



Centraal Planbureau

CPB Notitie | 29 februari 2012

Slechte en goede prestaties van scholen: is er een relatie met personeelskenmerken?

Op verzoek van het ministerie van Onderwijs, Cultuur en Wetenschap en het ministerie van Economische Zaken, Landbouw en Innovatie



CPB Notitie

Aan: Ministerie van Onderwijs, Cultuur en Wetenschap
Ministerie van Economische Zaken, Landbouw en Innovatie
T.a.v. OCW, Directie Kennis, André de Moor

Centraal Planbureau
Van Stolkweg 14
Postbus 80510
2508 GM Den Haag

T (070) 3383 380
I www.cpb.nl

Contactpersoon
Riemer Faber
Karen van der Wiel

Datum: 29 februari 2012

Betreft: De prestaties en eigenschappen van scholen in het voortgezet onderwijs

Slechte en goede prestaties van scholen: is er een relatie met personeelskenmerken?

1 Inleiding en samenvatting

Het ministerie van Onderwijs, Cultuur en Wetenschap en het ministerie van Economische Zaken, Landbouw en Innovatie hebben het Centraal Planbureau gevraagd om onderzoek te doen naar de determinanten van heel slechte en heel goede prestaties van scholen in het voortgezet onderwijs. Is er een relatie tussen de beschikbare indicatoren voor prestaties en de beschikbare eigenschappen van middelbare scholen in Nederland? Deze CPB Notitie analyseert schoolprestaties over de periode 1996-2008 en relateert deze prestaties aan een reeks eigenschappen van scholen. Aangezien er vooral personeelsgegevens van scholen bekend zijn, ligt het zwaartepunt van de analyse op dit type informatie.

Dit onderzoek sluit aan bij de recente interesse van het ministerie van Onderwijs, Cultuur en Wetenschap en het ministerie van Economische Zaken, Landbouw en Innovatie in excellente scholen. De ministeries willen excellente prestaties van scholen graag aanmoedigen door deze te gaan belonen. Daarnaast willen de ministeries graag leren van succesvolle scholen om zo het niveau van het onderwijs verder te kunnen verhogen.¹ Echter, om excellente scholen aan te kunnen wijzen is het nodig om prestaties van scholen te meten. En om de prestaties van niet-excellente scholen te kunnen verbeteren, is het noodzakelijk om te weten welke eigenschappen van excellente scholen er precies voor zorgen dat deze zo goed zijn. Deze Notitie probeert bij te dragen aan de kennis over deze twee vraagstukken.

In deze CPB Notitie worden bestaande gegevens uit verschillende bronnen met elkaar gecombineerd om zo tot nieuwe inzichten te komen over de prestaties van scholen in het voortgezet onderwijs. De informatiebron van prestaties van scholen is de jaarlijkse opbrengstenkaart van de Inspectie van het

¹ Zie, bijvoorbeeld, de kamerbrief 'Aanbieding actieplannen primair onderwijs, voortgezet onderwijs en leraren' en pagina 10-11 van de bijlage 'Actieplan beter presteren: opbrengst gericht en ambitieus. Het beste uit leerlingen halen' (ministerie van Onderwijs, Cultuur en Wetenschap, 23 mei 2011). Zie ook 'Jaarwerkplan 2012. Activiteiten Inspectie van het Onderwijs' (Inspectie van het Onderwijs, oktober 2011).

Onderwijs. Deze organisatie houdt toezicht op scholen. De meeste data over eigenschappen van scholen zijn verkregen via de Dienst Uitvoering Onderwijs (DUO). DUO is de uitvoeringsorganisatie van de Rijksoverheid voor onderwijs. De gebruikte data zijn dan ook grotendeels administratieve data, bijvoorbeeld afkomstig uit de salarisadministratie van scholen. Zowel de Inspectie van het Onderwijs als DUO is onderdeel van het ministerie van Onderwijs, Cultuur en Wetenschap.

Deze CPB Notitie begint met het bestuderen van de verschillende prestatie-indicatoren die de Inspectie van het Onderwijs verzamelt voor het voortgezet onderwijs. Dit zijn het gemiddelde cijfer op het centraal eindexamen, de mate waarin leerlingen op een schoolsoort eindigen die boven het oorspronkelijke schooladvies ligt (het onderbouwrendement), de mate waarin leerlingen met een diploma van school gaan (het bovenbouwrendement) en het verschil tussen het gemiddelde cijfer voor het centraal eindexamen en de schoolonderzoeken.² Deze prestatie-indicatoren worden vervolgens zo goed mogelijk gecorrigeerd voor verschillen in het soort leerling dat een school heeft. Bij het vergelijken van de prestatie-indicatoren blijkt dat de verschillende maatstaven niet heel sterk op elkaar lijken. Een school die goed presteert op één bepaald onderdeel, presteert dus niet noodzakelijkerwijs goed op andere onderdelen. Dit is een belangrijke reden om vervolgens met een combinatiemaatstaf te werken, een maatstaf die de prestaties op meerdere dimensies samenneemt.

Uit de data-analyse blijkt verder dat de prestaties van een middelbare school zoals gemeten door de indicatoren niet altijd constant zijn door de tijd: een school die in een bepaald jaar zeer slecht (of goed) presteert, presteert niet noodzakelijkerwijs zeer slecht (of goed) in de daaropvolgende jaren. Wat betreft de combinatiemaatstaf, is één jaar na het meetmoment nog maar 48% van het oorspronkelijke verschil tussen de gemiddelde score van de 10% van de scholen met de slechtste prestaties en de 10% van de scholen met de beste prestaties over. Wel blijft er op de langere termijn een verschil bestaan tussen de slechtst en best presterende scholen: na zes jaar is nog steeds 19% van het verschil over. Er zijn maar enkele scholen die consistent bij de groep scholen met de slechtste of beste prestaties behoren. Slechts 1,2% van alle scholen behoort drie jaar achter elkaar tot de 10% slechtst presterende scholen en maar 1,6% van alle scholen behoort drie jaar achter elkaar tot de 10% best presterende scholen. Echter, deze groepen zouden een stuk kleiner zijn als prestaties volledig willekeurig zouden zijn.

Vervolgens probeert deze Notitie de door de indicatoren gemeten prestaties van scholen te verklaren door middel van (vooral) personeelsgegevens van deze scholen. Er worden analyses uitgevoerd waarbij de prestatie van een school de te verklaren variabele is, maar ook analyses waarbij de te verklaren variabele aangeeft of een school tot de slechtst presterende 10% scholen of tot de best presterende 10% scholen behoort. In eerste instantie wordt gebruik gemaakt van schattingsmethoden die veranderingen in de prestaties en eigenschappen van een school door de tijd heen benutten. Voorbeelden van beschikbare personeelsgegevens van een school zijn: het aantal leerlingen per medewerker, het percentage vrouwelijke medewerkers op een school, het percentage medewerkers per leeftijdscategorie, het percentage medewerkers dat vertrekt, het percentage medewerkers dat nieuw is, een benadering van het percentage bevoegde leraren, het percentage leraren dat meerdere vakken geeft en het gemiddelde salaris van medewerkers. Het blijkt echter dat de geobserveerde eigenschappen de prestaties van een school zoals gemeten door de indicatoren niet goed kunnen verklaren. Voor geen enkele eigenschap is er een overtuigend robuust en statistisch significant verband met de prestatie-indicatoren, niet voor heel slecht presterende scholen, maar ook niet voor heel goed presterende scholen.

Naast schattingsmethoden die variatie in de data over tijd benutten, wordt ook gebruik gemaakt van de verschillen tussen scholen. Dit is om twee redenen. Voor veel personeelskenmerken zijn de verschillen tussen scholen groter dan de verschillen binnen een school door de tijd heen (het gemiddelde salaris dat een school betaalt is bijvoorbeeld redelijk constant door de tijd). Bovendien zijn er aanwijzingen dat veranderingen in de prestaties van een school (zoals gemeten door de indicatoren) door de tijd heen minder

² Zie voor andere toepassingen van deze prestatiedata bijvoorbeeld Dijkstra et al. (2001) (in het bijzonder Dronkers en Veenstra (2001)), Van Alphen en Dronkers (2011) en Koning en Van der Wiel (2012). Voor een algemene discussie over het meten van prestaties van scholen zie bijvoorbeeld Kane en Staiger (2002), Todd en Wolpin (2003) en Rothstein (2010).

informatief zijn, omdat er weinig scholen consistent heel slecht of heel goed presteren. Echter, ook via schattingsmethoden die verschillen tussen scholen gebruiken, kunnen de prestaties van scholen maar beperkt worden verklaard door de geobserveerde eigenschappen. De conclusie van deze Notitie is dat voor geen enkele van de geobserveerde eigenschappen kan worden bewezen dat er een relatie bestaat met de door de indicatoren gemeten prestaties van scholen. Dit betekent niet noodzakelijkerwijs dat er geen relatie bestaat tussen de prestaties en eigenschappen van scholen.

Hoewel dit een opvallende conclusie is, is deze conclusie wel in lijn met ander onderzoek. Veel eerdere studies, naar andere landen en periodes, hebben ook geen relatie gevonden tussen prestaties van scholen en relatief eenvoudig observeerbare eigenschappen van scholen, zoals personeelsgegevens. Daarnaast laten deze eerdere studies ook zien dat prestaties van scholen (en leraren) wel substantieel van elkaar verschillen. Zie bijvoorbeeld Coleman et al. (1966) voor een vroege studie, Rivkin et al. (2005) en Hanushek en Rivkin (2010). Voor overzichtsstudies, zie bijvoorbeeld Hanushek (1997), Scheerens en Bosker (1997) en Hanushek en Rivkin (2006). Deze Notitie bevestigt dus het beeld dat het met de standaard observeerbare eigenschappen niet eenvoudig is vast te stellen wat er precies voor zorgt dat een school slecht of goed presteert.

Een beperking van deze Notitie is dat niet de algemene kwaliteit van scholen wordt gemeten en verklaard. De gebruikte prestatie-indicatoren omvatten slechts een gedeelte van de prestaties van scholen. De indicatoren kijken bijvoorbeeld niet naar de invloed van een school op de lange-termijn-doelen van onderwijs (zoals het niveau van de vervolgopleidingen van leerlingen en hun latere inkomen). Een algemene kwaliteitsmaatstaf zou een ander beeld kunnen geven dan de gebruikte prestatie-indicatoren. Daarnaast is een belangrijke beperking dat veel eigenschappen van leerlingen en scholen niet worden geobserveerd (zoals de aanleg van leerlingen voor het halen van hoge cijfers). Hierdoor zijn de prestatie-indicatoren imperfect gecorrigeerd voor de verschillende leerlingen die scholen hebben. De niet-geobserveerde eigenschappen zouden ook eventueel de schattingsresultaten kunnen beïnvloeden, doordat er geen goed onderscheid kan worden gemaakt tussen het effect op de prestaties van de niet-geobserveerde eigenschappen en het effect van de wel-geobserveerde eigenschappen.

De rest van deze Notitie is als volgt opgebouwd. Sectie 2 bespreekt de verschillende maatstaven voor prestaties van scholen. Sectie 3 verklaart de prestaties van scholen door de eigenschappen van scholen. Sectie 4 sluit af met enkele conclusies.

2 Prestatie-indicatoren

2.1 Prestaties per schoolsoort gemeten op meerdere dimensies

De prestaties in het voortgezet onderwijs worden in Nederland bijgehouden door de Inspectie van het Onderwijs. De Inspectie van het Onderwijs verzamelt hiervoor een aantal kwantitatieve indicatoren per schoolvestiging en aangeboden schoolsoort. De term schoolsoort verwijst hier naar de vijf (groe) niveaus waarop een leerling eindexamen kan doen: voorbereidend middelbaar beroepsonderwijs basisberoepsgerichte leerweg (vmbo-b), voorbereidend middelbaar beroepsonderwijs kaderberoepsgerichte leerweg (vmbo-k), voorbereidend middelbaar beroepsonderwijs theoretische en gemengde leerweg (vmbo-gt), hoger algemeen vormend onderwijs (havo) en voorbereidend wetenschappelijk onderwijs (vwo).

De Inspectie van het Onderwijs kijkt naar meerdere dimensies van prestaties in plaats van één, vanwege de moeilijkheden rondom het meten van prestaties. Ten eerste is het bijvoorbeeld zo dat middelbare scholen met verscheidene soorten leerlingen te maken hebben. Niet alle leerlingen kunnen even makkelijk met een goed cijfer een diploma halen. Een tweede voorbeeld is dat het gemiddelde eindexamencijfer niet altijd dezelfde betekenis heeft op verschillende scholen. De groep leerlingen die opgaat voor een bepaald eindexamen is namelijk niet gegeven, maar tot op zekere hoogte een keuze van de school. Een school kan er voor kiezen om alleen de beste leerlingen examen te laten doen of om leerlingen op een lagere schoolsoort examen te laten doen dan de schoolsoort waarvoor zij zouden kunnen slagen. Op deze manier kan een school een hoog gemiddeld cijfer behalen. Verder bestaan eindexamencijfers niet alleen uit gestandaardiseerde toetsen die voor alle leerlingen in een jaar hetzelfde zijn (centraal eindexamen), maar ook voor de helft uit toetsen die door de school worden opgesteld (schoolonderzoeken). Het is voor een school dus mogelijk om relatief 'makkelijke' schoolonderzoeken te maken en daardoor meer eindexamens uit te reiken.

In deze sectie worden alle continue prestatie-indicatoren geïntroduceerd zoals de Inspectie van het Onderwijs die verstrekt in de (voormalige) opbrengstenkaarten. Via DANS-KNAW heeft de Inspectie van het Onderwijs het CPB toestemming verleend om met de inspectiebestanden te werken. Tabel 1 bevat per schoolsoort de gemiddelde waarde van de prestatie-indicatoren voor scholen met één vestiging voor de jaren 2002-2008. Er is gekozen voor een focus op scholen met één vestiging vanwege databeperkingen die hierna uitgebreid worden besproken. Op de volgende pagina's worden de indicatoren gedefinieerd.

Cijferindicatoren per schoolsoort en schoolvestiging

Cijfers kernvakken centraal eindexamen (CE)

Dit getal is een gemiddelde van de cijfers die zijn behaald door alle examenkandidaten voor (alle vormen van) de vakken Nederlands, Engels en wiskunde tijdens het centraal eindexamen. Voor de schoolsoorten havo en vwo wordt vanaf 2011/2012 speciaal gestuurd op deze basisvakken. De indicator 'Cijfers kernvakken CE' is goed te vergelijken binnen schoolsoorten, omdat het overgrote merendeel van de leerlingen deze vakken volgt. Wat in Tabel 1 opvalt zijn de hoge eindexamencijfers op het vmbo-b en de lage cijfers op de havo.

Cijfers alle vakken centraal eindexamen (CE)

Deze cijferindicator is een gewogen gemiddelde van alle eindexamencijfers behaald door alle eindexamenkandidaten op een vestiging tijdens het centraal eindexamen. Alle vakken zijn verwerkt in deze indicator, zij het gewogen met het aantal leerlingen dat in dit vak eindexamen deed. Aangezien de aangeboden eindexamenvakken verschillen tussen schoolvestigingen, is de indicator 'Alle cijfers CE' minder goed te vergelijken tussen vestigingen. Voor de schoolsoorten vmbo-b en vmbo-k lijkt het beter om naar het

Tabel 1: Gemiddelde prestatie-indicatoren per schoolsoort in 2002-2008 (voor scholen met één vestiging)

Schoolsoort	(I) Vmbo-b	(II) Vmbo-k	(III) Vmbo-gt	(IV) Havo	(V) Vwo
Cijfers kernvakken CE	6,77	6,46	6,39	6,28	6,49
Alle cijfers CE	6,69	6,39	6,35	6,24	6,39
Vershil cijfers kernvakken CE-SO	0,20	-0,05	-0,16	-0,24	-0,41
Onderbouwrendement	100,3	100,3	100,9	101,9	101,8
Bovenbouwrendement	89,7	85,8	85,8	63,8	67,3
Aantal observaties	772	776	1505	1300	1642

gemiddelde cijfer op alle vakken te kijken dan naar het gemiddelde cijfer op de kernvakken, gezien het belang van praktijkgericht onderwijs in deze schoolsoorten. Uit Tabel 1 blijkt dat het gemiddelde cijfer over alle eindexamenvakken lager ligt dan het cijfer over de kernvakken Nederlands, Engels en wiskunde.

Vershil cijfers kernvakken centraal eindexamen en schoolonderzoeken (SO)

Dit getal is het gemiddelde verschil tussen de cijfers behaald bij het centraal eindexamen en de schoolonderzoeken voor (alle vormen van) de vakken Nederlands, Engels en wiskunde. Aangezien de schoolonderzoeken voor ongeveer de helft meetellen bij het vaststellen of een leerling is geslaagd, kan een school met relatief ‘makkelijke’ schoolonderzoeken het percentage geslaagden verhogen. De indicator ‘Vershil cijfers kernvakken CE-SO’ zegt dus tot op zekere hoogte iets over de mate waarin scholen het hun leerlingen ‘gemakkelijk’ maken. Hoe kleiner (sterker negatief) dit verschil is, hoe makkelijker de schoolonderzoeken ten opzichte van het centraal eindexamen waarschijnlijk zijn geweest. Dit hoeft niet noodzakelijkerwijs zo te zijn: de schoolonderzoeken kunnen moeilijk zijn geweest, maar de stof van de schoolonderzoeken kan hebben afgeweken van de geëxamineerde stof op het centraal eindexamen en scholen kunnen relatief veel onderwijstijd aan de stof voor de schoolonderzoeken hebben besteed. Uit Tabel 1 komt naar voren dat de gemiddelde cijfers voor schoolonderzoeken in het algemeen hoger liggen dan voor het centraal eindexamen, behalve op het vmbo-b. Bovendien is het verschil tussen de cijfers voor het centraal eindexamen en de schoolonderzoeken met name op het vwo groot.

