

Kwaliteitsmeting middelbare scholen beloont selectie

Erik Schut en Margreet den Haan

De huidige kwaliteitsbeoordeling van middelbare scholen houdt onvoldoende rekening met kwaliteitsverschillen tussen leerlingen. Hierdoor wordt selectie beloond en worden onderwijsprestaties vertekend. Vanaf 2016 wordt een nieuwe methode om leerresultaten te meten geïntroduceerd. Deze nieuwe beoordelingsmethode zal de genoemde problemen echter niet verkleinen en zal bovendien leiden tot een ongelijk speelveld tussen gemengde en categorale middelbare scholen. Ook het afschaffen van de Cito-toets als selectie-instrument in 2015 kan deze problemen niet oplossen en spant het paard achter de wagen. Effectiever is om de Cito-toets te hanteren om de kwaliteitsmeting van middelbare scholen te corrigeren voor selectie-effecten. Uit een analyse op basis van cijfers van het centraal eindexamen van vwo-leerlingen blijkt namelijk dat deze positief samenhangen met hun Cito-score van de basisschool.

1 Inleiding

Het kabinet wil selectie van leerlingen door middelbare scholen op basis van hun score op de Cito-eindtoets van de basisschool onmogelijk maken. Vanaf 2015 mogen middelbare scholen de Cito-score niet langer hanteren als toelatingscriterium. Voortaan mogen zij de toelating van nieuwe leerlingen uitsluitend baseren op het advies van de basisschool. Een middelbare school moet een leerling die bijvoorbeeld een havo-advies krijgt minimaal op havo-niveau plaatsen en mag de leerling niet weigeren als zij de Cito-score of andere scores te laag vindt. De Cito-eindtoets blijft wel bestaan¹, maar dient uitsluitend nog als *second opinion* om een mogelijk te laag basisschooladvies te ontdekken. Als een leerling de eindtoets namelijk beter maakt dan verwacht op grond van het advies, moet de basisschool het schooladvies heroverwegen. Wanneer het resultaat van de eindtoets minder goed is dan verwacht, mag de basisschool het schooladvies echter niet aanpassen.

¹ Vanaf het schooljaar 2014-2015 is het voor alle leerlingen van groep 8 in het reguliere basisonderwijs zelfs verplicht om een Centrale eindtoets Primair Onderwijs te maken. De overheid stelt hiervoor aan scholen de Centrale Eindtoets – beter bekend als de Cito-eindtoets – beschikbaar. Scholen kunnen ook kiezen voor twee andere, door de minister van Onderwijs toegelaten eindtoetsen. De ervaring uit eerdere jaren wijst uit dat veruit de meeste scholen kiezen voor de Cito-eindtoets (College voor Toetsen en Examens 2014).

Hoewel het voor middelbare scholen hierdoor lastiger wordt om de beste leerlingen te selecteren, blijft selectie wel onverminderd lonend. Bij de kwaliteitsmeting van middelbare scholen wordt namelijk nauwelijks gecorrigeerd voor selectie-effecten. Scholen die betere leerlingen aantrekken – hetzij actief, hetzij passief (via zelfselectie) – of minder goede leerlingen ontmoedigen om een zo hoog mogelijke opleiding te volgen, kunnen beter scoren op de prestatie-indicatoren die door de Inspectie van het Onderwijs worden gehanteerd. Aangezien deze prestatie-indicatoren ook leidend zijn voor de openbare keuze-informatie over middelbare scholen, kunnen scholen die succesvol selecteren meer leerlingen werven. Met het wegnemen van de Cito-toets als selectie-instrument wordt feitelijk het paard achter de wagen gespannen. Immers, zolang selectie lonend blijft, blijft het voor middelbare scholen aantrekkelijk om betere leerlingen te selecteren. Begin 2015 bleek dan ook dat tal van middelbare scholen naar wegen zoeken om goede leerlingen te selecteren en werd staatssecretaris Dekker hierover door de Tweede Kamer ter verantwoording geroepen (Van Beeten 2015, Kuiper 2015, Tweede Kamer 2015, Vasterman 2015a). Maar wellicht belangrijker is dat scholen worden ontmoedigd om de minder goede leerlingen te stimuleren tot het volgen van een zo hoog mogelijke opleiding door het maximale uit zichzelf te halen. Uit het recente jaarlijkse Onderwijsverslag van de Onderwijsinspectie blijkt dat de striktere selectie van middelbare scholen ertoe leidt dat leerlingen steeds minder ‘opstromen’ naar het vwo en dat bredere brugklassen met meerdere niveaus geleidelijk verdwijnen (Inspectie van het Onderwijs 2015, Vasterman 2015b). De kansen op een betere opleiding voor laatbloeiers en kinderen die van huis uit minder gestimuleerd worden, lijken af te nemen.

In dit artikel zetten wij allereerst uiteen hoe de onderwijsinspectie de kwaliteit van middelbare scholen beoordeelt. Hierna analyseren wij waarom deze kwaliteitsmeting wordt vertekend door selectie-effecten door te onderzoeken in hoeverre de prestaties worden beïnvloed door meetbare kwaliteitsverschillen tussen leerlingen. Vervolgens bespreken wij de gevolgen hiervan voor de schoolkeuze en de concurrentieverhoudingen tussen gemengde en categorale scholen. Ten slotte geven wij aan hoe de kwaliteitsmeting kan worden verbeterd door de resultaten van de eindtoets van de basisschool te gebruiken voor correctie van de prestatie-indicatoren.

