracht: het onderwijsvoorrangsbeleid. Het doel van het onderwijsvoorrangsbeleid is het verminderen van onderwijsachterstand die ontstaat ten gevolge van economische, sociale en culturele factoren. Het gaat om allochtone en autochtone leerlingen wier startpositie in het onderwijs ongunstiger is op grond van de genoemde factoren. Voor de toekenning van faciliteiten ter realisering van ovb-doelen werd een nieuw systeem geïntroduceerd. Basisscholen kregen personeel naar rato van het aandeel leerlingen uit de ovb-doelgroepen onder de schoolbevolking.⁴¹ Bij de vaststelling van het personeel werden leerlingen uit ovb-doelgroepen daarbij zwaarder meegewogen:

⁴⁰ zie bijvoorbeeld Mulder (1996).

⁴¹ Bij de toekenning van faciliteiten werd een drempel gehanteerd. Scholen met een klein aantal achterstandsléerlingen werden geacht dit te kunnen opvangen.

leerlingen uit minderheden van wie de ouders een laag opleidingsniveau hebben kregen een wegingsfactor 1,9, autochtone leerlingen die op grond van het lage opleidings- of beroepsniveau van hun ouders tot een doelgroep behoren kregen een wegingsfactor 1,25. Bovendien zijn er nog leerlingen met gewicht 1,4 (schipperskinderen) en leerlingen met gewicht 1,7 (woonwagen- en zigeunerkinderen). Naast de extra formatie voor scholen zijn er ook faciliteiten voor onderwijsvoorrangsgebieden. Dit zijn plaatselijke of regionale samenwerkingsverbanden tussen scholen en welzijnsinstellingen in gebieden waar sprake is van een cumulatie van achterstandskenmerken. Deze gebieden krijgen jaarlijks geld voor de coördinatie van voorrangsactiviteiten en voor het uitvoeren van specifieke projecten.

Begin jaren negentig werden de eerste resultaten bekend van de landelijke evaluatieonderzoeken naar de effecten van het onderwijsvoorrangsbeleid, de LEO-onderzoeken. De prestaties van allochtone leerlingen bleven ver achter bij die van Nederlandse kinderen, dit gold met name voor Turkse en Marokkaanse leerlingen. Uit de onderzoeken bleek ook dat er maar weinig scholen waren die serieuze maatregelen namen om het prestatieniveau van de allochtone leerlingen te verbeteren. Mulder (1996) kwam tot de conclusie dat het onderwijsvoorrangsbeleid er niet in was geslaagd de prestaties van leerlingen uit de minderheden in het basisonderwijs te verbeteren⁴².

Gemeentelijk onderwijsachterstandenbeleid

Met de invoering van de Wet Gemeentelijk Onderwijsachterstandenbeleid (GOA) per 1 augustus 1998 is het onderwijsvoorrangsbeleid gedecentraliseerd naar de gemeenten, in samenwerking met de schoolbesturen (GOA is ten dele al in 1997 in werking gesteld). Doelstellingen en verwachtingen hierbij zijn:

- de regie over een samenhangende aanpak van onderwijsinstellingen en andere instellingen bij de bestrijding van onderwijsachterstanden wordt gevoerd door gemeenten;
- het beter kunnen inspelen op variëteit, namelijk op lokale verschillen in specifieke behoeften, situaties en kwaliteit.

Er wordt tevens verwacht/ gehoopt dat gemeenten zullen overgaan tot het inzetten van (meer) eigen middelen en dat onderwijs door de decentralisatie een duidelijker plaats verwerft op de lokale politieke agenda (SCP, 1999). In dit nieuwe naar gemeenteniveau gedecentraliseerd beleid fungeert een landelijk beleidskader (LBK) waarin de doelstellingen van het onderwijsachterstandenbeleid worden aangegeven. Het GOA moet gericht zijn op:

- de voor- en vroegschoolse periode;

⁴² Naar onze mening kan deze conclusie op basis van de beschikbare gegevens niet worden getrokken, zie paragraaf 5.3 voor nadere uitleg.

- beheersing van de Nederlandse taal;
- vermindering van verwijzing naar het speciaal onderwijs;
- vermindering schooluitval;
- evenredige deelname van doelgroepen aan het onderwijs;
- monitoring.

De GOA-wet laat de gewichtenregeling voor basisscholen in tact, terwijl de OVB-gebiedsmiddelen zijn omgezet in een grotere gemeentelijke doeluitkering. Zowel de gewichtenregeling van de basisscholen als de doeluitkering van de gemeenten worden conform de doelstellingen van het LBK ingezet op basis van gezamenlijke afspraken van gemeenten en schoolbesturen.

Begin 2001 verscheen een rapport van de Algemene Rekenkamer over de bestrijding van onderwijsachterstanden. In het rapport wordt de beleidsinformatie uit de tot dan toe verschenen studies en rapporten over het achterstandenbeleid geanalyseerd. Geconstateerd wordt dat ‘in de ruim 20 jaar dat er een onderwijsachterstandenbeleid, in welke vormgeving dan ook, wordt gevoerd, er geen blijvende resultaten zijn geboekt. De onderwijsachterstanden zijn niet merkbaar verminderd.’ (Rekenkamer, 2001, p. 7) (zie ook paragraaf 5.3).

5.2 Hoe worden de middelen voor achterblijvers besteed?

De timing van de inzet van achterstandsmiddelen

In Nederland worden middelen voor preventie van onderwijsachterstanden ingezet in het voorschoolse traject, het basisonderwijs en het voortgezet onderwijs (tabel 5.1).

Tabel 5.1 Uitgaven Onderwijsachterstandenbeleid en Voor en Vroegschoolse Educatie 1996-2001
(x f1 miljoen)

	1996	1997	1998	1999	2000	2001
Voor- en vroegschoolse educatie	2,4	3	3	3	78	183
gewichtenformatie (PO)	477,2	497,8	499,4	509,9	513,8	514,8
Cumi-faciliteiten (VO)	91,1	88,4	94,7	87,9	104,1	112
Onderwijsvoorrangsgebieden	74,1	77,1				
Specifieke uitkering GOA			116,5	173,1	174,3	180,4
overig OVB/ GOA	179,1	193,5	204,8	212	335,2	328
Totaal PO+VO+GOA+overig	821,5	856,7	915,4	982,8	1127,4	1135,3
BVE						
Bestrijding voortijdig schoolverlaten	6	6	6	6	61	71
ESF-bestrijding schoolverlaten	35	35	45	35	61*	61*

* voor projecten gereserveerde middelen

bron: Rekenkamer (2001) en OCW

Het grootste deel van de achterstandsmiddelen wordt besteed in het basisonderwijs in de vorm van gewichtenformatie die grotendeels bestemd zijn voor 1.25 leerlingen (16 % van alle leerlingen in 1999/ 2000) en 1.9 leerlingen (12,8 van alle leerlingen in 1999/ 2000). Het voortgezet onderwijs krijgt verhoudingsgewijs veel minder middelen. Deze zogenoemde Cumi-gelden zijn alleen bestemd voor leerlingen uit culturele minderheden, dit betreft 8,8 procent van alle leerlingen in 1999/ 2000. In de BVE-sector neemt de inzet van achterstandsmiddelen de laatste jaren toe. Deze middelen worden ingezet ter bestrijding van voortijdig schoolverlaten. De laatste jaren wordt substantieel meer uitgegeven aan voor- en vroegschoolse educatie. De intensivering in 2000 en 2001 is onderdeel van het recente 'Onderwijskansenplan 2000-2005'. De voor- en vroegschoolse fase is bestemd voor kinderen van 3 tot 6 jaar.

Aan welke interventies worden de middelen besteed?

