

CPB Document

No 49

Januari, 2004

Prestatieprikkel in het Nederlandse onderwijs

Wat kunnen we leren van recente buitenlandse ervaringen?

Erik Canton, Dinand Webbink

Centraal Planbureau
Van Stolkweg 14
Postbus 80510
2508 GM Den Haag

Telefoon (070) 338 33 80
Telefax (070) 338 33 50
Internet www.cpb.nl

ISBN 90-5833-157-1

Inhoud

Ten geleide	5
Samenvatting	7
1 Inleiding	9
2 Prestatieprikkel in het onderwijs: Achtergrond	13
2.1 Prestaties en bekostiging	14
2.2 Prestaties en accountability	15
2.3 Prestaties en beloning	16
3 Beoordeling van onderwijsprestaties: Meten is niet altijd weten	21
4 Internationale ervaringen met prestatieprikkel in het onderwijs	25
4.1 Prestaties en bekostiging	25
4.2 Prestaties en accountability	28
4.3 Prestaties en beloning	30
4.4 Samenvatting	36
5 Prestatieprikkel in het Nederlandse onderwijs: Opties voor beleid	37
5.1 Nederland versus buitenlandse ervaringen	37
5.2 Prestaties en bekostiging	39
5.3 Prestaties en accountability	40
5.4 Prestaties en beloning	41
Bijlage: Andere initiatieven met prestatieprikkel in het Verenigd Koninkrijk	43
Literatuur	45
Abstract	48

Ten geleide

Deze studie richt zich op prestatieprikkels in het Nederlandse onderwijs. Een aantal landen experimenteert met prikkels voor onderwijsinstellingen en docenten. In Nederland bestaat terughoudendheid bij het gebruik van prikkels in het onderwijs. Is dit terecht? Wat kunnen we leren van buitenlandse ervaringen?

De auteurs danken de leden van de begeleidingscommissie voor hun constructieve inbreng. Deze commissie bestond uit Erica van Alfen (Algemene Onderwijsbond), Paul de Bijl en Paul Reuter (Ministerie van Financiën), Herman Scholten (Ministerie van Onderwijs, Cultuur en Wetenschappen), Simone Doolaard (Universiteit Twente), Trudy Moerkamp (Sectorbestuur Onderwijsarbeidsmarkt), en Ib Waterreus (SCHOLAR, Universiteit van Amsterdam). Ook gaat dank uit naar diverse collega's van het CPB, met name Wout de Bruin en Pierre Koning, voor nuttig commentaar.

Gesprekken met deelnemers aan een conferentie over prestatiebeloning in het onderwijs (Chicago, 21-22 november 2002) waren nuttig bij de beschrijving van de ervaringen in de Verenigde Staten. De paragraaf over prestatiebeloning in het Verenigd Koninkrijk is mede gebaseerd op gesprekken met specialisten (gevoerd in Londen, 3-4 oktober 2002). Speciale dank gaat uit naar Stephen Grounds (Office for Standards in Education, OFSTED), Andrew Morris en David Powell (National Union of Teachers, NUT), en Jason Tarsh (Department for Education and Skills, DfES).

F.J.H. Don, directeur

Samenvatting

In dit rapport staan de prestaties van het Nederlandse onderwijs centraal. Wat zijn de knelpunten, en hoe kunnen deze worden aangepakt? Prestatieprikkelers kunnen een belangrijk instrument zijn om doelen te bereiken. Verschillende landen experimenteren de laatste jaren met prestatieprikkelers voor onderwijsinstellingen en docenten. Nederland is op dit punt terughoudend. Is dit terecht? Wat kunnen we leren van buitenlandse ervaringen?

Het doel van de studie is om na te gaan in hoeverre het gebruik van prestatieprikkelers de resultaten van het onderwijs kan verbeteren. Daarvoor wordt allereerst bekeken welke prikkels er bestaan in het Nederlandse onderwijs. We maken daarbij onderscheid tussen prikkels verbonden aan de bekostiging van instellingen, aan het systeem van publieke verantwoording, en aan de beloning van docenten.

De overheid kan instellingen direct prikkelen via de bekostiging, bijvoorbeeld door de hoogte van de bekostiging afhankelijk te maken van het aantal afgestudeerden. *Accountability*, publieke verantwoording van onderwijsinstellingen, impliceert een meer indirecte prikkel, voortkomend uit het keuzegedrag van leerlingen en ouders ('stemmen met de voeten'). Tenslotte kunnen prikkels worden ingebouwd in de beloning van docenten, bijvoorbeeld door middel van individuele prestatiebeloning.

Het toepassen van prestatieprikkelers vereist meting van prestaties. In het onderwijs is het vaststellen van de prestaties van instellingen en docenten niet eenvoudig, omdat leerlingen niet alleen op school leren maar ook daarbuiten. Het 'productieproces' in het onderwijs wijkt daarmee af van dat in de meeste economische sectoren. Een geschikte prestatie-indicator meet de bijdrage van de school / docent aan de leerprestaties van studenten, en is moeilijk te manipuleren.

Het Verenigd Koninkrijk, de Verenigde Staten en Israël experimenteren de laatste jaren met verschillende beloningsvormen voor docenten en scholen. Deze ervaringen hebben laten zien dat prikkels werken, maar ook dat *you get what you pay for*. Verder blijkt dat manipulatie van de prestatie-indicatoren op de loer ligt. Wat kunnen we hiervan leren, uitgaande van de Nederlandse situatie en de mogelijkheden van het meten van prestaties van instellingen en docenten?

Een intensiever gebruik van prestatiebekostiging in Nederland lijkt zinvol, maar niet zonder risico's. De prioriteit kan worden gelegd bij het voorkomen van voortijdig schoolverlaten in de BVE-sector (beroepsonderwijs en volwasseneneducatie). Dit wordt immers door velen gezien als het grootste knelpunt in het beroepsonderwijs. De huidige bekostiging voor het BVE is deels

al afhankelijk van het aantal geslaagden. Het bonussysteem zou dan kunnen aangrijpen op het percentage uitvallers vergeleken met dat bij andere instellingen, rekeninghoudend met verschillen in achtergrondkenmerken van leerlingen.

Ranglijsten op basis van prestaties van scholen kunnen een krachtige prikkel geven voor verbetering van de kwaliteit van het onderwijs. Recente studies laten echter zien dat ranglijsten onbetrouwbaar zijn, zelfs als zij geconstrueerd zijn op basis van 'state of the art' technieken. *Accountability* behelst meer dan ranglijsten. Het publiek informeren over de resultaten op scholen, zoals dat bijvoorbeeld gebeurt via de kwaliteitskaarten, geeft een prikkel aan scholen om zich in te spannen om de kwaliteit van het onderwijs te verbeteren. In de huidige Nederlandse context is er nog de nodige winst te halen bij de informatievoorziening over de resultaten op scholen en onderwijsinstellingen. Voor instellingen in de BVE-sector en het hoger onderwijs bestaan (nog) geen kwaliteitskaarten. Het ontwikkelen van een kwaliteitskaart of vergelijkbaar instrument voor deze sectoren van het onderwijs kan de *accountability*-prikkel versterken.

Over de effecten van individuele prestatiebeloning in het onderwijs is tot dusver weinig bekend. Het experiment dat in Israël is uitgevoerd laat overtuigend zien dat de prestaties van leerlingen verbeteren. Dit suggereert dat toepassing van individuele prestatiebeloning in Nederland veelbelovend kan zijn. Binnen de huidige beleidscontext, die gericht is op het versterken van de autonomie van instellingen, is het echter aan het management van scholen om te beslissen over het toepassen van individuele prestatiebeloning.

1 Inleiding

De laatste jaren neemt de belangstelling voor de prestaties van de overheid sterk toe. Dit komt onder meer tot uitdrukking in de invoering van de beleidsbegroting nieuwe stijl: Van Beleidsbegroting tot Beleidsverantwoording (VBTB). Bij deze nieuwe stijl van uitvoering van publieke taken worden de overheid en (semi-)publieke organisaties beoordeeld op de uitkomsten van hun beleid. Dit geldt ook voor het onderwijs. In de Rijksbegroting 2004 zijn concrete streefwaarden geformuleerd op vele beleidsterreinen. Zo wordt bijvoorbeeld gestreefd naar een reductie van de taalachterstand van doelgroepopleerlingen in 2006 met 25% ten opzichte van de achterstand in 2001-2002. Prestatieprikkels kunnen een belangrijk instrument zijn om doelen te bereiken. Verschillende landen experimenteren de laatste jaren met prestatieprikkels voor onderwijsinstellingen en docenten. Nederland is op dit punt terughoudend. Is dit terecht of laten we hierdoor kansen liggen?

Deze studie analyseert het gebruik van prestatieprikkels in het onderwijs. Het doel van de studie is om na te gaan in hoeverre het gebruik van prestatieprikkels de resultaten van het Nederlandse onderwijs kan verbeteren. Daarvoor wordt allereerst bekeken welke prikkels er bestaan in het Nederlandse onderwijs. Dit kunnen financiële prikkels zijn via de bekostiging van instellingen, maar ook meer indirecte prikkels die voortkomen uit het keuzegedrag van leerlingen en ouders ('stemmen met de voeten'). Vervolgens wordt ingegaan op het meten van prestaties van instellingen en docenten. Het toepassen van prestatieprikkels vereist immers meting van prestaties. In het onderwijs is het vaststellen van de prestaties van instellingen en docenten niet eenvoudig omdat leerlingen niet alleen op school leren maar ook daarbuiten. Het 'productieproces' in het onderwijs wijkt daarmee af van dat in de meeste economische sectoren. De recente buitenlandse ervaringen met prestatieprikkels in het onderwijs komen daarna aan bod. Engeland, de Verenigde Staten en Israël experimenteren de laatste jaren met verschillende beloningsvormen voor docenten en voor scholen. Wat kunnen we leren van deze ervaringen, uitgaande van de Nederlandse situatie en de mogelijkheden van het meten van prestaties van instellingen en docenten?

Prestatieprikkels kunnen verschillende vormen aannemen. Bij sommige vormen, zoals de bekostiging van instellingen, heeft de overheid directe invloed. Bij andere vormen, zoals individuele prestatiebeloning van docenten, is dat niet het geval. In de huidige beleidscontext is de beloning van individuele docenten namelijk een zaak van het management van instellingen. In deze studie worden prestatieprikkels in het onderwijs ingedeeld in drie typen. De indeling is gebaseerd op de organisatie van het Nederlandse onderwijs (zie paragraaf 2) en, in het verlengde daarvan, de mate waarin de overheid invloed heeft op de prikkels. Er wordt onderscheid gemaakt tussen:

- prestaties en bekostiging;
- prestaties en *accountability*;
- prestaties en beloning.

De overheid kan instellingen direct prikkelen via de bekostiging, bijvoorbeeld door de hoogte van de bekostiging afhankelijk te maken van het aantal afgestudeerden. Deze prestatieprikkel wordt momenteel toegepast in het hoger onderwijs. Instellingen ontvangen meer middelen als meer studenten afstuderen.

Bij prestaties en *accountability* is sprake van een indirecte prikkel. In Nederland zijn leerlingen en ouders vrij om een school te kiezen die aansluit bij hun voorkeuren. Scholen worden bekostigd op basis van het aantal leerlingen. Dit geeft scholen een prikkel om goede prestaties te leveren omdat deze prestaties waarschijnlijk een rol spelen bij de schoolkeuze van leerlingen en ouders. Scholen met goede prestaties zullen gemakkelijker leerlingen aantrekken en scholen met zwakke prestaties kunnen leerlingen verliezen. De overheid kan deze indirecte prikkel versterken door de bekostiging sterker afhankelijk te maken van het aantal leerlingen. Een andere mogelijkheid is het vergroten van de transparantie van de onderwijsmarkt. Dit kan door meer inzicht te bieden in de prestaties van scholen, hetgeen momenteel beoogd wordt met het systeem van kwaliteitskaarten in het voortgezet onderwijs. Essentieel is dat de prikkel voorkomt uit het gedrag van leerlingen en ouders.

Het derde type prestatieprikkel in deze studie is prestatiebeloning. Bij prestatiebeloning is de beloning van de docent of directeur (voor een deel) direct afhankelijk van de prestaties. Het is bijvoorbeeld mogelijk om individuele docenten een toeslag te geven als een relatief groot deel van hun leerlingen het eindexamen haalt. Als de prikkel wordt gegeven aan een team van docenten spreken we over teambeloning. Als de prikkel aan de hele school wordt gegeven lijkt teambeloning sterk op de reguliere bekostiging. De beloning van individuele docenten of teams van docenten binnen een school is momenteel een zaak van het management. De overheid heeft hierop geen directe invloed.

Het Nederlandse onderwijs kent verschillende sectoren die sterk kunnen verschillen qua organisatie en gebruik van prestatieprikkels. Deze studie is niet in de eerste plaats gericht op een analyse van de verschillen tussen onderwijssectoren, maar op het in kaart brengen van algemene principes over prestatieprikkels in het onderwijs. Deze inzichten kunnen voortkomen uit de economische theorie en de empirie over recente internationale ervaringen. Op basis van de algemene inzichten wordt bekeken in welke sectoren het toepassen van prestatieprikkels het meest kansrijk is.