2 Kwaliteitsindicatoren onderwijsinspectie

In Nederland is de Inspectie van het Onderwijs verantwoordelijk voor het toezicht op de kwaliteit van het onderwijs. Voor een effectief toezicht is het van belang om goed inzicht te hebben in de onderwijskwaliteit die scholen leveren. Wereldwijd is er toenemende aandacht voor het meten van schoolprestaties (Loeb en Figlio 2011, Timmermans et al. 2015). De drijfveer hiervoor is dat belanghebbenden bij goed onderwijs – leerlingen, ouders, bedrijven en beleidsmakers – vaak onvoldoende zicht hebben op de kwaliteit van de schoolprestaties (het bekende principaal-agentprobleem). Het meten, vergelijken en beoordelen van de kwaliteit van scholen

is echter lastig vanwege het multidimensionale karakter van kwaliteit en de veelheid aan factoren die de kwaliteit beïnvloeden (Loeb en Figlio 2011). Idealiter wordt de kwaliteit van het onderwijs op een school gemeten aan de hand van de bijdrage die deze school levert aan de ‘leerwinst’ op alle relevante dimensies die leerlingen boeken gedurende hun schoolopleiding. In de praktijk blijkt de toegevoegde waarde van scholen aan de leerwinst moeilijk goed te bepalen. Ten eerste omdat niet alle relevante dimensies (met name de non-cognitieve vaardigheden) goed meetbaar zijn. Maar ook omdat een groot aantal ‘buitenschoolse’ factoren de leerwinst bepalen, waarvan niet alle gegevens op individueel niveau beschikbaar zijn of het aantal waarnemingen uit oogpunt van statistische betrouwbaarheid (te) beperkt is. Wanneer scholen worden vergeleken op basis van meetbare toegevoegde waarde, bestaat het risico dat scholen zich vooral gaan richten op de meetbare kwaliteitsdimensies (‘teaching to the test’) of dat scholen gaan selecteren op buitenschoolse factoren die effect hebben op de leerwinst waarvoor niet wordt gecorrigeerd, bijvoorbeeld de sociaal-economische achtergrond van de leerlingen (Loeb en Figlio 2011). In hoeverre gegevens over ‘leerwinst’ en ‘toegevoegde waarde’ bruikbaar zijn voor het beoordelen van de kwaliteit van scholen is dan ook omstreven. Het toenemend gebruik van ‘Value-Added Models’ (VAM) bij de beoordeling van leraren en scholen in de Verenigde Staten vormde voor de American Statistical Association (2014) de aanleiding om een Statement te publiceren over de mogelijkheden en beperkingen van deze modellen. In Nederland is door Timmermans et al. (2011) onderzocht in hoeverre verschillende VAM-modellen de verschillen in eindexamenresultaten tussen vmbo(tl)-scholen en tussen havoscholen kunnen verklaren. De onderzoekers vinden dat de resultaten van de verschillende modellen aanzienlijk uiteenlopen en afhankelijk zijn van de keuze van de verklarende variabelen. Dit maakt het in hun ogen riskant om dergelijke modellen te gebruiken om de prestaties van scholen te vergelijken. Op basis van hun bevindingen adviseren zij om in elk geval te corrigeren voor het aanvangsniveau van de scholieren (de belangrijkste verklarende variabele), indicatoren van hun sociaal-economische status en kenmerken van de samenstelling van de schoolpopulatie. Recent concludeerde de Onderwijsraad (2014) dat leerwinst en toegevoegde waarde weliswaar goede instrumenten zijn om scholen te helpen hun onderwijs te verbeteren, maar dat zij niet bruikbaar zijn voor een oordeel over de kwaliteit van de school als geheel. Op grond van een pilot onder 25 scholen concludeerde ook de Inspectie van het Onderwijs (2014a) dat leerwinst goed te onderzoeken is, maar dat de invloed van de school daarop (de toegevoegde waarde) lastiger is te bepalen.

In weerwil van deze discussie over de mogelijkheid om de toegevoegde waarde van scholen adequaat te *kunnen* meten, *worden* in Nederland de opbrengsten of onderwijsprestaties van een school door de Inspectie van het Onderwijs gemeten, waarbij een viertal ‘opbrengstindicatoren’ worden gehanteerd (zie kader).

Kwaliteitscriteria onderwijsinspectie

De Inspectie van het Onderwijs (2011, 2013) hanteert de volgende vier opbrengstindicatoren voor de beoordeling van scholen in het voortgezet onderwijs (VO):

1. *Rendement onderbouw: gebaseerd op (i) de positieve of negatieve afwijking van de positie leerling in derde leerjaar ten opzichte van het basisschooladvies, en (ii) het aandeel leerlingen dat in leerjaar 1 en 2 is gedoubleerd.*
2. *Rendement bovenbouw: gebaseerd op de kans op een onvertraagde doorstroom van leerjaar 3 tot diploma, ofwel de kans dat leerlingen zonder zitten-blijven hun diploma halen.*
3. *Gemiddeld cijfer Centraal Examen (CE);*
4. *Verskil tussen het gemiddelde cijfer van het schoolexamen (SE) en dat van het Centraal Examen (CE);*

De indicatoren worden berekend als gewogen driejaargemiddelde. Per opleiding per school wordt bij de eerste drie indicatoren door middel van een relatieve score (5-puntsschaal) aangegeven hoe deze score ten opzichte van dezelfde opleidingen op vergelijkbare scholen. Er worden 7 vergelijkingsgroepen onderscheiden, waarbij havo/vwo-vestigingen en categorale vwo-vestigingen elk een aparte groep vormen. Bij de eerste drie indicatoren beschouwt de Inspectie een score van 2,0 of minder (laagste kwartiel) als onvoldoende, terwijl een score op de laatste indicator onvoldoende is als het verschil meer dan 0,5 punt bedraagt. Wanneer een opleiding op meer dan één indicator onvoldoende scoort, is het totaaloordeel onvoldoende en krijgt de betreffende opleiding van de Inspectie het predicaat “zwak” of “zeer zwak” (indien ook het onderwijsproces als onvoldoende of de school langer dan twee jaar als “zwak” wordt beoordeeld). Het oordeel van Inspectie over de opleidingen per school is openbaar toegankelijk via de website van de Inspectie www.onderwijsinspectie.nl.