Rendement bovenbouw per schoolsoort en schoolvestiging

Het bovenbouwrendement is een procesindicator die per schoolsoort op een vestiging meet of leerlingen in de bovenbouw goed doorstromen. De indicator is een schatting van het percentage leerlingen dat vanaf het derde leerjaar zonder zittenblijven een diploma heeft ontvangen. Dit geschatte percentage maakt gebruik van de bevorderings- en slagingspercentages van de huidige derde-, vierde-, (en waar mogelijk) vijfde- en zesdeklassers van het afgelopen jaar. Door de bevorderings- en slagingspercentages van verschillende klassen te gebruiken, wordt het effect van een cohort met slechte leerlingen gedempt in de indicator. De overige prestatie-indicatoren gebruiken gegevens van één (eindexamen)klas, waardoor de verschillen tussen jaren groter kunnen zijn. Aangezien het aantal jaren bovenbouw sterk uiteenloopt tussen het vmbo (één jaar) en het vwo (drie jaar), verschilt het bovenbouwrendement ook sterk tussen de verschillende schoolsoorten (zie Tabel 1). Daarom is het belangrijk om de prestaties van scholen op deze indicator te vergelijken met de prestaties van scholen van hetzelfde soort.

Rendement onderbouw per schoolvestiging

Het onderbouwendement is een procesindicator die per schoolvestiging meet of leerlingen in de eerste drie jaar goed doorstromen binnen de vestiging. De indicator geeft daarbij aan welke schoolsoort leerlingen volgen in het derde leerjaar en vergelijkt deze schoolsoort met het schooladvies van de basisschool. Daarnaast houdt de indicator rekening met het aantal jaren dat een leerling over de eerste twee leerjaren heeft gedaan. Het gaat hierbij om de situatie van het cohort leerlingen dat in het eerste en tweede leerjaar op de betreffende school zat. De berekening per leerling is als volgt:

$$100\% * (\text{positie jaar 3} - \text{advies} - \text{doublure} + 2)/2$$

Daarnaast geldt nog een extra 1/4 punt in de teller voor leerlingen met een indicatie voor leeuwondersteuning (lwoo). De waarde van iedere schoolsoort hoger is één punt en de waarde van iedere schoolsoort lager is ook één punt. Een leerling met een vmbo-(g)t-advies die zonder zittenblijven op het vwo eindigt, levert dus een score van 200% op. Vervolgens wordt van alle leerlingen in de derde klas het gemiddelde genomen per schoolvestiging. Het onderbouwendement is de indicator die het best is ontworpen om de toegevoegde waarde van een school te meten. In tegenstelling tot de andere indicatoren meet het onderbouwendement namelijk niet het prestatieniveau van leerlingen, maar de verandering in dit niveau.

Prestatie-indicatoren van schoolsoorten optellen tot een waarde voor de vestiging

Het vervolg van deze Notitie gaat verder in op de prestaties van scholen, in plaats van op de prestaties van de aparte schoolsoorten op een school. Het valt namelijk aan te nemen dat veel beleidsbeslissingen van een directie voor alle schoolsoorten binnen een vestiging gelden, net als dat veel personeelsleden binnen een vestiging betrokken zijn bij meer dan één schoolsoort. Behalve het onderbouwendement zijn alle indicatoren wel per schoolsoort beschikbaar. Dit betekent dat er voor bijvoorbeeld een brede scholengemeenschap - waar alle schoolsoorten worden aangeboden - per indicator vijf scores bekend zijn. Om vervolgens tot een indicator per schoolvestiging te komen, wordt een gewogen gemiddelde genomen van de schoolsoorten met het aantal leerlingen per schoolsoort als gewicht.

Tabel 2 geeft een overzicht van deze ruwe optelsommen tussen 2002 en 2008. In de kolommen I-VI staan de gemiddelde prestatie-indicatoren per type scholengemeenschap, variërend van scholen waar alleen praktijkonderwijs wordt gegeven tot zelfstandige gymnasia. Het hoge eindexamencijfer voor gymnasia (scholen met alleen vwo) valt op, net als het hoge onderbouwendement bij lycea en het hoge bovenbouwendement bij praktijkscholen. Een deel van deze verschillen heeft te maken met de opbouw van schoolsoorten binnen een scholengemeenschap. Een optelsom van prestaties pakt qua bovenbouwendement bijvoorbeeld ongunstig uit voor scholengemeenschappen met veel vwo-leerlingen, die scoren door het hogere aantal jaar bovenbouw namelijk per definitie lager. Om rekening te houden met dergelijke verschillen in niveaus tussen schoolsoorten, wordt verderop in deze Notitie een optelsom gemaakt waarin per jaar de niveauverschillen tussen schoolsoorten worden geneutraliseerd. Deze optelsom vormt de basis van de verdere berekeningen.

Focus op scholen met één vestiging

In de een-na-laatste kolom van Tabel 2 bevinden zich de gemiddelden over alle scholen in Nederland die maar één schoolvestiging hebben. Deze selectie wordt gemaakt, omdat in het vervolg van deze Notitie voornamelijk scholen met één vestiging worden behandeld. Van alle administratieve schooleenheden (de zogenaamde BRIN's) bestaat 59% uit scholen met één vestiging, terwijl 17% twee vestigingen heeft, 11% drie vestigingen en de overige 13% meer dan drie vestigingen. De keuze voor scholen met één vestiging is niet ideaal, maar wel noodzakelijk, omdat de informatie over school- en personeelseigenschappen die wordt gebruikt in Sectie 3 slechts op BRIN-niveau beschikbaar is. Dit betekent dat voor een school met meerdere vestigingen de personeelsgegevens de gemiddelde waarde weergeven over alle vestigingen. Echter,

Tabel 2: Gemiddelde prestatie-indicatoren per scholengemeenschap in 2002-2008

Scholengemeen- schap	(I) Vmbo- school (b&k)	(II) Vmbo- school (gt)	(III) Brede gemeen- schap	(IV) Smalle gemeen- schap	(V) Lyceum	(VI) Gymna- sium	(VII) Alle	(VIII) Alle
Aantal vestigingen	1	1	1	1	1	1	1	>1
Cijfers kernvakken CE	6,53	6,45	6,44	6,36	6,41	6,85	6,47	6,41
Alle cijfers CE	6,45	6,40	6,38	6,29	6,35	6,67	6,39	6,37
Vershil cijfers kernvakken CE-SO	-0,07	-0,21	-0,19	-0,29	-0,29	-0,22	-0,22	-0,20
Onderbouw- rendement	100,6	97,4	100,2	101,9	104,9	100,8	101,3	101,8
Bovenbouw- rendement	86,6	85,0	75,1	69,2	66,6	74,4	74,3	76,4
Aantal observaties	170	122	475	699	262	265	2129	1703

schoolvestigingen opereren soms zeer zelfstandig. Meestal is de dagelijkse leiding van scholen op vestigingsniveau en is personeel maar op één vestiging werkzaam. Door de selectie van scholen met één vestiging is de bestudeerde eenheid altijd één school, en niet soms één school en soms het gemiddelde van wat in de praktijk eigenlijk twee scholen zijn. Om te controleren of de selectie van scholen met één vestiging niet al te selectieve resultaten oplevert, is de laatste kolom in Tabel 2 opgenomen. Hierin worden dezelfde prestatie-indicatoren weergegeven, maar dan voor alle scholen met meer dan één vestiging. Er blijken kleine verschillen te bestaan tussen de scholen met één vestiging en de scholen met meer vestigingen. Cijfers op scholen met één vestiging liggen iets hoger en het bovenbouwrendement ligt iets lager. Dit kan beide verklaard worden doordat er vooral veel vmbo-scholen meerdere vestigingen hebben. Opvallend is bovendien dat het aantal scholen met één vestiging, vooral door fusies, is afgenomen over de tijd. Ook in Sectie 3 is een robuustheidsanalyse opgenomen om het effect van de selectie van scholen met één vestiging te bestuderen. Al met al blijken de verschillen beperkt te zijn en heeft de selectie geen invloed op de kwalitatieve conclusies.

2.2 Prestatie-indicatoren gecorrigeerd voor instroom en type scholengemeenschap

Tot nu toe zijn alleen bruto prestatie-indicatoren besproken. Dit zijn indicatoren die geen rekening houden met de verschillende omstandigheden waarin scholen opereren. Echter, middelbare scholen hebben bijvoorbeeld te maken met verschillende soorten leerlingen. Daarnaast laat de vorige subsectie zien dat prestaties van een school afhangen van het type scholengemeenschap. Een lyceum in 't Gooi heeft dus een andere uitgangspositie dan een brede scholengemeenschap in Rotterdam Charlois. Dit betekent dat de prestatie-indicatoren van deze twee scholen niet zomaar met elkaar kunnen worden vergeleken.

Het effect op de prestatie-indicatoren van omstandigheden waar scholen geen invloed op hebben moet zoveel mogelijk worden uitgeschakeld. Dit om te voorkomen dat bijvoorbeeld verschillen tussen leerlingen de prestatieverschillen voor een groot deel verklaren. Om de prestatie-indicatoren zo goed mogelijk te

kunnen vergelijken, worden deze voor een aantal duidelijke verschillen tussen schoolpopulaties gecorrigeerd door middel van regressietechnieken. In feite worden eerst de invloed op de prestaties van bijvoorbeeld het percentage vmbo-scholieren dat de theoretische leerweg volgt, het percentage leerlingen uit achterstandsgebieden en de grootte van de gemeente geneutraliseerd. De correctie is, vooral gezien het beperkte aantal omstandigheden waarvoor wordt gecorrigeerd, niet ideaal. Het is echter beter om de imperfect gecorrigeerde maatstaven te analyseren dan de niet-gecorrigeerde maatstaven. De rest van deze Notitie gebruikt de gecorrigeerde maatstaven. Daarom volgt eerst een uitleg van de correctiemethode.

Cijfers en bovenbouwrendement: correctie op het niveau van schoolsoort

De correctie op de prestatie-indicatoren wordt waar mogelijk apart uitgevoerd per schoolsoort en per jaar. Dit betekent dat de invloed van bijvoorbeeld leerlingen uit achterstandsgebieden op het bovenbouwrendement kan verschillen tussen het vmbo-b in 2003 en het vmbo-gt in 2008. Per jaar, schoolsoort en indicator wordt een regressie geschat met zes beschikbare variabelen die mogelijk samenhangen met het niveau van leerlingen. In deze regressies worden ook de observaties van de individuele locaties van scholen met meerdere vestigingen gebruikt. De gebruikte variabelen zijn het aantal inwoners in een gemeente (en het kwadraat hiervan), het aantal vestigingen dat een school heeft, of de schoolsoort wordt aangeboden op een vmbo-school, een brede scholengemeenschap, een smalle scholengemeenschap, een lyceum of een gymnasium, het percentage vmbo-leerlingen dat de theoretische of gemengde leerweg volgt, het percentage vmbo-leerlingen met een indicatie voor leerwegondersteuning (lwoo) en het percentage leerlingen uit achterstandsgebieden of in eerdere jaren uit een culturele minderheid (en het kwadraat hiervan). Het inwoneraantal zegt niet direct iets over het niveau van leerlingen, maar is wel een omgevingsvariabele die een school niet onder controle heeft. Omdat dit kenmerk gecorreleerd is met de prestaties, wordt ook hiervoor gecorrigeerd. Aangezien Tabel 2 laat zien dat er grote verschillen zijn in prestaties tussen de verschillende scholengemeenschappen, is het type scholengemeenschap ook onderdeel van de correctie. Zo kan het bijvoorbeeld zijn dat leerlingen met een sterke sociaaleconomische achtergrond eerder kiezen voor een zelfstandig gymnasium. Waarschijnlijk zijn er naast de zes bovengenoemde eigenschappen van leerlingen veel meer eigenschappen van leerlingen die een rol spelen en idealiter zouden deze worden opgenomen in de regressies. Helaas is meer informatie over eigenschappen van leerlingen niet beschikbaar.³

In Tabel B1-a in de bijlage zijn de regressiecoëfficiënten voor de correctie op de prestatie maatstaf cijfers op het centraal eindexamen te vinden en in Tabel B1-b zijn de regressiecoëfficiënten voor de correctie op de prestatie maatstaf bovenbouwrendement te vinden. De tabellen geven een indicatie van de invloed van de controlevariabelen. Deze resultaten geven aan welke kenmerken scholen hebben die goed presteren, maar tonen daarmee nog geen causale verbanden aan. Ten behoeve van de interpretatie zijn alle percentages uitgedrukt als fracties (dat wil zeggen, de waarden liggen tussen 0 en 1). Ook met het oog op de interpretatie zijn deze regressies, in tegenstelling tot de daadwerkelijk toegepaste correcties, voor een aantal jaren tegelijkertijd geschat. De verschillende kolommen staan voor de verschillende schoolsoorten. Uit Tabel B1-a blijkt dat de grootte van een gemeente, los van alle andere invloeden, bij de meeste schoolsoorten samenhangt met hogere cijfers. Hoe groter de gemeente, hoe beter er gescoord wordt op het centraal eindexamen. Dit effect is het zwakst bij het vmbo-b en duidelijk het sterkst bij het vwo. Uit Tabel B1-b blijkt dat de relatie tussen de grootte van een gemeente en het bovenbouwrendement anders in elkaar zit. In gemeenten met een bevolking tot ongeveer 350.000 mensen wordt er naarmate de gemeente groter wordt steeds slechter gescoord op deze indicator, terwijl in de nog grotere gemeenten weer beter wordt gepresteerd. Het hebben van meerdere vestigingen hangt over het algemeen niet positief samen met zowel de behaalde cijfers op het eindexamen als met het bovenbouwrendement. Uitzondering hierop vormt het vmbo-b: in deze schoolsoort worden juist betere cijfers gehaald en meer diploma's uitgedeeld op scholen met meerdere vestigingen.

³ Door de recente invoering van het onderwijsnummer in het voorgezet onderwijs kan in de toekomst beter rekening worden gehouden met de verschillen tussen leerlingen op scholen. Hierdoor zou het bijvoorbeeld mogelijk moeten zijn om de Cito-score die leerlingen haalden in groep acht van de basisschool te gebruiken.

De tabellen geven verder blijk van verschillen in de cijferprestaties en het bovenbouwrendement tussen de verschillende scholengemeenschappen. Cijfers en bovenbouwrendement op de havo en het vwo zijn duidelijk beter op schoolvestigingen waar geen vmbo-scholieren op de schoolvestiging aanwezig zijn. Ook de afwezigheid van havo-leerlingen op gymnasia hangt positief samen met de prestaties van vwo-leerlingen op zowel cijfers als bovenbouwrendement. Op de diverse vmbo-soorten lijkt het echter qua cijfers niet uit te maken of deze schoolsoort deel uitmaakt van een grotere gemeenschap of niet. Qua bovenbouwrendement van vmbo-leerlingen is er een positief verband met de afwezigheid van andere schoolsoorten.

Hoe meer vmbo-leerlingen op een schoolvestiging verhoudingsgewijs deelnemen aan de theoretische leerweg, hoe beter de cijferprestaties zijn in het vmbo-b en het vmbo-gt en hoe beter het bovenbouwrendement is in het vmbo-b. Het percentage lwoo-leerlingen heeft een tegenovergestelde werking. Hoe meer leerlingen met een leerwegondersteuning, hoe lager de cijfers en het bovenbouwrendement in het vmbo-b en vmbo-k. Ook het hebben van veel leerlingen uit achterstandsgebieden (de zogenaamde APC-leerlingen, voorheen de CUMI-leerlingen) is negatief gecorreleerd met de behaalde cijfers op het centraal eindexamen en het bovenbouwrendement voor alle schoolsoorten. De correlatie met cijfers is verwaarloosbaar voor scholen met tussen de 0% en 10% APC-leerlingen. Pas bij hogere concentraties scholieren uit achterstandswijken is de correlatie met cijfers substantieel. De correlatie met het bovenbouwrendement is al negatief bij kleine percentages APC-leerlingen. Het effect is het sterkst op de havo en het minst sterk op het vmbo-b.

Onderbouwrendement: correctie op het niveau van schoolvestiging

De hierboven beschreven correctie is toegepast op alle cijferindicatoren en op de indicator voor het bovenbouwrendement. Het onderbouwrendement wordt echter per schoolvestiging gemeten, waardoor een correctie per schoolsoort niet mogelijk is. Daarom is er voor gekozen om de correctie op het onderbouwrendement per jaar en type scholengemeenschap uit te voeren. In Tabel B1-c in de bijlage staan de resultaten van de uitgevoerde regressies (wederom voor een aantal jaren bij elkaar). Voor alle duidelijkheid: ook hier kunnen de resultaten niet worden geïnterpreteerd als causale verbanden.

Er vallen drie dingen op bij de regressies van het onderbouwrendement. Ten eerste is er een duidelijk positief verband tussen veel leerlingen met een indicatie voor een leerwegondersteuning (lwoo) en het onderbouwrendement op scholen waar vmbo-b en/of vmbo-k wordt aangeboden. Dit heeft waarschijnlijk direct te maken met de extra 'punten' die scholen krijgen toegewezen voor dit soort leerlingen in de berekening van het onderbouwrendement (zie Sectie 2.1). Blijkbaar schiet deze aanpassing zijn doel enigszins voorbij. Ten tweede is de negatieve samenhang van leerlingen uit achterstandswijken en het onderbouwrendement minder sterk dan het negatieve verband met cijfers voor het centraal eindexamen en het bovenbouwrendement. Alleen bij smalle scholengemeenschappen en gymnasia is er een correlatie met het percentage APC-leerlingen. Ten derde is de verklarende kracht van de regressies, met name in het geval van brede en smalle scholengemeenschappen, een stuk lager dan bij de regressies voor de examencijfers. Blijkbaar zeggen de eigenschappen van leerlingen minder over deze indicator. Dit komt hoogstwaarschijnlijk doordat er in deze prestatie-indicator rekening wordt gehouden met de schooladviezen die leerlingen vanuit hun basisschool meekrijgen. Hiermee is al gecorrigeerd voor belangrijke initiële verschillen tussen leerlingen.