De voorliggende studie is onderdeel van een breder CPB-project naar prestatieprikkels in de semi-publieke sector (zie Koning e.a., 2004). Daarvoor wordt gekeken naar prestatieprikkels in zes sectoren: sociale zekerheid, politie, TNO, onderwijs, wetenschappelijk onderzoek en medisch specialisten. Met de brede aanpak wordt beoogd sectoroverstijgende lessen te kunnen trekken.

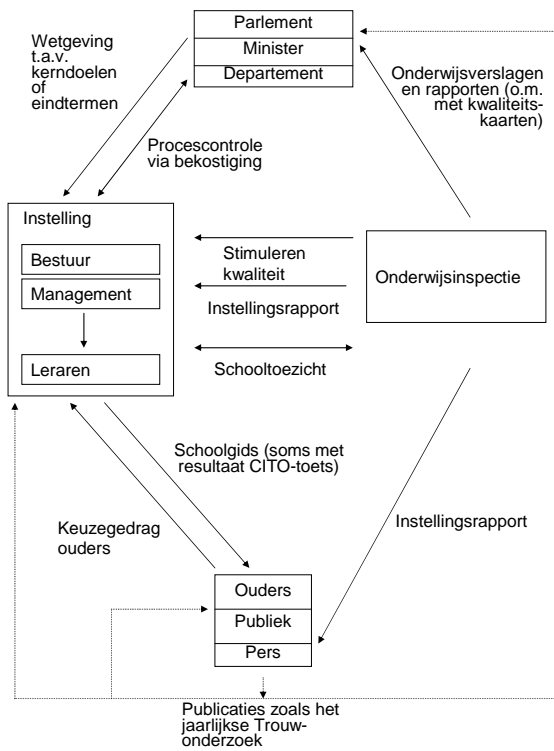
De opbouw van deze studie is als volgt. In paragraaf 2 introduceren we het analytische kader. Wat zijn de voor- en nadelen van prestatieprikkels (prestatiebeloning, *accountability* en prestatiebekostiging), en wat zijn mogelijke manieren om prestatieprikkels in te zetten? De organisatie van het Nederlandse onderwijs komt ook aan bod in deze paragraaf. We zullen hierbij eveneens stilstaan bij de recente initiatieven in de richting van competentiebeloning. In paragraaf 3 gaan we uitgebreid in op het meten van onderwijsprestaties en de problemen daarbij. Paragraaf 4 bespreekt recente internationale ervaringen met prestatieprikkels. Tenslotte presenteren we beleidsopties voor Nederland in paragraaf 5.

2 Prestatieprikkel in het onderwijs: Achtergrond

Waarom kunnen prikkels leiden tot betere prestaties? Bij economische analyses van prikkels wordt veelal uitgegaan van het principaal-agent model. De principaal is de opdrachtgever en de agent is de uitvoerder. De kernvraag in het model is: *hoe kan de principaal de agent ertoe bewegen de door de principaal gewenste taak uit te voeren, tegen minimale kosten voor de principaal?* De principaal kan het natuurlijk vragen aan de agent. Maar het is niet zeker of dit tot de gewenste uitkomst leidt omdat de agent mogelijk zijn eigenbelang nastreeft. In de praktijk beschikt de uitvoerder (de agent) vaak over meer informatie dan de opdrachtgever (de principaal). De uitvoerder kan dit informatievoordeel uitbuiten ten faveure van zichzelf en ten koste van de principaal. Als bijvoorbeeld een manager geen zicht heeft op de inspanningen van de werknemers, dan kunnen deze werknemers besluiten het rustig aan te doen. Door gebruik te maken van prestatieprikkel kan deze belangentegenstelling worden opgelost. Zo kan de beloning afhankelijk worden gesteld van de geleverde prestaties. Denk hierbij aan prestatiebekostiging voor instellingen en prestatiebeloning voor docenten en directeurs van instellingen. Dit vereist dan wel dat prestaties goed meetbaar zijn. Naast prestatiebeloning zijn er ook andere instrumenten beschikbaar om prestaties van werknemers te verbeteren. Een mogelijkheid is om lonen te bieden boven het niveau van het marktloon (*efficiency wages*). Dit kan werknemers aanzetten tot een grotere inspanning. Geringe inspanning kan namelijk tot ontslag leiden en in dat geval gaat de werknemer er echt op achteruit omdat het marktloon lager ligt dan het loon in de huidige baan. Impliciete prikkels kunnen ook leiden tot een grote inspanning. De literatuur over impliciete prikkels benadrukt dat werknemers niet alleen kijken naar expliciete compensatie, maar ook bijvoorbeeld naar promotiekansen en status. Uitgebreide overzichten zijn te vinden in Prendergast (1999) en Burgess en Metcalfe (1999). Een algemeen overzicht van prestatieprikkel in de semi-publieke sector is te vinden in Dixit (2002) en Koning e.a. (2004).

Figuur 2.1 geeft een schematisch overzicht van de organisatie van het Nederlandse onderwijs. Vier verschillende actoren staan centraal: de instellingen, de Onderwijsinspectie, de leerlingen / studenten en hun ouders, en het openbaar bestuur. Onderwijsinstellingen hebben te maken met twee belangrijke principalen: de overheid en de leerlingen c.q. studenten en hun ouders. De overheid bekostigt de instellingen, en ziet toe op de kwaliteit. Ook vertegenwoordigt het parlement en de regering het bredere maatschappelijk belang van goed onderwijs (de uiteindelijke principaal van het onderwijs is de maatschappij). Leerlingen en studenten zijn de 'klanten' van de onderwijsinstellingen, en zijn direct gebaat bij kwalitatief goed onderwijs. In paragraaf 2.1 beschouwen we de relatie tussen de onderwijsinstelling als agent en de overheid als principaal, en in 2.2 kijken we naar de relatie tussen scholen en het publiek. Tenslotte onderzoeken we in paragraaf 2.3 de relatie tussen directie en docenten binnen de onderwijsinstelling.

Figuur 2.1 Organisatie van het Nederlandse onderwijs



2.1 Prestaties en bekostiging

De relatie tussen overheid en instellingen kan worden beschreven aan de hand van het principaal-agent model. De overheid (de principaal) delegeert onderwijstaken naar de instellingen (de agenten). Hierbij dient de overheid de kwaliteit van het onderwijs te waarborgen. Hoe worden schoolprestaties beoordeeld? Scholen worden bezocht door de Onderwijsinspectie, die beoordeelt hoe de kerndoelen in de leerstof zijn verwerkt en het instructieproces evalueert. In de BVE-sector worden eindtermen opgesteld. In het hoger onderwijs daarentegen wordt de kwaliteit door de sector zelf vastgesteld door middel van visitatiecommissies.

De overheid kan op een aantal manieren onderwijsprestaties van scholen laten meewegen bij de bekostiging. Een gebruikelijke manier is het stellen van minimumeisen aan de kwaliteit van het onderwijs. Wanneer instellingen onder de maat presteren dan kan de Onderwijsinspectie aanwijzingen geven en de overheid ingrijpen. Hierbij kan men denken aan intensievere controle

en het bieden van hulp (financiële of personele steun, advies), vervanging van het management, en in het uiterste geval sluiting van de school.

Ook kan de principaal gebruik maken van prestatiebekostiging. In een aantal onderwijssectoren is de bekostiging afhankelijk van prestaties. Zo hangt de bekostiging van instellingen in het hoger beroepsonderwijs (HBO) en wetenschappelijk onderwijs (WO) voor een groot deel af van het aantal afgestudeerden (zie CPB en CHEPS (2001) voor een beschrijving van de bekostigingsmodellen). In het BVE hangt ruwweg 20 procent van de bekostiging af van het aantal gediplomeerden.¹

Bij het in kaart brengen van de prestaties van scholen dient rekening gehouden te worden met de samenstelling van de leerlingpopulatie. Zo blijken kinderen van welgestelde en hoog opgeleide ouders gemiddeld genomen beter te presteren op school dan kinderen uit probleemwijken. Vanuit het perspectief van de overheid dient het functioneren van scholen dan ook te worden afgemeten aan de bijdrage van scholen aan de ontwikkeling van leerlingen: wat voegt de onderwijsinstelling toe aan de vaardigheden van leerlingen, gegeven hun beginniveau?² We gaan in paragraaf 3 verder in op het meten van onderwijsprestaties.

2.2 Prestaties en accountability

De relatie tussen onderwijsinstellingen en publiek staat centraal in de discussie over wat in het Engels *accountability* wordt genoemd. Een Nederlandse vertaling die wellicht het beste de lading dekt is publieke verantwoording. Hierbij worden schoolprestaties openbaar gemaakt, zodat de instelling zich moet verantwoorden richting overheid, publiek en media. De voordelen van *accountability* hangen dus met name samen met het openbare karakter. Dit werkt als een extra prikkel voor scholen om goed te functioneren, omdat ouders dan beter zijn geïnformeerd over het reilen en zeilen van de school: slechtere prestaties kunnen worden afgestraft door dalende leerlingenaantallen.

Accountability is in Nederland op gang gebracht door de media. Dagblad Trouw heeft, zich beroepend op de Wet Openbaarheid Bestuur, toegang gekregen tot informatie over het functioneren van scholen van de Onderwijsinspectie. Trouw gebruikt deze informatie om ranglijsten te construeren. De jaarlijkse editie waarin de nieuwe resultaten worden gepubliceerd vindt gretig aftrek. Ook weekblad Elsevier publiceert met regelmaat over de “beste scholen” en de “beste universiteiten en hogescholen”. Als reactie hierop is de Onderwijsinspectie begonnen met het samenstellen van kwaliteitskaarten van scholen in het primair onderwijs (PO) en

¹ Een uitgebreide beschrijving van de diverse financieringsstromen is te vinden in Eurydice (2002).

² Ouders daarentegen lijken meer geïnteresseerd in het niveau van de school. Zij gaan er kennelijk vanuit dat hun kinderen voordeel hebben van de omgang met goede leerlingen.

voortgezet onderwijs (VO) , met onder meer als doel om de ouders te helpen bij de schoolkeuze.

Naast voordelen kunnen ook negatieve effecten optreden. Zo kan *accountability* leiden tot ‘samenspanning’ binnen de scholen: docenten en schooldirectie proberen de cijfers rooskleuriger voor te stellen door bijvoorbeeld zwakkere leerlingen buiten de officiële beoordelingen te houden.³ Aanwijzingen hiervoor zijn gevonden in de Verenigde Staten (we komen hierop terug bij de bespreking van internationale ervaringen). Om soortgelijke redenen worden op sommige Nederlandse scholen leerlingen afgeraden om een zwaar β -profiel te kiezen.

2.3 Prestaties en beloning

Bij de instellingen is onderscheid gemaakt tussen bestuur en management enerzijds en leraren anderzijds. Ook hier is sprake van een principaal-agent relatie. Het management (de principaal) delegeert beslissingsbevoegdheden omtrent de inrichting van het onderwijsproces grotendeels naar de leraren (de agenten). Hoewel verwacht kan worden dat de doelstellingen van leraren zullen overlappen met die van het management, is het evenwel mogelijk dat leraren daarnaast eigen doelstellingen nastreven die niet in het maatschappelijk belang hoeven te zijn.

De beloningsstructuur van leraren ligt grotendeels vast. In de huidige beloningsstructuur van het Nederlandse onderwijs zijn de salarissen van leraren onderverdeeld in vier schalen, van LA tot LD. In Tabel 2.1 staan de minimum- en maximumsalarissen vermeld (met uitzondering van speciale regelingen, eenmalige toeslagen, etc.).⁴

Schooltype	Schaal	Beginsalaris	Eindsalaris
Basisonderwijs	LA	2073 euro	2991 euro
VO 2 ^e graads	LB	2150 euro	3285 euro
VO 1 ^e graads	LD	2171 euro	4361 euro
BVE	LB / LC	2150 euro	3834 euro

Noot: VO staat voor voortgezet onderwijs; BVE staat voor beroepsonderwijs en volwasseneneducatie.
Bron: www.leraar.nl.

³ Tegenvoorbeelden zijn ook bekend. Heckman e.a. (1996) vinden dat medewerkers van reïntegratiekantoren voorrang geven aan moeilijk plaatsbare mensen, terwijl het systeem van prestatiebekostiging de kantoren juist prikkelt om zich te richten op makkelijk plaatsbare mensen. De gevreesde *cream-skimming* is achterwege gebleven omdat medewerkers ook worden gedreven door de behoefte om de zwakkere groepen te helpen.

⁴ Niet alle eerstegraads docenten worden ingeschaald op LD-niveau, zie Hassink (2000).

Normaal gesproken wordt het salaris ieder jaar met één periodiek verhoogd (in het VO duurt het dan 18 jaar totdat de eindschaal is bereikt). Als een leraar uitstekend functioneert, bij bijzondere prestaties of bij moeilijk te vervullen vacatures bestaat de mogelijkheid om leraren extra te belonen, bijvoorbeeld met een eenmalig bedrag of een maandelijkse toelage.