Sinds 2011 publiceert de Inspectie ook het absolute gemiddeld CE-cijfer (driejaargemiddelde) als nieuwe indicator. Hoewel deze indicator nog geen rol speelt bij het bepalen van het opbrengstenoordeel wordt wel aangegeven of een opleiding boven of onder de norm scoort (voor vwo vastgesteld op 6,0). In tegenstelling tot bij de relatieve score wordt hierbij geen onderscheid gemaakt tussen verschillende type vestigingen (zoals havo/vwo en categoriaal vwo).

Vanaf 2016 wordt een nieuw onderwijsresultatenmodel ingevoerd waarbij de huidige relatieve scores worden vervangen door absolute scores (Ministerie van OCW 2015). Vanaf dat moment worden opleidingen beoordeeld op basis van hun gemiddelde CE-cijfer, waarbij niet alleen wordt gekeken naar het al dan niet voldoen aan de norm, maar ook of een school “ruim boven de norm” presteert. Vwo-opleidingen scoren onder de norm bij een gemiddeld CE-cijfer lager dan 6,30 en ruim boven de norm bij een gemiddelde CE-cijfer hoger dan 6,61 (Inspectie van het Onderwijs 2014b).

Het opbrengstenoordeel van de Inspectie over een school is openbaar en wordt jaarlijks gepubliceerd in de vorm van opbrengstenoverzichten (over meerdere jaren) en opbrengstenkaarten (over één jaar) op de website www.onderwijsinspectie.nl. De gedachte achter de vier gehanteerde indicatoren is dat zij gezamenlijk een evenwichtig beeld geven van de onderwijsprestaties van een school (‘balanced score model’). Zo geven de eerste drie indicatoren scholen tegengestelde prikkels waardoor strategisch gedrag wordt tegengegaan. Een school kan bijvoorbeeld goede resultaten halen in de onderbouw door veel leerlingen te laten doorstromen naar de bovenbouw, maar als dit te gemakkelijk gebeurt, zal dit leiden tot tegenvallende resultaten in de bovenbouw of slechte examencijfers. Omgekeerd zal een strenge selectie in de onder- en bovenbouw leiden tot betere eindexamenresultaten, maar lagere doorstroomrendementen. De vierde indicator

staat meer op zichzelf en dient om te voorkomen dat scholen via te eenvoudige schoolexamens de eindexamenresultaten opkrikken.

De indicatoren nemen echter lang niet alle prikkels voor strategisch gedrag weg. Zowel het doorstroomrendement als de eindexamencijfers kunnen namelijk door scholen positief worden beïnvloed door selectie van goede leerlingen. De prikkel voor selectie wordt door de Onderwijsinspectie wel enigszins gemitigeerd door het bovenbouwrendement en het gemiddelde CE-cijfer te corrigeren voor het aandeel leerlingen uit achterstandswijken – zogeheten armoede-probleem-cumulatiegebieden (apcg) – en het aandeel zij-instromers (leerlingen die bij een andere school zijn weggegaan of weggestuurd). De ervaring wijst namelijk uit dat deze ‘risicoleerlingen’ gemiddeld meer moeite hebben om zonder vertraging en met mooie cijfers het eindexamen te halen. Zo wordt voor vwo-scholen met minimaal 30 procent leerlingen uit achterstandswijken het normatieve gemiddelde CE-cijfer verlaagd met 0,34 punt vermenigvuldigd met het aandeel van deze leerlingen in het totaal aantal eindexamenleerlingen. De correcties voor risicoleerlingen houden echter maar zeer ten dele rekening met de capaciteitsverschillen tussen leerlingen bij instroom in het voortgezet onderwijs. Wanneer met deze capaciteitsverschillen onvoldoende rekening wordt gehouden, kunnen de kwaliteitsscores van de inspectie worden vertekend door selectie-effecten, hetzij zelfselectie door ouders en leerlingen hetzij selectie door scholen. De indicatoren meten dan geen goede scholen, maar goede leerlingen. Bij gelijke onderwijskwaliteit zullen scholen met goede leerlingen dan beter scoren dan scholen met minder goede leerlingen.

3 Cito-score als indicator

Een belangrijke algemene indicator van capaciteitsverschillen tussen leerlingen is hun score op de Cito-toets aan het eind van de basisschool (Kuyper en Van der Werf 2012). Samen met het advies van de basisschool over de leerprestaties van de leerling vormde de Cito-toets tot 2014 een belangrijk toelatingscriterium voor scholen in het voortgezet onderwijs. Zo werd een Cito-score van 538 – 545 doorgaans gehanteerd voor toelating tot een gemengde havo/vwo brugklas, terwijl een score van 545 – 550 werd gehanteerd voor toelating tot ongedeelde vwo (atheneum, gymnasium). Scholen konden zelf hun toelatingseisen bepalen. Zo waren er categorale gymnasia die een minimumscore van 548 vereisten, terwijl op sommige scholen met een gemengde brugklas van vmbo tot vwo in het extreme geval zelfs een leerling met een score van 528 nog in het vwo terecht kon komen (Elzinga 2011). Wanneer een hogere Cito-score samenhangt met een hoger doorstroomrendement en hogere eindexamencijfers zullen opleidingen met leerlingen met een hogere Cito-score op de kwaliteitsindicatoren van de inspectie hoger scoren dan vergelijkbare opleidingen met leerlingen met een lagere Cito-score, zonder dat hun onderwijsprestaties beter hoeven te zijn. Hoewel deelname aan de Cito-toets vrijwillig was, nam in 2014 85 procent van de basisscholen