Berekening gecorrigeerde indicatoren

Na de serie regressieanalyses wordt een nieuwe waarde van de prestatie-indicatoren geconstrueerd, bestaande uit het niet door de regressie te verklaren gedeelte van de prestaties (de foutterm). Binnen deze gecorrigeerde prestatie-indicatoren spelen de invloeden van bijvoorbeeld achterstandsléerlingen, de schoolsoort en het type scholengemeenschap dus geen rol meer. Vervolgens wordt er voor elke gecorrigeerde prestatie-indicator per jaar en per schoolsoort een rangorde gemaakt van de scholen. Deze rangordes worden vervolgens verdeeld in tien even grote groepen. In de eerste groep bevinden zich dan (ongeveer) de 10% scholen die het slechtst presteren op die indicator in die schoolsoort in dat jaar en in de tiende groep bevinden zich dan (ongeveer) de 10% scholen die het best presteren op dezelfde indicator in die

schoolsoort in dat jaar. De groepen zijn mogelijk niet altijd exact 10% van het totaal, omdat het mogelijk is dat scholen dezelfde positie in de rangorde hebben doordat zij precies dezelfde score hebben gehaald (dus scholen met dezelfde score vallen altijd in dezelfde groep).

Belang van correctie verschilt per prestatie-indicator

Hoe groot de correctie op de prestatie-indicatoren is, komt tot uitdrukking in de correlaties tussen de oorspronkelijke prestatie-indicatoren en de gecorrigeerde indicatoren. Tabel 3 presenteert een overzicht van deze mate van samenhang. Hoe dichterbij 1 ligt de correlatie, hoe minder sterk de correctie is geweest. En andersom: hoe dichterbij 0 ligt de correlatie, hoe groter de invloed was van de variabelen waarvoor is gecorrigeerd. In Kolom I is de correlatie te vinden tussen de oorspronkelijke indicatoren en een indicator die alleen corrigeert voor de niveaoverschillen tussen de verschillende schoolsoorten en tussen de verschillende jaren. Hieruit blijkt dat het met name voor het bovenbouwrendement veel uitmaakt of er rekening gehouden wordt met de verschillende schoolsoorten die op een school worden aangeboden. Ook voor het verschil tussen de cijfers voor het centraal eindexamen en de schoolonderzoeken maakt deze correctie veel uit. Wat betreft de cijfers voor het centraal eindexamen is het kennelijk minder belangrijk om voor de niveaoverschillen tussen schoolsoorten en jaren te corrigeren.

Kolom II in Tabel 3 geeft vervolgens de samenhang weer tussen de prestatie-indicatoren die al gecorrigeerd zijn voor de verschillen in schoolsoorten en jaren, en de prestatie-indicatoren die daarnaast ook gecorrigeerd zijn voor de bekende kenmerken van leerlingen. Hoe lager deze correlaties, hoe belangrijker de correctie op eigenschappen van leerlingen is. De samenhang tussen de gecorrigeerde en ongecorrigeerde indicator is opvallend hoog bij het onderbouwrendement. Dit betekent dat het initiële niveau van leerlingen geen belangrijke verklaring is voor de verschillen tussen scholen op deze indicator. Met andere woorden: het ongecorrigeerde onderbouwrendement is, voor zover beoordeeld kan worden, een goede benadering van de toegevoegde waarde van scholen (op deze dimensie). De laagste correlatie tussen de twee indicatoren is te vinden voor cijfers voor het centraal eindexamen. Of scholen goed scoren op het eindexamen is blijkbaar voor een substantieel deel te verklaren door het type leerling dat een school aantrekt.

Tabel 3: Samenhang ongecorrigeerde en gecorrigeerde maatstaven in 1996-2008 (op vestigingsniveau voor scholen met één vestiging)

	(I) Correlatie tussen oorspronkelijke indicator en indicator gecorrigeerd voor schoolsoorten en jaren (stap 1)	(II) Correlatie tussen indicator gecorrigeerd voor schoolsoorten en jaren (stap 1) en indicator gecorrigeerd voor alle kenmerken van scholen en leerlingen (stap 2)
Cijfers kernvakken CE	0,94	0,85
Alle cijfers CE	0,95	0,82
Vershil cijfers kernvakken CE-SO	0,82	0,92
Onderbouwrendement	0,90	0,95
Bovenbouwrendement	0,66	0,86

2.3 Combinatiemaatstaf noodzakelijk, want prestatie-indicatoren lopen sterk uiteen

Deze subsectie bespreekt de combinatiemaatstaf die in het vervolg van deze Notitie gebruikt wordt. Bovendien komt de samenhang tussen de verschillende (gecorrigeerde) prestatie-indicatoren aan bod. Verderop wordt het gebrek aan samenhang tussen de indicatoren besproken. Dit versterkt de noodzaak voor een maatstaf die rekening houdt met meerdere prestatiedimensies.

Combinatiemaatstaf van cijfers centraal eindexamen, onder- en bovenbouwrendement

Om de prestaties van scholen op de diverse dimensies samen te vatten, wordt een combinatiemaatstaf berekend met behulp van de cijfers voor het centraal eindexamen, het onderbouwrendement en het bovenbouwrendement. Deze combinatiemaatstaf houdt dus zowel rekening met het eindniveau van het onderwijs (gemeten via de cijfers voor het centraal eindexamen), als met de doelmatigheid van het onderwijs (gemeten via het onderbouwrendement en het bovenbouwrendement). Het uitgangspunt in deze Notitie is dat al deze dimensies van prestaties belangrijk zijn en dat de individuele maatstaven elkaar aanvullen.

Er zijn ontelbaar veel manieren om een combinatiemaatstaf samen te stellen. Deze Notitie pretendeert niet de optimale manier gevonden te hebben. De toegevoegde waarde van deze studie zit vooral in de combinatie van prestatie-indicatoren met personeelsgegevens van scholen (Sectie 3). Er is daarom gekozen voor een eenvoudige, maar complete, maatstaf die dicht ligt bij de methode van de Inspectie van het Onderwijs.

In de combinatiemaatstaf worden voor de schoolsoorten vmbo-gt, havo en vwo alleen de eindexamencijfers voor de kernvakken (Nederlands, Engels en wiskunde) gebruikt. Dit deels omdat er op de prestaties voor deze vakken wordt gestuurd vanaf 2011/2012 en deels omdat verschillen in vakkenpakketten de vergelijking dan niet bemoeilijken. Voor de schoolsoorten vmbo-b en vmbo-k worden echter alle vakken gebruikt om het gemiddelde eindexamencijfer te berekenen. Dit vanwege het grote belang van praktijkgerichte vakken bij deze schoolsoorten.

De enige prestatie-indicator die niet wordt meegenomen in de combinatiemaatstaf is het verschil tussen de cijfers bij het centraal eindexamen en het schoolonderzoek. Dit besluit heeft twee redenen: een positief verschil kan weliswaar betekenen dat er 'gesjoemeld' wordt met het niveau van de schoolonderzoeken, maar er zijn ook verklaringen mogelijk in de tegenovergestelde richting. Bovendien speelt deze indicator ook bij de Inspectie van het Onderwijs een kleine rol.

De berekening van de combinatiemaatstaf vindt in eerste instantie per schoolsoort plaats. Allereerst wordt van de drie indicatoren cijfers centraal eindexamen, onderbouwrendement en bovenbouwrendement per schoolsoort en per jaar een rangorde gemaakt. Vervolgens worden per indicator, jaar en schoolsoort de scholen opgedeeld in tien groepen van (ongeveer) gelijke grootte, van de slechtst presterende (1) tot de best presterende scholen (10). Dit betekent dat per schoolsoort elk jaar 10% van de scholen het predicaat slechtst presterend ontvangt en 10% van de scholen het predicaat best presterend. Ten slotte wordt voor iedere vestiging per schoolsoort en jaar de gemiddelde score over de drie indicatoren berekend. Er wordt gekozen voor de simpelste berekening van een combinatiemaatstaf: met gelijke gewichten voor de drie indicatoren. Voor jaren waar het onderbouwrendement niet beschikbaar is, wordt het gemiddelde van de overige twee maatstaven genomen. Om vervolgens tot een combinatiemaatstaf per school te komen, wordt het gewogen gemiddelde berekend over de verschillende schoolsoorten. De gewichten zijn daarbij het aantal leerlingen per schoolsoort.

Samenhang prestatie-indicatoren is beperkt

Tabel 4 geeft de correlaties weer tussen de individuele prestatie-indicatoren onderling en de combinatiemaatstaf. De correlaties tussen de drie verschillende cijferindicatoren zijn relatief hoog. Dit is

Tabel 4: Correlaties tussen de gecorrigeerde prestatie-indicatoren in 1996-2008 (voor scholen met één vestiging)

	(I) Cijfers kern CE	(II) Alle cijfers CE	(III) Verschil CE-SO	(IV) Onderbouw -rendement	(V) Bovenbouw -rendement	(VI) Combinatie -maatstaf
Cijfers kernvakken CE	1,00					
Alle cijfers CE	0,75	1,00				
Verschil cijfers kernvakken CE-SO	0,70	0,54	1,00			
Onderbouw- rendement	-0,13	-0,13	-0,10	1,00		
Bovenbouw- rendement	0,27	0,39	0,07	-0,02	1,00	
Combinatie- maatstaf	0,52	0,50	0,29	0,51	0,63	1,00

niet verwonderlijk gezien de gelijkenissen tussen deze maatstaven. De correlaties blijken echter laag te zijn tussen de cijferindicatoren en het onder- en bovenbouwrendement. Er bestaat zelfs een negatief verband tussen het onderbouwrendement en de drie cijferindicatoren. Scholen die goed scoren op het onderbouwrendement, scoren dus vaak slecht op eindexamencijfers. Dit lijkt logisch. Immers, leerlingen die een hogere schoolsoort doen dan geadviseerd is door de basisschool, zullen daar gemiddeld genomen lagere cijfers halen dan leerlingen die onder de geadviseerde schoolsoort uitkomen. Daarnaast is er een licht negatieve relatie tussen het onderbouwrendement en het verschil tussen cijfers op het centraal eindexamen en schoolonderzoekcijfers. Mogelijk maken scholen die veel leerlingen een hogere schoolsoort laten doen vaker 'makkelijke' schoolonderzoeken. Verder blijken het onderbouwrendement en het bovenbouwrendement ook maar beperkt samen te hangen. Het is dus niet vanzelfsprekend dat scholen die goed zijn in de doorstroming in de onderbouw deze prestatie herhalen in de bovenbouw (hoewel hier natuurlijk ook het bovengenoemde effect speelt van de schoolsoort die een leerling volgt).

De enige prestatie-indicator die (per definitie) een positieve relatie van behoorlijke omvang heeft met alle andere prestatie-indicatoren is de combinatiemaatstaf. Het lijkt daarom een juiste keuze om in het vervolg van deze Notitie vooral de samengestelde prestatie maatstaf te bestuderen aangezien deze een uitgebreider beeld geeft van de prestaties dan één individuele maatstaf.

2.4 Slechtst en best presterende scholen: behoorlijke verschillen die na één jaar voor de helft zijn ingelopen

De Notitie draait in het bijzonder om de prestaties van scholen aan de onderkant en aan de bovenkant van de verdeling. Daarmee worden in dit geval de scholen bedoeld die vallen binnen de 10% slechtst presterende scholen op een indicator en de scholen die vallen binnen de 10% best presterende scholen op een indicator. De eerste kolom in Tabel 5 bevat de verschillen tussen de gemiddelde waarde van de gecorrigeerde prestatie-indicatoren van de groep slechtst presterende scholen en het gemiddelde van de groep best presterende scholen. Wederom worden alleen scholen met één vestiging bekeken. De verschillen zijn vooral op de

rendementsindicatoren aanzienlijk. Het verschil in het onderbouwrendement bedraagt ongeveer een kwart van de gemiddelde waarde in Tabel 1, terwijl het verschil in het bovenbouwrendement ongeveer een derde van de gemiddelde waarde bedraagt.

Met behulp van de data over prestaties kan voor de periode 1996-2008 worden bestudeerd hoe de prestaties van scholen van jaar op jaar veranderen. De ontwikkelingen worden in deze subsectie in kaart gebracht vanaf het moment dat een school als slechtst of als best presterend wordt geclassificeerd. In Tabel 5 zijn de samengevatte resultaten te vinden. Uit de tweede kolom van Tabel 5 volgt dat een aanzienlijk deel van de verschillen tussen de onderkant en de bovenkant van de verdeling na één jaar tijd is verdwenen. Bij zowel het onderbouwrendement als de combinatiemaatstaf is één jaar na de constatering dat de ene groep scholen tot de slechtst presterende behoort en de andere groep tot de best presterende, minder dan de helft van het oorspronkelijke verschil tussen de gemiddelde scores van de twee groepen over. Bij de indicatoren voor cijfers en het bovenbouwrendement zijn de verschillen persistenter. De meest volhardende maatstaven zijn alle cijfers op het centraal eindexamen en het verschil tussen de cijfers op het centraal eindexamen en de schoolonderzoeken. Van beide blijft 61% van het verschil over na één jaar. Het is wel goed om te realiseren dat als de prestaties van scholen volledig willekeurig zouden zijn, de percentages in de tweede kolom van Tabel 5 rond de 0% zouden liggen. Er is dan namelijk geen reden om aan te nemen dat voorheen gemiddeld presterende scholen een jaar later anders zouden scoren dan voorheen slecht of goed presterende scholen. Gezien het feit dat de percentages in Kolom II ver boven de 0% uitkomen, zit er informatie in de predicaten van een jaar eerder.

Het is niet verrassend dat de verschillen tussen de groep slechtst presterende scholen en de groep best presterende scholen na zes jaar nog verder zijn verdwenen (derde kolom Tabel 5). Toch bestaat er zes jaar nadat een groep scholen het predicaat slechtst of best presterend heeft ontvangen, nog tussen de 16% (voor onderbouwrendement) en 35% (voor alle cijfers op het centraal eindexamen) van de oorspronkelijke verschillen. De gemiddeld slechtst presterende school en de gemiddeld best presterende school blijven dus op de lange termijn wel iets van elkaar verschillen.

Tabel 5: Verschillen tussen de slechtst en best presterende scholen op de gecorrigeerde prestatie-indicatoren in 1996-2008 (voor scholen met één vestiging)

	(I) Vergelijking slechtst en best presterende scholen		(III)	(IV) (V) Vergelijking met gemiddeld presterende scholen	
	Verskil gemiddelden	% verschil over na 1 jaar	% verschil over na 6 jaar	% verschil over na 6 jaar	
				Slechtst prest. scholen	Best prest. scholen
Cijfers kernvakken CE	0,84	57%	29%	23%	35%
Alle cijfers CE	0,75	61%	35%	45%	41%
Verskil cijfers kernvakken CE-SO	0,87	61%	23%	31%	12%
Onderbouwrendement	23,8	47%	16%	16%	16%
Bovenbouwrendement	28,6	58%	31%	28%	34%
Combinatiemaatstaf	5,9	48%	19%	16%	23%

Het valt verder op dat de verschillen in de combinatiemaatstaf net zo snel terugvallen als die in het onderbouwrendement. Het samengestelde karakter van de combinatiemaatstaf lijkt hiervoor van belang: de prestaties op drie verschillende dimensies hangen per definitie minder samen dan de prestaties op een enkele dimensie, en de optelsom van de drie indicatoren zal dus ook meer verschillen van jaar tot jaar. Ook het feit dat het onderbouwrendement sneller terugvalt dan het bovenbouwrendement is (gedeeltelijk) te verklaren door de betreffende definities: de eerste indicator gebruikt gegevens van één cohort leerlingen, terwijl de tweede indicator gegevens van twee tot vier cohorten bevat. Omdat dezelfde groep leerlingen dus een aantal jaren op rij invloed heeft op het bovenbouwrendement, zal deze indicator minder snel terugzakken voor de beste presterende scholen en minder snel omhoogschieten voor de slechtst presterende scholen.