Er zijn momenteel enkele veranderingen gaande in de beloningsstructuur van leraren. Mede naar aanleiding van een onderzoek van Berenschot (Berenschot, 1998), wordt momenteel gewerkt aan invoering van zogenoemde competentiebeloning in het onderwijs. Het uitgangspunt bij het ontwerp van een competentiemodel is dat leraren een loopbaan moeten kunnen doorlopen van juniorleraar tot ervaren leraar of zelfs excellente leraar. Daartoe heeft Berenschot een lijst van competenties opgesteld. Aan de hand van een persoonlijk ontwikkelingsplan (POP) werken docenten aan de ontwikkeling van hun competenties. Tijdens het functioneringsgesprek komt de ontwikkeling van de leraar en de realisatie van het persoonlijk ontwikkelingsplan aan de orde. Bij het beoordelingsgesprek wordt vastgesteld in hoeverre de in het functioneringsgesprek gemaakte afspraken zijn nagekomen, en of de gestelde doelen zijn bereikt. De salarisgroei is afhankelijk van deze beoordeling. De verantwoordelijkheid voor implementatie van competentiebeloning ligt bij de onderwijsinstellingen. Bijna een kwart van de scholen in het primair onderwijs werkt met competentieprofielen, en bij 21% van de scholen wordt gewerkt met persoonlijk ontwikkelingsplannen. Zo'n 37% van de scholen in het voortgezet onderwijs geeft aan met vormen van beloningsdifferentiatie te werken (zie Inspectie van het Onderwijs, 2003). Competentiebeloning is vooral inputgericht (op inspanningen van de docent), en het is niet duidelijk in hoeverre competentiebeloning kan bijdragen aan betere prestaties.

Een andere verandering betreft de ongelijkheid in beloning van docenten die samenhangt met de werving van zij-instromers in verband met de huidige schaarste aan leraren. Het salaris van zij-instromers hangt af het salaris in de vorige functie en kan daardoor afwijken van dat van andere docenten met vergelijkbare taken.

De principaal-agent relatie binnen de school tussen management en docenten kan aanleiding zijn om prestatiebeloning te overwegen. Prestatiebeloning kan tot een beter resultaat leiden wanneer het management goed zicht heeft op de geleverde prestaties van werknemers. Herziening van het arbeidscontract en introductie van prestatiebeloning zorgt ervoor dat managers en werknemers dezelfde doelstellingen nastreven. Met prestatiebeloning bedoelen we een beloningsstructuur waarbij het loon op één of andere manier is gekoppeld aan prestaties. Het kan dan gaan om prestaties van de medewerker, of van een afdeling, of van grotere eenheden, en zelfs van de prestaties van de gehele organisatie.

Een inmiddels klassiek te noemen voorbeeld betreft invoering van prestatiebeloning in een Amerikaans bedrijf voor het installeren van autoruiten (Lazear, 1996). Na invoering van prestatiebeloning steeg de productiviteit met 36%. De helft van de winst kon worden toegerekend aan grotere inspanningen van werknemers. De andere helft werd veroorzaakt door een verandering van de samenstelling van het personeel. Het gebruik van prestatiebeloning bleek in deze studie meer productieve werknemers aan te trekken.

Prestatiebeloning in het onderwijs is een controversieel onderwerp. Problemen rond het meten van de 'output' van scholen of individuele docenten maken het gebruik van prestatiebeloning veel lastiger dan in sommige andere economische sectoren. Bovendien is prestatiebeloning in het algemeen niet populair bij betrokkenen in het veld. Met name vakbonden zijn vaak fel gekant tegen het gebruik ervan (zie bijvoorbeeld Ballou, 2001).

De vraag of prestatiebeloning een waardevol instrument kan zijn voor het onderwijs lijkt evenwel geenszins van tafel. Zo zijn de laatste jaren in diverse landen nieuwe beleidsinitiatieven met prestatiebeloning gestart. Hierbij zijn met name de ervaringen in Israël, Engeland en de Verenigde Staten interessant, omdat voor deze landen ook evaluatiestudies beschikbaar zijn. Door de experimentele opzet overtuigen vooral twee Israëlische studies, waarin de effecten van teambeloning (Lavy, 2002a) en individuele prestatiebeloning (Lavy, 2002b) worden onderzocht (deze studies komen uitgebreid aan de orde in hoofdstuk 4). De resultaten van beide studies suggereren dat prestatiebeloning in Israël heeft geleid tot betere leerprestaties en minder uitval. Met name de gunstige resultaten voor zwakkere leerlingen zijn een reden om nader te kijken naar deze beleidsinterventies. Ongeveer een kwart van de Nederlandse leerlingen verlaat immers het onderwijs zonder een startkwalificatie te behalen.

Er zijn diverse manieren om prestatiebeloning in te voeren. Een gangbaar onderscheid is:

- individuele prestatiebeloning;
- teambeloning.

Verder kan onderscheid worden gemaakt in prestatiebeloning op basis van absolute en relatieve prestatie maatstaven, en objectieve en subjectieve indicatoren. Bij individuele absolute prestatiebeloning is de variabele beloningscomponent gekoppeld aan de prestaties van de individuele werknemer. Een voorbeeld is een bonus wanneer een verkoper een bepaalde omzet heeft bereikt. Met individuele relatieve prestatiebeloning bedoelen we een beloningsstructuur waarbij het variabele beloningsdeel afhangt van de relatieve prestaties van de werknemer, dus de prestatie in vergelijking met die van zijn of haar collega's. Hierbij kan bijvoorbeeld gedacht worden aan een bonus voor de kwart best presterende werknemers. Bij teambeloning is de

variabele beloning afhankelijk van de prestaties van een groep werknemers (waarbij ook weer onderscheid kan worden gemaakt naar absolute en relatieve prestaties van het team). De omvang van het team kan variëren van een kleine groep werknemers die nauw verwante werkzaamheden verricht tot, in het uiterste geval, alle werknemers binnen een organisatie. Deze laatste vorm wordt overigens veelvuldig toegepast in het bedrijfsleven, denk bijvoorbeeld aan tantièmeregelingen.

Het gebruik van prestatiebeloning is echter niet zonder gevaren. Ten eerste kunnen er problemen optreden met het specifieke ontwerp van het prestatiebeloningssysteem. Zo heeft absolute individuele prestatiebeloning het karakter van een open einde regeling. Wanneer iedereen betere prestaties levert dan zullen de totale salarisuitgaven stijgen.⁵ Om de kosten beter beheersbaar te maken kunnen werkgevers het systeem van prestatiebeloning baseren op relatieve prestaties. Hiermee ligt het percentage van de werknemers dat in aanmerking komt voor een bonus vast. Individuele relatieve prestatiebeloning kan er toe leiden dat collega's minder goed samenwerken, of elkaar zelfs gaan tegenwerken (roddelen bij de baas, informatie achterhouden). Het gaat dan immers om de eigen prestatie ten opzichte van die van anderen, oftewel een toernooi (zie Lazear en Rosen, 1981). Teambeloning kan een antwoord zijn op zulk non-coöperatief gedrag. Daar staat echter tegenover dat teambeloning leidt tot het zogenaamde liftersgedrag. Dit liftersgedrag verzwakt de prikkels om goede prestaties te leveren. Hoe groter het team, des te sterker is de verleiding om mee te liften met de inspanningen van collega's (vgl. Lazear, 1999).⁶

Op de tweede plaats kan prestatiebeloning nadelig uitpakken wanneer werknemers intrinsiek gemotiveerd zijn. Uit experimenteel onderzoek (zie o.a. Kreps (1997), en Frey en Jegen (2001)) blijkt dat extrinsieke (financiële) prikkels ten koste kunnen gaan van intrinsieke motivatie, zelfs van die activiteiten waarvoor de opdrachtnemer beloond wordt. Het is niet verrassend dat gerichte financiële prikkels kunnen leiden tot veronachtzaming van activiteiten waar geen (extra) beloning op rust. Moeilijker verklaarbaar is dat (per saldo) ook minder inspanning geleverd zou worden voor de activiteiten waar prestatiebeloning wordt toegepast. De reden hiervoor is dat, op zijn minst in de perceptie van de opdrachtnemer, door prestatiebeloning het karakter van de te verrichten activiteit verandert. De opdrachtnemer vat prestatiebeloning dan op als een inbreuk op zijn beleidsdiscretie, zijn mogelijkheden om naar eigen inzicht tot oplossingen te komen worden beperkt. Bovendien kan hij – gezien de beloning die er op staat – tot het inzicht komen dat zijn taak veel zwaarder is dan hij dacht.

⁵ Het is dan de vraag of de verbeterde prestaties opwegen tegen de extra salarisuitgaven.

⁶ Zulk liftersgedrag kan worden tegengegaan wanneer collega's elkaar controleren en corrigeren. Teambeloning kan dan toch gunstige effecten sorteren (zie Knez en Simester, 2001).

Op de derde plaats kunnen doelstellingen onderling strijdig zijn. Zo is de doelstelling van kwaliteit mogelijk strijdig met die van deelname: het is waarschijnlijk dat de gemiddelde leerprestaties zullen dalen als de zwakkere leerlingen het onderwijs niet verlaten. Het is in dat geval niet wenselijk om sterke prikkels voor beide doelstellingen te hanteren.

Tenslotte ontstaan moeilijkheden wanneer output niet goed meetbaar is. Dit komt uitgebreid aan de orde in de volgende paragraaf.

3 **Beoordeling van onderwijsprestaties: Meten is niet altijd weten**

“Niet alles wat gemeten wordt is belangrijk, en niet alles wat belangrijk is kan gemeten worden”,

Albert Einstein

Het gebruik van prestatieprikkels vereist meting van prestaties. In het onderwijs doen zich hierbij verschillende problemen voor. Deze paragraaf gaat in op meetproblemen die bij het gebruik van prestatiebekostiging, *accountability*, en prestatiebeloning een rol spelen.

Productie in het onderwijs: meetbaar – niet-meetbaar / gemeten – niet-gemeten

Onderwijs is bedoeld om leerlingen tal van vaardigheden en kennis bij te brengen van belang voor de latere participatie in de samenleving en voor de individuele ontwikkeling. Het ligt voor de hand om daarbij allereerst te denken aan cognitieve vaardigheden zoals taal en rekenen in het basisonderwijs of wiskunde en Engels in het voortgezet onderwijs. Daarnaast levert het onderwijs ook bijdragen aan andere vaardigheden zoals discipline, tolerantie, zelfvertrouwen, ambities en creativiteit. Dit type vaardigheden is veel lastiger te meten, en misschien zelfs helemaal niet. Bovendien moet bij de gemeten output in het algemeen een keuze worden gemaakt welk deel van de output te toetsen. Voor het toepassen van prestatieprikkels betekent dit dat delen van de output van het onderwijs, namelijk het niet-meetbare en het niet-gemeten deel, buiten beschouwing blijven. Het gevolg daarvan kan zijn dat inspanningen van docenten of management van scholen zich vooral zullen richten op de gemeten output onder verwaarlozing van andere taken. Daar staat tegenover dat het gebruik van prestatie-indicatoren scholen er juist toe kan aanzetten zich in te spannen voor die kennis en vaardigheden waaraan de samenleving grote waarde hecht. Prestatiebeloning c.q. prestatiebekostiging kan dan een instrument zijn voor het aanpakken van de maatschappelijke prioriteiten binnen het onderwijs, zoals het voorkomen van voortijdig schoolverlaten.

Wat is de bijdrage van de school of van de docent aan de prestaties van leerlingen?

De prestaties van leerlingen zijn het resultaat van de inspanningen van verschillende actoren, zoals de school, de docent, de leerling en de ouders. Leerlingen leren niet alleen op school maar ook in tal van maatschappelijke verbanden buiten het onderwijs. Daarmee wijkt het productieproces in het onderwijs in belangrijke mate af van dat in de meeste andere economische sectoren. Ook belangrijk is dat selectie een essentieel kenmerk is van het onderwijs. Ouders kiezen bijvoorbeeld een school voor hun kinderen, en binnen scholen worden leerlingen ingedeeld in groepen op grond van hun eigenschappen.

De consequentie daarvan is dat het lastig is om de bijdrage van de school aan de prestaties van de leerling te bepalen. Gebruikelijk is om hiervoor zogenoemde onderwijsproductiefuncties te schatten waarin de prestaties van leerlingen gerelateerd worden aan kenmerken van scholen, docenten, leerlingen en ouders (zie box 'Onderwijsproductiefuncties: Niveau en toegevoegde waarde'). Aangezien vele factoren van invloed kunnen zijn op de prestaties van leerlingen is het lastig om de bijdrage van de school te isoleren. Onderzoekers zullen geen rekening kunnen houden met alle mogelijke factoren. Daardoor bestaat altijd de mogelijkheid dat een bijdrage wordt toegeschreven aan een bepaalde school terwijl die in feite wordt veroorzaakt door niet-waargenomen factoren. Hetzelfde geldt voor het bepalen van de bijdrage van een docent. Het is immers heel goed mogelijk dat de prestaties van een leerling in een bepaald vak ook beïnvloed worden door de lessen in andere vakken. De bijdrage van een individuele docent is in dat geval moeilijk met zekerheid vast te stellen.