Zoals hiervoor is opgemerkt, neemt de inzet van middelen in het voor- en vroegschoolse traject recent sterk toe. Aan welke interventies deze middelen worden besteed is nog niet bekend. Ze zijn in de eerste plaats bestemd voor capaciteitsuitbreiding van de peuterspeelzalen. Daarnaast

zijn de middelen bestemd voor scholing, programma-ontwikkeling en afstemming tussen peuterspeelzalen en basisscholen.

Over de besteding van de middelen in het basisonderwijs zijn geen directe rapportages beschikbaar. Wel is in het zogenoemde PRIMA-onderzoek aan directeuren in het basisonderwijs gevraagd hoe de achterstandsmiddelen worden ingezet. Driekwart van de directeuren, die aangeven dat zij achterstandsmiddelen ontvangen, antwoordt dat de middelen ofwel geheel aan klassenverkleining worden uitgegeven ofwel deels aan klassenverkleining en deels aan specifieke activiteiten. Dit antwoordpatroon wordt gevonden in de drie tot dusver uitgevoerde golven van het PRIMA-onderzoek in '94/'95, '96/'97 en '98/'99.⁴³

Over de besteding van de Cumi-gelden in het voortgezet onderwijs zijn evenmin directe rapportages beschikbaar. Door de Rekenkamer is in 1998 een enquête uitgevoerd onder scholen betrokken bij de bestrijding van de onderwijsachterstanden in het voortgezet onderwijs.⁴⁴ Daarin is gevraagd welke activiteiten de scholen ondernomen hebben met de aanvullende middelen. Voor de CUMI-gelden wordt een groot aantal verschillende activiteiten genoemd: extra lessen (genoemd door 75 % van de scholen), huiswerkbegeleiding (76 %), deskundigenbegeleiding (63 %), contacten met ouders (51 %), begeleiding docenten (55 %), aanpassen onderwijs (40 %), andere activiteiten (18 %) en activiteiten voor alle leerlingen (29 %).

Ten aanzien van de 'timing van middelen' kan geconcludeerd worden dat het zwaartepunt van de inzet van middelen voor bestrijding onderwijsachterstanden ligt bij het basisonderwijs. De belangrijkste interventie bij de besteding van deze middelen daar is het verkleinen van de klassen.⁴⁵ De uitgaven aan voor- en vroegschoolse educatie nemen recent sterk toe, in 2001 wordt hieraan meer uitgegeven dan het totaal van de Cumi-gelden in het voortgezet onderwijs. Met andere woorden, tot voor kort was de *timing* van de inzet van middelen gericht op het basisonderwijs en niet op de allervroegste fase. Daarbij kiezen scholen in het basisonderwijs vooral voor de interventie '*het verkleinen van de groepsgrootte*'.

⁴³ Ook volgens toenmalig minister Ritzen werden de achterstandsmiddelen grotendeels gebruikt voor het verkleinen van klassen. Zie Ritzen e.a. (1997), *School finance and school choice in the Netherlands*, *Economics of Education Review*, vol. 16, no. 3, p. 331.

⁴⁴ Rekenkamer (1998), *Bestrijding onderwijsachterstand in het voortgezet onderwijs*, p. 18.

⁴⁵ Het gaat hier om de verkleining van klassen voortvloeiend uit het onderwijsachterstandenbeleid en niet om het project klassenverkleining in de onderbouw van het basisonderwijs.

5.3 Wat weten we over de resultaten?

In Nederland zijn diverse evaluatiestudies uitgevoerd naar de effecten van het achterstandsbeleid. Geen van deze studies voldoet aan de eisen die in paragraaf 2 zijn gesteld: (1) een experimentele of quasi-experimentele benadering, (2) waarbij ook lange termijn effecten worden gemeten. Dit betekent dat alle tot dusver gerapporteerde resultaten vertekend kunnen zijn door niet-waargenomen kenmerken van leerlingen, ouders of scholen. Anders gezegd, 'harde' resultaten over de effecten van de besteding van achterstandsmiddelen zijn niet beschikbaar⁴⁶. Hieronder bespreken we de belangrijkste projecten met als kanttekening dat de resultaten mogelijk vertekend kunnen zijn door endogeniteit.

Piramide en Kaleidoscoop

Op advies van de Commissie Voorschoolse Educatie (1994) is in 1995 begonnen met de ontwikkeling van twee 'experimentele' voor- en voerschoolse educatieve programma's. Met 'experimentele' wordt hier bedoeld sterk gestructureerde programma's op peuterspeelzalen en basisscholen. De programma's worden uitgevoerd in peuterspeelzalen en kleutergroepen van basisscholen onder gunstige pedagogische kwaliteitsvoorwaarden. Kaleidoscoop is gebaseerd op het huidige High/Scope-programma dat voortbouwt op het Perry-Preschool programma. Piramide is gebaseerd op onderdelen van het programma Succes for All. Beide programma's bieden een intensief programma met brede educatieve doelen voor kinderen van 3 tot 6 jaar in maatschappelijke achterstandssituaties. Het programma duurt 3 jaar en wordt gedurende vier dagdelen per week uitgevoerd. In het schooljaar '96/'97 zijn de programma's op een aantal plaatsen gestart. Het evaluatie-onderzoek is voorjaar 1997 gestart. Inmiddels is een eindrapportage beschikbaar over drie metingen⁴⁷. De laatste meting is derhalve afgenomen in groep 2 van de basisschool.

resultaten

In het evaluatieonderzoek is bij elke experimentele peuterspeelzaal of kleutergroep van de basisschool een vergelijkbare locatie gezocht waarbij in de analyse zoveel mogelijk rekening wordt gehouden met waargenomen verschillen. De leerlingen zijn derhalve niet door toeval toegewezen aan de experimentele groep en aan de controlegroep. Met een dergelijke opzet is het onmogelijk om te controleren voor niet waargenomen factoren die mogelijk gecorreleerd zijn met deelname aan het interventieprogramma.

⁴⁶ De Algemene Rekenkamer (2001) komt tot eenzelfde conclusie.

⁴⁷ zie Veen, A., J. Roeleveld en P. Leseman (2000), Evaluatie van Kaleidoscoop en Piramide, eindrapportage, SCO-Kohnstamm Instituut, Amsterdam.

In het onderzoek wordt gevonden dat Kaleidoscoop op alle vier woordenschattoetsen hoger scoort dan de controlegroep. Het gemiddelde effect is zwak tot matig. Piramide scoort hoger op twee van de vier woordenschattoetsen en op 'Ordenen'. De gevonden effecten worden overwegend zwak tot bescheiden genoemd met een enkele uitschieter naar boven en naar beneden. Geconcludeerd wordt dat de gevonden effecten in overeenstemming zijn met vergelijkbaar buitenlands onderzoek⁴⁸.