hieraan deel, terwijl 10 procent koos voor een andere eindtoets en 5 procent aan geen enkele eindtoets deelnam (College voor Toetsen en Examens 2014). Gelet op de hoge deelname aan de Cito-toets is het opmerkelijk dat met de verschillen in Cito-score van de leerlingen bij het bepalen van de kwaliteit van de onderwijsprestaties geen rekening wordt gehouden. Nog opmerkelijker is dat er over het verband tussen Cito-score, doorstroomrendement en eindexamencijfers maar weinig bekend lijkt te zijn.

4 Cito-score en doorstroomrendement

Het enige onderzoek naar het verband tussen Cito-score en doorstroomrendement is uitgevoerd door het Gronings Instituut voor Onderzoek van Onderwijs (Kuyper en Van der Werf 2012). Uit dit onderzoek op basis van gegevens van circa 1400 leerlingen die in 1999 aan het voortgezet onderwijs begonnen (het zogeheten VOCL'99-cohort), blijkt dat er tussen Cito-score en rendement een sterke samenhang bestaat. Uit Tabel 1, waarin de bevindingen van het GION zijn samengevat, blijkt dat de kans dat een leerling zonder vertraging het vwo-diploma haalt sterk toeneemt met de Cito-score: voor elke extra punt boven de 545 neemt de onvertraagde slaagkans 7 tot 10 procentpunt toe. Opvallend is voorts dat vooral het onderbouwrendement sterk toeneemt met een stijgende Cito-score. Het effect van de Cito-score op het doorstroomrendement wordt kleiner naarmate leerlingen in een hoger leerjaar terechtkomen. Bij de leerlingen die zonder vertraging het eindexamenjaar halen lijkt de Cito-score nog maar van beperkte invloed op hun uiteindelijke slaagkans.

Tabel 1 Percentages leerlingen op koers voor vwo-diploma, uitgesplitst naar Cito-score

Cito-score	Op koers in leerjaar (%)						Vwo-diploma
	1	2	3	4	5	6	
545	90,3	79,7	63,8	52,8	44,5	39,5	36,3
546	92,6	78,7	65,5	57,3	49,4	45,5	43,1
547	94,5	84,9	74,6	66,9	55,8	54,0	50,7
548	95,6	89,1	80,2	72,3	68,1	63,3	60,4
549	98,1	96,2	91,4	85,4	79,7	73,9	71,5
550	98,6	97,2	94,4	91,1	87,7	82,3	80,9

Bron: Kuyper en Van der Werf (2012), p.36 (Tabel 3.16).

Duidelijk is dus dat scholen hun doorstroomrendement – en dus hun kwaliteitsscore – kunnen verhogen door nieuwe leerlingen te selecteren op een hoge Cito-score, of door leerlingen met een relatief lage Cito-score te ontmoedigen vanuit de havo of gemengde brugklas naar het vwo ‘op te stromen’. Tot op zekere hoogte is dit wenselijk, omdat dit een te ruimhartig acceptatiebeleid ontmoedigt. Een te soepel acceptatie- en opstroombeleid kan leiden tot onnodige uitval en vertraging en is

niet alleen kostbaar (Van Vuuren et al. 2015) maar leidt ook tot frustratie bij zowel leerlingen als docenten. Een nadeel van een strenge selectie is echter dat leerlingen die ondanks een lagere Cito-score wel de potentie hebben om het vwo succesvol af te ronden (zoals de ‘laatbloeiers’ en ‘vechters’) hiertoe de kans wordt ontnomen. Onderzoek wijst uit dat vroege selectie een negatief effect heeft op de deelname aan en het afronden van het hoger onderwijs (Van Elk et al. 2011). Ongeveer een kwart van de leerlingen blijkt in de eerste drie jaren van het VO te veranderen van onderwijsniveau en dit geldt in grotere mate voor leerlingen die in het eerste jaar uitsluitend zijn geplaatst op grond van hun Cito-score (Feron et al. 2015; Ter Weel 2015). Veranderen van onderwijsniveau is vooral mogelijk op scholen met een gemengde brugklas.

5 Cito-score en eindexamencijfers

Ook onderzoek naar het verband tussen Cito-score en CE-cijfers is uiterst schaars. Timmermans et al. (2011) onderzochten op basis van leerlingen die in 1999 aan de vmbo(tl) en havo begonnen (het VOCL’99-cohort) de relatie tussen hun eindexamencijfer en hun score op de entree Cito-toets in groep 7 van de basisschool (‘prior achievement’) en andere leerlingkenmerken, zoals intelligentie, leeftijd, geslacht, SES, ethniciteit en tweede taal. Bij vmbo-scholen blijkt dat slechts 12% van de variatie in de eindexamenresultaten tussen deze scholen kan worden toegeschreven aan de scholen zelf, terwijl een veel groter deel van de variatie (42%) kan worden verklaard door verschillen tussen scholieren, waarbij het leeuwendeel (35%) voor rekening komt van de verschillen in hun Cito-score op de entreetoets. Ook op het niveau van de individuele vmbo-scholieren blijkt veruit het grootste deel van de verklaarde variatie toe te schrijven aan verschillen in Cito-score (17% van de in totaal 27%). Naast Cito-score (aanvangsniveau) hadden ook intelligentie, SES en leeftijd een significant effect op de eindexamenresultaten. Voor havo-scholen blijkt hetzelfde te gelden maar ligt de verklaarde variantie lager. Ook uit studies in een aantal andere landen bleken aanvangsniveau (‘prior achievement’) en sociaal-economische status (SES) de belangrijkste contextuele variabelen voor het verklaren van verschillen in studieprestaties (Timmermans et al. 2015).