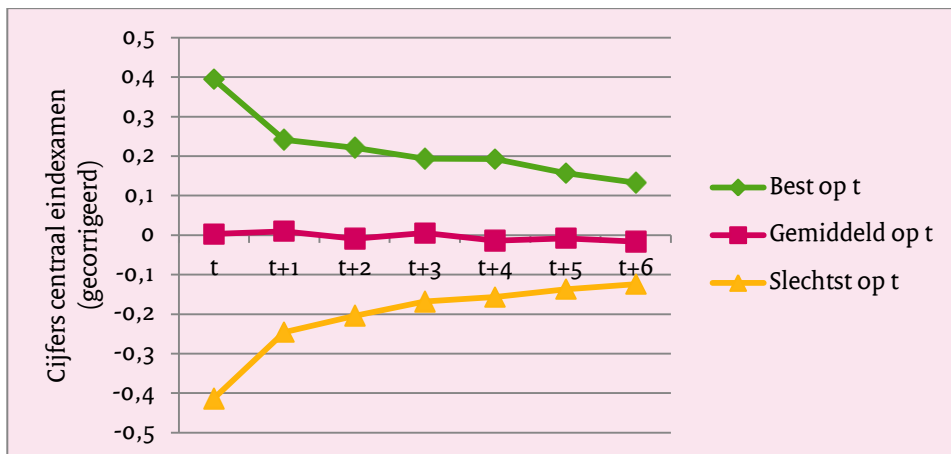
In de vierde en vijfde kolom van Tabel 5 worden aparte resultaten gepresenteerd voor de onderkant en voor de bovenkant van de prestatieverdeling. Hiervoor is een vergelijking gemaakt met de 10% van de scholen die op $t=0$ gemiddeld hebben gepresteerd. Het gaat dus om de verschillen tussen de groep gemiddeld scorende scholen aan de ene kant en de groepen slechtst en best presterende scholen aan de andere kant. De kolommen geven het percentage van het verschil aan dat zes jaar na deze constatering nog over is tussen deze groepen. Ze beschrijven dus ook hoe de persistentie van de slechtst presterende scholen verschilt van die van de best presterende scholen. Er is geen duidelijk patroon te ontwaren in de mate waarin een indicator iets zegt over de lange termijn. Bij het onderbouwrendement is geen verschil tussen de convergentie aan de onder- en de bovenkant. Bij de cijfers voor kernvakken op het centraal eindexamen en bij het bovenbouwrendement gaat de convergentie sneller tussen slecht presterende scholen en gemiddeld presterende scholen. Maar bij de andere cijferindicatoren zijn het juist de best presterende scholen die sneller naar het gemiddelde terugveren.

De volgende figuren laten grafisch zien wat ook in Tabel 5 staat, en bieden tevens een blik op het tijdsverloop van prestatieverschillen tussen scholen. Figuur 1 geeft een overzicht van de gemiddelde waarde van de combinatiemaatstaf voor alle jaren nadat een groep scholen als slechtst, als gemiddeld of als best presterend is bestempeld. Deze groepen zijn dus een aantal jaren na deze beoordeling gevolgd. De combinatiemaatstaf varieert daarbij tussen de 1 (alle schoolsoorten van de school presteren op alle indicatoren het slechtst) en de 10 (alle schoolsoorten van de school presteren op alle indicatoren het best). Het verschil in de gemiddelde scores van de slechtst en van de best presterende groep vervalft duidelijk het meeste gedurende het eerste jaar na de classificatie. In de jaren twee tot en met zes zet de trend naar een gemiddeld presterende school door, zij het op een veel gematigder voet.

Figuur 1: Gemiddelde combinatiemaatstaf in jaren na de kwalificatie slechtst, gemiddeld of best presterend



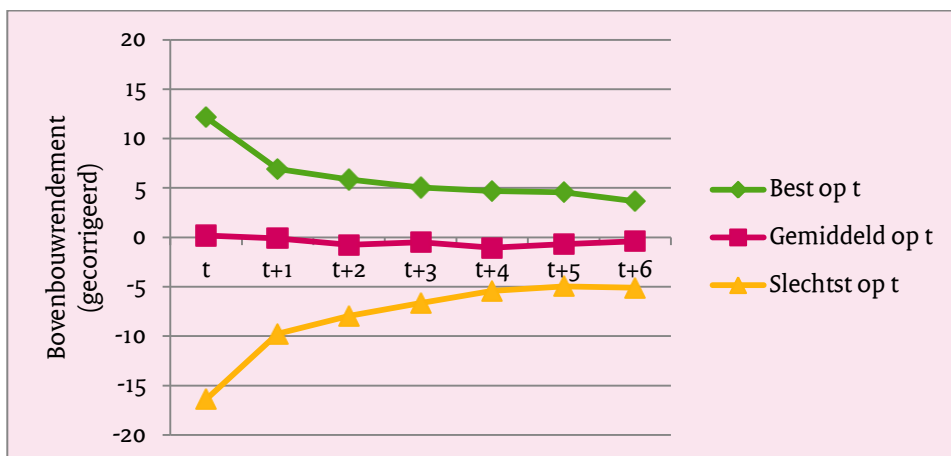
Figuur 2: Gemiddelde cijfers CE in jaren na de kwalificatie slechtst, gemiddeld of best presterend



Figuur 3: Gemiddelde onderbouwrendement in jaren na de kwalificatie slechtst, gemiddeld of best presterend



Figuur 4: Gemiddelde bovenbouwrendement in jaren na de kwalificatie slechtst, gemiddeld of best presterend



De figuren 2, 3 en 4 presenteren dezelfde gegevens voor de individuele prestatie-indicatoren. Figuur 2 gaat in op de (gecorrigeerde) cijfers voor het centraal eindexamen. Figuur 3 geeft een beeld van het onderbouwrendement, terwijl Figuur 4 dit doet voor het bovenbouwrendement. Doordat hier gecorrigeerde maatstaven worden gepresenteerd, heeft het niveau van de indicator - in tegenstelling tot de ontwikkeling in de indicator - geen betekenis meer. Uit de figuren valt duidelijk op te maken dat de verschillen in het onderbouwrendement sneller afvlakken dan de verschillen in de cijfers of het bovenbouwrendement.

Weinig scholen presteren jaar in jaar uit het slechtst of best

Tot dusver werden de prestaties van de slechtst en best presterende scholen vertegenwoordigd door de gemiddelde score per groep per tijdsperiode. Maar deze groepen bestaan uit individuele scholen die ook weer kunnen verschillen in hun prestaties.

Tabel 6 gaat na hoeveel slechtst presterende scholen en hoeveel best presterende scholen consequent dit predicaat ontvangen. De waarden in de tabel geven het percentage aan van de scholen dat drie of zes jaar achter elkaar als slechtst of als best presterend wordt beoordeeld. Uitgaande van de combinatiemaatstaf, behoort 1,21% van alle scholen drie jaar achter elkaar tot de 10% slechtst presterende scholen en 1,56% van alle scholen drie jaar achter elkaar tot de 10% best presterende scholen. Slechts een selecte groep scholen weet deze status zes jaar lang vol te houden: maar 0,05% van de scholen presteert continu het slechtst op de combinatiemaatstaf en maar 0,13% van de scholen presteert continu het best. Wel dient te worden opgemerkt dat deze groepen een stuk kleiner zouden zijn als de prestaties volledig willekeurig waren. In dat geval zou 0,1% van de scholen drie jaar lang en 0,0001% van de scholen zes jaar lang in één van de twee extreme categorieën vallen.

Als wordt gekeken naar de individuele prestatie-indicatoren, dan vallen twee dingen op. Ten eerste presteren er meer scholen consequent heel slecht of heel goed op een enkele indicator dan op de combinatiemaatstaf. Dit hangt samen met het feit dat de combinatiemaatstaf een gemiddelde is van drie individuele indicatoren die niet sterk samenhangen (zie de eerdere uitleg in deze subsectie). Ten tweede presteren er meer scholen consequent aan de bovenkant van de verdeling, dan dat er scholen consequent aan de onderkant daarvan presteren. Dit valt mogelijk te verklaren doordat de Inspectie van het Onderwijs scholen aan de onderkant van de verdeling dwingt tot verbetering, maar zich niet bemoeit met de bovenkant.

Tabel 6: Percentage scholen dat persistent als slechtste of beste presteert (voor scholen met één vestiging)

	(I)	(II)	(III)	(IV)
	Slechtst presterend 3 jaar lang	6 jaar lang	Best presterend 3 jaar lang	6 jaar lang
Cijfers CE	2,0%	0,8%	2,5%	0,8%
Onderbouwrendement	1,4%	0,7%	2,4%	1,0%
Bovenbouwrendement	1,7%	0,4%	2,1%	0,9%
Combinatiemaatstaf	1,21%	0,05%	1,56%	0,13%

3 De relatie tussen prestaties en personeelskenmerken

Deze sectie legt met behulp van regressiemethoden een relatie tussen de prestaties van scholen in het voortgezet onderwijs en hun (personeels)kenmerken. Hierbij krijgen de slechtst en best presterende scholen de meeste aandacht. Het zal blijken dat de geobserveerde eigenschappen van scholen nauwelijks kunnen verklaren of een school zich aan de onderkant of bovenkant van de prestatieverdeling bevindt.

3.1 Veel personeelskenmerken van scholen zijn bekend

In deze subsectie worden de data over de eigenschappen van scholen beschreven. De data over prestaties zijn al aan bod gekomen in Sectie 2. Tabel B2 in de bijlage bevat een lijst van alle eigenschappen in de dataset en voor al deze kenmerken de gemiddelde waarde en standaarddeviatie. Veel van de kenmerken zeggen iets over de medewerkers van een school. Om een overzicht te geven van de eigenschappen, zijn deze opgedeeld in acht groepen.

Allereerst is er een groep met vrij algemene kenmerken van een school, zoals het aantal leerlingen, het aantal leerlingen in een samengestelde brugklas (dit is een klas die meerdere schoolsoorten bevat) en het aantal medewerkers per leerling. De tweede groep bevat (demografische) eigenschappen van medewerkers, zoals de leeftijd en het geslacht. Daarnaast is er een groep met eigenschappen die de verandering in het personeelsbestand van een school weergeven (het percentage medewerkers dat de school verlaat in een jaar en het percentage medewerkers dat nieuw is in een jaar). De vierde groep is een groep met karakteristieken die de bijdrage van leraren en directieleden aan de verschillende eigenschappen meten (bijvoorbeeld het percentage van de vrouwen of nieuwe medewerkers dat leraar is). Ook is er een groep met eigenschappen over de bevoegdheid van leraren (met bijvoorbeeld informatie over het percentage bevoegde leraren en het type bevoegdheid) en een groep met salariskarakteristieken (zoals het gemiddelde salaris van medewerkers). Bovendien is van sommige personeelskarakteristieken niet alleen de waarde voor de groep van alle medewerkers bekend, maar ook de waarde voor vier functiegroepen: directieleden, leraren, ondersteunende medewerkers en leraren in opleiding (LIO's). De groep met deze functiegroepspecifieke eigenschappen bevat bijvoorbeeld per school het percentage vertrekkende directieleden, het aantal leerlingen per leraar en het gemiddelde salaris van ondersteunende medewerkers. Ten slotte, is er een groep met kenmerken die informatie geven over de onderwijsmethode en religieuze achtergrond van een school.

De eigenschappen verschillen genoeg van elkaar om de individuele bijdrage van een variabele te kunnen bepalen (met andere woorden, er zijn geen aanwijzingen dat multicolineariteit een probleem is). Voor slechts één combinatie van eigenschappen die tegelijkertijd worden opgenomen in een regressie is de correlatie hoger dan 0,7 (de correlatie tussen 'omvang baan leraren/omvang baan directie' en 'salaris leraren/salaris directie' is 0,75).

Voor sommige kenmerken bestaat er een voor de hand liggende hypothese voor de relatie tussen de eigenschap en de prestaties van een school. De derde kolom van Tabel B2 in de bijlage laat per kenmerk deze hypothese zien. Zo is de verwachting dat er een positief verband bestaat tussen de prestaties van een school en respectievelijk het salaris dat een school betaalt, het percentage bevoegde leraren en het aantal andere scholen in de gemeente (concurrentie). Ook is bijvoorbeeld de verwachting dat er een negatief verband bestaat tussen de prestaties en het aantal leerlingen per medewerker. Echter, voor veel kenmerken zijn meerdere verbanden mogelijk. Het aangegeven verband is de relatie waarvan het waarschijnlijk is dat deze overheerst. Zo is in Tabel B2 de verwachting dat er een positief verband bestaat tussen prestaties en nieuwe medewerkers, omdat een school een nieuwe medewerker aanstelt als dit een verbetering is ten opzichte van de oude situatie. Maar het is natuurlijk ook mogelijk dat prestaties juist achteruit zullen gaan omdat de leraar en leerlingen eerst moeten wennen aan de nieuwe situatie. Daarnaast is er voor veel kenmerken op voorhand geen kennis over hoe de relatie eruit ziet. De leeftijd van leraren zou bijvoorbeeld positief kunnen bijdragen aan de prestaties van een school omdat oudere leraren over het algemeen meer ervaring zullen

hebben dan jongere leraren. Het effect zou echter ook negatief kunnen zijn omdat jongere leraren wellicht recenter hun opleiding hebben afgerond en daarom de nieuwste kennis beter beheersen. Andere kenmerken waarvan het effect op voorhand niet duidelijk is, zijn bijvoorbeeld de aanwezigheid van veel samengestelde brugklassen, het geslacht van medewerkers, het belang van leraren voor prestaties versus het belang van directieleden voor prestaties en de schoolmethode. Voor al deze eigenschappen kunnen geen expliciete hypothesen worden geformuleerd, maar zal de data moeten spreken. Het bestaan van een relatie tussen deze eigenschappen en de prestaties kan niet a priori worden uitgesloten, maar er is geen kennis over hoe deze relatie eruit zou zien. Bijvoorbeeld, gaan relatief hoge salarissen voor leraren of relatief hoge salarissen voor directieleden samen met goede prestaties?

De data over eigenschappen bestrijken de periode 1994-2008. Echter, enkele interessante kenmerken zijn slechts voor een gedeelte van deze periode beschikbaar, bijvoorbeeld voor het begin of het einde van de periode. De laatste kolom van Tabel B2 bevat voor elk kenmerk het maximum aantal observaties dat beschikbaar is voor een school. Het gemiddelde salaris is bijvoorbeeld voor maximaal zeven jaren aanwezig. De uiteindelijke analyses bevatten ongeveer 300 scholen als alleen scholen met één vestiging worden gebruikt en ongeveer 500 scholen als alle scholen worden gebruikt. Het totale databestand bevat 975 unieke registratienummers van scholen, maar niet voor al deze registratienummers zijn alle variabelen aanwezig. De data komen uit verschillende bronnen en niet voor alle scholen kunnen deze bronnen aan elkaar worden gekoppeld. Daarnaast zijn er voor sommige scholen weinig observaties beschikbaar. De scholen die niet worden gebruikt door een gebrek aan data lijken willekeurig te zijn verdeeld over de populatie. Echter, zoals in Sectie 2 al is besproken, bestaan er wel verschillen tussen de scholen met slechts één vestiging en scholen met meerdere vestigingen. Om het effect hiervan te bestuderen, worden de schattingsresultaten voor zowel de populatie van alle scholen als de populatie van scholen met één vestiging gepresenteerd (het zal blijken dat de belangrijkste kwalitatieve conclusies niet afhangen van de keuze van de populatie).

Zoals hierboven al is vermeld, komt de gebruikte informatie uit verschillende bronnen. De meeste data zijn afkomstig van DUO. DUO is de uitvoeringsorganisatie van de Rijksoverheid voor onderwijs en beschikt daarom over veel (administratieve) gegevens. Het eerste databestand van DUO dat in deze Notitie wordt gebruikt, komt voort uit de salarisadministratie van scholen. Dit bestand bevat eigenschappen zoals het gemiddelde salaris van medewerkers, de leeftijd van medewerkers en het verloop van medewerkers. Het tweede databestand van DUO dat wordt gebruikt, bevat informatie over de bevoegdheid van leraren en de (hoeveelheid) vakken die ze geven. Dit bestand komt voort uit personeelstellingen (IPTO). Ten slotte zijn enkele eigenschappen afkomstig uit de opbrengstenkaarten van de Inspectie van het Onderwijs. Sectie 2 beschrijft deze databron uitgebreider aangezien dit ook de bron is voor de data over prestaties. Gebruikte informatie van de Inspectie van het Onderwijs is bijvoorbeeld het aantal leerlingen en de religieuze achtergrond van een school. Overigens zijn in de originele databronnen een paar eigenschappen niet gemeten per school, maar op een lager niveau (bijvoorbeeld schoolvestiging). Deze kenmerken zijn geaggregeerd tot een eigenschap op schoolniveau.

3.2 Prestaties worden beperkt verklaard door eigenschappen

Allereerst wordt de score van een school op de combinatiemaatstaf verklaard. Hiervoor wordt de ‘fixed effects within estimator’ gebruikt. Deze schattingsmethode benut de veranderingen in de prestaties en eigenschappen van een school door de tijd heen. De methode kijkt dus of en in welke mate de prestaties van een school veranderen als de eigenschappen van deze school veranderen. Daarnaast houdt deze schattingsmethode rekening met niet-geobserveerde eigenschappen. De geobserveerde eigenschappen van scholen zijn voornamelijk die van de leraren op deze scholen. Echter, naast deze eigenschappen van leraren (zoals salaris), zijn ook niet-geobserveerde eigenschappen van leraren (zoals motivatie) en niet-geobserveerde eigenschappen van leerlingen (bijvoorbeeld talent voor het halen van hoge cijfers) van belang voor de prestaties van een school. Bij het construeren van de prestatie-indicatoren is geprobeerd om zo goed mogelijk te corrigeren voor niet-geobserveerde eigenschappen van leerlingen (zie Sectie 2). Echter, deze correctie is niet perfect. Een voordeel van de gebruikte schattingsmethode is dat niet-geobserveerde factoren die variëren tussen scholen, maar die constant zijn door de tijd, geen effect hebben op de

schattingsresultaten. Dus als sommige scholen structureel meer leerlingen binnen krijgen waarmee het makkelijker is om goed te presteren, dan heeft dit geen invloed op de geschatte relatie tussen de prestaties en geobserveerde eigenschappen.