De gevonden resultaten van de programma's zijn, in het licht van de internationale bevindingen, enigermate teleurstellend. Mogelijk heeft dit te maken met de korte voorbereiding voor het uitvoeren van beide programma's. Ook zou de peergroup een rol kunnen spelen. Van zowel de experimentele als de controlegroep is de 'peergroup' niet volledig bekend, omdat niet alle kinderen in de gekozen locaties meededen aan het onderzoek. Afgaande op de samenstelling van de steekproeven komen in de experimentele groepen meer Turkse en Marokkaanse leerlingen voor dan in de controlegroepen. Dit kan ongunstige peergroup effecten opleveren met betrekking tot de taalontwikkeling in het Nederlands. Ook is niet bekend wat er gebeurt bij de controlegroepen (welke programma's worden toegepast). Verder wordt de evaluatie geplaagd door een grote uitval bij de overgang van de peutergroepen naar de kleutergroepen op de lagere school. Ook blijken er belangrijke verschillen te bestaan tussen de basisscholen van de leerlingen uit de experimentele groepen en van de leerlingen uit de controle groep. De leerlingpopulatie van de basisscholen van de controlegroepen heeft een hogere sociaal-economische achtergrond. Niet duidelijk is in hoeverre de gevonden resultaten door het bovenstaande wordt beïnvloed. De kosten van de programma's lijken relatief gering. Voor de gehele driejarige periode (1 jaar peuterspeelzaal en 2 jaar basisonderwijs) bedragen de kosten naar schatting ongeveer 4000 gulden per deelnemer.⁴⁹

Het Onderwijsvoorrangsbeleid/ GOA

Diverse studies zijn en worden uitgevoerd om de effecten van het onderwijsvoorrangsbeleid, thans GOA, te bepalen. Het grote probleem bij deze studies is dat er geen controlegroep beschikbaar is, dat wil zeggen een groep door toeval aangewezen leerlingen waarvoor geen achterstandsmiddelen beschikbaar zijn of een groep waarvoor juist extra middelen beschikbaar werden gesteld. Ook is tot dusver geen quasi-experimentele benadering toegepast in Nederland⁵⁰. Dit betekent dat we op dit moment de 'counterfactual' niet kennen (wat zou er

⁴⁸ Het vergelijkingsmateriaal is daarvoor ontleend aan een metastudie over 18 studies, zie Leseman e.a. (1998).

⁴⁹ Gebaseerd op Schonewille, e.a. (2000), p. 78-79. De aanloopkosten en uitvoeringskosten zijn gedeeld door het maximale aantal leerlingen per groep.

⁵⁰ Binnen het NWO-prioriteitsprogramma SCHOLAR wordt op dit moment gewerkt aan een studie waarin de toevallige variatie in OVB-middelen tussen onderwijsvoorrangsgebieden wordt gebruikt als natuurlijk experiment.

gebeurt zijn zonder onderwijs-voorrangsmiddelen). Het algemene beeld uit de tot dusver uitgevoerde studies is dat de achterstanden van leerlingen uit de doelgroepen van het Onderwijsachterstandenbeleid gedurende het basisonderwijs nauwelijks verminderen. Een uitzondering hierop is de recent verschenen studie naar de prestaties van minderheden in het Nederlands onderwijs.⁵¹ In deze studie wordt geconstateerd dat de achterstand van minderheden in het onderwijs tussen 1988 en 1998 is verminderd, vooral in rekenen en in veel mindere mate in taal. In hoeverre de verbetering van de resultaten gerelateerd is aan het gevoerde beleid dan wel veroorzaakt wordt door buitenschoolse factoren is niet duidelijk. Geen van de uitgevoerde studies rond het onderwijsvoorrangbeleid/ GOA werpt licht op het causale effect van de inzet van achterstandsmiddelen.

KEA

Op enkele scholen in het basisonderwijs worden gunstige resultaten gemeld bij het verminderen van onderwijsachterstanden. Dit betreft scholen die deelnemen aan het KEA-project (Kleinschalig Experiment Achterstandsbestrijding). Het belangrijkste middel van KEA is de verandering van het instructiegedrag van leraren. Leraren worden actief begeleid, geobserveerd en geëvalueerd (in de vorm van minimaal 30 klassenconsultaties per jaar). Op deze scholen wordt ook veel aandacht besteed aan de vorderingen van de leerlingen en wordt zo nodig extra instructie aangeboden aan de leerlingen. Ook wordt voorrang gegeven aan het wegwerken van taalachterstanden. De gerapporteerde evaluaties duiden op duidelijke verbeteringen van de prestaties van de leerlingen. Bij de evaluaties is evenwel geen sprake van een experimentele opzet. Daarom is niet duidelijk in hoeverre de gerapporteerde verbetering van de leerprestaties toegeschreven kunnen worden aan de KEA-aanpak.⁵²

Onderzoek naar uitval in het voortgezet onderwijs en BVE

In het onderzoek naar uitval in het voortgezet onderwijs en BVE in Nederland wordt in het algemeen de verschillen tussen uitvallers en 'blijvers' in kaart gebracht alsmede de redenen van uitval zoals aangegeven door de uitvallers. Onderzoek naar de effectiviteit van een bepaalde interventie/programma gericht op preventie van uitval op basis van een experimentele opzet is niet gedaan in Nederland.

⁵¹ Tesser, P.T.M en J.Iedema (2001)

⁵² Het Rotterdamse CED heeft toegezegd de gegevens over de leerprestaties op de KEA-scholen aan het CPB beschikbaar te stellen. Het voornemen is om deze gegevens in een later stadium te analyseren.

6 Samenvatting en conclusies

Hoe kunnen de middelen voor preventie van onderwijsachterstanden het best worden ingezet? Deze vraag staat centraal in dit onderzoek. Daarbij worden twee deelvragen onderscheiden. In welke onderwijsfase kunnen de middelen het best worden besteed (de timing van de inzet van middelen) en welke interventie verdient daarbij de voorkeur (waaraan moeten de middelen worden besteed)?

De Nederlandse aanpak

In Nederland wordt het merendeel van de middelen voor preventie van onderwijsachterstanden ingezet in het basisonderwijs. De middelen worden daarbij voornamelijk gebruikt voor het verkleinen van groepen. Worden de middelen op deze wijze optimaal ingezet? Op grond van Nederlands onderzoek weten we dat niet. De effecten van de inzet van achterstandsmiddelen in het Nederlands onderwijs zijn niet bekend. Het evaluatieonderzoek dat tot dusver is uitgevoerd kent geen (quasi-)experimentele benadering en vaak ook geen controlegroep⁵³. Het is dan moeilijk om vast te stellen of veranderingen in leerprestaties veroorzaakt worden door de inzet van middelen voor een bepaalde onderwijsinterventie of door andere factoren die niet worden waargenomen door de onderzoeker, bijvoorbeeld buitenschoolse factoren. Dit betekent dat op dit moment de causale effecten van het Nederlands onderwijsachterstandenbeleid niet bekend zijn⁵⁴. In het buitenland zijn de afgelopen jaren diverse studies uitgevoerd naar de causale effecten van onderwijsinterventies. De resultaten van dit buitenlands onderzoek zijn in het voorliggende onderzoek op een rijtje gezet en vergeleken met de Nederlandse aanpak.

Een overzicht van buitenlands experimenteel onderzoek naar onderwijsinterventies

Lange tijd bestond bij economen een grote scepsis over de effecten van extra middelen voor onderwijsinterventies. Deze scepsis was gebaseerd op een klassiek overzichtsartikel van diverse interventies in het onderwijs, zoals het verkleinen van groepen of het verhogen van het salaris van docenten⁵⁵. De conclusie daarin was dat geen systematische relatie gevonden was tussen de diverse interventies en de resultaten van leerlingen. Het beeld verandert evenwel als strengere

⁵³ Een quasi-experimentele benadering is gebaseerd op toevallige situaties in de werkelijkheid vergelijkbaar met 'echte' experimenten. Deze toevallige situaties worden ook vaak natuurlijke experimenten genoemd.

⁵⁴ Regelmatig wordt de indruk gewekt dat het Nederlands onderwijsachterstandenbeleid niet succesvol is. De feitelijke situatie is echter dat de causale effecten niet bekend zijn.