Het meten van de toegevoegde waarde

Een bekende methode voor het vaststellen van de bijdrage van een school is het meten van de toegevoegde waarde. Hiervoor wordt een leerling op twee verschillende momenten getoetst. Het verschil tussen de beide meetmomenten is de toegevoegde waarde (zie box 'Onderwijsproductiefuncties: Niveau en toegevoegde waarde'). Door het invoeren van een instroomtoets, zoals voorgenoemen door het kabinet Balkenende 1, komt het meten van de toegevoegde waarde een stap dichterbij. Het is echter een misvatting om de toegevoegde waarde gelijk te stellen aan de bijdrage van een school. De toegevoegde waarde wordt namelijk bepaald door alle inputs vanuit het gezin, de klasgenoten en de school tussen de beide meetmomenten (zie vergelijking (2) in box 'Onderwijsproductiefuncties: Niveau en toegevoegde waarde'). Bij het werken met de toegevoegde waarde hoeft alleen geen rekening gehouden te worden met de aangeboren aanleg. Die is immers gelijk op beide meetmomenten. Als de toegevoegde waarde geheel wordt toegeschreven aan de school wordt in feite verondersteld dat bijvoorbeeld ouders geen invloed hebben op de prestaties van hun kinderen. Deze veronderstelling lijkt weinig realistisch. Aannemelijk is dat ouders inspelen op de studieresultaten van hun kinderen. Bij goede prestaties, bijvoorbeeld door grote inspanningen van de docent, is er geen reden zich met het onderwijs te bemoeien. Bij tegenvallende prestaties is het mogelijk dat ouders gaan meehelpen met het huiswerk, of dat zij bijles organiseren voor hun kinderen. De compenserende inspanningen van ouders kunnen aldus de toegevoegde waarde beïnvloeden. Door de toegevoegde waarde geheel aan de school toe te rekenen worden in dit geval de prestaties van scholen niet goed geschat (zie Todd en Wolpin, 2003).

Een ander probleem bij het meten van de toegevoegde waarde door scholen is dat de schattingen zeer gevoelig zijn voor meetfouten. Bij het bepalen van de toegevoegde waarde

worden meefouten opgeblazen.⁷ Een recente studie van Kane en Staiger (2001) laat zien dat dit een groot probleem is bij scholen van beperkte omvang. Zij laten ook zien dat een rangschikking van scholen op basis van toegevoegde waarde in feite neerkomt op een loterij. Dit probleem kan gereduceerd worden door te kijken naar prestaties over een langere periode (Kane en Staiger, 2002).

Manipulatie van metingen

Tenslotte kan meting van onderwijsprestaties worden gemanipuleerd. Zo vinden Ladd en Walsh (2002), Jacob (2002) en Figlio en Getzler (2002) aanwijzingen dat scholen proberen de gemiddelde testcores te verhogen door zwakke leerlingen buiten de tests te houden en bijvoorbeeld sneller door te sturen naar speciale programma's. Het is ook mogelijk dat er specifiek op de test wordt getraind, zonder dat dit direct leidt tot verbetering van leerprestaties. Het oefenen van eerdere toetsen, creëren van optimale testcondities, en leerlingen aansporen om hun best te doen tijdens de test kan leiden tot verbetering van de testresultaten zonder dat de leerprestaties daadwerkelijk zijn verbeterd. In enkele recente studies wordt inderdaad bewijs gevonden voor *teaching-to-the-test* (zie o.a. Jacob, 2002). Bovendien kunnen instellingen leerlingen ontmoedigen om een zwaar vakkenpakket te kiezen (vgl. het eerder besproken voorbeeld van instellingen die leerlingen afraden om een β -profiel te volgen).

De conclusie is dat het bepalen van de bijdrage van scholen of docenten aan de prestaties van leerlingen lastig is. Door de toegevoegde waarde te meten worden de datavereisten sterk verminderd. Daarmee zijn de problemen echter geenszins opgelost: de toegevoegde waarde is niet hetzelfde als de bijdrage van de school. Bovendien kunnen meefouten dan een grotere rol spelen. Ook kunnen de uitkomsten worden gemanipuleerd. Als objectief meten lastig is kunnen subjectieve methoden overwogen worden. Bij subjectieve beoordelingen kan ook rekening worden gehouden met andere factoren en waarnemingen die niet goed te kwantificeren zijn. Zulke subjectieve methoden zijn echter ook niet probleemloos. Persoonlijke verhoudingen zouden bijvoorbeeld een rol kunnen spelen bij de beoordeling. Wellicht geeft een combinatie van objectieve en subjectieve beoordeling het meest betrouwbare beeld (zie Baker e.a., 1994). Bij de huidige kennis ten aanzien van het meten van prestaties van scholen en docenten wordt de belangrijke vraag dan: hoe bruikbaar zijn prestatieprikkels bij een niet-perfecte meting van prestaties van scholen of docenten? Om deze vraag te beantwoorden zullen we in de volgende paragraaf een aantal internationale ervaringen bespreken.

⁷ De toegevoegde waarde wordt gevonden door schoolprestaties te meten op twee tijdstippen. Het gedeelte van de verandering van schoolprestaties dat kan worden toegeschreven aan de scholen, is de toegevoegde waarde. Bij het meten van schoolprestaties worden fouten gemaakt, en bij het meten van de verandering in schoolprestaties worden deze meefouten opgeblazen, omdat twee keer gemeten moet worden. Voorbeeld: wanneer op tijdstip 1 de schoolprestaties verkeerd worden gemeten en op tijdstip 2 correct, dan wijkt de gemeten verandering in schoolprestaties af van de feitelijke verandering.

Onderwijsproductiefuncties: Niveau en toegevoegde waarde

Een standaard onderwijsproductiefunctie kent de volgende vorm:

$$(1) Y_{it} = f(F_i^{(t)}, P_i^{(t)}, S_i^{(t)}, A_i) + \varepsilon_{it}$$

met Y_{it} = prestatie van leerling i op tijdstip t , $F_i^{(t)}$ = inputs vanuit het gezin gecumuleerd tot tijdstip t , $P_i^{(t)}$ = bijdragen van klasgenoten (*peers*) gecumuleerd tot tijdstip t , $S_i^{(t)}$ = school inputs gecumuleerd tot tijdstip t , A_i = aangeboren aanleg, en een stochastische storingsterm ε_{it} (zie Hanushek, 2002).

Deze formulering laat in de eerste plaats zien dat een groot aantal factoren van invloed is op de leerprestaties, zowel factoren van binnen als van buiten het onderwijs. In de tweede plaats is het productieproces cumulatief, dat wil zeggen het bouwt voort op een serie van inputs over de tijd. Het meten van de toegevoegde waarde is gebaseerd op dit cumulatieve aspect. Een algemene formulering is dan:

$$(2) Y_{it2} - Y_{it1} = f(F_i^{(t2-t1)}, P_i^{(t2-t1)}, S_i^{(t2-t1)}) + \varepsilon_{it2} - \varepsilon_{it1}$$

met $F_i^{(t2-t1)}$ = inputs vanuit het gezin tussen tijdstip $t2$ en $t1$, etc. Het aantrekkelijke van vergelijking (2) is dat de dataverrekeningen voor het schatten worden beperkt. Alle gecumuleerde inputs tot tijdstip $t1$ verdwijnen namelijk als het verschil wordt genomen. Ook de aangeboren aanleg verdwijnt uit deze vergelijking omdat die constant wordt verondersteld in de tijd. De toename van de prestaties hangt dan af van de inputs vanuit gezin, klasgenoten en school tussen $t2$ en $t1$. Deze formulering laat ook zien dat het geheel toeschrijven van de toegevoegde waarde aan de school veronderstelt dat er geen inputs zijn vanuit het gezin of de *peers* tussen $t2$ en $t1$. Vergelijking (2) veronderstelt wel dat het niveau op tijdstip $t1$ geen effect heeft op de studievoortgang. Een minder restrictief model zou zijn:

$$(3) Y_{it2} = f(F_i^{(t2-t1)}, P_i^{(t2-t1)}, S_i^{(t2-t1)}, Y_{it1}) + \varepsilon_{it2} - \varepsilon_{it1}$$

Bij lineaire effecten kan (3) namelijk geschreven worden als:

$$(4) Y_{it2} = \alpha + \beta F_i^{(t2-t1)} + \delta P_i^{(t2-t1)} + \gamma S_i^{(t2-t1)} + \lambda Y_{it1} + \varepsilon_{it2} - \varepsilon_{it1}$$

waarbij (3) alleen gelijk is aan (2) als wordt verondersteld dat $\lambda = 1$.

Bij het schatten van onderwijsproductiefuncties is van belang dat leerlingen geclusterd zijn binnen scholen of binnen klassen. Dit betekent dat zij te maken hebben met dezelfde niet-waargenomen factoren. Een algemeen probleem bij het schatten van onderwijsproductiefuncties is dat een veelheid van factoren van invloed is. Daardoor kunnen niet-waargenomen factoren (opgepikt door ε_{it}) de schattingen vertekenen. In recente studies wordt daarom vaak gezocht naar natuurlijke experimenten die leiden tot variatie in school inputs onafhankelijk van niet-waargenomen factoren.

4 Internationale ervaringen met prestatieprikkels in het onderwijs

De laatste jaren zijn in diverse landen nieuwe beleidsinitiatieven met prestatieprikkels gestart. Hierbij zijn met name de ervaringen in Israël, Engeland en de Verenigde Staten interessant, omdat voor deze landen ook evaluatiestudies beschikbaar zijn. In dit onderdeel bespreken we deze internationale ervaringen. We volgen hierbij de eerder gemaakte driedeling: i. prestaties en bekostiging, ii. prestaties en *accountability*, en iii. prestaties en beloning.

4.1 Prestaties en bekostiging

Bij de eerste relatie die we onderzoeken, kijken we naar een speciale variant van prestatie-bekostiging, waarbij gebruik wordt gemaakt van teambeloning. Bij goed presteren ontvangen de instellingen extra financiële middelen van de overheid, en deze middelen worden (geheel of ten dele) doorgesluisd naar de docenten. Met andere woorden, het betreft hier prestatiebekostiging waarbij de eventuele additionele financiering is geormerkt.

In een aantal landen is ervaring opgedaan met teambeloning in het onderwijs. Veelal zijn deze initiatieven zo recent dat er nog geen evaluatiestudies beschikbaar zijn. We zullen ons hier beperken tot bespreking van de ervaringen met teambeloning die zijn geëvalueerd.⁸

4.1.1 Dallas

Beschrijving

In 1991 is in Dallas een teambeloningsprogramma ingevoerd.⁹ Personeel in de scholen die het hoogst scoren op de *School Effectiveness Index* ontvangen een financiële beloning (zie box 'Teambeloning in Dallas: Hoe wordt output gemeten?'). Het totale beschikbare budget is zo'n \$2,5 miljoen per jaar.

Effecten

De evaluatie van Ladd (1999) laat zien dat slaagpercentages in Dallas meer zijn toegenomen dan in andere grote Texaanse steden als rekening wordt gehouden met stad- en schoolspecifieke verschillen. Verder blijkt dat de impact van het programma behoorlijk verschilt tussen etnische groepen, waarbij voor Latijns-Amerikaanse en blanke leerlingen de meest positieve resultaten worden gevonden. Er is geen bewijs dat het Dallas-programma gunstig heeft uitgedaakt voor Afro-Amerikaanse kinderen.

⁸ En we beperken ons tot landen die qua niveau van ontwikkeling vergelijkbaar zijn. Zie Glewwe e.a. (2003) voor een studie naar de effecten van teambeloning in het onderwijs in Kenia.

⁹ In de VS zijn diverse initiatieven met teambeloning in het onderwijs ontplooid. Zie de homepage van het *Consortium for Policy Research in Education* (<http://www.wcer.wisc.edu/cpre/>) voor nadere informatie.

Teambeloning in Dallas: Hoe wordt output gemeten?

De methodologie om scholen te rangschikken wordt beschreven in Ladd (1999) en Mendro e.a. (1999). De bijdrage van de school aan de individuele leerprestaties wordt geschat aan de hand van karakteristieken van studenten zoals etniciteit, geslacht en sociaal-economische achtergrond. In de tweede stap worden de residuen van de regressie verklaard uit de residuen van het voorafgaande jaar. De residuen leveren een schatting voor de verbetering in leerprestaties. Tenslotte worden gemiddelden berekend op het niveau van een school. Met deze procedure wordt gepoogd te vermijden dat schattingen van de toegevoegde waarde worden vertekend ten gunste van de scholen met meer bevoorrechte kinderen.

Zo'n 20% van de scholen komt elk jaar in aanmerking voor een financiële bonus. Schoolhoofden en leraren ontvangen \$1000 en andere medewerkers (zoals conciërges en secretaresses) krijgen een bonus van \$500. Verder ontvangen de winnende scholen \$2000 ter besteding aan activiteiten.