Er moet een aanname worden gemaakt over hoe lang het duurt voordat een verandering in een eigenschap invloed heeft op de prestaties van een school. Immers, als een school vandaag allemaal bevoegde leraren aanstelt, dan heeft dat nog geen effect op de examencijfers van morgen. Daarom wordt als waarde van de eigenschap de gemiddelde waarde van de eigenschap in het jaar van de prestatie en de twee voorgaande jaren gebruikt. Deze waarde wordt gekozen omdat het onderbouwendement wordt bepaald door de prestaties van leerlingen die in principe drie jaar op een school zitten. Een nadeel van een lange(re) vertraging is dat veel data niet kunnen worden gebruikt voor het schatten. Als een gemiddelde over twee of vier jaar wordt gebruikt, dan veranderen van geen enkele van de gerapporteerde resultaten de kwalitatieve conclusies. Het gebruiken van vertraagde eigenschappen lost ook mogelijke endogeniteitsproblemen op: het maakt het bestaan van een causale relatie tussen eigenschappen en prestaties aannemelijker. Het lijkt namelijk onwaarschijnlijk dat de huidige prestaties van scholen invloed hebben op eigenschappen van deze scholen in het verleden.

Tabel 7 rapporteert de schattingsresultaten voor drie verschillende specificaties. De specificaties verschillen in het aantal opgenomen schooleigenschappen. Omdat niet alle data voor dezelfde periode en in dezelfde mate beschikbaar zijn, bestaat er een afweging tussen het aantal observaties dat gebruikt kan worden en het aantal eigenschappen dat kan worden opgenomen in de vergelijking. Bovendien geven de drie verschillende specificaties direct een test op de robuustheid van de resultaten. De eerste kolom van de tabel laat de specificatie zien met de minste eigenschappen, maar met de meeste observaties. Deze specificatie bevat het aantal leerlingen op een school, het aantal andere scholen in de gemeente, het percentage leerlingen in een samengestelde brugklas, het aantal leerlingen per medewerker, het percentage vrouwelijke medewerkers op een school, het percentage medewerkers per leeftijdscategorie, het percentage medewerkers per omvang van de aanstelling en het percentage medewerkers dat vertrekt en erbij komt in een jaar. Daarnaast bevat het de verhouding tussen het aantal leraren en het aantal directieleden en variabelen die meten of er relatief veel leraren of directieleden vrouw zijn, in deeltijd werken, vertrekken of erbij komen. Ten behoeve van de interpretatie zijn alle percentages uitgedrukt als fracties (dat wil zeggen, de waarden liggen tussen 0 en 1). In de tweede kolom is informatie over de bevoegdheid van leraren toegevoegd (een benadering van het percentage bevoegde leraren, het percentage bevoegde leraren dat tweedegraads bevoegd is, het percentage meervoudig bevoegde leraren en het percentage leraren dat meerdere vakken geeft) en in de derde kolom is informatie over de salarissen van personeel toegevoegd (niveau, spreiding en de verhouding tussen het gemiddelde leraren- en directiesalaris).

De opgenomen eigenschappen kunnen de prestaties van een school niet goed verklaren, zie bijvoorbeeld de ter illustratie in de onderste rij opgenomen R^2 . Bovendien blijken de resultaten sterk af te hangen van de specificatie en de exacte definitie van de variabelen. In de eerste specificatie heeft geen enkele variabele een robuust statistisch significante impact op de combinatiemaatstaf. Het percentage nieuwe medewerkers is hier weliswaar statistisch significant, maar dit is geen robuust resultaat. Kleine wijzigingen in de geschatte vergelijking beïnvloeden namelijk dit resultaat. Zo is deze eigenschap bijvoorbeeld niet significant in de tweede en derde kolom van de tabel. Ook de tweede specificatie levert geen inzichten op. De gerapporteerde significantie van het percentage bevoegde leraren verdwijnt namelijk als de specificatie marginaal wordt gewijzigd. Ten slotte heeft ook de specificatie waarin de salarissenmerken zijn opgenomen weinig verklarende kracht. Echter, uit deze vergelijking volgt wel de redelijk robuuste bevinding dat grote verschillen in de salarissen binnen een school een negatieve impact hebben op de prestaties. Daarnaast is het opvallend dat, hoewel slechts eenmaal significant, in alle specificaties personeel tussen de 45 en 55 jaar (de categorie waarmee de andere categorieën worden vergeleken) en ouder dan 55 jaar het beste lijkt voor de prestaties van een school.

Tabel 7: Schattingsresultaten voor de combinatiemaatstaf

	(I)	(II)	(III)
	Specificatie 1	Specificatie 2	Specificatie 3
Algemene eigenschappen			
aantal leerlingen (per 100)	-0,041	-0,027	0,006
aantal scholen in gemeente	0,042	0,025	-0,097
% leerlingen in samengestelde brugklas	-0,352	0,040	-0,039
aantal leerlingen per medewerker	-0,055	-0,046	-0,145
Eigenschappen van medewerkers			
% vrouwelijke medewerkers	1,314	3,127	-4,987
% medewerkers < 25 jaar	-6,518	-6,054	1,968
% medewerkers 25-35 jaar	-0,899	-2,403	-2,829
% medewerkers 35-45 jaar	-1,464	-2,333*	-1,522
% medewerkers > 55 jaar	1,911	1,186	2,230
% medewerkers < 0,2 fte	-5,306	3,154	7,703
% medewerkers 0,2-0,4 fte	-0,322	0,512	6,016
% medewerkers 0,4-0,6 fte	2,375	1,955	2,343
% medewerkers 0,6-0,8 fte	-1,059	-1,930	-0,003
Verandering van personeelsbestand			
% nieuwe medewerkers	3,223*	2,027	2,280
% vertrekkende medewerkers	-2,397	0,074	-3,045
Bijdrage van leraren aan eigenschappen (ten opzichte van bijdrage directieleden)			
% leraren van de medewerkers	0,834	-2,716	-7,116
% leraren van de vrouwen	0,486	2,805	4,092
omvang baan leraren/omvang baan directie	-0,045	0,072	-0,474
% leraren van de nieuwe medewerkers	-0,179	-0,791*	-0,220
% leraren van de vertrekkende medewerkers	0,385	0,911	0,122
Bevoegdheid			
% bevoegde leraren (benadering)		1,817*	
% 2e-graads bevoegd van bevoegde leraren		-0,914	
% leraren dat meervoudig bevoegd is		0,270	
% leraren dat meerdere vakken geeft		0,631	
Salaris			
gemiddeld salaris (per 100)			0,021
standaarddeviatie salaris (per 100)			-0,747**
salaris leraren/salaris directie			-0,300
Aantal observaties	2653	1849	1231
Aantal scholen	335	311	289
R ² (within)	0,03	0,04	0,04

Notities: de symbolen *, ** en *** geven aan dat een variabele significant is op respectievelijk 5%, 1% en 0,1%. De (niet-gerapporteerde) standaardfouten zijn geclusterd op schoolniveau. Schoolspecifieke en jaarspecifieke 'fixed effects' zijn niet gerapporteerd.

Tabel B3 en B4 in de bijlage presenteren enkele van de uitgevoerde robuustheidsanalyses. De tabellen bevatten alleen de alternatieve analyses voor de specificatie met de meeste observaties (de eerste kolom van Tabel 7), maar de andere specificaties geven gelijke uitkomsten. Als eerste robuustheidsanalyse is voor de individuele prestatie maatstaven een aparte vergelijking geschat. Tabel B3 rapporteert de resultaten voor de drie componenten van de combinatiemaatstaf. Sectie 2 laat zien dat de individuele maatstaven sterk van elkaar verschillen. Dit komt terug in de regressieresultaten omdat de geschatte (niet-significante) coëfficiënten ook behoorlijk van elkaar verschillen. Zo leidt het vertrek van relatief veel leraren (en dus van relatief weinig directieleden) tot een significant hogere score op het onderbouwrendement, maar niet tot hogere cijfers voor het centraal eindexamen of een beter bovenbouwrendement. Verder bieden deze schattingsresultaten weinig aanknopingspunten. De schattingsresultaten voor de andere individuele maatstaven zijn niet gerapporteerd, omdat de kwalitatieve conclusies die uit deze resultaten volgen hetzelfde zijn.

De eerste kolom van Tabel B4 bevat de schattingsresultaten voor de vergelijking waar de drie componenten van de combinatiemaatstaf niet meer gecorrigeerd zijn voor verschillen tussen leerlingen (zie Sectie 2). De correctie zou er voor moeten zorgen dat verschillen tussen scholen in prestaties worden verminderd. Immers, goede en slechte prestaties worden (deels) verklaard door de leerlingen die naar een school toe gaan. Sectie 2 laat al zien dat deze correctie de maatstaf beïnvloedt. De schattingsresultaten bevestigen deze kennis, want voor de meeste (soms nu significante) eigenschappen is het effect zonder correctie net iets sterker positief of negatief dan het effect met correctie. De belangrijkste conclusie van deze vergelijking is echter dat de resultaten niet drastisch verschillen en dat de conclusies niet direct volgen uit de manier waarop de maatstaven worden gecorrigeerd.

In de overige kolommen van Tabel B4 wordt het effect van aggregatie over scholen en schoolsoorten bestudeerd. Aggregatie heeft, tot zover met behulp van de beschikbare data kan worden beoordeeld, geen grote invloed op de resultaten. Eerst wordt de aggregatie over vestigingen besproken. Tot nu toe zijn alleen scholen met één vestiging gebruikt in de analyse. Dit omdat 'vestigingen' van scholen soms zeer zelfstandig opereren en in de praktijk soms alleen onder hetzelfde schoolbestuur vallen. Door alleen scholen met één vestiging te gebruiken, zijn de scholen beter vergelijkbaar. Een nadeel is echter dat ongeveer 40% van de data niet kan worden gebruikt en dat de scholen die wel worden gebruikt niet helemaal representatief zijn (zie Sectie 2). Om het effect hiervan te bestuderen, wordt ook de hele analyse uitgevoerd voor de populatie van alle scholen. De resultaten die hieruit volgen bevestigen het beeld dat volgt uit de populatie van scholen met één vestiging. De tweede kolom van Tabel B4 rapporteert het schattingsresultaat voor de eerste specificatie. Allereerst valt op dat ook in deze vergelijking (die veel meer observaties gebruikt) geen enkele eigenschap een statistisch significante invloed heeft op de prestaties. Daarnaast verschillen de geschatte coëfficiënten niet sterk van de eerder geschatte coëfficiënten (zie de eerste kolom van Tabel 7). Als tweede wordt het effect van aggregatie over schoolsoorten besproken. Hoewel zo goed mogelijk rekening wordt gehouden met de verschillende schoolsoorten die scholen aanbieden, blijft het natuurlijk altijd onzeker in hoeverre het bestaan van verschillende schoolsoorten de vergelijking van prestaties beïnvloedt. Bovendien kan de relatie tussen prestaties en een eigenschap verschillen tussen schoolsoorten en dit komt mogelijk niet aan het licht bij geaggregeerde data (eerstegraads bevoegde leraren kunnen bijvoorbeeld een ander effect hebben op het vwo dan op het vmbo). Om de invloed van aggregatie over schoolsoorten te bestuderen, worden dezelfde vergelijkingen geschat als eerder, maar nu met als populatie scholen met één vestiging die alleen vmbo aanbieden en scholen met één vestiging die alleen vwo aanbieden (dit zijn respectievelijk ongeveer 60 en 40 scholen). In deze populaties is dus geen sprake van aggregatie over vestigingen en geen sprake van aggregatie over schoolsoorten (en scholengemeenschappen). De resultaten bieden eigenlijk geen nieuwe inzichten. De derde en vierde kolom van Tabel B4 laten zien dat ook in dit geval de schoolkenmerken de prestaties slechts beperkt kunnen verklaren. In de specificatie voor het vwo is geen enkele variabele significant, in de specificatie voor het vmbo zijn het percentage nieuwe en vertrekkende medewerkers significant, maar dit is geen robuust resultaat.

Overigens is ook gezocht naar niet-lineaire verbanden tussen prestaties en eigenschappen van scholen (bijvoorbeeld, het optimale aantal leerlingen op een school of de man-vrouwverhouding van het personeel), maar deze resultaten worden niet gerapporteerd omdat ze geen aanvullende inzichten geven.

3.3 Ook de slechtste en beste prestaties worden beperkt verklaard

In het bijzonder zijn de eigenschappen van scholen die het slechtst en het best presteren interessant. Een school wordt in een bepaald jaar gedefinieerd als slechtst presterend (best presterend) als deze school in dat jaar bij de 10% slechtst (best) presterende scholen hoort. Er worden twee verschillende vergelijkingen geschat. Eén vergelijking waarbij de te verklaren variabele 1 is als de school slecht presteert en 0 als dit niet het geval is. En één vergelijking waarbij de te verklaren variabele 1 is als de school bij de best presterende scholen hoort en 0 als dit niet het geval is. Op deze manier kan worden getest welke eigenschappen de kans verhogen dat een school bij de groep hoort met de slechtst presterende scholen en welke eigenschappen de kans verhogen dat een school hoort bij de groep met de best presterende scholen. Als verklarende variabelen voor de binaire prestatievariabelen worden dezelfde groepen van eigenschappen gebruikt als in de vorige subsectie. De 'logit fixed effects estimator' wordt gebruikt om de vergelijkingen te schatten. Ook deze methode houdt rekening met niet-geobserveerde factoren die constant zijn door de tijd, maar verschillen tussen scholen.

Tabel 8 bevat de schattingsresultaten. In plaats van de geschatte coëfficiënten, rapporteert de tabel de 'odds ratios'. Een 'odds ratio' is de kans dat een voorval gebeurt, gedeeld door de kans dat het voorval niet gebeurt. Dus als een 'odds ratio' gelijk is aan 2, dan is de kans twee keer zo groot dat het voorval gebeurt dan dat het voorval niet gebeurt. En een 'odds ratio' van een 0,5 betekent dat de kans twee keer zo groot is dat het voorval niet gebeurt dan dat het voorval wel gebeurt. In termen van de gerapporteerde resultaten geeft een 'odds ratio' dus aan hoe de kans verandert dat een school bij de slechtst of best presterende groep hoort als de eigenschap met 1 eenheid (of in het geval van procenten, met 100%) stijgt. De oneven kolommen laten de vergelijkingen zien waar de groep met slechtst presterende scholen wordt verklaard. Geen enkele eigenschap is significant van belang (en robuust) voor het verklaren van de kans dat een school slecht presteert. Bovendien hangen de resultaten wederom sterk af van de exacte specificatie (zie in de gepresenteerde specificaties bijvoorbeeld hoe sterk de geschatte 'odds ratios' verschillen voor het percentage vrouwelijke medewerkers). De even kolommen laten dezelfde resultaten zien voor de vergelijkingen waar de groep met best presterende scholen wordt verklaard. Ook hier helpt geen van de eigenschappen bij het verklaren van de kans dat een school hoort bij de groep met best presterende scholen. Al met al kunnen de slechtste en beste prestaties van scholen maar matig worden verklaard door de geobserveerde eigenschappen.

Tabel B5 en B6 in de bijlage presenteren enkele robuustheidsanalyses. De gerapporteerde analyses zijn allemaal op basis van de specificatie waarvoor de meeste observaties beschikbaar zijn. Eerst rapporteert Tabel B5 de schattingsresultaten van de twee vergelijkingen voor de drie componenten van de combinatiemaatstaf. Aangezien deze maatstaven substantieel verschillen, verschillen ook de waarden van de geschatte 'odds ratios'. Het lijkt echter niet mogelijk om algemene lessen te trekken uit deze resultaten. Voor de overige individuele maatstaven is dezelfde analyse uitgevoerd, maar de resultaten hiervan worden niet gerapporteerd omdat er dezelfde kwalitatieve conclusies uit volgen. Het blijft opvallend dat de eigenschappen van scholen niet of nauwelijks helpen bij het verklaren van de groep met slechtst en best presterende scholen.

De eerste en tweede kolom van Tabel B6 rapporteren de resultaten van de schattingen waarbij de combinatiemaatstaf bestaat uit de normale waarden van de drie componenten (in plaats van de waarden die zijn gecorrigeerd voor de eigenschappen van leerlingen). In dit geval lijkt de correctie toch een behoorlijke invloed te hebben. Sommige eigenschappen zijn opeens significant en andere, niet-significante, eigenschappen krijgen het tegenovergestelde effect. De derde en vierde kolom van Tabel B6 rapporteren de schattingsresultaten voor de hele populatie van scholen (dus inclusief de scholen met meer dan één vestiging). Ondanks de stijging van het aantal observaties, is hier in geen van de twee vergelijkingen een eigenschap van belang voor het verklaren van de prestaties. Daarnaast valt op dat deze schattingsresultaten minder lijken op de resultaten voor de populatie van scholen met één vestiging dan in de vorige subsectie.