⁵⁵ Hanushek (1986)

eisen gesteld worden aan de kwaliteit van de studies⁵⁶. Verschillende studies waarin het causale effect van onderwijsinterventies wordt geanalyseerd aan de hand van een (quasi-)experimentele benadering laten zien dat meer middelen voor bepaalde interventies wel degelijk kunnen leiden tot betere leerprestaties. Voor het overzicht zijn daarom studies geselecteerd uit de internationale economische literatuur die een experimentele of quasi-experimentele opzet kennen. De deelnemers aan een bepaald programma worden in deze studies door toeval toegewezen aan een experimentele of een controle groep. In dat geval kan het causale effect van een bepaalde interventie worden vastgesteld.

De resultaten van de experimentele literatuur zijn in tabel 6.1 samengevat. Per onderwijsinterventie zijn de effecten gegeven op de leerprestaties van alle leerlingen (kolom 2) en van achterblijvers (kolom 3). Met het oog op de deelvraag naar de 'timing' van de onderwijsinterventies zijn de resultaten van de vroegste interventies het eerst gegeven. Om vast te kunnen stellen of de middelen efficiënt worden ingezet dienen de effecten te worden gezien in relatie tot de kosten verbonden aan de interventie. Dit gebeurt in de meest rechtse kolom waarin een kwalitatieve typering is gegeven van de kosten-effectiviteit van de interventie. Voor elke interventie wordt de kosten-effectiviteit *vergeleken met de kosten-effectiviteit van het verkleinen van de groepsgrootte*. Deze kolom geeft aan of een gulden besteed aan een bepaalde interventie meer of minder leerwinst voor de leerling oplevert dan een gulden besteed aan het verkleinen van de groepsgrootte. Dit is als referentie gekozen omdat het verkleinen van de groepsgrootte de belangrijkste interventie is bij de preventie van onderwijsachterstanden in Nederland. De kwalitatieve typering is gebaseerd op de resultaten van de buitenlandse studies en dit geldt ook voor de schatting van de kosten-effectiviteit van het verkleinen van de groepsgrootte.

Het overzicht in tabel 6.1 laat zien dat het verkleinen van groepen leidt tot een verbetering van de leerprestaties van achterblijvers in de kenniseconomie. Met het verkleinen van groepen zijn echter in het algemeen veel middelen gemoeid. In buitenlandse studies wordt dan ook voor zeven soorten interventies een gunstiger kosten-effectiviteit gevonden dan voor het verkleinen van groepen. Dit zijn: (1) voor- en vroegschoolse programma's, (2) vergroting van het aantal lessen, (3) de scholing van docenten, (4) voucherprogramma's, (5) vergroting van concurrentie, (6) teambeloning en (7) verlenging van de leerplicht.

⁵⁶ Krueger (2000)

Tabel 6.1 Schatting van de kosten en effecten van interventies voor preventies van onderwijsachterstanden volgens buitenlandse studies

Interventie <i>voor-/ vroegschoolse</i>	Effecten		kosten-effectiviteit
	algemeen	achterblijvers	t.o.v. verkleinen groepsgrootte
Perry Preschool	?	+	>>
Head Start*	?	+	>
Chicago Child Parents Center*	?	+	>>
<i>schoolse traject</i>			
verkleining groepsgrootte	+	+	=
meer lesuren / schooltijd	+	+	>
scholing docent	+	+	>
initiële opleiding docent	?	?	?
ervaring docent	?	?	?
beloning docent	?	?	?
ICT	= / -	= / -	<
peergroup (spreidingsbeleid)	?	?	?
vouchers	?	+	>
vergroting concurrentie	+	?	>
teambeloning	+	+	>
standaarden	+ / -	+ / -	?
verlenging leerplicht	+	+	>
specifieke projecten rond uitval	?	+	<

+ positief effect, = geen effect, - negatief effect, + / - wisselend effect, = / - geen of negatief effect, ? is onbekend effect; > kosten-effectiviteit gunstiger dan bij verkleinen van groepsgrootte, >> veel gunstiger, < ongunstiger; * niet experimenteel.

De meest gunstige kosten-effectiviteit is gevonden bij de voor-/ vroegschoolse projecten (Perry Preschool en Chicago Child Parents Center). Dit wijst erop dat middelen besteed aan vroege interventies het meest opbrengen. Buitenlandse studies suggereren derhalve een zo vroeg mogelijke 'timing' van de interventie. Waaraan de middelen het best kunnen worden besteed wordt niet direct duidelijk uit buitenlandse studies. Wel laten deze studies zien dat het verkleinen van groepen een relatief dure interventie is.

Wat betekenen de buitenlandse resultaten voor het Nederlands onderwijs?

De resultaten uit buitenlandse studies zijn niet zonder meer geldig voor het Nederlands onderwijs. Dit hangt allereerst af van de inrichting van het onderwijs. Bovendien beperkt het Nederlands beleid voor preventie van onderwijsachterstanden zich niet uitsluitend tot het verkleinen van groepen en is recent nieuw beleid gestart dan wel voorgenomen. Als de buitenlandse resultaten gezien worden als indicaties voor een kosten-effectieve preventie van onderwijsachterstanden kan geconstateerd worden dat diverse elementen al terug te vinden zijn in de huidige inrichting van het Nederlands onderwijs. Dit geldt met name voor het aantal lesuren en de interventies rond concurrentie. Nederlandse leerlingen hebben internationaal

gezien al relatief veel uren. Daar komt bij dat door recent beleid met betrekking tot de brede school, zaterdagschool en dergelijke, getracht wordt de leertijd van vooral achterstandsleerlingen uit te breiden. Ook op het terrein van het vergroten van concurrentie en vouchers is de huidige Nederlandse situatie duidelijk anders en, als uitgegaan wordt van de buitenlandse bevindingen, gunstiger dan in veel andere landen. Nederland kent in het basis en voortgezet onderwijs in feite al een soort vouchersysteem. Scholen worden namelijk gefinancierd op basis van het aantal leerlingen en leerlingen kunnen een school kiezen (de financiering volgt de leerling). Dit betekent ook dat in het Nederlands onderwijs, anders dan in de Verenigde Staten, al lange tijd concurrentie tussen scholen bestaat. Dit geldt echter niet in dezelfde mate voor alle onderwijssectoren. Zo zijn de laatste jaren onder invloed van de ROC-vorming in het secundair beroepsonderwijs regionale monopolies ontstaan. Het vergroten van de concurrentie in deze sector is derhalve wel een reële optie en zou bijvoorbeeld kunnen door private aanbieders in aanmerking te laten komen voor bekostiging.⁵⁷

Binnen het Nederlands beleid gericht op preventie van onderwijsachterstanden neemt het gewicht van voor- en vroegschoolse programma's momenteel sterk toe. De komende jaren zal een substantiële intensivering van de inspanningen op dit terrein plaats vinden. Deze beleidsontwikkeling is duidelijk in lijn met de resultaten uit buitenlandse studies. Ook de voorgenomen aanpassing van de leerplicht in het wetsvoorstel 'voor uitbreiding leerplicht voor vierjarigen' is daarmee in lijn. Deze maatregel is vooral bedoeld voor achterstandsleerlingen die soms niet direct op vierjarige leeftijd naar school gaan. Met de nieuwe wet Wet op de beroepen in het Onderwijs wordt beoogd de bekwaamheid van docenten op peil te houden door bijscholing. Buitenlandse studies laten zien dat scholing van docenten een kosten-effectieve aanpak van onderwijsachterstanden kunnen zijn.