Een probleem bij de interpretatie van de resultaten is dat er voor de invoering van teambeloning in Dallas al sprake is van positieve effecten. Deze effecten zijn wellicht veroorzaakt door de publiciteit rondom de hervormingsplannen. Zulke 'Hawthorne'-effecten zullen na verloop van tijd verdwijnen, maar Ladd vindt een blijvend positief effect op studieprestaties. Dit kan duiden op een positieve bijdrage van het teambeloningsprogramma.

Voor schooluitval wordt een soortgelijke analyse uitgevoerd, en er wordt een belangrijke reductie van het aantal voortijdig schoolverlaters gevonden. Bovendien worden deze effecten sterker na verloop van tijd. Voorts laat de studie zien dat na invoering van teambeloning het verloop onder schooldirecteuren sterk toenam. Het aantrekken van meer geschikte schooldirecteuren is mogelijk één van de oorzaken van de verbetering van prestaties van leerlingen.

Beoordeling

De evaluatiestudie komt tot de conclusie dat men voorzichtig optimistisch kan zijn over de mogelijkheden om met een teambeloningsmodel de leerprestaties te verbeteren. Of de prestaties daadwerkelijk zijn verbeterd door teambeloning is aanvechtbaar. Ook is niet bekend in hoeverre het teambeloningsprogramma tot ongewenste effecten heeft geleid.

4.1.2 Israël

Beschrijving

In februari 1995 kondigde het Israëlijs Ministerie van Onderwijs een experiment aan met teambeloning in het secundair onderwijs, met als doel verlaging van de uitval en verbetering van de leerprestaties. In eerste instantie namen 62 scholen deel aan het experiment, en later is hier nog een aantal scholen bijgekomen.

De scholen concurreerden om een totaalbedrag van ongeveer \$1,44 miljoen in 1996. De prestatie-indicatoren zijn het gemiddeld aantal behaalde studiepunten per leerling, het aantal

leerlingen dat het *matriculation* certificaat¹⁰ heeft behaald, en het uitvalpercentage per school. Schoolprestaties werden als volgt gemeten. Gemiddelde uitkomsten per school werden genormaliseerd (voor het jaar van het experiment en het voorafgaand jaar) ten opzichte van het verwachte effect gebaseerd op regressie-analyse, waarbij rekening werd gehouden met de sociaal-economische achtergrondkenmerken van de leerlingenpopulatie. Vervolgens werden de scholen gerangschikt op basis van verbetering van de onderwijsprestaties.¹¹ De beste 33% kwam in aanmerking voor een bonus. De verdeling van de bonus over de winnende scholen gebeurde op basis van hun plaats op de ranglijst. Het grootste bedrag dat in 1996 aan een school werd uitgekeerd bedroeg \$105.000 en het laagste bedrag \$13.250. Driekwart van het gewonnen bedrag werd direct uitgekeerd aan de leraren (naar rato van hun bruto-inkomen), en de rest werd gebruikt voor verbetering van de arbeidsomstandigheden. In 1996 bedroeg de hoogste bonus per leraar \$1000, en de laagste \$250 (bij een gemiddeld jaarsalaris van zo'n \$30000).

Effecten

Het effect van teambeloning is geschat via een voor- en nameting van de prestaties van leerlingen op scholen die deelnamen aan het experiment en op scholen die de controlegroep vormden (Lavy, 2002a). Aan het experiment namen alleen scholen deel die de enige van hun soort waren in de regio. Als controlegroep werden scholen gekozen waarvan er twee in een regio waren. De resultaten van prestatiebeloning voor seculiere scholen kunnen als volgt worden samengevat:

- het aandeel leerlingen dat het *matriculation* examen aflegt neemt enigszins toe;
- de gemiddelde score neemt toe;
- het aantal behaalde studiepunten neemt toe;
- het aandeel voortijdig schoolverlaters is gedaald.

Voor religieuze scholen worden zelfs nog grotere effecten gevonden, en de effecten zijn in het eerste jaar dat teambeloning wordt toegepast al zichtbaar.

Beoordeling

De evaluatie laat zien dat prestatieprikkels tot significante verbeteringen op de prestatie-indicatoren hebben geleid. Merk op dat de gemiddelde score is toegenomen ondanks de daling van het aantal voortijdig schoolverlaters. Er is dan ook geen onderlinge tegenstrijdigheid tussen de doelstellingen van kwaliteit en deelname.

¹⁰ Het *matriculation* certificaat is een noodzakelijke maar niet een voldoende voorwaarde voor toelating tot een universitaire opleiding.

¹¹ Het is ons niet bekend hoe de prestatie-indicatoren werden gewogen.

De prestaties van scholen zijn gemeten door rekening te houden met achtergrondkenmerken van leerlingen en de prestaties in eerdere jaren. Bekend is dat deze wijze van outputmeting niet perfect is (zie o.a. Kane and Staiger, 2001). De resultaten voor Israël kunnen er dan op wijzen dat teambeloning ook bij een niet-perfecte meting van prestaties positieve effecten kan sorteren. Een kanttekening bij het onderzoek is dat de bonussen relatief laag zijn. Er is dan ook gesuggereerd dat niet het systeem van teambeloning de oorzaak is van de verbetering van de onderwijsprestaties, maar dat de media-aandacht voor het experiment een meer plausibele oorzaak is. Voorts wordt in de studie geen aandacht besteed aan *gaming*: slagen scholen erin de resultaten rooskleuriger voor te stellen dan ze in werkelijkheid zijn? Tenslotte, omdat het programma alleen is geïmplementeerd op scholen in relatief kleine gemeenschappen is voorzichtigheid bij de generalisatie van de resultaten geboden.

4.1.3 Verenigd Koninkrijk

Beschrijving

In het VK is met ingang van het schooljaar 2000-'01 jaarlijks £60 miljoen beschikbaar voor scholen die goede vooruitgang boeken (*improvement awards*) en scholen die zeer goed presteren (*excellence awards*). Zo'n 75% van deze *School Achievement Awards* gaat naar *improvement awards* en de resterende 25% naar *excellence awards*. De bonus hangt af van de grootte van de school (het aantal leerlingen), en de school ontvangt ongeveer £30 per leerling. Het bestuur van de school bepaalt op welke wijze het geld wordt verdeeld over de medewerkers. De toekenning van de *awards* gebeurt als volgt. Voor de derde ronde van de *School Achievement Awards* wordt gekeken naar (i) het verschil in de resultaten tussen 1999 en 2002 (bonus bij snelle vooruitgang) en (ii) de resultaten in 2002 vergeleken met scholen in gelijke omstandigheden (bonus bij goede resultaten).¹² Helaas is er nog geen evaluatie van de effecten gepubliceerd. Interviews met direct betrokkenen duiden erop dat de effecten zwak zijn. Dit kan wellicht worden verklaard door gebrekkig functioneren van het stelsel. In de praktijk ontvangt in het ene jaar de ene 50% van de scholen een bonus, en in het andere jaar de andere 50%.

4.2 Prestaties en accountability

Hoewel in veel landen wel een bepaalde vorm van *accountability* in het onderwijs aanwezig is, is tot dusver slechts één evaluatie beschikbaar. In dit onderdeel beperken we ons dan ook tot bespreking van dit onderzoek.

¹² Er wordt gekeken naar scholen die onderling min of meer vergelijkbaar zijn. Hierbij worden globale maatstaven gebruikt om rekening te houden met sociaal-economische achtergrond van kinderen (bijvoorbeeld het aantal kinderen dat in aanmerking komt voor een gratis lunch).

Beschrijving

In januari 2002 werd in de VS de *No Child Left Behind Act* aangenomen, waarmee een nieuwe impuls werd gegeven aan *accountability* in het onderwijs. De nieuwe wet verplicht staten om leerlingen van groep 3 tot en met 8 te testen, en de testresultaten te gebruiken om de prestaties van scholen in kaart te brengen. Leerlingen krijgen de mogelijkheid om naar een andere school over te stappen als scholen onder de maat presteren. Bij voortdurend slechte prestaties kunnen scholen worden gesloten, of kan het personeel worden vervangen. In Chicago is zo'n beoordelingssysteem gebaseerd op testresultaten van leerlingen reeds in 1996-'97 ingevoerd.

Effecten

De evaluatie richt zich op drie vragen (Jacob, 2002):

- Leidt openbaarheid van schoolprestaties tot betere studieresultaten?
- Als er een positief effect op studieprestaties wordt gevonden, komt dit dan door verbeteringen in algehele studieprestaties, verbeteringen in testspecifieke vaardigheden, tijdelijke extra inspanningen van studenten, of een combinatie hiervan?
- Reageren docenten en management strategisch op de prikkels inherent aan het testsysteem?

Ten eerste, Jacob vindt sterke verbeteringen in scores op wiskunde- en leesvaardigheidstoetsen na invoering van het beoordelingssysteem. Deze verbeteringen zijn veel groter dan verwacht op basis van trends uit het verleden in Chicago, en ook groter dan verbeteringen in andere grote steden in de regio. Ook wordt gevonden dat de relatief slecht presterende scholen grotere verbeteringen laten zien dan andere scholen.

Voor wat betreft de tweede vraag blijkt uit het onderzoek van Jacob dat de verbeteringen met name worden veroorzaakt door betere testspecifieke vaardigheden en meer inspanningen van studenten. Scores op een vergelijkbare toets die niet gebruikt wordt bij de beoordeling van scholen, te weten het *Illinois Goals Assessment Program* (IGAP), laten geen verbetering zien. Er zijn echter ook geen aanwijzingen voor verwaarlozing van bepaalde taken. Leerlingen scoren namelijk niet slechter op de IGAP-test. Ook blijkt dat de grootste verbeteringen in testcores worden waargenomen op het einde van het examen. Dit duidt erop dat leerlingen meer inspanningen leveren tijdens het examen (Jacob noemt dit een *test stamina* effect). Wellicht heeft dit te maken met een andere verandering. Met de invoering van het beoordelingssysteem is ook een einde gemaakt aan het beleid om alle leerlingen elk jaar te laten doorstromen naar het volgende jaar ('*social promotion*'). In het nieuwe systeem kunnen zwakke leerlingen worden verplicht om extra cursussen gedurende de zomervakantie te volgen, en, als de resultaten daarna nog steeds onbevredigend zijn, een jaar te doubleren. Dit moedigt leerlingen aan om meer inspanningen te leveren.

Tenslotte, het blijkt dat docenten op een aantal manieren strategisch reageren op de prikkels verbonden aan het beoordelingssysteem. Er wordt een toename van het aandeel leerlingen in speciale klassen waargenomen, en een toename van leerlingen die niet getest hoeven te worden. Dit kan erop duiden dat scholen proberen hun prestaties rooskleuriger voor te stellen door zwakke leerlingen buiten de officiële beoordelingen te houden. Verder blijven meer kinderen zitten in de klassen voorafgaand aan een jaar met een belangrijke test. En tot slot zijn de studieprestaties in andere vakken (natuurwetenschappen en maatschappijleer) minder verbeterd dan bij wiskunde en leesvaardigheid.

Beoordeling

Jacob (2002) trekt zelf een voorzichtige conclusie: “... *these results suggest that high-stakes testing has the potential to substantially improve student learning, but needs to be approached with caution*”. De evaluatie laat zien dat testcores sterk verbeteren. Het blijft evenwel onduidelijk of dit betekent dat leerlingen meer geleerd hebben. Scores op een vergelijkbare toets stijgen namelijk niet. Een andere belangrijke les uit deze studie is dat scholen proberen de gemiddelde testresultaten te manipuleren door zwakke leerlingen buiten de tests te houden.¹³ Deze vorm van strategisch gedrag kan worden tegengegaan door minimumeisen te stellen aan het aantal leerlingen dat wordt getest (zo moet in Dallas minstens 95% van de leerlingen worden getest).

4.3 Prestaties en beloning

Individuele prestatiebeloning in het onderwijs is zeldzamer dan teambeloning. Ondanks de sterkere prestatieprikkel (bij teambeloning zijn prikkels zwakker en kan liftersgedrag optreden), kunnen mogelijke nadelen het gebruik ervan in de weg staan (meer inkomensrisico, gevaar van niet-coöperatief gedrag). Verder is er een grijs gebied waar individuele prestatiebeloning overgaat in competentiebeloning. Zo ligt evaluatie van didactische vaardigheden (bij competentiebeloning) dicht bij evaluatie van studievoortgang (bij prestatiebeloning). In dit onderdeel gaan we in op enkele relevante en interessante buitenlandse ervaringen.