Tabel 8: Schattingsresultaten voor de binaire combinatiemaatstaf

	(I)	(II)	(III)	(IV)	(V)	(VI)
	Binaire combinatiemaatstaf					
	Specificatie 1		Specificatie 2		Specificatie 3	
	Slechtst	Best	Slechtst	Best	Slechtst	Best
Algemene eigenschappen						
aantal leerlingen (per 100)	1,018	0,861	1,049	0,819	0,918	0,860
aantal scholen in gemeente	0,857	1,054	0,691	0,975	1,706	1,568
% leerlingen in samengestelde brugklas	1,537	0,691	0,738	1,330	0,064	0,856
aantal leerlingen per medewerker	1,148	0,973	1,215	1,244	1,466	0,811
Eigenschappen van medewerkers						
% vrouwelijke medewerkers	2,405	145,309	0,009	>1000	>1000	0,000
% medewerkers < 25 jaar	>1000	27,552	>1000	>1000	>1000	>1000
% medewerkers 25-35 jaar	0,516	0,149	>1000	0,007	0,351	0,000
% medewerkers 35-45 jaar	39,840	0,012	62,923	0,000	>1000	0,000
% medewerkers > 55 jaar	0,059	11,667	0,004	18,893	0,174	6,698
% medewerkers < 0,2 fte	463,210	0,000	0,139	23,654	>1000	>1000
% medewerkers 0,2-0,4 fte	0,313	0,002	12,866	0,000	0,000	0,000
% medewerkers 0,4-0,6 fte	0,034	25,182	0,010	422,438	0,126	0,000
% medewerkers 0,6-0,8 fte	0,433	0,040	6,670	0,003	0,028	0,000
Verandering van personeelsbestand						
% nieuwe medewerkers	0,000*	59,246	0,000	12,400	0,000	0,001
% vertrekkende medewerkers	0,220	0,032	0,047	82,770	358,424	>1000
Bijdrage van leraren aan eigenschappen (ten opzichte van bijdrage directieleden)						
% leraren van de medewerkers	0,009	24,850	18,215	>1000	0,000	0,000
% leraren van de vrouwen	9,575	0,450	32,764	>1000	0,800	0,355
omvang baan leraren/omvang baan directie	1,252	0,331	0,732	0,019	1,451	0,000
% leraren van de nieuwe medewerkers	1,250	0,844	1,355	0,109	0,805	2,616
% leraren van de vertrekkende medewerkers	0,551	1,299	0,074	1,960	0,156	1,832
Bevoegdheid						
% bevoegde leraren (benadering)			1,611	13,354		
% 2e-graads bevoegd van bevoegde leraren			0,040	0,051		
% leraren dat meervoudig bevoegd is			0,565	1,228		
% leraren dat meerdere vakken geeft			0,825	5,091		
Salaris						
gemiddeld salaris (per 100)					0,772	0,390
standaarddeviatie salaris (per 100)					1,257	0,071
salaris leraren/salaris directie					1,489	8,610

	(I)	(II)	(III)	(IV)	(V)	(VI)
	Binaire combinatiemaatstaf					
	Specificatie 1		Specificatie 2		Specificatie 3	
	Slechtst	Best	Slechtst	Best	Slechtst	Best
Aantal observaties	1132	1048	604	631	344	331
Aantal scholen	127	120	83	91	74	70
Pseudo-R ²	0,04	0,03	0,10	0,08	0,09	0,17

Notities: de symbolen *, ** en *** geven aan dat een variabele significant is op respectievelijk 5%, 1% en 0,1%. De (niet-gerapporteerde) standaardfouten zijn berekend via de 'bootstrap'-methode. Schoolspectifieke en jaarspectifieke 'fixed effects' zijn niet gerapporteerd. Als de teller van een 'odds ratio' stijgt, dan daalt per definitie de noemer (en vice versa). Daarom worden 'odds ratios' relatief snel heel groot of heel klein. Met het oog op de presentatie worden 'odds ratios' die groter zijn dan 1000 gerapporteerd als '>1000'.

3.4 Een andere aanpak verandert de conclusies niet

In deze subsectie wordt de variatie in de data tussen scholen gebruikt. Er wordt gekeken of en in welke mate de eigenschappen van scholen die relatief het best en het slechtst presteren verschillen. De vorige subsecties bevatten schattingsmethodes die alleen gebruik maken van de verandering in de eigenschappen van een school door de tijd heen. Echter, voor veel eigenschappen is de variatie tussen scholen groter dan de variatie van een school door de tijd heen. Voor een school zijn bijvoorbeeld de eigenschappen van leraren redelijk constant door de tijd, omdat het merendeel van de leraren voor een langere periode bij deze school werkt. Daarom varieert bijvoorbeeld het gemiddelde salaris dat een school betaalt relatief weinig tussen jaren, wat het moeilijk maakt om het effect van veranderingen in het salaris vast te stellen. Bovendien is het de vraag in hoeverre jaar op jaar veranderingen in de prestatie-indicatoren informatief zijn. Het zou kunnen dat de jaarlijkse prestatie van een school slechts een toevallige afwijking is van de gemiddelde prestaties van een school over een langere periode (toevallig wil hier zeggen dat de afwijking wordt veroorzaakt door niet-geobserveerde oorzaken, zoals eigenschappen van leerlingen). Immers, de eerdere schattingen kunnen prestaties slechts zeer gedeeltelijk verklaren. Ten slotte, kan door het gebruiken van variatie tussen scholen worden gekeken naar het effect van eigenschappen die constant zijn door de tijd, zoals de lesmethode of religieuze achtergrond van een school.

Onderbouwrendement is de enige maatstaf waarvoor variatie tussen scholen kan worden gebruikt. Schattingsmethodes die de variatie tussen scholen gebruiken (zoals de 'random effects estimator' en de 'between estimator') zijn alleen correct onder vrij sterke condities. Voor deze methodes is het namelijk noodzakelijk dat de niet-observeerbare factoren die de prestatie verklaren (en die constant zijn door de tijd), niet gecorreleerd zijn met de verklarende variabelen. Bij veel prestatie maatstaven gaat dit waarschijnlijk niet op. Bijvoorbeeld, in een regressie van cijfers van leerlingen op eigenschappen van leraren is een belangrijke niet-geobserveerde factor de eigenschappen van de leerlingen (bijvoorbeeld talent of motivatie). Vaak zullen deze eigenschappen van leerlingen gecorreleerd zijn met de eigenschappen van leraren. Denk bijvoorbeeld aan de motivatie van de leraar en de redenen voor een leraar om de school te verlaten. Echter, in de maatstaf onderbouwrendement spelen de niet-observeerbare factoren van leerlingen een minder grote rol omdat het immers de prestaties van leerlingen afzet tegen het basisschooladvies van deze leerlingen (Sectie 2 vindt ook dat de correctie minder invloed heeft op het onderbouwrendement dan op de andere maatstaven). Daarom lijkt alleen voor deze maatstaf de vereiste aanname voor het gebruiken van variatie tussen scholen minder problematisch. Een Hausman-test bevestigt dit: alleen voor de maatstaf onderbouwrendement kan in alle drie de specificaties niet de nulhypothese worden verworpen dat er geen relatie is tussen de observeerbare en niet-observeerbare factoren die constant zijn door de tijd. Dus volgens deze test kunnen de 'random effects estimator' en de 'between estimator' worden gebruikt om het onderbouwrendement te verklaren.

Tabel 9: Schattingsresultaten voor het onderbouwrendement

	(I)	(II)	(III)
	Onderbouwrendement		
	Specificatie 1	Specificatie 2	Specificatie 3
Algemene eigenschappen			
aantal leerlingen (per 100)	-0,008	-0,011	-0,007
aantal scholen in gemeente	0,018	0,010	0,019
% leerlingen in samengestelde brugklas	-0,134	0,072	0,603
aantal leerlingen per medewerker	-0,012	-0,068	-0,065
Eigenschappen van medewerkers			
% vrouwelijke medewerkers	-4,773	-7,760	-6,164
% medewerkers < 25 jaar	-21,927	-23,554	-32,761
% medewerkers 25-35 jaar	5,294	8,478	14,172
% medewerkers 35-45 jaar	1,321	-2,900	-3,811
% medewerkers > 55 jaar	-5,477	-8,956	-8,877
% medewerkers < 0,2 fte	-30,199	-17,582	-46,849
% medewerkers 0,2-0,4 fte	-1,725	-7,712	4,152
% medewerkers 0,4-0,6 fte	-3,894	5,861	7,997
% medewerkers 0,6-0,8 fte	0,761	-6,812	2,614
% nieuwe medewerkers	3,075	7,785	-5,354
Verandering van personeelsbestand			
% vertrekkende medewerkers	4,204	-2,047	3,996
% leraren van de medewerkers	-27,087	-39,564*	-32,590*
Bijdrage van leraren aan eigenschappen (ten opzichte van bijdrage directieleden)			
% leraren van de vrouwen	6,384	8,400	4,283
omvang baan leraren/omvang baan directie	0,602	-2,310	-1,142
% leraren van de nieuwe medewerkers	1,600	6,667	6,605
% leraren van de vertrekkende medewerkers	8,636	1,544	-1,948
Bevoegdheid			
% bevoegde leraren (benadering)		-0,165	
% 2e-graads bevoegd van bevoegde leraren		2,734	
% leraren dat meervoudig bevoegd is		-0,702	
% leraren dat meerdere vakken geeft		-0,831	
Salaris			
gemiddeld salaris (per 100)			0,070
standaarddeviatie salaris (per 100)			-0,246
salaris leraren/salaris directie			4,727
Methode en religie			
Daltonschool	1,399	1,321	1,296
Montessorischool	-1,116	-2,250	-1,674
gereformeerde school	-4,416	2,639	
reformatorische school	2,463	-3,541	
katholieke school	-0,044	-2,812	-1,299
interconfessionele school	0,190	-0,750	-0,125

	(I)	(II)	(III)
	Onderbouwrendement		
	Specificatie 1	Specificatie 2	Specificatie 3
samenwerkingsverband met andere visie	0,604	-0,148	-0,004
Aantal observaties	4127	3199	2213
Aantal scholen	506	482	491
R ² (between)	0,09	0,08	0,08

Notities: de symbolen *, ** en *** geven aan dat een variabele significant is op respectievelijk 5%, 1% en 0,1%. De (niet-gerapporteerde) standaardfouten zijn berekend via de 'bootstrap'-methode. Schoolspectifieke en jaarspectifieke 'fixed effects' zijn niet gerapporteerd.

Omdat onderbouwrendement voor een kortere periode beschikbaar is dan de andere individuele maatstaven (vanaf 1999 in plaats van 1996) en er daarom relatief weinig observaties beschikbaar zijn, worden de resultaten gepresenteerd van de analyse waarin alle scholen zijn gebruikt (en niet alleen de scholen met één vestiging). De conclusies die volgen uit deze resultaten zijn grotendeels hetzelfde. Waar ze afwijken, wordt dit vermeld.

Eerst wordt de absolute score van een school op het onderbouwrendement verklaard. In Tabel 9 worden de resultaten van de 'between estimator' gerapporteerd om een zo sterk mogelijk contrast te krijgen met de resultaten uit Tabel 7 (de 'random effects estimator' geeft een gewogen gemiddelde van de 'between estimator' en de eerder gebruikte methode). De resultaten bieden weinig houvast voor de verklaring van de prestaties. Het is opvallend dat veel variabelen (zoals salaris, bevoegdheid, het percentage van het personeel dat vertrekt en het percentage van het personeel dat er nieuw bij komt) geen duidelijke, significante impact hebben op het onderbouwrendement van een school. De ter indicatie in de onderste rij gerapporteerde R² (gebaseerd op de variatie tussen scholen) laat voor elke specificatie zien dat de verklarende kracht relatief beperkt is, maar toch wat groter is dan in de eerder gepresenteerde specificaties. De gepresenteerde vergelijkingen geven wel aan dat scholen met een goed onderbouwrendement vaker scholen zijn waar relatief veel directieleden werken (en dus relatief weinig leraren). Hoewel niet-statistisch significant, is opvallend dat hier, in tegenstelling tot bij de eerdere resultaten, relatief veel personeel met een leeftijd tussen de 25 en 35 jaar leidt tot een hogere score (uit de eerdere resultaten bleek dat veel personeel van 45 jaar en ouder de score op de combinatiemaatstaf verbeterd).

Ten tweede wordt onderzocht welke eigenschappen kunnen verklaren dat een school hoort bij de groep met scholen die het slechtst presteren of juist bij de groep met scholen die het best presteren op het onderbouwrendement. Tabel 10 rapporteert de resultaten (hier is de vergelijking geschat met de 'logit random effects estimator'). Hoewel ook in deze exercitie veel eigenschappen geen verklarende kracht hebben voor de prestaties, brengt de exercitie toch enkele nieuwe inzichten naar voren. Het eerste dat opvalt, is dat in alle specificaties het aantal leerlingen van een school een significante invloed heeft op de prestaties: een stijging van het aantal leerlingen verkleint de kans dat een school hoort bij de groep met slechtst presterende scholen, maar verkleint tegelijkertijd ook de kans dat een school hoort bij de groep met best presterende scholen. Dit suggereert dat scholen met weinig leerlingen relatief vaak het slechtst en het best scoren op het onderbouwrendement. Echter, dit resultaat volgt niet uit de analyse gebaseerd op de populatie van scholen met slechts één vestiging. Een verklaring hiervoor zou kunnen zijn dat het aantal leerlingen op deze scholen minder verschilt (deze leerlingen moeten immers in één vestiging passen) en dat het daarom moeilijker is om de gevonden relatie vast te stellen. Een tweede resultaat dat volgt uit Tabel 10 is dat scholen met meer andere scholen in hun gemeente een kleinere kans hebben om bij de groep met slechtst presterende scholen te horen. Dit zou kunnen betekenen dat concurrentie voorkomt dat een school hoort bij de slechtst presterende scholen (maar er is geen bewijs dat het een school helpt bij het bereiken van de groep met best presterende scholen). Het laatste resultaat dat volgt uit de tabel is dat veel vrouwelijke

Tabel 10: Schattingsresultaten voor het binaire onderbouwrendement

	(I)	(II)	(III)	(IV)	(V)	(VI)
	Binair onderbouwrendement					
	Specificatie 1		Specificatie 2		Specificatie 3	
	Slechtst	Best	Slechtst	Best	Slechtst	Best
Algemene eigenschappen						
aantal leerlingen (per 100)	0,955***	0,926***	0,956**	0,923***	0,954*	0,912***
aantal scholen in gemeente	0,963*	1,007	0,978	1,007	0,946*	1,012
% leerlingen in samengestelde brugklas	1,671	1,055	1,562	1,101	2,117	1,800
aantal leerlingen per medewerker	0,971	1,016	0,973	1,002	0,895	1,015
Eigenschappen van medewerkers						
% vrouwelijke medewerkers	3,042	0,662	2,720	0,225	0,353	0,338
% medewerkers < 25 jaar	321,759	0,018	>1000	0,014	145,622	0,002
% medewerkers 25-35 jaar	0,100	6,991	0,044	14,386	0,002	57,136
% medewerkers 35-45 jaar	0,327	0,319	1,211	0,166	0,003	0,097
% medewerkers > 55 jaar	2,882	0,245	5,409	0,203	0,019	0,113
% medewerkers < 0,2 fte	1,039	0,005	1,274	0,386	>1000	0,003
% medewerkers 0,2-0,4 fte	0,891	7,418	0,007	3,006	6,477	164,510
% medewerkers 0,4-0,6 fte	0,437	0,362	4,567	0,693	3,646	299,793
% medewerkers 0,6-0,8 fte	2,329	0,083	2,051	0,019	28,596	0,342
Verandering van personeelsbestand						
% nieuwe medewerkers	0,684	7,134	0,476	2,972	0,132	3,076
% vertrekkende medewerkers	6,552	0,036	1,463	0,062	21,588	0,005
Bijdrage van leraren aan eigenschappen (ten opzichte van bijdrage directieleden)						
% leraren van de medewerkers	>1000	2,175	>1000	8,138	>1000*	0,006
% leraren van de vrouwen	3,910	0,004*	0,865	0,001*	0,296	0,029
omvang baan leraren/omvang baan directie	0,278	0,476	0,223	0,363	0,152	3,356
% leraren van de nieuwe medewerkers	2,144	1,236	2,059	0,752	0,887	0,442
% leraren van de vertrekkende medewerkers	0,644	4,873	0,256	4,667	2,044	4,418
Bevoegdheid						
% bevoegde leraren (benadering)			1,191	1,948		
% 2e-graads bevoegd van bevoegde leraren			0,634	1,234		
% leraren dat meervoudig bevoegd is			0,489	0,519		
% leraren dat meerdere vakken geeft			1,998	0,600		
Salaris						
gemiddeld salaris (per 100)					1,036	1,289
standaarddeviatie salaris (per 100)					1,227	0,826
salaris leraren/salaris directie					2,382	0,069

	(I)	(II)	(III)	(IV)	(V)	(VI)
	Binair onderbouwrendement					
	Specificatie 1		Specificatie 2		Specificatie 3	
	Slechtst	Best	Slechtst	Best	Slechtst	Best
Methode en religie						
Daltonschool	1,288	2,006	1,481	1,915	1,876	1,096
Montessorischool	1,127	0,171	0,889	0,097	1,649	0,125
gereformeerde school	2,292	0,837	3,889	1,356		
reformatorische school	1,571	0,982	1,391	1,017		
katholieke school	3,136	1,109	2,622	0,622	5,648	0,668
interconfessionele school	1,069	0,639	0,916	0,741	1,106	0,627
samenwerkingsverband met andere visie	0,768	0,987	0,602	1,006	0,830	0,914
Aantal observaties	4127	4127	3199	3199	2213	2213
Aantal scholen	506	506	482	482	491	491
Pseudo-R ²	0,02	0,03	0,03	0,04	0,05	0,04

Notities: gelijk aan de notities bij Tabel 8.

leraren (en weinig vrouwelijk directieleden) de kans verkleint dat een school hoort bij de best presterende scholen. Dit impliceert dat op de best presterende scholen relatief veel vrouwen directielid zijn (en dus relatief weinig vrouwen leraar).

3.5 Functiegroepspecifieke kenmerken verklaren niet veel beter

Tot slot wordt een vierde specificatie gebruikt die ook enkele eigenschappen per functiegroep bevat. De vier functiegroepen zijn directie, leraren, ondersteunend personeel en leraren in opleiding (LIO's). Voor elk van deze functiegroepen bevat de specificatie het aantal leerlingen per medewerker, het percentage vrouwen, het percentage nieuwe medewerkers, het percentage vertrekkende medewerkers en het salarisoniveau. Een nadeel van deze specificatie is dat er relatief weinig observaties kunnen worden gebruikt voor de schattingen. Eerst wordt de score van een school op de combinatiemaatstaf verklaard. Deze regressie sluit aan bij de in Tabel 7 gepresenteerde vergelijkingen (zie Sectie 3.2). De eerste kolom van Tabel 11 presenteert het resultaat. Het blijkt dat van alle functiegroepspecifieke kenmerken slechts één statistisch significant en redelijk robuust is: het aantal leerlingen per leraar. Meer leerlingen per leraar leidt tot een lagere score op de combinatiemaatstaf. Een stijging van drie leerlingen per leraar maakt van een gemiddeld presterende (50%) school ongeveer een 5%-school, dus een school die hoort bij de groep met slechtst presterende scholen. De significantie van deze eigenschap is opvallend omdat in Tabel 7 het geschatte effect van het aantal leerlingen per medewerker wel consistent negatief is, maar niet significant. Er zijn geen andere robuuste en significante resultaten.