Geconstateerd kan worden dat in het Nederlands beleid gericht op preventie van onderwijsachterstanden lange tijd de nadruk lag op het verkleinen van groepen. Dit lijkt een relatief dure interventie die bovendien de huidige schaarste aan onderwijspersoneel vergroot. Ook kan geconstateerd worden dat de inrichting van het Nederlands onderwijs diverse kosten-effectieve elementen lijkt te kennen. Bovendien lijkt het recente beleid in de richting te gaan van een kosten-effectieve aanpak. Zekerheid daarover bestaat evenwel niet omdat 'hard' empirisch bewijs voor Nederland niet beschikbaar is. Dit bewijs kan worden verkregen door te experimenteren met nieuwe interventies gericht op preventie van onderwijsachterstanden. Daarmee kan kennis worden verworven waarvan alle komende generaties kunnen profiteren. Experimenteren bevordert een kosten-effectieve aanpak van onderwijsachterstanden. De keerzijde van experimenteren is de ongelijke behandeling van leerlingen. Deelnemers aan een

⁵⁷ zie CPB (2001), Centraal Economische Plan 2001, p. 136.

experiment kunnen profijt hebben van een bepaalde interventie, niet deelnemers wordt dit voordeel onthouden. Door niet te experimenteren wordt dit voordeel echter onthouden aan komende generaties leerlingen. Met andere woorden, bij experimenteren speelt het dilemma tussen ongelijke behandeling en kennis verwerven. Verschillende trends in de kenniseconomie duiden er op dat de consequenties van onderwijsachterstanden ernstiger zullen worden. Het verwerven van kennis door te experimenteren wordt dan belangrijker.

Literatuur

- Akerhielm, K. (1995), Does class size matter?, in: *Economics of Education Review*, vol. 14, no. 3, pp. 229-241.
- Angrist, J.D. and A.B. Krueger (1991), Does compulsory schooling attendance affect schooling and earnings? in: *Quarterly Journal of Economics*, vol. 106, no. 4, pp. 979-1014.
- Angrist, L. and V. Lavy (1999), Using Maimonides Rule to estimate the effect of class size on scholastic achievement, *Quarterly Journal of Economics*, CXIV, p. 533-575.
- Angrist, L. and V. Lavy (1998), Does teacher training affect pupil learning? Evidence from matched comparisons in Jerusalem Public Schools, NBER Working Paper 6781
- Angrist, L. and V. Lavy (1999), New evidence on classroom computers and pupil learning, NBER Working paper 7424.
- Barnett, S. (1995), Long-term effects of Early Childhood Programs on Cognitive and School Outcomes, in: *The future of children*, 5:3, p. 25-30
- Betts, J.R. (1996), Is there a link between school inputs and earnings? Fresh scrutiny of an old literature, in: Burtless, G. (1996), *Does money matter?, The effect of school resources on student achievement and adult success*, Brookings Institution Press, Washington D.C., pp. 141-191.
- Betts, J.R. and J. Grogger (2000), The impact of grading standards on student achievement, educational attainment and entry-level earnings, NBER Working Paper 7875.
- Bishop, J. (1997), The effect of national standards and curriculum-based exams on achievement, in: *American Economic Review*, vol. 87, no. 2, pp. 260-264.
- Blok, H. en P. M. Leseman (1996), Effecten van voorschoolse stimuleringsprogramma's: een review van reviews, in: *Pedagogische Studiën*, 73, p. 184-197.
- Boozer, M and C. Rouse (1995), Intraschool variation in class size: patterns and implications, NBER working paper 5144.
- Burtless G. (1996), *Does Money Matter? The effect of School Resources on Student Achievement and Adult Success*, Washington D.C.: Brookings Institution
- Card, D. and A. B. Krueger (1992), Does school quality matter? Returns to education and the characteristics of public schools in the United States, *Journal of Political Economy*, 100 (february): 1-40
- Card, D and A. B. Krueger (1996), Labor Market Effects of School Quality: Theory and Evidence, in: Gary Burtless, editor, *Does Money Matter? The effect of School Resources on Student Achievement and Adult Success*, Washington D.C.: Brookings Institution, pp. 97-140
- Card, D. (1999), The causal effect of education on earnings, in: O. Ashenfelter & D. Card, *Handbook of labour economics* (pp. 1-90), Amsterdam: Elsevier.

- Card, D. and Lemieux (2001), Can falling supply explain the rising return to college for younger men? A cohort based analysis, in: *Quarterly Journal of Economics*, vol. 116, 2- mai 2001, p. 705-746
- Chubb, J.E. and T.M. Moe (1990), *Politics, Markets & America's Schools*, Washington, D.C.: The Brookings Institution.
- Coleman, J. S. and others (1966), *Equality of educational opportunity*, Government Printing Office.
- Currie, J. and D. Thomas (1995), Does Head Start make a difference? in: *American Economic Review*, 85:3, p. 341-364.
- Currie, J. and D. Thomas (1998), School quality and the longer-term effects of Head Start, NBER Working Paper 6362.
- Currie, J. and D. Thomas (1999), Does Head Start help Hispanic Children? in: *Journal of Public Economics*.
- Currie, J. (2001), Early childhood intervention programs: what do we know?, in: *Journal of Economic Perspectives*, vol. 15, no. 2 - spring 2001, p. 213-238.
- Dobbelsteen, S., J. Levin and H. Oosterbeek (1999), The causal effect of class size on scholastic achievement: distinguishing the pure class size effect from the effect of changes in class composition, Scholar working paper 04/99, University of Amsterdam
- Dynarski, M. and P. Gleason (1998), *How can we help? What we have learned from evaluations of federal dropout-prevention programs*, Mathematica Policy Research, Princeton, NJ.
- Evans, Oates and Schwab (1992), Measuring peer group effects: a study of teenage behaviour, *Journal of Political Economy* 100, p. 966-991.
- Figlio, D.N. and M.E. Lucas (2000), Do high grading standards affect student performance, NBER Working Paper 7985.
- Figlio, D.N. and M.E. Page (2000), School choice and the distributional effects of ability tracking: Does separation increase equality?, NBER Working Paper 8055.
- Garces, E., D. Thomas and J. Currie (2000), Longer term effects of Head Start, NBER working paper 8054.
- Glass, G.V. and Smith, M.L. (1979), Meta-analysis of research on class size and achievement, *Educational Evaluation and Policy Analysis*, 1, p. 2-16.
- Goldhaber and Brewer (1997), Why don't schools and teachers seem to matter?, Assessing the impact of unobservables on educational productivity, in: *Journal of Human Resources*, vol. 32, no. 3, pp. 505-523.
- Guryan, J. (2001), Does money matter? Regression-discontinuity estimates from education finance reform in Massachusetts, NBER working paper 8269.

- Hanushek, E. A. (1986), The economics of schooling: production and efficiency in public schools, *Journal of Economic Literature*, 24 (september): 1141-77.
- Hanushek, E.A. (1989), The impact of differential expenditures on school performance, *Educational Researcher*, 18 (may): 45-51, 62.
- Hanushek, E.A. (1996a), Measuring investment in education, *Journal of Economic Perspectives*, vol. 10, no. 4, p. 9-30
- Hanushek, E.A. (1996b), School resources and student performance, in: Burtless, G. (1996), *Does money matter?, The effect of school resources on student achievement and adult success*, Brookings Institution Press, Washington D.C., pp. 43-73
- Hanushek, E.A., S.G. Rivkin and L.L. Taylor (1996), Aggregation and the estimated effects of school resources, in: Review of Economics and Statistics.
- Hanushek, E.A. (1997), Assessing the effects of School Resources on Student Performance: An update, *Educational Evaluation and Policy Analysis*, 19, 2, p. 141-64
- Hanushek, E.A., J.F. Kain, S.G. Rivkin (1998), Teachers, Schools and Academic Achievement, NBER Working Paper 6691.
- Hanushek, E.A., J.F. Kain, S.G. Rivkin (1999), Do higher salaries buy better teachers?, NBER Working Paper 7082.
- Hanushek, E.A. (2000), Evidence, Politics, and the Class Size Debate, University of Rochester, New York
- Harmon, C. and I. Walker (1995), Estimates of the economic return to schooling for the United Kingdom, in: *American Economic Review*, vol. 85, no. 5, pp. 1278-1286.
- Heckman, J.J., A. Layne-Farrar and P. Todd, Does measured school quality really matter? An examination of the earnings-quality relationship, in: Burtless, G. (1996), *Does money matter?, The effect of school resources on student achievement and adult success*, Brookings Institution Press, Washington D.C., pp. 192-289.
- Heckman, J.J. (2000), Policies to foster human capital (with discussion), Res. Economics. 54 (March 2000): 3-82.
- Hedges, L.V., R.D. Laine and R. Greenwald (1994), Does money matter? A meta-analysis of studies on the effects of differential school inputs on student outcomes, in: *Educational Researcher*, 23 (april), pp. 5-14.
- Heek, F. van (1968), *Het Verborgen Talent, milieu schoolkeuze en schoolgeschiktheid*, J.A. Boom en zoon, Uitgevers te Meppel
- Hoxby, M. C. (1998), The effects of class size and composition on student achievement: new evidence from natural population variation, NBER working paper 6869.
- Hoxby, C. (2000), Peer effects in the classroom: learning from gender and race variation, NBER Working Paper 7867.