4.3.1 Michigan

Beschrijving

In Michigan is in 1996 individuele prestatiebeloning voor docenten toegepast op een middelbare school (Eberts e.a., 2002). Het systeem van prestatiebeloning biedt twee mogelijkheden voor extra inkomsten bovenop het basissalaris. Ten eerste wordt er een *retention bonus* uitgekeerd als ten minste 80% van de leerlingen de lessen blijft volgen. De bonus is gelijk voor alle docenten, en bedraagt ongeveer 12% van het basissalaris. Op de tweede plaats kunnen docenten een bonus verdienen op basis van een evaluatie door studenten. De studenten

¹³ Ook Figlio en Getzler (2002) vinden aanwijzingen voor deze vorm van *gaming* in Florida.

beoordelen 15 factoren op een schaal van 1 tot 5, en als een docent gemiddeld¹⁴ 4,65 of hoger scoort dan wordt het basissalaris met zo'n 5% verhoogd en de *retention bonus* met 10% (4,65 was de gemiddelde score in het voorafgaande schooljaar). Het jaarsalaris voor een leraar op Bachelor-niveau die alleen basissalaris ontvangt bedroeg in 1998-'99 \$22.848 (uitgaande van 9 maanden, 4 kwartalen met 7 groepen). Inclusief beide bonussen zou deze leraar uitkomen op een jaarlijks inkomen van \$27.412. Het gaat dan ook om substantiële bedragen die extra kunnen worden verdiend.

Effecten

In de evaluatie worden studieprestaties op de betrokken school vergeleken met die van een school met een vergelijkbare leerlingenpopulatie waar een meer traditioneel beloningssysteem wordt gebruikt. Er wordt gekeken naar gemiddelde cijfers, deelname aan lessen en slaagpercentages. De effecten van prestatiebeloning worden geschat door gebruik te maken van een voor- en nameting op beide scholen. Voor wat betreft uitvalpercentages vinden Eberts e.a. een daling in beide scholen, maar zoals verwacht is de daling het sterkst in de school waar prestatiebeloning wordt gehanteerd. Beide scholen scoren in de nameting slechter in termen van gemiddelde cijfers, maar de achteruitgang is het sterkst in de school waar prestatiebeloning wordt toegepast. Dit is in overeenstemming met de verwachtingen, omdat met prestatiebeloning wordt gepoogd om uitval van zwakkere leerlingen te verminderen.

De onderwijsdeelname van studenten is gemeten op een willekeurig gekozen dag in de laatste week van de lessen. De gemiddelde dagelijkse aanwezigheid van studenten gedurende de lessen is erg laag, en is in de school met prestatiebeloning zelfs licht gedaald. Tenslotte, het slaagpercentage van studenten (met uitzondering van de voortijdig schoolverlaters) daalde sterker in de school met prestatiebeloning. Verklaring hiervoor is wederom dat er in deze school minder voortijdig schoolverlaters zijn.

Beoordeling

De resultaten in deze studie suggereren dat prikkels 'werken'. Prestatiebeloning heeft inderdaad tot een reductie in het aantal voortijdig schoolverlaters geleid. Verbeteringen op andere outputmaatstaven zijn echter uitgebleven. Er worden geen positieve effecten gevonden op studieprestaties en gemiddelde aanwezigheid. Bovendien zijn er aanwijzingen voor *gaming* in termen van manipulatie van de meting van aanwezigheid. Het aantal voortijdig schoolverlaters is volgens de officiële meting gedaald, maar de gemiddelde dagelijkse aanwezigheid van leerlingen is ook gedaald in de school met prestatiebeloning. Eberts e.a. (blz. 926, 2002) schrijven in hun conclusie:

¹⁴ Gemiddeld over alle klassen waarin de leraar doceert, en gewogen met klassengrootte.

“... the results suggest that pay for performance incentives can motivate agents to produce outcomes that are directly rewarded. However, the study also suggests that incentive systems within complex organizations, such as schools, with multiple tasks and outcomes, team production, and multiple stakeholders, may produce results that are unintended and, at times, misdirected – unless carefully constructed and implemented.”

De cruciale vraag bij dit onderzoek is in hoeverre het gevonden verschil in ontwikkeling van studieprestaties tussen beide scholen wordt veroorzaakt door prestatiebeloning. De onderzoeksopzet vereist dat de studentenpopulaties op beide scholen vergelijkbaar zijn. Toch blijken er belangrijke (en statistisch significante) verschillen te zijn in gemiddelde cijfers, en daarmee in de vergelijkbaarheid van scholen. Dit laat ruimte voor alternatieve verklaringen voor het gevonden verschil in de ontwikkeling van studieprestaties op deze scholen.

4.3.2 Israël

Beschrijving

In september 2001 kondigde het Israëliësch Ministerie van Onderwijs een nieuw bonusprogramma voor leraren (Engels, Hebreeuws, Arabisch en wiskunde) aan in 49 middelbare scholen. Het betrof een driejarig experiment op basis van vrijwillige deelname. Selectie van de scholen was niet door toeval bepaald: het programma was gericht op scholen die in het recente verleden relatief slechte prestaties hebben laten zien (minder dan 45% van de leerlingen behaalt het *matriculation* examen). Leraren ontvangen een financiële bonus voor verbeteringen van de uitslagen van leerlingen op het *matriculation* examen. In onderstaande box wordt uitgelegd hoe prestaties worden gemeten, en op welke wijze de bonussen worden bepaald. Er is gekozen voor relatieve in plaats van absolute prestatiebeloning omdat de bonussen zodoende binnen budget blijven en omdat er geen duidelijke standaarden zijn die als basis kunnen dienen voor absolute prestatiebeloning. Het bonusprogramma kan dan ook worden omschreven als een rangordetoernooi.

Individuele prestatiebeloning in Israël: Hoe wordt output gemeten?

Twee maatstaven van studieprestaties worden gebruikt als indicator voor de prestaties van leraren: het slaagpercentage en de gemiddelde score op het *matriculation* examen. De prestaties worden bepaald aan de hand van een vergelijking van de resultaten van de leerlingen van een bepaalde leraar met de voorspelling van een regressie. In het regressiemodel wordt gecorrigeerd voor sociaal-economische achtergrondkenmerken van de leerlingen, het bereikte niveau in het vak, en een *school fixed effect* om rekening te houden met niet-waargenomen heterogeniteit tussen scholen. De regressies werden uitgevoerd voor zowel het slaagpercentage als de gemiddelde score op het *matriculation* examen. Leraren die voor beide maatstaven een positief residu (feitelijke uitkomst minus voorspelde uitkomst) laten zien, worden gerangschikt en ontvangen punten voor hun plaats op de ranglijst: 16 punten voor de eerste plaats, 12 voor de tweede plaats, 8 voor de derde plaats, en 4 voor de vierde plaats. Het slaagpercentage krijgt een extra gewicht van 25% in de beoordeling (20, 15, 10 en 5 punten, respectievelijk). Het totaal aantal punten voor de twee ranglijsten wordt gebruikt om leraren een rangorde toe te kennen in het toernooi. De bonussen worden als volgt bepaald: 30-36 punten \$7500; 21-29 punten \$5750; 10-20 punten \$3500; en 9 punten \$1750. Dit zijn significante bedragen ten opzichte van het gemiddelde bruto jaarinkomen van een leraar op een middelbare school van ongeveer \$30.000. Bovendien kunnen leraren meerdere bonussen winnen. Van de 629 leraren hebben er 302 bonussen ontvangen, waarvan 16 docenten twee bonussen hebben ontvangen (en één docent heeft zelfs twee eerste plaatsen bereikt, en ontving een totaal geldbedrag van \$15.000).

Effecten

Lavy (2002b) onderzoekt of dit programma met individuele prestatiebeloning heeft geleid tot hogere inspanningen van leraren en betere studieprestaties. Zoals opgemerkt is selectie van de scholen niet willekeurig. Dit bemoeilijkt de evaluatie. Om het causale effect van prestatiebeloning op de studieprestaties vast te stellen worden drie methoden gebruikt.¹⁵ De resultaten zijn vergelijkbaar. Bij wiskunde heeft het programma geleid tot een verbetering van 19%. Voor Engels wordt een verbetering van 21% gevonden. Bij twee vakken waarop niet werd afgerekend worden geen effecten gevonden. Vervolgens is gekeken hoe de leerwinst is gerealiseerd. Zijn de verbeteringen het gevolg van grotere inspanningen van docenten, of zijn lesmethoden gewijzigd? Om deze vraag te beantwoorden is een telefonische enquête uitgevoerd onder de docenten Engels en wiskunde die deelnamen aan het programma. Voor wat betreft de onderwijsmethode gaven de docenten Engels aan dat ze intensiever gebruik maakten van individuele instructie, groepering van leerlingen volgens aanleg, en aanpassing van de leermethoden aan de aanleg van leerlingen. Wiskundedocenten maakten alleen vaker gebruik van groepering van leerlingen volgens aanleg. Ook zijn er aanwijzingen dat docenten hogere inspanningen hebben geleverd. Zowel docenten Engels als wiskundedocenten gaven aan extra

¹⁵ Op drie manieren wordt een controlegroep gevormd. De eerste manier betreft een natuurlijk experiment waarbij gebruik wordt gemaakt van willekeurige meetfouten bij de selectie van scholen voor het programma. Selectie gebeurt namelijk op basis van voorlopige informatie. Sommige scholen worden aldus ten onrechte tot het programma toegelaten, terwijl anderen onterecht worden afgewezen. Deze meetfouten kunnen gebruikt worden bij identificatie van de effecten van prestatiebeloning. De tweede manier waarop getracht wordt een geschikte controlegroep te vinden is gebaseerd op discontinuïteit in het selectiecriteria. Scholen met een *matriculation rate* hoger dan maar in de buurt van de drempelwaarde van 45% zullen vergelijkbaar zijn met scholen met een *matriculation rate* van net onder de 45%. Eerstgenoemde groep kan in dat geval dienen als controlegroep. Tenslotte, de derde methode die Lavy gebruikt is *propensity score matching*. Met deze techniek wordt een controlegroep gevormd die sterk vergelijkbaar is op observeerbare kenmerken.

lesuren bovenop het normale programma gedurende het jaar te hebben gegeven. Verder blijkt dat deze extra inspanningen intensiveren voor het Bagrut-examen. Tenslotte richten docenten Engels hun aandacht meer op zwakke studenten, terwijl wiskundedocenten zich iets meer richten op de gemiddelde leerling.

Beoordeling

Het Israëliësch experiment met individuele prestatiebeloning voor docenten heeft geleid tot een verbetering van de prestaties van leerlingen. Hoewel het experiment een looptijd van enkele jaren zou hebben is het echter, in verband met politieke ontwikkelingen, al na het eerste jaar beëindigd. Hierdoor blijft onduidelijk of de gevonden positieve resultaten incidenteel dan wel structureel zijn. Het is namelijk de vraag hoe leraren die de bonus mislopen zich in de daarop volgende jaren gaan gedragen. Zullen ze zich opnieuw extra inspannen of vallen ze terug op hun oude werkwijze?

De snelle beëindiging van het experiment maakt het ook onmogelijk om vast te stellen of zich selectie-effecten bij de werving van nieuwe docenten voordoen. Als dit effect zich ook in het onderwijs zou voordoen, prestatiebeloning trekt betere docenten (zie paragraaf 2.3), zijn de gevonden effecten een ondergrens van de mogelijke winst die kan worden behaald. Daarbij dient echter wel aangetekend te worden dat het aantrekken van goede werknemers door het onderwijs elders in de samenleving een keerzijde kan hebben. Het is mogelijk dat goede werknemers worden onttrokken aan andere beroepen / sectoren.

4.3.3 Verenigd Koninkrijk

Beschrijving

In 2000 werd in het Verenigd Koninkrijk een vorm van individuele prestatiebeloning in het onderwijs ingevoerd, de zogenoemde *Performance Threshold*. Het *Performance Threshold* systeem is uitgebreid beschreven in o.a. Burgess e.a. (2001). Het beloningsstelsel voor leraren omvat 9 salarisschalen. Een leraar in de eindschaal kan een aanvraag indienen voor een 'bonus'. Om hiervoor in aanmerking te komen, moet worden aangetoond dat de leraar aan bepaalde standaarden voldoet. Als dit het geval is, dan ontvangt de leraar een bonus van £2000 (op jaarbasis). Deze bonus wordt ook in daaropvolgende jaren uitgekeerd, zonder dat de leraar opnieuw een aanvraag hiervoor hoeft in te dienen. Het is niet geheel helder aan welke standaarden voldaan moet zijn, maar het lijkt erop dat vooral naar competenties wordt gekeken. Het is dan ook wellicht beter om te spreken van competentiebeloning in plaats van prestatiebeloning.