De vraag is of ook binnen het vwo de Cito-score van leerlingen die het zesde leerjaar hebben bereikt voorspellende waarde heeft voor hun CE-cijfer. Om deze vraag te beantwoorden hebben wij op basis van gegevens over Cito-scores en eindexamenresultaten van een school met een gemengde havo/vwo-brugklas over de periode 2010-2014 onderzocht of er een verband bestaat tussen het gemiddeld CE1-cijfer (CE zonder herexamen) per leerling en hun Cito-score van de basisschool. In vergelijking met de studie van Timmermans et al. (2011) heeft dit onderzoek betrekking op een ander schooltype en een andere (maar vergelijkbare) Cito-toets, op slechts één school (waardoor verschillen tussen scholen geen rol kunnen spelen bij de variatie in eindexamencijfers) en op vijf opeenvolgende jaren (in plaats van slechts één jaar).

In de periode 2010-2014 deden 230 leerlingen op deze school het vwo-eindexamen (gemiddeld 46 per jaar), waarbij van 200 leerlingen hun Cito-score in groep acht van de basisschool bekend was (87 procent, vergelijkbaar met het landelijk gemiddelde). Het verband tussen CE-cijfers en Cito-score is geschat door middel van lineaire regressie, zowel voor elk jaar afzonderlijk als voor de vijf jaren tezamen. De resultaten zijn weergegeven in Tabel 2.

Tabel 2 Verband tussen Cito-score en gemiddeld cijfer centraal examen vóór herexamen (CE1)

	2010	2011	2012	2013	2014	2010-2014
Effect Cito-score op CE1-cijfer ^a	0,076**	0,096*	0,076*	0,068*	0,063**	0,076*
2010 (=1) ^b	(0,032)	(0,022)	(0,022)	(0,024)	(0,024)	(0,012)
2011 (=1) ^b						-0,528**
2012 (=1) ^b						(0,178)
2013 (=1) ^b						-0,560**
						(0,178)
						-0,325
						(0,176)
						0,136
						(0,175)
N	40	40	42	43	35	200
R ²	0,12	0,34	0,23	0,16	0,17	0,26
Gemiddeld CE1-cijfer	5,66	5,60	5,75	6,27	6,09	5,87
Gemiddelde Cito-score eindexamen-kandidaten	542,6	542,2	541,1	542,0	541,4	541,8

* Significant op 1-procentniveau; **significant op 5-procentniveau.

^a Standaardfout tussen haakjes.

^b De jaardummies geven de verschillen in gemiddeld CE1-cijfer in de verschillende jaren aan ten opzichte van 2014 bij een gepoolde regressie over de gehele periode (laatste kolom). Zo lag het gemiddelde CE1-cijfer in 2010 en 2011 circa een halve punt lager dan in 2014. Vanaf 2012 moeten leerlingen voor hun CE (na herexamen) gemiddeld minimaal 5,5 halen om te kunnen slagen.

Uit Tabel 2 blijkt een significant positief verband tussen de Cito-score van de eindexamenleerling op de basisschool en zijn of haar gemiddelde CE1-cijfer. Over de vijfjaarperiode 2010-2014 gaat een toename van de Cito-score met één punt gepaard met een 0,076 punt hoger CE1-cijfer. Gemiddeld wordt ongeveer een kwart van de variantie in CE1-cijfers verklaard door de verschillen in Cito-score, waarmee verklaarde variantie vergelijkbaar is met die van de eerdere studie van Timmermans et al. (2011) onder vmbo(tl)- en havo-scholieren. Het verband tussen CE1-cijfer en Cito-score blijkt bovendien tamelijk robuust over de vijf eindexamenjaren. De mate waarin het CE1-cijfer toeneemt per extra Cito-punt varieert in de onderzoeksperiode van 0,063 tot 0,096 punt. Zoals ook blijkt uit de tabel ligt de gemiddelde Cito-score van de eindexamenkandidaten op deze school met ongeveer 542 ruim beneden de ondergrens van 545 die door veel ongedeelde vwo-scholen tot

nu toe wordt gehanteerd. Bij een gemiddelde van 545 zou op basis van de schattingresultaten het gemiddelde CE1-cijfer naar verwachting 0,23 punt ($3 * 0,076$) hoger liggen. Wanneer, zoals tot voor kort bij sommige categorale gymnasia, alleen leerlingen met een Cito-score van 548 en hoger zouden zijn toegelaten, zou het gemiddelde CE1-cijfer op deze school naar verwachting ongeveer een halve punt hoger zijn uitgekomen.