- Hoxby, C. (2000), Does competition among public schools benefit students and taxpayers?, in: *American Economic Review*, vol. 90, no. 5, pp. 1209-1238.
- Karoly, L. e.a. (1998), Investing in our children: what we know and don't know about the costs and benefits of early childhood interventions, Santa Monica: RAND.
- Krueger, A.B. (1999), Experimental estimates of education production functions, in: *Quarterly Journal of Economics*, 114 (2), p. 497-532.
- Krueger, A.B. (2000), Economic considerations and Class Size, Princeton University, working paper 447, september 2000
- Krueger, A.B. and D.M. Whitmore (2001), The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from project Star, in: *Economic Journal*, 41, 468.
- Kulik, C.C. and J.A. Kulik (1991), Effectiveness of computer-based instruction: an updated analysis, in: *Computers in human behavior*, 7, p. 75-94.
- Ladd, H.F. (1999), "The Dallas School Accountability and Incentive Program: an Evaluation of its Impacts on Student Outcomes," *Economics of Education Review*, (18) pp 1-16.
- Lavy, V. (1999), Using dual natural quasi-experimental designs to evaluate the effect of school hours and class size on student achievement, Department of Economics, Hebrew University of Jerusalem.
- Lavy, V. (1999), "Evaluating the Effect of Teachers' Performance Incentives on Pupils' Achievements," mimeo, Hebrew University, Jerusalem.
- Lavy, V. (1999), Externalities and efficiency effects of class heterogeneity on student scholastic achievement, mimeo, Hebrew University, Jerusalem.
- Leseman, P.P.M., M.F. Otter, H. Blok, P. Deckers (1998), Effecten van voor- en vroegschoolse educatieve centrumprogramma's. Een meta-analyse van studies gepubliceerd tussen 1985 en 1996, in: *Nederlands Tijdschrift voor Opvoeding, Vorming en Onderwijs*, 14., nr. 3, p. 134-154.
- Leuven, E. en H. Oosterbeek (2000), Rendement van onderwijs stijgt, in: *Economische Statistische Berichten*, 23-6-2000, p. 523-524.
- Lindahl, M. (2000), Summer learning and the effect of schooling: evidence from schooling, in: M. Lindahl (2000), *Studies of causal effects in empirical labor economics*, Swedish Institute for social research, 43.
- Link, C. and J. Mulligan (1986), The merits of a longer school day, in: *Economics of Education Review*, 5, 4, pp. 373-381.
- Loeb, S. and J. Bound (1995), The effect of measured school inputs on academic achievement: Evidence from the 1920s, 1930s and 1940s Birth Cohorts, NBER Working Paper 5331.

- Meghir, C. and M. Palme (1999), Assessing the effect of schooling on earnings using a social experiment, unpublished paper, Stockholm School of Economics, Sweden.
- Mulder, L. (1996), *Meer voorrang, minder achterstand?, Het Onderwijsvoorrangsbeleid getoetst*, proefschrift, ITS, Nijmegen
- Murnane, R.J. and F. Levy (1996), Evidence from fifteen schools in Austin, Texas, in: Burtless, G. (1996), *Does money matter?, The effect of school resources on student achievement and adult success*, Brookings Institution Press, Washington D.C., pp. 93-96.
- Olson and Ackerman (2000), High school inputs and labor market outcomes for male workers in their mid-thirties: new data and new estimates from Wisconsin, Institute for Research on Poverty, Discussion Paper no. 1205-00.
- Peterson, Howell, Wolf and Campbell (2001), School vouchers results from randomized experiments, NBER
- Rekenkamer (1998), *Bestrijding onderwijsachterstand in het voortgezet onderwijs*, Den Haag.
- Rekenkamer (2001), *Bestrijding van onderwijsachterstanden*, Den Haag.
- Reynolds, A. (1998), Extended early childhood intervention and school achievement: Age thirteen findings from the Chicago Longitudinal Study, in: *Child Development*, 69:1, p. 231-246.
- Rivkin, S.G. (2000), School desegregation, academic attainment and earnings, in: *Journal of Human Resources*, 2000.
- Ritzen, J.M.M., J. van Dommelen and F.J. de Vijlder (1997), School finance and school choice in the Netherlands, *Economics of Education Review*, vol. 16, no. 3, pp. 329-335.
- Rouse, C.E. (2000), "Private School Vouchers and Student Achievement: an Evaluation of the Milwaukee Parental Choice Program," *Quarterly Journal of Economics*, 113 (2), pp. 553-602.
- Sacerdote, B. (2000), Peer effects with random assignment: Results for Dartmouth Roommates, in: *Quarterly Journal of Economics*, vol. 116, 2, mai 2001, p. 681-704.
- Schonewille, B., J.J.J. Kloprogge en A. van der Leij (2000), *Kaleidoscoop en Piramide, samenvattend evaluatierapport*, Evaluatie Begeleidingscommissie projecten voor- en vroegschoolse educatie voor de Ministeries VWS en OCenW.
- Schweinhart, L., H. Barnes and D. Weikart (1993), Significant benefits: The High/Scope Perry Preschool Study Through Age 27, Monograph of the High/Scope Educational Research Foundation, Number 10, Ypsilanti, Michigan: High-Scope Educational Research Foundation.
- Spence Boocock, S. (1995), Early Childhood Programs in other nations: Goals and Outcomes, in: *The future of children*, 5:3.
- Summers, A.A. and B.L. Wolfe (1977), Do schools make a difference? in: *American Economic Review*, vol 67, no. 4, pp. 639-52.

- Tyler, J.H., R.J. Murnane, J.B. Willett, Do the cognitive skills of school dropouts matter in the labor market?, in: *Journal of Human Resources*, vol. 35, no. 4, pp. 748-754.
- Van der Klauw, W. (2001), An evaluation of the effectiveness of Title I compensatory education programs, University of North Carolina at Chapel Hill, in mimeo.
- Veen, A., J. Roeleveld en P. Leseman (2000), Evaluatie van Kaleidoscoop en Piramide, eindrapportage, SCO-Kohnstamm Instituut, Amsterdam
- Welch, F. (1973), Black-white differences in returns to schooling, *American Economic Review*, 63 (december); 893-907.