Effecten

De vereisten om voor salarisverhoging in aanmerking te komen zijn niet zeer strikt gedefinieerd¹⁶, en de *assessment* van aspirant-leraren aan het einde van hun opleiding is in feite veel strenger dan de test voor de *Performance Threshold*. Omdat er geen strikte eisen worden gesteld is het lastig voor het schoolhoofd om een negatieve beslissing te verdedigen. Bovendien worden de extra salarisuitgaven betaald door het Ministerie van Financiën. Het schoolhoofd heeft dientengevolge weinig reden om de salarisverhoging niet toe te kennen. De beoordeling wordt dan ook gezien als een middel ter verdediging van extra onderwijsuitgaven bij het Ministerie van Financiën.¹⁷

In 2000 zat zo'n 75% van de leraren in hun eindschaal. Ongeveer 80% van deze groep diende een aanvraag in, waarvan 97% werd gehonoreerd (cijfers ontleend aan Burgess en Croxson, 2001).¹⁸ Vanwege het permanente karakter van de bonus, komt het in feite neer op promotie naar een hogere salarisschaal. Over de effecten van de *Performance Threshold* op de leerprestaties kunnen nog geen harde uitspraken worden gedaan. Maar als de *Performance Threshold* wordt gezien als een soort van algemene salarisverhoging, dan is het onwaarschijnlijk dat leraren meer inspanningen gaan leveren of hun lesmethoden gaan aanpassen om de studieprestaties van de leerlingen te verbeteren. Er zijn geen duidelijke aanwijzingen dat de *Performance Threshold* heeft geleid tot meer inspanningen van de leraar in het klaslokaal. Het blijkt dat de schoolhoofden de leraren veelal van te voren hebben ingelicht dat vrijwel iedereen in aanmerking zou komen voor de salarisverhoging, dus dan is het niet vreemd dat leraren dan geen extra inspanningen gaan leveren.

Op de vraag of de werkdruk is toegenomen door invoering van de *Performance Threshold* kwam een ontkennend antwoord. Overigens is het invullen van de formulieren en het verzamelen van de benodigde gegevens wel een tamelijk grote klus, waaraan sommige docenten een week hebben gewerkt (er zijn echter ook verdenkingen van het overschrijven van de formulieren van anderen). Daar staat tegenover dat dit hun vrijwel zeker een permanente salarisverhoging opleverde. Wel is de administratie van veel leraren verbeterd, omdat informatie moet worden aangeleverd over de studieprestaties van de leerlingen.

¹⁶ Bureau Hay McBer heeft standaarden ontwikkeld ter beoordeling van de prestaties van leraren, maar deze methode bleek te ingewikkeld en is nooit ingevoerd.

¹⁷ Op macroniveau is er geen gebrek aan geld voor onderwijs, en, volgens één der geïnterviewden, "money for education is considered to be a good thing".

¹⁸ De volgende redenen werden genoemd door de 20% die geen aanvraag heeft ingediend: (i) men is principieel tegen prestatiebeloning, (ii) men had geen idee wat de kans was om voor promotie in aanmerking te komen, en (iii) zelfselectie.

Beoordeling

Het is opmerkelijk dat alleen docenten die in hun eindschaal zitten, voor de bonus in aanmerking kunnen komen, en dat er slechts eenmalig wordt gekeken naar het functioneren van de leraren (het betreft een permanente bonus). Het Engelse *Performance Threshold* systeem heeft dan ook meer weg van een algemene salarisverhoging dan van een systeem van individuele prestatiebeloning. Naast de *Performance Threshold* zijn in het VK enkele andere interessante initiatieven met prestatieprikkels in het onderwijs gestart. In de bijlage van deze studie komen deze kort aan de orde.

4.4 Samenvatting

Tabel 4.1 laat zien dat de meeste studies gunstige effecten op studieresultaten en uitval vinden. Er zijn echter ook aanwijzingen dat invoering van prestatieprikkels gepaard gaat met bijwerkingen. Soms zijn deze bijwerkingen makkelijk te ondervangen. Door bijvoorbeeld minimumeisen te stellen aan het aantal leerlingen dat wordt getest, kan vermeden worden dat scholen de resultaten rooskleuriger proberen voor te stellen door zwakke leerlingen niet te testen. Het probleem van *teaching-to-the-test* is echter lastiger te ondervangen. De belangrijkste lessen van de buitenlandse ervaringen zijn:

- prikkels werken;
- *you get what you pay for*;
- *gaming* is realiteit;
- meetproblemen zijn belangrijk in het onderwijs.

Tabel 4.1 **Overzicht van internationale ervaringen**

Type prestatieprikkel	Land	Studie	Studieresultaten	Uitval	Bijwerkingen
Teambeloning	VS, Dallas	Ladd (1999)	omhoog (?)	omlaag	?
	Israël	Lavy (2002a)	omhoog	omlaag	?
<i>Accountability</i>	VS, Chicago	Jacob (2002)	omhoog	?	<i>teaching-to-the-test</i> en meer leerlingen in speciale programma's
Individuele prestatiebeloning	VS, Michigan	Eberts e.a. (2002)	omlaag door minder uitval	omlaag	gegevens over deelname worden gemanipuleerd
	Israël	Lavy (2002b)	omhoog	?	<i>teaching-to-the-test</i>

5 Prestatieprikkel in het Nederlandse onderwijs: Opties voor beleid

5.1 Nederland versus buitenlandse ervaringen

Het voorgaande overzicht laat zien dat het toepassen van prestatieprikkel (prestatiebekostiging, *accountability*, en prestatiebeloning) in het onderwijs gecompliceerd is en maatwerk vereist. Allereerst kent het onderwijs een veelheid aan doelen die soms strijdig zijn. Voorwaarde voor het effectief toepassen van directe prestatieprikkel is het stellen van prioriteiten. Van directe prestatieprikkel is immers bekend dat *'you get what you pay for'*. Willen we effectief gebruik maken van prestatieprikkel, dan is overeenstemming nodig over wat er geleerd moet worden in het onderwijs en welke onderdelen daarbij het belangrijkste zijn. In de tweede plaats is het lastig om de resultaten van leerlingen toe te schrijven aan scholen of aan individuele docenten. Dit komt doordat, anders dan in de meeste andere sectoren, onderwijsproductie ook buiten school plaatsvindt. Bovendien wordt het onderwijs gekenmerkt door (zelf-)selectie die maar voor een deel kan worden waargenomen door de onderzoeker. Bij de huidige kennis is het objectief vaststellen van prestaties van scholen en docenten lastig. De economische theorie stelt in dat geval voor om te zoeken naar een mix van objectieve en subjectieve meting. In de derde plaats bestaat in het onderwijs, evenals in andere sectoren, de mogelijkheid om resultaten te manipuleren. Scholen kunnen bijvoorbeeld zwakkere leerlingen weren van de 'allesbeslissende' toets. De oplossing hiervoor is meer toezicht, of het zoeken naar indicatoren die lastig te manipuleren zijn. Ondanks deze problemen laten de buitenlandse ervaringen zien dat een zorgvuldig gebruik van prikkels de prestaties kan verbeteren.

Bij de vertaling van de buitenlandse ervaringen naar de Nederlandse situatie dient rekening gehouden te worden met de specifieke Nederlandse context. Met name in de Verenigde Staten is nauwelijks sprake van directe concurrentie tussen scholen in primair en voortgezet onderwijs. In diverse staten zijn ouders gebonden aan de school in de wijk, terwijl de financiering van het onderwijs is gekoppeld aan de lokale onroerend goed belasting. Dit kan leiden tot aanzienlijke verschillen in onderwijskwaliteit. Door de vrijheid van schoolkeuze, en ook door andere verschillen, wijkt het Nederlandse systeem hiervan af. In Tabel 5.1 is het gebruik van prestatieprikkel in het Nederlandse onderwijs op hoofdlijnen weergegeven.

Tabel 5.1 Prestatieprikkel in het Nederlandse onderwijs

	Prestatie- bekostiging	Accountability		Prestatiebeloning
Onderwijssector		aantal aanbieders	transparantie	
Primair onderwijs	nee	groot	kwaliteitskaart	nee
Voortgezet onderwijs	nee	groot	kwaliteitskaart	nee
BVE	20%	beperkt	beperkt	nee
Hoger onderwijs	50%	beperkt	beperkt, keuzegids / media	nee
Invloed overheid	ja	deels	ja	nee

Prestatiebekostiging komt niet voor in het primair en voortgezet onderwijs. De recente discussie over de kleutertoets en de toegevoegde waarde van scholen doet verwachten dat dit de komende jaren ook niet zal veranderen. In de BVE-sector is ongeveer 20% van de financiering gebaseerd op prestaties, in het hoger onderwijs is dat ongeveer 50%. De concurrentie tussen scholen varieert sterk tussen sectoren. In het primair en voortgezet onderwijs is het aantal aanbieders groot en valt er derhalve veel te kiezen voor leerlingen en ouders. In de BVE-sector en in het hoger onderwijs is het aantal aanbieders daarentegen beperkt en is regelmatig sprake van regionale monopolies. De concurrentie tussen scholen wordt bevorderd door de transparantie van de markt. Op dit punt zijn alleen voor het PO en VO kwaliteitskaarten beschikbaar. Voor de andere sectoren bestaan dergelijke instrumenten nog niet. Wel zijn instellingen verplicht om informatie te verschaffen aan leerlingen en ouders, bijvoorbeeld in de vorm van een schoolgids. Het hoger onderwijs kent een keuzegids en in de media worden regelmatig ranglijsten van instellingen gepubliceerd. Individuele prestatiebeloning wordt in het Nederlandse onderwijs niet toegepast.

In de laatste rij van Tabel 5.1 is aangegeven in hoeverre de overheid invloed heeft op een bepaalde prestatieprikkel. De overheid kan het instrument van prestatiebekostiging direct toepassen. Ook kan de overheid instrumenten (laten) ontwikkelen om de transparantie van de markt te vergroten. De overheid heeft echter geen directe invloed op de inzet van individuele prestatiebeloning. Dit is momenteel een taak voor het management van scholen. In principe kan de overheid het aantal aanbieders in een bepaalde sector beïnvloeden. In de BVE-sector en in het HBO is het aantal aanbieders bijvoorbeeld sterk afgenomen onder invloed van beleid gericht op schaalvergroting. Het aantal aanbieders kan toenemen door te streven naar schaalverkleining. Een andere manier om het aantal aanbieders te vergroten is het streven naar een ‘open bestel’: private aanbieders in aanmerking laten komen voor publieke bekostiging. De implicaties van een ‘open bestel’ zijn op dit moment niet duidelijk en vallen buiten het bereik van deze studie.

Vergelijking van de sectoren in Tabel 5.1 laat zien dat er relatief weinig prestatieprikkels zijn in de BVE-sector. Ook valt op dat er de transparantie van de onderwijsmarkt vaak beperkt is en dat prestatiebeloning voor docenten niet wordt toegepast.

5.2 Prestaties en bekostiging

In het licht van de buitenlandse ervaringen lijkt het toepassen van teambeloning in Nederland zinvol maar niet zonder risico. De toepassing is het meest kansrijk in delen van het onderwijs waar heldere prioriteiten kunnen worden gesteld, meting van de onderwijsproductie relatief weinig problemen geeft en weinig risico bestaat op manipulatie. Toepassing lijkt vooral kansrijk in delen van het onderwijs met weinig prestatieprikkels zoals concurrentie tussen instellingen. Op grond van deze criteria zou de BVE-sector een kansrijk terrein kunnen zijn voor het toepassen van teambeloning. De prioriteit kan worden gelegd bij het voorkomen van voortijdig schoolverlaten. Dit wordt immers door velen gezien als het grootste knelpunt in het beroepsonderwijs. De huidige bekostiging voor het BVE is deels afhankelijk van het aantal geslaagden. Het systeem van teambeloning zou dan kunnen aangrijpen op het percentage uitvallers vergeleken met dat bij andere instellingen, rekeninghoudend met verschillen in achtergrondkenmerken van leerlingen. Teambeloning ligt in dat geval eerder voor de hand dan individuele prestatiebeloning, omdat de problemen van voortijdig schoolverlaters waarschijnlijk zelden zijn toe te schrijven aan individuele docenten. Instellingen of opleidingen binnen instellingen die het meest succesvol zijn in het terugdringen van voortijdig schoolverlaten ontvangen meer geld via de bekostiging of via teambeloning. Deze middelen kunnen vrij worden besteed, bijvoorbeeld voor extra beloning van docenten.

Het meten van de bijdrage van scholen aan de resultaten van leerlingen lijkt minder lastig dan in andere sectoren vanwege de monopoliepositie van veel BVE-instellingen. Omdat er weinig te kiezen valt is het risico op vertekening van de prestaties van scholen door niet-waargenomen factoren (selectie) minder groot. Anders gezegd, leerlingpopulaties op BVE-instellingen in verschillende regio's zullen niet sterk variëren, en na controle voor observeerbare kenmerken van leerlingen zou een redelijk betrouwbare schatting gemaakt kunnen worden van de prestaties van BVE-instellingen. Aangezien BVE-instellingen niet kunnen selecteren aan de poort speelt *gaming* op dit punt geen rol.

Beleids optie I: Gebruik sterkere financiële prikkels voor instellingen of opleidingen ter bestrijding van voortijdig schoolverlaten in de BVE-sector, hetzij via de bekostiging hetzij via teambeloning.