6 Onderwijsprestaties vertekend door selectie

Hoewel het onderzoek betrekking heeft op de gegevens vwo-scholieren van één school, komen de resultaten overeen met de bevindingen van Timmermans et al. (2011) onder een omvangrijk cohort leerlingen van verschillende vmbo(tl)- en havo-scholen. Hoewel nader onderzoek zal moeten uitwijzen of de resultaten ook gelden voor vwo-leerlingen op andere scholen, is er op voorhand geen reden waarom het verband tussen Cito-score en eindexamencijfers op die scholen sterk zou verschillen. Wanneer hetzelfde verband ook voor andere scholen geldt, is duidelijk dat scholen die in het verleden streng hebben geselecteerd een hoger gemiddeld CE-cijfer kunnen halen dan scholen waarbij ook leerlingen met een lagere Cito-score een kans krijgen. Ook wanneer scholen bij instroom vanaf heden niet meer effectief kunnen selecteren, zullen zij minimaal nog zes jaar profiteren van deze strenge selectie omdat pas dan de meeste leerlingen van het in 2014 ingestroomde cohort eindexamen zal hebben gedaan. Wanneer de resultaten ook gelden voor andere scholen geldt bovendien dat scholen die leerlingen met een relatief lage Cito-score aanmoedigen om ‘op te stromen’ naar het vwo zichzelf in de vingers snijden omdat dit naar verwachting leidt tot lagere eindexamencijfers.

De kans dat scholen die tot op heden een streng acceptatiebeleid hebben gevoerd door de Inspectie als ‘zwak’ worden beoordeeld is gering. Ook zal het voor deze scholen relatief eenvoudig zijn om in het met ingang van 2016 nieuw in te voeren beoordelingsmodel van de Inspectie (zie kader) de komende jaren ‘ruim boven de norm’ te scoren. Bovendien worden de kwaliteitsindicatoren van de Inspectie ook gehanteerd voor de samenstelling van de verschillende ranglijsten van scholen die beogen om ouders en leerlingen in staat te stellen in hun omgeving de beste school te kiezen. Tot 2014 konden ouders en leerlingen hiervoor terecht bij de jaarlijkse ranglijsten. De Volkskrant (en daarvoor Trouw) en het weekblad Elsevier. Vanaf 2014 is het stokje van De Volkskrant overgenomen door het Centrum voor Hoger Onderwijs Informatie (CHOI). Met een Keuzegids Middelbare Scholen (www.keuzegids.org/middelbarescholen) en de gerelateerde website www.kiesjeschool.nl beoogt het CHOI-ouders en -leerlingen te informeren over de kwaliteit van scholen en een overzicht te bieden van ‘de beste scholen’ per provincie. Beide ranglijsten zijn gebaseerd op de kwaliteitsindicatoren van de Inspectie (inclusief het absolute gemiddelde CE-cijfer) en geven dus een door selectie-effecten vertekend beeld van de onderwijsprestaties van de betreffende scholen. Door de uiteenlopende wijze waarop de verschillende indicatoren worden gewo-

gen, geselecteerd en geïnterpreteerd leveren de verschillende ranglijsten overigens geen eenduidig beeld op (Abels, 2014). In elk geval geven de ranglijsten een vertekend beeld van de feitelijke onderwijsprestaties van scholen, waardoor ouders en leerlingen een verkeerde schoolkeuze kunnen maken.

7 Ongelijk speelveld tussen gemengde en categorale scholen

Met ingang van 2015 mogen middelbare scholen de toelating niet meer laten afhangen van het resultaat van de eindtoets maar uitsluitend van het advies van de basisschool. Dit advies moeten basisscholen uiterlijk 1 maart afgeven, terwijl de eindtoets voortaan pas tussen 15 april en 15 mei wordt afgenomen.

De eindtoets is bovendien niet langer vrijwillig maar verplicht voor alle basisschoolleerlingen, waarbij basisscholen kunnen kiezen tussen een Cito-eindtoets en een tweetal andere, door de overheid toegelaten, eindtoetsen. Middelbare scholen mogen leerlingen geen extra toetsen laten afleggen om het niveau te bepalen en evenmin de toelating baseren op andere toetsen die leerlingen op de basisschool maken (bijvoorbeeld de entreetoets in groep 7). De basisschool mag deze gegevens overigens wel betrekken bij het schooladvies. Ook bij overaanmelding mogen toetsscores niet als toelatingscriterium worden gebruikt, maar moet er bijvoorbeeld worden geloot.

De nieuwe regels maken selectie lastiger, maar niet onmogelijk. Zo hoeven categorale vwo-scholen alleen leerlingen toe te laten met een enkelvoudig vwo-advies en mogen zij leerlingen met een gemengd havo/vwo-advies weigeren (Tweede Kamer 2015). Daarentegen moeten scholen met een gemengde havo/vwo-brugklas in principe ook leerlingen met een enkelvoudig havo-advies toelaten. Bovendien kunnen middelbare en basisscholen over de onderbouw van de basisschooladviezen allerlei afspraken maken waarbij toetsresultaten een rol kunnen spelen. Voor een enkelvoudig vwo-advies hebben bijvoorbeeld alle Rotterdamse basisscholen afgesproken om bepaalde minimumscores op de Cito-toetsen uit het leerlingvolgsysteem van groep 6, 7 en 8 te hanteren (www.erasmiaans.nl/aanmelding-23-78.html), waardoor het eerste selectiemoment feitelijk wordt vervroegd. In antwoord op Kamervragen hierover antwoordde staatssecretaris Dekker (OC&W) er geen probleem mee te hebben wanneer basisscholen met het voortgezet onderwijs afspreken hoe en met welke gegevens zij het schooladvies onderbouwen, en bijvoorbeeld afspraken maken over hoe de basisschool toetsgegevens gebruikt bij de totstandkoming van het advies (Tweede Kamer 2015). Alleen mag een middelbare school geen leerlingen weigeren als een basisschool niet met de lokale afspraken meedoet.