Bijlage 1 Het middelendebat

Kritiek op Hanushek

Hedges, Laine&Greenwald (1994) wijzen erop dat tussen 1970 en 1990 belangrijke verschuivingen hebben plaatsgevonden in de leerlingenpopulaties: een veel groter deel van de leerlingen is afkomstig uit achterstandsgroepen hetgeen een neerwaarts effect heeft op de gemiddelde testcores. Hun belangrijkste kritiek betreft de wijze waarop Hanushek de resultaten van 377 studies samenbrengt. Met de methode van Hanushek wordt als het ware elke studie geteld en gescoord naar een positief of negatief effect⁵⁸. Een heranalyse van deze studies met een formele statistische procedure leert dat wel degelijk sprake is van een betekenisvol positief verband tussen onderwijsmiddelen en leerlingprestaties.

Murmane&Levi (1996) wijzen op de resultaten van een natuurlijk experiment waarbij 15 scholen in Austin (Texas) substantieel meer middelen kregen als gevolg van een rechtszaak. Op twee van deze vijftien scholen werd een significante verbetering van de leerprestaties bereikt, de overige dertien scholen boekten geen voortgang. Dit resultaat kan op twee manieren worden uitgelegd. Allereerst dat additionele middelen wel degelijk leiden tot verbetering van resultaten op 2 scholen. Maar ook dat gemiddeld genomen de additionele middelen nauwelijks leiden tot vooruitgang van leerresultaten. Dit duidt erop dat vooral gekeken moet worden naar de wijze waarop de additionele middelen besteed worden: waarom zijn twee scholen wel succesvol?

Kritiek op Card&Krueger

In een replicatiestudie bleek dat studies waarin de onderwijsuitgaven gemeten werden als gemiddelde van een staat de resultaten van C&K bevestigen, bij desaggregatie van deze uitgavenmaat naar bijvoorbeeld het niveau van het schooldistrict of naar schoolniveau verdwijnt het positieve effect van uitgaven op latere lonen (Betts, 1995)⁵⁹. In de tweede plaats blijkt het positieve effect van uitgaven op de latere beloning vooral gevonden te worden in studies waarin leerlingen onderwijs volgden voor 1960. In studies waarin leerlingen afstuderen na 1960 wordt weinig bewijs gevonden voor het effect van onderwijsmiddelen op beloning. Een mogelijk verklaring voor dit verschil is dat na 1960 de variatie in beschikbare onderwijsmiddelen voor scholen aanzienlijk kleiner is geworden. Betts suggereert ook dat door specifieke veranderingen in de schoolorganisaties, zoals bureaucratisering of grotere invloed van vakbonden, het

⁵⁸ Verschillende auteurs noemen dit 'a vote counting exercise'.

⁵⁹ Deze 'aggregation bias' wordt mogelijk veroorzaakt doordat de geaggregeerde uitgavenmaat samenhangt met verschillen op het niveau van staten, zoals andere regels voor het curriculum, certificering van docenten, aanstellingen van docenten enz. (Hanushek, Rivkin en Taylor, 1996).

moeilijker is geworden om onderwijsmiddelen om te zetten in hogere opbrengsten voor leerlingen⁶⁰.

Heckman, Layne-Farrar & Todd (1996) laten zien dat het model van C&K verschillende belangrijke veronderstellingen bevat die de uitkomsten kunnen beïnvloeden. Als bijvoorbeeld de veronderstelling van een lineaire relatie tussen onderwijsmiddelen en beloning wordt losgelaten blijkt het effect van onderwijsmiddelen op lonen sterk af te nemen⁶¹. Zij concluderen dat het identificeren en toetsen van de veronderstellingen uit het model van C&K leidt tot een verzwakking van de conclusie dat onderwijsuitgaven een systematisch en meetbaar effect hebben op de latere lonen.

Naar een synthese van de empirische bevindingen?

Kan er een synthese worden bereikt uit deze schijnbaar tegengestelde posities? Allereerst de opvatting van de hoofdrolspelers.

C&K beweren dat bepaalde resultaten van het onderwijs, die belangrijk zijn voor de latere lonen, slecht gemeten worden met de standaard toetsen in het onderwijs. Met andere woorden, beter uitgeruste scholen slagen er wel degelijk in de 'verdienmogelijkheden' van leerlingen te verbeteren maar slagen er niet in om de gemeten leerprestaties te verbeteren. Het toevoegen van de precieze leerresultaten (bijvoorbeeld eindexamencijfers) aan een model waarin loonverschillen verklaard worden uit verschillen in het voltooide onderwijs levert nauwelijks een extra verklaring van de loonverschillen.

Hanushek suggereert dat de resultaten van C&K vertekend worden door dataproblemen. De kwaliteit van scholen is gecorreleerd met niet-waargenomen kenmerken van ouders die van invloed zijn op de latere lonen. In de analyse wordt daardoor ten onrechte een effect toegekend aan de kwaliteit van scholen.

De 'puzzle' kan volgens Burtless voor een deel worden opgelost door te kijken naar verschillen tussen perioden. De resultaten van Hanushek hebben vooral betrekking op het onderwijs vanaf 1960, het onderzoek van C&K richt zich op de arbeidsmarktprestaties van personen die het onderwijs gevolgd hebben voor 1960. Doordat in een meer recente periode de verschillen in beschikbare middelen voor scholen sterk zijn afgenomen is het mogelijk dat de opbrengsten van

⁶⁰ Olson&Ackerman (2000) wijzen erop dat studies waarbij onderwijsuitgaven gerelateerd worden aan relatief jonge werknemers (niet ouder dan begin dertig jaar) geen significant effect vinden. Studies over werknemers vanaf 35 jaar vinden wel significante effecten. Het is mogelijk dat werknemers die in een 'rijkere onderwijsomgeving' geleerd hebben meer kans hebben op het volgen van post-initiele trainingen en hiervoor betalen met een lager loon in het eerste deel van hun loopbaan.

⁶¹ Door rekening te houden met 'sheep-skin' effecten op beloning (het halen van een diploma levert een loonpremie).

additionele middelen lager is geworden. Met andere woorden, als arme scholen rijker worden neemt het effect van additionele middelen af.

Burtless (1996) hecht de meeste waarde aan de bevindingen van Hanushek. Het argument van de beperkte meting van de resultaten van het onderwijsproces acht hij twijfelachtig. In feite betekent dit argument namelijk dat scholen er wel in slagen doelen te bereiken die ze niet kunnen meten (arbeidsmarktprestaties) en dat ze er niet in slagen de doelen te bereiken die ze direct kunnen observeren (leerprestaties). Zijn conclusie van het hele debat in 1996 is dat additionele middelen voor scholen geen garantie zijn voor betere prestaties van scholen. Een verhoging van de onderwijsuitgaven zonder aanpassing in de schoolorganisatie geeft weinig hoop voor verbetering van prestaties van leerlingen. Recente ontwikkelingen wijzen echter in een andere richting (zie paragraaf 2.2).

Bijlage 2 Specifieke projecten rond onderwijsuitval

Ohio's LEAP-programma

Het LEAP-programma in Ohio biedt 'tiener-ouders' voor elke maand dat ze onderwijs volgen 62 dollar extra. Als ze daarentegen regelmatig verzuimen of het onderwijs voortijdig verlaten wordt hun uitkering met \$62 verlaagd (ruim 20% van de uitkering). Uit een evaluatie blijkt dat het programma alleen effect heeft op tieners die bij instroom in het programma nog op school zaten. Twee jaar na deelname aan het programma zijn geen effecten meer waarneembaar (Heckman, 2000).

Sponsor a scholar

Doel van dit programma uit Philadelphia (VS) is om leerlingen vanuit de high school te laten doorstromen naar vervolgonderwijs (colleges). De leerlingen krijgen mentoren, academische steun, hulp bij aanmelding voor het vervolgonderwijs en additionele financiële steun (\$ 6000). Voor de deelnemers aan het programma wordt een toename van de leerprestaties van de leerlingen en van de instroom in het vervolgonderwijs gevonden (Heckman, 2000).