Vervolgens wordt het onderbouwrendement verklaard door de specificatie met functiegroepspecifieke eigenschappen. Deze regressie sluit aan bij de in Tabel 9 gepresenteerde vergelijkingen (zie Sectie 3.4). De tweede kolom van Tabel 11 laat zien dat geen enkele functiegroepspecifieke variabele significant is. Wel blijkt dat een personeelsbestand met veel medewerkers die ouder zijn dan 55 jaar samengaat met een lage score op het onderbouwrendement. Dit sluit aan bij de eerdere bevindingen (zie Tabel 9).

Tabel 11: Schattingsresultaten met functiegroepspecifieke eigenschappen

	(I) Combinatiemaatstaf Specificatie 4 Scholen met 1 vestiging	(II) Onderbouwrendement Specificatie 4 Alle scholen
Algemene eigenschappen		
aantal leerlingen (per 100)	0,177	0,036
aantal scholen in gemeente	-1,357*	0,023
% leerlingen in samengestelde brugklas	1,329	1,194
Eigenschappen van medewerkers		
% medewerkers < 25 jaar	-19,029	-39,198
% medewerkers 25-35 jaar	-8,776	13,810
% medewerkers 35-45 jaar	-6,762	-30,268
% medewerkers > 55 jaar	3,089	-28,365*
% medewerkers < 0,2 fte	-13,074	33,536
% medewerkers 0,2-0,4 fte	-13,813	7,119
% medewerkers 0,4-0,6 fte	-1,915	14,969
% medewerkers 0,6-0,8 fte	-1,067	-1,226
Eigenschappen per functiegroep		
aantal leerlingen per directielid	0,001	-0,003
aantal leerlingen per leraar	-0,607***	-0,137
aantal leerlingen per ondersteunende medewerker	-0,091*	0,055
aantal leerlingen per LIO	-0,000	0,000
% vrouwelijke directieleden	0,704	-2,183
% vrouwelijke leraren	-5,215	11,081
% vrouwelijke ondersteunende medewerkers	-1,013	-6,872
% vrouwelijke LIO's	-0,273	-2,500
% nieuwe directieleden	-1,941	-3,781
% nieuwe leraren	9,170	-0,979
% nieuwe ondersteunende medewerkers	-4,780	29,493
% nieuwe LIO's	-1,097	-8,272
% vertrekkende directieleden	1,630	-0,652
% vertrekkende leraren	4,920	25,214
% vertrekkende ondersteunende medewerkers	-1,796	-0,604
% vertrekkende LIO's	-0,902	12,324
gemiddeld salaris directieleden (per 100)	-0,058	-0,278
gemiddeld salaris leraren (per 100)	-0,111	0,924
gemiddeld salaris ondersteunende medewerkers (per 100)	-0,180	-0,319
gemiddeld salaris LIO's (per 100)	0,120	-0,049
Methode en religie		
Daltonschool		-0,848
Montessorischool		0,515

	(I) Combinatiemaatstaf Specificatie 4 Scholen met 1 vestiging	(II) Onderbouwrendement Specificatie 4 Alle scholen
katholieke school		0,541
interconfessionele school		1,340
samenwerkingsverband met andere visie		-0,304
Aantal observaties	302	710
Aantal scholen	135	259
R ² (within) / R ² (between)	0,22	0,18

Notities: de symbolen *, ** en *** geven aan dat een variabele significant is op respectievelijk 5%, 1% en 0,1%. In Kolom I zijn de (niet-gerapporteerde) standaardfouten geclusterd op schoolniveau en in Kolom II zijn deze berekend via de 'bootstrap'-methode. Schoolspecifieke en jaarspecifieke 'fixed effects' zijn niet gerapporteerd.

3.6 Interpretatie en beperkingen

In de voorgaande subsecties wordt met behulp van verschillende schattingsmethoden gezocht naar een relatie tussen de prestaties en personeelskenmerken van scholen. Hierbij worden zowel verschillen in de data tussen scholen gebruikt, als verschillen in de data door de tijd heen. De score van een school op de combinatiemaatstaf en op de individuele maatstaven wordt verklaard. Daarnaast is er in het bijzonder aandacht voor scholen die het slechtst en het best presteren en wordt een groot aantal verschillende specificaties bestudeerd. Als de resultaten uit al deze analyses worden samengenomen, dan blijkt dat er voor geen enkele eigenschap een robuust statistisch significant verband is met de prestaties. Formeel gesteld: voor geen enkele eigenschap kan de hypothese dat er geen relatie is met de prestaties worden verworpen. De uitkomst is dus dat er geen resultaten zijn.

Er zitten twee belangrijke beperkingen aan de gebruikte data. Ten eerste worden de prestaties van scholen gemeten via het gemiddelde cijfer voor het centraal eindexamen, het bovenbouwrendement, het onderbouwrendement en het verschil tussen het gemiddelde cijfer op het centraal eindexamen en de schoolonderzoeken. Dit is een smalle definitie van schoolprestaties. Deze prestaties staan niet gelijk aan de algemene kwaliteit van een school. Immers, een gedeelte van de prestaties van scholen wordt niet gemeten. Er wordt bijvoorbeeld niet gemeten wat scholen bijdragen aan de lange-termijn-doelen van onderwijs, zoals het niveau van de vervolgopleidingen van leerlingen en het latere inkomen van leerlingen. Ook bevatten de gebruikte prestatie-indicatoren geen informatie over het 'zachtere' gedeelte van de prestaties, zoals de sociale ontwikkeling van leerlingen, hoe goed leerlingen kunnen samenwerken, het plezier waarmee leerlingen naar school gaan en het ziekteverzuim. Dat er geen verband wordt gevonden tussen prestaties en eigenschappen wil daarom niet zeggen dat de geobserveerde eigenschappen niet van belang zijn voor schoolprestaties. Immers, het zou kunnen dat de eigenschappen van invloed zijn op het niet-gemeten gedeelte van de prestaties (zie bijvoorbeeld Card en Krueger (1992) en Betts (1995)).

Ten tweede zijn er beperkingen aan de beschikbare data over leerlingen en scholen. Veel kenmerken van deze twee zijn niet bekend. Van leerlingen is het bijvoorbeeld niet bekend wat leerlingen al hebben geleerd op de basisschool, wat leerlingen leren van hun ouders, wat het talent van leerlingen is voor het halen van hoge cijfers en wat het doorzettingsvermogen van leerlingen is. Ook van scholen is veel informatie niet bekend. Zo is van de leraren bijvoorbeeld weliswaar de bevoegdheid bekend, maar er is geen tot weinig informatie over andere, moeilijker te meten, facetten van hun kwaliteit. Denk bijvoorbeeld aan hoe goed

een leraar kan vertellen of orde houden. Of hoe gemotiveerd een leraar is en hoe betrokken hij is bij de school en leerlingen. Daarnaast is er bijvoorbeeld niets bekend over de kwaliteit van de sfeer op een school.

Het eerste gevolg van dit gebrek aan data is dat de prestatie maatstaven imperfect worden gecorrigeerd voor de omstandigheden waar een school geen invloed op heeft. Vooral de onbekende eigenschappen van leerlingen lijken hier van belang. Deze imperfecte correctie heeft natuurlijk een directe invloed op de rangorde van scholen: scholen met veel leerlingen waarmee het relatief eenvoudig is om goed te presteren blijven, ondanks de correctie, in het voordeel. En scholen met veel leerlingen waarmee het relatief moeilijk is om goed te presteren, blijven in het nadeel. Alleen voor de prestatie-indicator onderbouwrendement lijkt de imperfecte correctie minder belangrijk omdat deze de verandering in de prestaties van leerlingen meet.

Ten tweede kunnen de niet-geobserveerde kenmerken van leerlingen en scholen de interpretatie van de schattingsresultaten beïnvloeden. Dit werkt als volgt. Door de imperfecte correctie zit er ruis in de prestatie maatstaven. Daarnaast bestaan er mogelijk ongeobserveerde kenmerken die van belang zijn voor het verklaren van de prestaties van scholen (denk aan de bovenstaande ongeobserveerde kenmerken van leraren). Als de ruis in de maatstaven of één van de ongeobserveerde relevante kenmerken gecorreleerd is met de geobserveerde eigenschappen van scholen, dan kan het effect van de geobserveerde eigenschappen op de prestaties niet goed worden geschat. Er kan in dat geval namelijk geen onderscheid worden gemaakt tussen het effect op de prestaties van de geobserveerde eigenschappen en het effect op de prestaties van de ruis of ongeobserveerde kenmerken ('omitted variable bias'). Hierdoor kan het geschatte effect van de geobserveerde eigenschappen zowel groter als kleiner zijn dan het in werkelijkheid is (zie ook bijvoorbeeld Oosterbeek en Webbink (2001) en Webbink (2005)).

In dit specifieke geval lijkt op voorhand een overschatting waarschijnlijker dan een onderschatting. Dit terwijl een overschatting minder invloed zou hebben op de conclusies. Elk van de bovengenoemde onbekende kenmerken van leerlingen en scholen zal waarschijnlijk een positief effect hebben op de prestaties van een school. Bovendien, als deze onbekende kenmerken zouden samenhangen met bijvoorbeeld het percentage bevoegde leraren (of bijvoorbeeld het gemiddelde salaris op een school), dan is het waarschijnlijk dat er een positieve samenhang is. Met andere woorden, de verwachting zou zijn dat leerlingen waarmee het makkelijker is om goed te presteren, vaker naar een school gaan met meer bevoegde leraren. En dat leraren die betere didactische vaardigheden hebben, gemiddeld gesproken, vaker bevoegd zijn. Daarom zou, grofweg gezegd, het niet opnemen van deze onbekende kenmerken betekenen dat het effect van het percentage bevoegde leraren sterker positief zou worden geschat dan het in werkelijkheid is. Het lijkt alsof bevoegdheid de prestaties verklaart, maar het zijn eigenlijk de ongeobserveerde kenmerken. Een dergelijke overschatting zou echter geen invloed hebben op de kwalitatieve conclusies van deze Notitie. Immers, er wordt zelfs met deze mogelijke overschatting al geen robuust significante bijdrage van bevoegdheid gevonden. Soortgelijke redeneringen gaan op voor de andere geobserveerde eigenschappen.

Hoe dan ook, een te groot of te klein geschat effect door de beperkte data over kenmerken van leerlingen en scholen valt nooit helemaal uit te sluiten. Mocht hier echter sprake van zijn, dan verandert alleen de interpretatie van de resultaten. In plaats van het causale effect van een eigenschap van een school op de prestaties (oftewel het 'directe effect'), wordt dan de correlatie tussen een eigenschap en de prestaties gemeten (het 'totale effect'). Dus in dit laatste geval zou de interpretatie van de resultaten zijn dat geen enkele geobserveerde eigenschap van scholen gecorreleerd is met de prestaties.

4 Conclusie: prestaties van scholen fluctueren door de tijd en kunnen beperkt worden verklaard door personeelskenmerken

In deze CPB Notitie worden eerst verschillende maatstaven vergeleken die de prestaties van scholen in het voortgezet onderwijs meten. Deze prestatie-indicatoren worden jaarlijks berekend door de Inspectie van het Onderwijs en omvatten ondermeer het gemiddelde cijfer voor het centraal eindexamen, de mate waarin leerlingen op een schoolsoort eindigen die boven het oorspronkelijke schooladvies ligt (het onderbouwrendement) en de mate waarin leerlingen met een diploma van school gaan (het bovenbouwrendement). Het blijkt dat de verschillende maatstaven behoorlijk van elkaar verschillen. In de eerste plaats verschillen zij in de mate waarin een correctie voor geobserveerde verschillen tussen leerlingen uitmaakt. Deze correctie wordt toegepast om de invloed op de prestatie-indicatoren van omstandigheden waar een school geen invloed op heeft zo klein mogelijk te maken. Zo is het onderbouwrendement slechts beperkt te verklaren door de kenmerken van leerlingen, terwijl de cijfers dit wel voor een groot deel zijn. Verder zijn de verbanden tussen de prestatie-indicatoren niet sterk. Zo correleren onderbouw- en bovenbouwrendement nauwelijks met elkaar. De keuze van de maatstaf heeft daarom een sterke invloed op welke scholen het predicaat slechtst presterend of het predicaat best presterend ontvangen. Een combinatiemaatstaf die de verschillende maatstaven samenvoegt, lost dit probleem gedeeltelijk op.

De prestaties van veel scholen (zoals gemeten door de indicatoren) zijn behoorlijk wisselend over de tijd. Na een jaar zijn de verschillen tussen de 10% slechtst presterende scholen en de 10% best presterende scholen ongeveer gehalveerd. Maar de prestaties zijn ook weer niet zo wisselend dat er niets aan ontleend kan worden. Zouden de prestaties volledig willekeurig zijn, dan zouden deze verschillen na een jaar niet meer bestaan. De prestaties aan de top lijken wel wat consistentier dan de prestaties aan de onderkant. Maar ook hier verschillen de prestatie-indicatoren van elkaar. Op onderbouwrendement zijn de meeste schommelingen in prestaties te vinden en op cijfers die leerlingen halen op het centraal eindexamen de minste. Er lijkt hier sprake te zijn van een afruil tussen hoe goed een maatstaf de toegevoegde waarde van een school meet en hoe persistent een maatstaf is door de tijd heen: een maatstaf zoals cijfers voor het centraal eindexamen waarop omstandigheden die de school niet kan beïnvloeden betrekkelijk veel invloed hebben, denk aan het type leerling dat een school heeft, varieert relatief weinig door de tijd heen. Terwijl een maatstaf zoals onderbouwrendement, waarop een betrekkelijk kleine invloed is van het type leerling, relatief veel veranderingen door de tijd heen laat zien.

Als tweede probeert deze CPB Notitie de prestaties van scholen volgens de verschillende maatstaven te verklaren door middel van (vooral) personeelskenmerken van deze scholen. Een behoorlijke hoeveelheid kenmerken is beschikbaar, zoals het salaris, de bevoegdheid van leraren en de instroom en uitstroom van personeel. Om een relatie te vinden tussen prestaties en eigenschappen, worden zowel econometrische methoden ingezet die gebruik maken van de verschillen in de prestaties en eigenschappen van een school door de tijd heen, als methoden die gebruik maken van de verschillen tussen scholen in de prestaties en eigenschappen. De geobserveerde eigenschappen kunnen de prestaties zoals gemeten door de indicatoren echter maar zeer beperkt verklaren. Voor geen enkele eigenschap is er een robuust statistisch significante relatie met de combinatiemaatstaf of met de afzonderlijke indicatoren. De resultaten ondersteunen niet de van tevoren verwachte relaties. De uitkomst is dan ook dat er geen resultaten zijn.

Overigens moet in ogenschouw worden genomen dat deze studie niet de algemene kwaliteit van scholen meet en verklaart. De prestatie-indicatoren meten slechts een gedeelte van de prestaties van een school. De indicatoren bevatten bijvoorbeeld geen informatie over de invloed van een school op de lange-termijndoelen van het onderwijs (zoals het latere inkomen van leerlingen en het niveau van hun vervolgopleidingen) en over het 'zachtere' gedeelte van de prestaties (zoals het plezier waarmee leerlingen naar school toe gaan). Het gebruiken van een brede kwaliteitsmaatstaf zou een ander beeld kunnen geven.

Dit is dan ook een interessante richting voor toekomstig onderzoek. Een belangrijke vraag hierbij is hoe de algemene kwaliteit van scholen het beste kan worden gemeten.

Daarnaast is een belangrijke vraag wat de variatie in de prestatie maatstaven wel kan verklaren. Immers, de geobserveerde eigenschappen van scholen doen dit niet en de variatie in de data over tijd is te klein om volledig aan toeval toe te kunnen schrijven. Toekomstig onderzoek zal deze oorzaken moeten identificeren. Voor dit onderzoek zullen echter meer data nodig zijn. Ten eerste is het mogelijk dat de variatie in de prestaties ontstaat doordat er slechts beperkte informatie beschikbaar is over de leerlingen van een school. Hierdoor kunnen de prestaties slechts imperfect worden geschoond voor verschillen tussen scholen die ontstaan door leerlingen. Data die zouden kunnen helpen bij het verder schonen zijn bijvoorbeeld de Cito-scores van leerlingen (beschikbaar via het onderwijsnummer) of meer informatie over de sociaaleconomische achtergrond van leerlingen. Ten tweede is het mogelijk dat de variatie ontstaat door niet-geobserveerde eigenschappen van scholen. Hoewel relatief veel eigenschappen wel beschikbaar zijn, zijn natuurlijk nog altijd veel eigenschappen niet beschikbaar. Zo is bijvoorbeeld relatief weinig bekend over de praktische vaardigheden van leraren zoals hoe goed een leraar kan presenteren, omgaan met leerlingen en orde houden. Onderwijsevaluaties zouden een mogelijke bron kunnen zijn voor deze meer subjectieve informatie. Extra data over eigenschappen van leerlingen en scholen kunnen ook helpen bij het beter schatten van het effect op de prestaties van de reeds geobserveerde eigenschappen. Zo kan in de toekomst de kennis over slecht en goed presterende scholen verder worden uitgebreid.