Categorale vwo-scholen kunnen dus nog steeds sterker selecteren dan gemengde havo/vwo-scholen. Bij de huidige kwaliteitsmeting zijn categorale en gemengde scholen ingedeeld in aparte vergelijkingsgroepen, waarbinnen de relatieve scores (op rendement en gemiddeld CE-cijfer) worden bepaald. Hierdoor wordt voorkomen dat categorale scholen dankzij een selectie van betere leerlingen op de relatie-

ve prestatie-indicatoren beter scoren dan de gemengde havo/vwo-scholen en er dus een ongelijk speelveld ontstaat. Sinds 2011 publiceert de Inspectie echter per opleiding ook absolute gemiddelde CE-cijfers en hanteert daarbij een uniforme norm, waarbij er geen onderscheid wordt gemaakt tussen vwo-opleidingen van gemengde en categorale scholen. Gelet op de gunstige leerlingselectie bij categorale opleidingen is de kans groot dat zij op deze absolute indicator beter scoren dan gemengde scholen. Hoewel de Inspectie deze absolute score nog niet betreft in haar oordeel over scholen, wordt deze wel gepubliceerd en gebruikt in de diverse rankings van middelbare scholen. Deze rankings vertekenen derhalve de feitelijke onderwijsprestaties ten nadele van gemengde scholen.

Problematischer is echter dat vanaf 2016 de Inspectie een nieuw kwaliteitsbeoordelingssysteem introduceert waarbij de relatieve scores per vergelijkingsgroep worden vervangen door absolute scores per opleiding (Ministerie van OCW 2015).² Bovendien gaat de Inspectie ook een bovennorm introduceren, waarbij scholen de kwalificatie ‘ruim boven de norm’ ontvangen als hun gemiddelde CE-cijfer deze norm overschrijdt. Het nieuwe model betekent dat het risico op een ongelijk speelveld tussen gemengde havo/vwo-scholen en categorale scholen sterk wordt vergroot. Gemengde havo/vwo-scholen worden hierdoor bovendien min of meer gedwongen om de ‘opstroom’ van minder goede leerlingen naar het vwo te ontmoedigen door daar strenge eisen aan te verbinden. Aldus worden deze leerlingen kansen op een betere opleiding ontnomen. Doordat in het nieuwe model veel meer nadruk komt te liggen op goede eindexamenresultaten en dus ook op het belang van een terughoudend opstroombeleid is van een evenwichtig ‘balanced score model’ geen sprake meer.

8 Conclusie en aanbeveling

Uit recent onderzoek blijkt dat concurrentie tussen scholen met gecombineerde opleidingen (havo/vwo of vmbo/havo) ertoe leidt dat deze scholen meer relatief zwakke leerlingen toelaten en beter presteren (Cabus en Cornelisz 2014). Voor categorale vwo- en havoscholen blijkt dit effect niet aanwezig. De onderzoekers trekken hieruit de conclusie dat concurrentie tussen scholen meer op basis van kwalitatief hoogwaardig onderwijs gevoerd zou moeten worden en minder gericht zou moeten zijn op het aantrekken van de relatief beste leerlingen. De beoordeling van scholen op basis van de kwaliteitsindicatoren van de Inspectie bevordert echter het omgekeerde. Om goed te scoren in de ranglijsten en voor het verkrijgen van een gunstig oordeel van de Inspectie (of voor het vermijden van een ongunstig oordeel) is het aantrekken van de beste leerlingen en het ontmoedigen van de minder goede leerlingen om een zo hoog mogelijke opleiding te volgen een effectieve strategie. Dit leidt tot een ongelijk speelveld tussen categorale scholen en scholen met een

² Zie kader. Voor meer details:

<http://publicaties.onderwijsinspectie.nl/xmlpages/page/onderwijsresultatenmodel-vo-2016/hoofdpijnen/een-nieuw-model/veranderingen>

gemengde brugklas, vermindert de prikkels tot het verbeteren van onderwijsprestaties en verlaagt de kansen van leerlingen die op de basisschool minder goed scoren, maar wel potentie hebben. Het in 2016 in te voeren nieuwe model voor de beoordeling van leerresultaten lost deze problemen niet op, maar zal het ongelijke speelveld tussen gemengde havo/vwo en categorale vwo-scholen juist vergroten en het opstromen van leerlingen verder ontmoedigen.

Het wegnemen van de Cito-toets als toelatingscriterium vermindert de mogelijkheden tot selectie. Hiermee wordt echter niet zozeer de kwaal bestreden maar een symptoom. Aanpak van de kwaal vereist het reduceren van de prikkel tot selectie, bijvoorbeeld door de prestatie-indicatoren beter te corrigeren voor de kwaliteit van de leerlingen. Wanneer effectief kan worden gecorrigeerd voor kwaliteitsverschillen tussen leerlingen, hoeft selectie zelfs helemaal geen probleem te zijn. De Cito-toets lijkt voor dit doel bij uitstek geschikt. Nederland beschikt met de Cito-score immers over een unieke, uniforme en gevalideerde kwaliteitsmaat over het grootste deel van – en vanaf 2015 zelfs van alle nieuwe – middelbare scholieren. Tot op heden is van deze informatie geen gebruik gemaakt. Door de scholen te verplichten de Cito-score en CE-cijfers op leerlingniveau aan te leveren, kunnen de gemiddelde CE-cijfers worden gecorrigeerd voor de Cito-score van de leerlingen. Op die manier krijgen ouders en leerlingen beter inzicht in welke scholen daadwerkelijk beter presteren in plaats van sterker selecteren. Uit onderzoek van Timmermans et al. (2011) blijkt bovendien dat de Cito-score weliswaar niet de enige, maar wel veruit de belangrijkste verklarende variabele vormt voor de verschillen in eindexamenresultaten, zowel tussen scholen als tussen leerlingen.