5.3 Prestaties en accountability

In Nederland is de discussie over *accountability* van onderwijsinstellingen in een stroomversnelling gebracht door de publicatie van een ranglijst van scholen uit het voortgezet onderwijs in het dagblad Trouw. Voorstanders van de publicatie betoogden dat ouders, kinderen en de overheid inzicht geboden moet worden in de prestaties van scholen. Voor ouders is dit belangrijk met het oog op de keuze van een school. Voor de overheid is de verantwoording over het gebruik van de beschikbaar gestelde publieke middelen van belang. Tegenstanders van de publicatie maakten zich zorgen over de kwaliteit van de ranglijsten. Worden de gemeten verschillen daadwerkelijk veroorzaakt door de school of spelen andere factoren, zoals de samenstelling van de leerlingenpopulatie, daarbij een rol? Vanuit het beleid is daarop gereageerd door het ontwikkelen van een kwaliteitskaart voor scholen in het voortgezet onderwijs. Kwaliteitskaarten vermelden verschillende resultaten van scholen, zoals het percentage geslaagden en de doorstroom van leerlingen. Er worden door de overheid geen ranglijsten gemaakt. Wel worden landelijke gemiddelden gepresenteerd van scholen die op grond van observeerbare leerlingkenmerken vergelijkbaar zijn.

Met het oog op het bevorderen van de prestaties van scholen via *accountability* is hier sprake van een dilemma. Enerzijds kunnen ranglijsten op basis van prestaties van scholen een krachtige prikkel geven voor verbetering van de kwaliteit van het onderwijs. Anderzijds laten recente studies zien dat zelfs ranglijsten die geconstrueerd zijn op basis van 'state of the art' technieken onbetrouwbaar zijn. Perverse prikkels liggen dan op de loer. *Accountability* behelst echter meer dan ranglijsten. Het publiek informeren over de resultaten op scholen, zoals dat bijvoorbeeld gebeurt via de kwaliteitskaarten, geeft een prikkel aan scholen om zich in te spannen om de kwaliteit van het onderwijs te verbeteren. Daarbij is het belangrijk om aan te geven dat *de resultaten op scholen* de uitkomst zijn van de prestaties van scholen, leerlingen en ouders, en derhalve niet gelijk hoeven te zijn aan *de prestaties van scholen*. In de huidige Nederlandse context is er nog de nodige winst te halen bij de informatievoorziening over de resultaten op scholen en onderwijsinstellingen. Voor instellingen in de BVE-sector en het hoger onderwijs bestaan (nog) geen kwaliteitskaarten. Het ontwikkelen van een kwaliteitskaart of vergelijkbaar instrument voor deze sectoren van het onderwijs kan de *accountability*-prikkels versterken. Daarnaast is niet duidelijk in hoeverre de huidige kwaliteitskaart voldoet. Wat is het bereik? Zijn de gegevens toegankelijk voor de doelgroepen? Bestaat er behoefte aan meer gegevens?

Beleidsopatie II: Verbeter de informatievoorziening aan ouders en leerlingen over de resultaten op scholen door het ontwikkelen van kwaliteitskaarten of soortgelijke instrumenten voor de BVE-sector en het hoger onderwijs.

5.4 Prestaties en beloning

Over de effecten van individuele prestatiebeloning in het onderwijs is tot dusver weinig bekend. Het experiment dat in Israël is uitgevoerd laat overtuigend zien dat de prestaties van leerlingen verbeteren. Dit suggereert dat toepassing van individuele prestatiebeloning in Nederland veelbelovend kan zijn. Binnen de huidige beleidscontext, die gericht is op het versterken van de autonomie van instellingen, is het echter aan het management van scholen om te beslissen over het toepassen van individuele prestatiebeloning. Daar komt bij dat scholen de afgelopen jaren gestart zijn met het invoeren van competentiebeloning en daarmee thans ervaringen opdoen. Mogelijk zullen onderwijsinstellingen er de voorkeur aan geven deze ervaringen af te wachten voor ze een nieuw traject willen starten met het invoeren van individuele prestatiebeloning. Dit betekent dat de mogelijkheden van de centrale overheid om beleid te voeren op dit terrein beperkt zijn. Beleidsopties bestaan uit het communiceren van de potentiële voordelen van individuele prestatiebeloning naar onderwijsinstellingen alsmede uit het inpassen van prestatiebeloning binnen de huidige competentiebeloning. Verder zouden zorgvuldig ontworpen experimenten meer inzicht kunnen verschaffen in de voor- en nadelen van individuele prestatiebeloning in het onderwijs. Gegeven dat instellingen hun eigen beloningsbeleid kunnen voeren, ligt het voor de hand om hierbij uit te gaan van vrijwillige deelname aan de experimenten.

Beleidsoptie III: Experimenteer met diverse vormen van individuele prestatiebeloning voor docenten.

Bijlage: Andere initiatieven met prestatieprikkels in het Verenigd Koninkrijk

In het VK wordt (naast de *Performance Threshold* voor leraren) een aantal andere interessante initiatieven ontplooid op het terrein van prestatieprikkels voor docenten en bestuurders in het onderwijs (zie ook Waterreus, 2001). De belangrijkste zijn:

- prestatiebeloning voor schoolhoofden;
- *Advanced Skills Teachers*;
- *fast track teachers*.

Genoemde initiatieven zijn pas korte tijd geleden gestart. Helaas zijn er (nog) geen grondige evaluatiestudies beschikbaar.

Prestatiebeloning voor schoolhoofden is ingevoerd begin jaren negentig. Ieder jaar evalueert het schoolbestuur de prestaties van het schoolhoofd, en bepaalt mede naar aanleiding hiervan het salaris. De prestaties worden beoordeeld aan de hand van vooraf opgestelde doelstellingen. Sinds kort wordt het bestuur bijgestaan door een externe deskundige. De vaststelling van het salaris heeft volgens een deskundige veel weg van ‘striking a deal’. Bij slecht functioneren kan het hoofd naar huis worden gestuurd door het schoolbestuur, maar dit komt slechts zeer zelden voor.

Advanced Skills Teachers (AST) zijn een soort ‘superdocenten’. Om te worden benoemd tot AST moet een uitgebreide *assessment* worden doorlopen, uitgevoerd door een externe deskundige die onder andere lessen bezoekt en met leerlingen praat. Iemand die als AST is aangesteld wordt geacht één dag in de week elders te gaan werken, om ook kennis te kunnen overdragen op leraren van andere scholen. Mede daarom zijn schoolhoofden er niet zo happig op om iemand als AST te benoemen. Scholen raken namelijk een goede leraar gedeeltelijk kwijt aan een andere school. Ook deze mogelijkheid om als AST te worden benoemd is wellicht eerder een vorm van competentiebeloning.

Fast Track Teachers zijn veelbelovende docenten met een min of meer uitgestippelde carrière. De loopbaanplanning houdt onder meer in dat leraren geregeld wisselen binnen of buiten de school, en de *Fast Track Teachers* ontvangen dubbele stijgingen van hun salaris ten opzichte van de andere docenten. Ook dit initiatief lijkt eerder op competentiebeloning dan prestatiebeloning.

Literatuur

- Baker, G., R. Gibbons, en K.J. Murphy, 1994, 'Subjective performance measures in optimal incentive contracts', *Quarterly Journal of Economics*, 109(4), 1125-1156.
- Ballou, D., 2001, 'Pay for performance in public and private schools', *Economics of Education Review*, 20, 51-61.
- Burgess, S., en B. Croxson, 2001, 'The impact of teacher pay reform on pupil attainment: An outline of the CMPO research project on the Performance Threshold', CMPO Working Paper, nummer 01/36.
- Burgess, S., B. Croxson, P. Gregg, en C. Propper, 2001, 'The intricacies of the relationship between pay and performance for teachers: Do teachers respond to performance related pay schemes?', CMPO Working Paper, nummer 01/35.
- Burgess, S., en P. Metcalfe, 2000, 'Incentives in organisations: A selective overview of the literature with application to the public sector', CMPO Working Paper, nummer 00/16.
- CPB, 2002, *De pijlers onder de kenniseconomie: Opties voor institutionele vernieuwing*, Centraal Planbureau en Koninklijke De Swart, Den Haag.
- CPB en CHEPS, 2001, *Higher education reform: Getting the incentives right*, Centraal Planbureau en Koninklijke De Swart, Den Haag.
- DfES, 2002, 'Education maintenance allowance: The first two years. A quantitative evaluation', Research Report, nummer 352, Department for Education and Skills, Londen.
- Dixit, A., 1997, 'Power of incentives in private versus public organizations', *American Economic Review*, 87(2), 378-382.
- Dixit, A., 2002, 'Incentives and organizations in the public sector: An interpretative review', *Journal of Human Resources*, 37(4), 696-727.
- Eberts, R., K. Hollenbeck, en J. Stone (2000), 'Teacher performance incentives and student outcomes', *Journal of Human Resources*, 37(4), 913-927.
- Figlio, D.N., en L.S. Getzler, 2002, 'Accountability, ability and disability: Gaming the system', NBER Working Paper, nummer 9307.

- Frey, B.S., en R. Jegen, 2001, 'Motivation crowding theory', *Journal of Economic Surveys*, 15(5), 589-611.
- Glewwe, P., N. Ilias, en M. Kremer, 2003, 'Teacher incentives', NBER Working Paper, nummer 9671.
- Hanushek, E.A., 2002, 'Publicly provided education', NBER Working Paper, nummer 8799.
- Hassink, W., 2000, 'Leraren, tekorten en salarissen; Een empirisch onderzoek voor het voortgezet onderwijs', CPB Werkdocument, nummer 129.
- Heckman, J.J., J.A. Smith, en C. Taber, 1996, 'What do bureaucrats do? The effects of performance standards and bureaucratic preferences on acceptance into the JTPA program', NBER Working Paper, nummer 5535.
- Hoxby, C., 2000, 'Peer effects in the classroom: Learning from gender and race variation', NBER Working Paper, nummer 7867.
- Inspectie van het Onderwijs, 2003, *Onderwijsverslag over het jaar 2002*, Inspectie van het Onderwijs, Utrecht.
- Jacob, B.A., 2002, 'Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago public schools', NBER Working Paper, nummer 8968.
- Kane, T.J., en D.O. Staiger, 2001, 'Improving school accountability measures', NBER Working Paper, nummer 8156.
- Kane, T.J., en D.O. Staiger, 2002, 'The promise and pitfalls of using imprecise school accountability measures', *Journal of Economic Perspectives*, 16(4), 91-114.
- Knez, M., en D. Simester, 2001, 'Firm-wide incentives and mutual monitoring at Continental Airlines', *Journal of Labor Economics*, 19(4), 743-772.
- Koning, P., E. Canton, M. Cornet, M. Pomp, J. van de Ven, R. Venniker, B. Vollaard, en D. Webbink (2004), 'Centrale doelen, decentrale uitvoering. Over de do's and don'ts van prestatieprikkels voor semi-publieke instellingen', CPB Document, nummer 45.
- Kreps, D.M., 1997, 'Intrinsic motivation and extrinsic incentives', *American Economic Review*, 87(2), 359-364.

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

Ladd, H.F., en R.P. Walsh, 2002, 'Implementing value-added measures of school effectiveness: Getting the incentives right', *Economics of Education Review*, 21, 1-17.

Lavy, V., 2002a, 'Evaluating the effect of teachers' group performance incentives on pupil achievement', *Journal of Political Economy*, 110(6), 1286-1317.

Lavy, V., 2002b, 'Paying for performance: The effect of teachers' financial incentives on students' scholastic outcomes', Mimeo, Universiteit van Jerusalem.

Lazear, E.P., en S. Rosen, 1981, 'Rank-order tournaments as optimum labor contracts', *Journal of Political Economy*, 89(5), 841-864.

Lazear, E.P., 1996, 'Performance pay and productivity', NBER Working Paper, nr. 5672.

Lazear, E.P., 1999, 'Personnel economics: Past lessons and future directions. Presidential address to the Society of Labor Economists, San Francisco, May 1, 1998', *Journal of Labor Economics*, 17(2), 199-236.

Mendro, B., R. Olivarez, M. Peters, T. Milanowski, E. Kellor, en A. Odden, 1999, 'A case study of the Dallas public school-based performance award program', Mimeo, Consortium for Policy Research in Education.

Osterloh, M., en B.S. Frey, 2000, 'Motivation, knowledge transfer, and organizational form', *Organization Science*, 11(5), 538-550.

Prendergast, C., 1999, 'The provision of incentives in firms', *Journal of Economic Literature*, 37(1), 7-63.

Todd, P.E., en K.I. Wolpin, 2003, 'On the specification and estimation of the production function for cognitive achievement', *Economic Journal*, 113(485), F3-F33.

Waterreus, I.M., 2001, *Incentives in Secondary Education*, Max Goote Kenniscentrum voor Beroepsonderwijs en Volwasseneneducatie, Universiteit van Amsterdam.

Waterreus, I.M., 2002, 'School performance measurement and teacher mobility in the Netherlands', Mimeo, Universiteit van Amsterdam.

Abstract

This study analyses performance incentives in the Dutch education sector. We explore existing incentives in the Dutch system connected with the design of the public funding system, accountability practices, and remuneration policies. Next we deal with performance measurement: what problems are encountered and what are the implications for the design of performance indicators? As several countries gained experience with incentives for schools and teachers, we review the literature and try to draw lessons. Introduction of team performance pay in the United States and Israel has been successful, especially with regard to reduction of early drop-out. The report concludes by sketching policy options for the Netherlands. More high-powered incentives for schools to retain students at risk, improvement of transparency, and experiments with individual performance pay can improve efficiency and educational attainment.