Big brothers/ big sisters (VS)

Leerlingen uit een eenoudergezin krijgen op school een oudere mentor toegewezen. Uit de evaluatie blijkt dat de leerlingen relatief minder drugs- en alcohol gebruiken, minder verzuimen en betere schoolresultaten behalen (Heckman, 2000).

Quantum opportunity program

Dit programma lijkt sterk op Sponsor a scholar. Leerlingen uit minderheidsgroepen in achterstandswijken krijgen hulp en steun bij onderwijsactiviteiten van een mentor. Bovendien krijgen leerlingen een beloning voor elk uur besteed aan het verbeteren van sociale of andere vaardigheden relevant voor de arbeidsmarkt (1 dollar voor de leerling en 1 dollar wordt gestort in een studiefonds). De kosten bedragen \$ 10600 per participant (over 4 jaar). Dit programma heeft het aantal schoolverlaters met een eindexamen en het aantal jongeren dat een vervolgonderwijs afmaakt sterk verhoogd. Een kosten-baten analyse op basis van een gerandomiseerd experiment geeft aan dat er netto sociale opbrengsten zijn (Heckman, 2000). Van alle preventieprogramma's tegen criminaliteit, die door het Amerikaanse Rand Corporation zijn onderzocht, bracht dit project de grootste reductie van criminaliteit teweeg. Het aantal arrestaties in de doelgroep halveerde ten opzichte van de controlegroep.

Teenage parent demonstration

Dit programma is uitgevoerd in Illinois en Chicago, over de periode 1987 tot 1991. Het project richt zich op jonge moeders tussen 17 en 19 jaar. Deze moeders hebben een kind, krijgen een uitkering en gaan niet meer naar school. Ze zijn voornamelijk afkomstig uit etnische minderheden en zijn ongehuwd. Ongeveer eenderde van de deelnemers heeft een diploma. Deelname aan het programma is verplicht, ongeacht de leeftijd van het kind. De deelnemers moeten onderwijs of trainingen volgen of werkervaring opdoen. Het project is in twee fasen geëvalueerd. In de eerste twee jaar na de intake voor het programma gingen de leerlingen vaker naar school, deden meer aan training en hadden meer werk dan personen die niet aan het programma hadden meegedaan. Op de korte termijn lijkt het programma op de leerprestaties effect te hebben gehad. De effecten van het programma zijn ongeveer 6 tot 7 jaar na de intake opnieuw gemeten. Circa 70 procent van de moeders die deelgenomen hadden aan het programma hadden nog een uitkering. De invloed van het programma is 3 tot 4 jaar na invoering van het programma verdwenen (Kisker, Rangarajan en Boller, 1998).

GED: certificering van dropouts

Voor leerlingen die het onderwijs zonder diploma hebben verlaten bestaat de mogelijkheid om alsnog een certificaat te halen, de zogenoemde "GED" (General Educational Development). De GED komt neer op een examen met vijf onderdelen: wiskunde, schrijven, sociale wetenschappen, wetenschap en tekstinterpretatie. In 1989 was de mediane voorbereidingstijd voor een GED examen ongeveer 30,5 uur. Dit geeft aan dat de GED vooral beoogt het reeds opgebouwde menselijk kapitaal te meten (en te certificeren) en niet leidt tot additionele opbouw van menselijk kapitaal. Deelname aan het GED is voor schoolverlaters niet verplicht, maar het wordt door de overheid wel gestimuleerd. In de VS is het GED behoorlijk populair onder voortijdig schoolverlaters. In 1996 probeerden 758500 voortijdig schoolverlaters een GED te behalen, 524500 van hen kregen uiteindelijk een GED. Recent is de invloed van de GED op het loon van onderwijsuitvallers geanalyseerd (Tyler e.a, 2000). Een directe vergelijking van het salaris van onderwijsuitvallers met en zonder GED levert niet het causale effect van de GED. Er kan namelijk selectie optreden bij de deelname aan de GED. Om het causale effect van de GED vast te stellen wordt de variatie tussen staten ten aanzien van de eisen (standaarden) bij het GED-examen gebruikt als natuurlijk experiment (sommige deelnemers zouden bijvoorbeeld in de ene staat het examen wel halen en in de andere niet). Gevonden wordt dat het verwerven van de GED het salaris van blanke 'dropouts' verhoogt met 10 tot 19 procent. Voor niet-blanke dropouts vinden ze daarentegen geen toename van de salarissen. Het gevonden effect kan volledig toe geschreven worden aan de signaalfunctie van de GED omdat in de schatting rekening wordt gehouden met verschillen in 'human capital'.

Choice and changes (VS)

Het programma "Choices and Changes" beoogt achterstandsleerlingen meer bewust te maken van de kosten en baten verbonden aan het nemen van beslissingen. Met een cursus, die op verschillende leerniveaus wordt gegeven wordt getracht de beslissingen over het toekomstige arbeidsaanbod te verbeteren en een beter begrip tot stand te brengen over de mogelijkheden om in de toekomst productiever te zijn. Het programma is in de VS gedurende 6 jaar onder meer dan 100 duizend leerlingen (elementary en high schools) uitgevoerd. In de evaluatie is gewerkt met een experimenteel design en gekeken naar de toename in cognitieve vaardigheden en attitude veranderingen. De effecten zijn gemeten na 1 jaar. Op alle leerniveaus wordt een verbetering van cognitieve vaardigheden gevonden. Alleen voor de laagste klassen op de elementary school wordt er ook een verbetering van de attitudes gemeten.

School dropout demonstration assistance program

Tussen 1991 en 1996 zijn in verschillende delen van de Verenigde Staten in het kader van het School Dropout Demonstration Assistance Program (SDDAP) diverse projecten uitgevoerd gericht op preventie van onderwijsuitval. Dit betrof ruim 20 projecten waaraan meer dan 10.000 studenten deelnamen. Twee typen projecten werden onderscheiden: (1) 'gerichte' projecten voor 'students at risk', (2) organisatorische herstructurering van scholen. De 'gerichte' projecten vonden plaats tijdens de 'middle school', de 'high school' en sommige na het vertrek uit het onderwijs (de zogenoemde GED-projecten die hiervoor al aan de orde kwamen). De herstructureringsprojecten vonden eveneens plaats op verschillende momenten van de onderwijsloopbaan. Deze projecten zijn geëvalueerd door Mathematica Policy Research (Dynarski & Gleason, 1998). De 'gerichte' projecten kenden een experimenteel design, bij de herstructureringsprojecten is dit niet mogelijk. De algemene conclusie van de totale evaluatie is dat de meeste programma's niet of nauwelijks bijdragen aan het verminderen van onderwijsuitval. Voor enkele programma's werden daarentegen wel positieve resultaten gevonden, en dit geldt met name voor GED-programma's. Het effect van herstructureringsprogramma's blijkt sterk af te hangen van de 'commitment' van de scholen. De onderzoekers concluderen dat de doelstelling om het aandeel 'high school' gediplomeerden te verhogen van 86 naar 90 procent niet te realiseren is met beperkte kosten: 'given what we know about impacts of dropout prevention programs and their costs, it does not look like we can get there from here'. Als oorzaak voor de geringe impact van uitvalprogramma's wordt genoemd de grote variatie aan factoren, zowel binnen als buitenschools, die van invloed kunnen zijn op uitval. Het verkleinen van groepen of scholen of het aanbieden van ondersteunende diensten zijn weinig effectief als daarmee niet wordt ingespeeld op de werkelijke oorzaken van de onderwijsuitval.