De voorgestelde correctie lijkt eenvoudig uitvoerbaar, zal scholen sterker aansporen tot betere onderwijsprestaties, zal hen aanmoedigen om leerlingen waar mogelijk te laten ‘opstromen’ naar een zo goed mogelijke opleiding en zal leiden tot eerlijker concurrentie tussen gemengde en categorale scholen. Aangezien de in 2016 nieuw in te voeren methode om de leerprestaties te meten selectie nog meer lonend maakt, is het van extra belang om voor selectie zo goed mogelijk te corrigeren.

Auteurs

Erik Schut (e-mail: schut@bmg.eur.nl) is als hoogleraar verbonden aan de Erasmus Universiteit Rotterdam; Margreet den Haan is werkzaam als natuurkundedocent op een middelbare school (havo/vwo).

Literatuur

Abels, R., 2014, Goede resultaten maken nog geen goede school, *Trouw*, 11 januari.

- American Statistical Association, 2014, ASA Statement on Value-Added Models for Educational Assessment, April 8, https://www.amstat.org/policy/pdfs/ASA_VAM_Statement.pdf
- Beeten, M. van, 2015, Gymnasia sturen 'zwakke' leerling nog altijd weg, *Algemeen Dagblad*, 5 februari.
- Cabus, S. en I. Cornelisz, 2014, Rendement van concurrentie tussen scholen, *ESB*, vol. 99(4682): 198-201.
- College voor Toetsen en Examens, 2014, De centrale eindtoets PO in 2015. Informatie voor scholen, CvTE, Utrecht.
- Elzinga, A., 2011, Eisen Cito-score verschillen per school, *J/M Ouders*, www.jmouders.nl/school/cito-en-toetsen/cito-en-eindtoetsen/eisen-cito-score-verschillen-per-school
- Elk, R. van, M. van der Steeg en D. Webbink, 2011, Does the timing of tracking affect higher education completion? *Economics of Education Review*, vol. 30(5): 1009-1021.
- Feron, E., T. Schils en B. ter Weel, 2015, *Does the teacher beat the test? The additional value of teacher assessment in predicting student ability*, CPB Discussion Paper 300, CPB, Den Haag.
- Inspectie van het Onderwijs, 2011, De beoordeling van opbrengsten in het voortgezet onderwijs. Hoe werkt het?, Publicatienummer: 22DW2011G018, Utrecht.
- Inspectie van het Onderwijs, 2013, Meerjaren opbrengsten VO 2013. Toelichting, Utrecht, mei.
- Inspectie van het Onderwijs, 2014a, Verslag eindbijeenkomst pilot leerwinst en toegevoegde waarde, 26 november, <http://www.onderwijsinspectie.nl/binaries/content/assets/Documents+algemeen/2014/eindbijeenkomst-leerwinst.pdf>
- Inspectie van het Onderwijs, 2014b, Onderwijsresultaten Voortgezet Onderwijs 2014 (Nieuw model – schaduwversie), Technische toelichting. Utrecht, oktober.
- Inspectie van het Onderwijs, 2015, De staat van het onderwijs. Onderwijsverslag 2013/2014, Utrecht, 15 april.
- Kuiper, R., 2015, Er is veel onenigheid tussen ouders, basis- en middelbare scholen over schooladvies, *De Volkskrant*, 24 februari.
- Kuyper, H. en G. van der Werf, 2012, Excellente leerlingen in het voortgezet onderwijs. Schoolloopbanen, risicofactoren en keuzen, GION, Groningen.
- Loeb, S. en D. Figlio, 2011, School accountability. In E. A. Hanushek, S. Machin en L. Woessmann (eds), *Handbook of the Economics of Education*, vol. 3: 383-423, San Diego, CA: North Holland.
- Ministerie van OCW, 2015, Nieuw model leerresultaten voortgezet onderwijs, Brief aan de voorzitter van de Tweede Kamer der Staten-Generaal, Referentie 785994, 25 juni.
- Onderwijsraad, 2014, Toegevoegde waarde: een instrument voor onderwijsverbetering - niet voor beoordeling, Den Haag, 7 april.
- Timmermans, A.C., S. Doolaard en I. de Wolf, 2011, Conceptual and empirical differences among various value-added models for accountability, *School Effectiveness and School Improvement*, vol. 22(4): 393-413.
- Timmermans, A.C. en S.M. Thomas, 2015, The impact of student composition on schools' value-added performance: a comparison of seven empirical studies, *School Effectiveness and School Improvement*, vol. 26(3): 487-498.
- Tweede Kamer, 2015, Omzeilen van het schooladvies, Stenogram van Plenair debat, 25 februari.
- Vasterman, J., 2015a, Vroeg schooladvies? Dat houden we zo, *NRC Handelsblad*, 26 februari.

Vasterman, J., 2015b, Leerlingen komen niet meer zo snel op een hoger niveau, *NRC Handelsblad*, 15 april.

Vuuren, D. van en K. van der Wiel, 2015, Zittenblijven in het primair en voortgezet onderwijs. Een inventarisatie van voor- en nadelen, CPB Policy Brief 2015/01, CPB, Den Haag.

Weel, B. ter, 2015, Wat weten we van cruciale onderwijskeuzes?

www.economie.nl/weblog/wat-weten-we-van-cruciale-onderwijskeuzes