Literatuur

Alphen, S. van en J. Dronkers, 2011, Trouw schoolprestaties 2011, www.vanalphendronkers.com.

Betts, J.R., 1995, Does school quality matter? Evidence from the National Longitudinal Survey of Youth, *Review of Economics and Statistics*, vol. 77, nr. 2, pag. 231-250.

Card, D. en A.B. Krueger, 1992, Does school quality matter? Returns to education and the characteristics of public schools in the United States, *Journal of Political Economy*, vol. 100, nr. 1, pag. 1-40.

Coleman, J.S., E.Q. Campbell, C.J. Hobson, J. McPartland, A.M. Mood, F.D. Weinfeld en R.L. York, 1966, *Equality of Educational Opportunity*, U.S. Government Printing Office, Washington, DC.

Dijkstra, A.B., S. Karsten, R. Veenstra en A.J. Visscher, red., 2001, *Het oog der natie: scholen op rapport*, Van Gorcum, Assen.

Dronkers, J. en R. Veenstra, 2001, Schoolprestatie-indicatoren in het voortgezet onderwijs: start, reacties en vervolg, in A.B. Dijkstra, S. Karsten, R. Veenstra en A.J. Visscher, red., *Het oog der natie: scholen op rapport*, pag. 21-35, Van Gorcum, Assen.

Hanushek, E.A., 1997, Assessing the effects of school resources on student performance: an update, *Educational Evaluation and Policy Analysis*, vol. 19, nr. 2, pag. 141-164.

Hanushek, E.A. en S.G. Rivkin, 2006, Teacher quality, in E.A. Hanushek en F. Welch, red., *Handbook of the Economics of Education*, vol. 2, pag. 1051-1078, North-Holland, Amsterdam.

Hanushek, E.A. en S.G. Rivkin, 2010, Generalizations about using value-added measures of teacher quality, *American Economic Review: Papers & Proceedings*, vol. 100, nr. 2, pag. 267-271.

Kane, T.J. en D.O. Staiger, 2002, The promise and pitfalls of using imprecise school accountability measures, *Journal of Economic Perspectives*, vol. 16, nr. 4, pag. 91-114.

Koning, P. en K. van der Wiel, 2012, Ranking the schools: how school-quality information affects school choice in the Netherlands, *Journal of the European Economic Association*, nog te verschijnen.

Oosterbeek, H. en D. Webbink, 2001, Risico's van indicatoren van schoolkwaliteit, in A.B. Dijkstra, S. Karsten, R. Veenstra en A.J. Visscher, red., *Het oog der natie: scholen op rapport*, pag. 111-120, Van Gorcum, Assen.

Rivkin, S.G., E.A. Hanushek en J.F. Kain, 2005, Teachers, schools, and academic achievement, *Econometrica*, vol. 73, nr. 2, pag. 417-458.

Rothstein, J., 2010, Teacher quality in educational production: tracking, decay, and student achievement, *Quarterly Journal of Economic*, vol. 125, nr. 1, pag. 175-214.

Scheerens, J. en R.J. Bosker, 1997, *The Foundations of Educational Effectiveness*, Pergamon, Oxford.

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

Webbink, D., 2005, Causal effects in education, *Journal of Economic Surveys*, vol. 19, nr. 4, pag. 535-560.

Bijlage

Tabel B1-a: Illustratie van correctieregressies per schoolsoort voor cijfers voor het centraal eindexamen (alle vakken bij vmbo-b en vmbo-k, kernvakken bij overige schoolsoorten)

Schoolsoort	(I) Vmbo-b	(II) Vmbo-k	(III) Vmbo-gt	(IV) Havo	(V) Vwo
Kenmerken van gemeente					
bevolking (per 100.000)	-0,001	-0,029	-0,046**	-0,031	-0,059***
bevolking in het kwadraat (per 100.000)	0,003	0,007*	0,007***	0,007**	0,009***
Kenmerken van school					
aantal vestigingen	0,006**	-0,004*	-0,007***	-0,013*	-0,012**
vmbo-school	0,033	0,011	-0,020	n.v.t.	n.v.t.
scholengemeenschap	referentiecategorie				
lyceum	n.v.t.	n.v.t.	n.v.t.	0,115***	0,059**
gymnasium	n.v.t.	n.v.t.	n.v.t.	n.v.t.	0,443***
Kenmerken van leerlingen					
% van vmbo-leerlingen per vestiging in theoretische leergang	0,174***	-0,014	0,243***	-0,010	0,034
% lwoo-leerlingen per vestiging (0-1)	-0,292***	-0,142**	n.v.t.	n.v.t.	n.v.t.
% APC-leerlingen per schoolsoort (0-1)	-0,039	-0,178	-0,133	0,103	0,201
% APC-leerlingen per schoolsoort in het kwadraat	-0,314*	-0,303*	-0,380***	-0,776***	-1,054***
Jaren					
jaar=2005	0,130***	0,483***	0,144***	0,087**	0,054**
jaar=2006	0,130***	0,198***	0,026	0,254***	-0,112***
jaar=2007	0,015	0,156***	-0,047**	0,346***	-0,109***
jaar=2008	referentiecategorie				
Aantal observaties	1942	1972	3140	1612	1989
R ²	0,11	0,28	0,20	0,22	0,30

Notities: de symbolen *, ** en *** geven aan dat een variabele significant is op respectievelijk 5%, 1% en 0,1%. Voor de schattingen zijn zowel scholen met één vestiging als scholen met meerdere vestigingen gebruikt. De geschatte effecten van de categorieën zijn uitgedrukt ten opzichte van het geschatte effect van de bijbehorende referentiecategorie (dus bijvoorbeeld de coëfficiënt voor 'jaar=2005' is het extra effect ten opzichte van het effect voor 'jaar=2008').

Tabel B1-b: Illustratie van correctieregressies per schoolsoort voor bovenbouwrendement

Schoolsoort	(I) Vmbo-b	(II) Vmbo-k	(III) Vmbo-gt	(IV) Havo	(V) Vwo
Kenmerken van gemeente					
bevolking (per 100.000)	-1,533**	-1,737**	-1,814***	-1,246*	-2,479***
bevolking in het kwadraat (per 100.000)	0,197**	0,194*	0,210***	0,250**	0,377***
Kenmerken van school					
aantal vestigingen	0,164***	-0,098	-0,131*	-0,410**	-0,389*
vmbo-school	1,298**	0,337	1,010**	n.v.t.	n.v.t.
scholengemeenschap	referentiecategorie				
lyceum	n.v.t.	n.v.t.	n.v.t.	1,732**	2,659***
gymnasium	n.v.t.	n.v.t.	n.v.t.	n.v.t.	9,952***
Kenmerken van leerlingen					
% van vmbo-leerlingen per vestiging in theoretische leergang	2,185*	-0,388	-0,829	-0,070	1,500
% lwoo-leerlingen per vestiging (0-1)	-2,822*	-1,605	n.v.t.	n.v.t.	n.v.t.
% APC-leerlingen per schoolsoort (0-1)	-8,083***	-14,962***	-9,576***	-21,483***	-10,283*
% APC-leerlingen per schoolsoort in het kwadraat	2,531	7,504	-3,974	3,584	-8,431
Jaren					
jaar=2005	-1,168*	1,885**	0,587	-0,079	1,773**
jaar=2006	-1,046*	0,217	0,197	0,759	1,832**
jaar=2007	-0,187	1,216	0,077	2,292***	3,945***
jaar=2008	referentiecategorie				
Aantal observaties	1797	1878	2917	1878	1957
R ²	0,10	0,10	0,15	0,13	0,16

Notities: gelijk aan de notities bij Tabel B1-a.

Tabel B1-c: Illustratie van correctieregressies per scholengemeenschap voor onderbouwrendement

Scholengemeenschap	(I) Vmbo- school (b&k)	(II) Vmbo- school (gt)	(III) Brede gemeen- schap	(IV) Smalle gemeen- schap	(V) Lyceum	(VI) Gymna- sium
Kenmerken van gemeente						
bevolking (per 100.000)	1,331	1,621	1,365	-0,232	-0,047	-2,618*
bevolking in het kwadraat (per 100.000)	-0,090	-0,273	-0,198	0,079	0,014	0,225
Kenmerken van school						
aantal vestigingen	0,181*	-0,213	0,314	0,290	-0,429*	-0,695
Kenmerken van leerlingen						
% havo-leerlingen per vestiging (0-1)	n.v.t.	n.v.t.	-1,246	-2,983	-8,796***	n.v.t.
% vwo-leerlingen per vestiging (0-1)	n.v.t.	n.v.t.	0,361	-0,222	n.v.t.	n.v.t.
% van vmbo-leerlingen per vestiging met theoretische leergang	n.v.t.	n.v.t.	2,482	n.v.t.	n.v.t.	n.v.t.
% lwoo-leerlingen per vestiging (0-1)	16,140***	39,430***	5,963*	n.v.t.	n.v.t.	n.v.t.
% APC-leerlingen per vestiging (0-1)	-0,873	-6,060	5,515	-14,347	16,500	32,450*
% APC-leerlingen per vestiging in het kwadraat	-22,552	10,016	-2,238	45,461**	-49,295	-62,174
Jaren						
jaar=2005	-2,769*	-1,722	-0,010	-0,066	2,084*	1,594
jaar=2006	-0,658	0,391	1,052	0,418	0,692	1,297
jaar=2007	0,667	-0,919	-0,070	0,871	0,184	0,611
jaar=2008	referentiecategorie					
Aantal observaties	806	370	825	1224	319	157
R ²	0,10	0,23	0,03	0,02	0,09	0,11

Notities: gelijk aan de notities bij Tabel B1-a.

Tabel B2: Samenvattende statistiek voor de eigenschappen van scholen

	(I) Gemiddelde	(II) Standaard- deviatie	(III) Verwachte relatie met prestaties	(IV) Max. aantal observaties per school	
Algemene eigenschappen					
1	aantal leerlingen (per 100)	15,04	9,64	+/-	13
2	aantal scholen in gemeente	6,68	7,90	+	13
3	% leerlingen in samengestelde brugklas	0,60	0,36	+/-	13
4	aantal leerlingen per medewerker	12,26	3,69	-	13
Eigenschappen van medewerkers					
5	% vrouwelijke medewerkers	0,36	0,11	+/-	15
6	% medewerkers < 25 jaar	0,02	0,03	+/-	15
7	% medewerkers 25-35 jaar	0,15	0,07	+/-	15
8	% medewerkers 35-45 jaar	0,26	0,10	+/-	15
9	% medewerkers 45-55 jaar	0,39	0,10	+/-	15
10	% medewerkers > 55 jaar	0,18	0,09	+/-	15
11	% medewerkers < 0,2 fte	0,03	0,04	+/-	15
12	% medewerkers 0,2-0,4 fte	0,08	0,05	+/-	15
13	% medewerkers 0,4-0,6 fte	0,14	0,06	+/-	15
14	% medewerkers 0,6-0,8 fte	0,15	0,06	+/-	15
15	% medewerkers > 0,8 fte	0,61	0,11	+/-	15
Verandering van personeelsbestand					
16	% nieuwe medewerkers	0,09	0,06	+	13
17	% vertrekkende medewerkers	0,08	0,05	-	15
Bijdrage van leraren aan eigenschappen (ten opzichte van bijdrage directieleden)					
18	% leraren van de medewerkers	0,92	0,04	+/-	15
19	% leraren van de vrouwen	0,97	0,05	+/-	15
20	omvang baan leraren/omvang baan directie	0,83	0,13	+/-	15
21	% leraren van de nieuwe medewerkers	0,90	0,22	+/-	14
22	% leraren van de vertrekkende medewerkers	0,87	0,24	+/-	15
Bevoegdheid					
23	% bevoegde leraren (benadering)	0,90	0,18	+	12
24	% 2e-graads bevoegd van bevoegde leraren	0,49	0,18	-	12
25	% leraren dat meervoudig bevoegd is	0,38	0,20	+	12
26	% leraren dat meerdere vakken geeft	0,39	0,24	+/-	12
Salaris					
27	gemiddeld salaris (per 100)	25,00	2,30	+	7
28	standaarddeviatie salaris (per 100)	11,47	1,33	+/-	7
29	salaris leraren/salaris directie	0,59	0,12	+/-	7

	(I) Gemiddelde	(II) Standaard- deviatie	(III) Verwachte relaties met prestaties	(IV) Max. aantal observaties per school	
Eigenschappen per functiegroep					
30	aantal leerlingen per directielid	266,35	228,44	-	13
31	aantal leerlingen per leraar	15,15	3,31	-	13
32	aantal leerlingen per ondersteunende medewerker	59,72	33,77	-	8
33	aantal leerlingen per LIO	2242,60	3563,74	-	8
34	% vrouwelijke directieleden	0,14	0,19	+/-	15
35	% vrouwelijke leraren	0,37	0,11	+/-	15
36	% vrouwelijke ondersteunende medewerkers	0,47	0,13	+/-	9
37	% vrouwelijke LIO's	0,60	0,37	+/-	8
38	% nieuwe directieleden	0,08	0,15	+	14
39	% nieuwe leraren	0,09	0,07	+	14
40	% nieuwe ondersteunende medewerkers	0,13	0,11	+	8
41	% nieuwe LIO's	0,99	0,09	+	8
42	% vertrekkende directieleden	0,10	0,15	-	15
43	% vertrekkende leraren	0,08	0,06	-	15
44	% vertrekkende ondersteunende medewerkers	0,10	0,10	-	9
45	% vertrekkende LIO's	0,99	0,09	-	9
46	gemiddeld salaris directieleden (per 100)	46,56	6,41	+	7
47	gemiddeld salaris leraren (per 100)	26,84	2,43	+	7
48	gemiddeld salaris ondersteunende medewerkers (per 100)	15,87	2,93	+	7
49	gemiddeld salaris LIO's (per 100)	6,84	2,78	+	7
Methode en religie					
50	Daltonschool	0,02	0,13	+/-	15
51	Montessorischool	0,01	0,10	+/-	15
52	gereformeerde school	0,02	0,14	+/-	15
53	reformatorische school	0,04	0,19	+/-	15
54	katholieke school	0,01	0,12	+/-	15
55	interconfessionele school	0,29	0,45	+/-	15
56	samenwerkingsverband met andere visie	0,31	0,46	+/-	15

Notities: in de eigenschappen in rij 18, 19, 21 en 22 is het aantal (vrouwelijke) medewerkers de som van het aantal (vrouwelijke) directieleden en leraren. In alle eigenschappen omvat de term bevoegd alleen eerstegraads en tweedegraads bevoegden.

Tabel B3: Schattingsresultaten voor de individuele maatstaven

	(I) Cijfers centraal eindexamen Specificatie 1	(II) Onderbouw- rendement Specificatie 1	(III) Bovenbouw- rendement Specificatie 1
Algemene eigenschappen			
aantal leerlingen (per 100)	0,001	-0,170	-0,168
aantal scholen in gemeente	0,008	0,041	0,182
% leerlingen in samengestelde brugklas	0,005	-2,478	-1,332
aantal leerlingen per medewerker	0,001	-0,150	-0,590*
Eigenschappen van medewerkers			
% vrouwelijke medewerkers	0,337	2,069	3,332
% medewerkers < 25 jaar	-0,665	-16,437	-19,055
% medewerkers 25-35 jaar	-0,384*	-2,131	-0,850
% medewerkers 35-45 jaar	-0,253	-2,461	-1,363
% medewerkers > 55 jaar	0,166	-5,864	17,359***
% medewerkers < 0,2 fte	0,066	-4,148	-25,292
% medewerkers 0,2-0,4 fte	-0,213	-3,713	-0,299
% medewerkers 0,4-0,6 fte	0,232	-1,615	8,190
% medewerkers 0,6-0,8 fte	-0,179	-14,181*	3,966
Verandering van personeelsbestand			
% nieuwe medewerkers	-0,043	14,129	-1,718
% vertrekkende medewerkers	0,198	-10,893	-3,134
Bijdrage van leraren aan eigenschappen (ten opzichte van bijdrage directieleden)			
% leraren van de medewerkers	0,365	-1,286	2,032
% leraren van de vrouwen	0,202	-15,541*	10,477
omvang baan leraren/omvang baan directie	0,127**	-1,483	-2,708
% leraren van de nieuwe medewerkers	0,013	-1,061	1,493
% leraren van de vertrekkende medewerkers	0,080	4,272**	-1,403
Aantal observaties			
Aantal observaties	2641	2278	2623
Aantal scholen	334	323	333
R ² (within)	0,02	0,02	0,03

Notities: gelijk aan de notities bij Tabel 7.

Tabel B4: Robuustheid van de schattingsresultaten voor de combinatiemaatstaf

	(I) Ongecorrigeerde combinatie- maatstaf Specificatie 1 Scholen met 1 vestiging	(II) Alle scholen	(III) Combinatiemaatstaf Specificatie 1 Scholen met 1 vestiging en alleen vmbo	(IV) Scholen met 1 vestiging en alleen vwo
Algemene eigenschappen				
aantal leerlingen (per 100)	-0,054	-0,016	0,037	0,088
aantal scholen in gemeente	-0,042	0,025	-0,026	-0,087
% leerlingen in samengestelde brugklas	-0,470*	-0,146	0,186	-0,503
aantal leerlingen per medewerker	-0,044	-0,020	-0,126	-0,049
Eigenschappen van medewerkers				
% vrouwelijke medewerkers	1,189	0,526	4,958	-3,987
% medewerkers < 25 jaar	-7,286*	-3,534	-6,282	2,322
% medewerkers 25-35 jaar	-1,490	-0,711	-3,615	-0,174
% medewerkers 35-45 jaar	-1,797*	-1,389	-3,093	0,982
% medewerkers > 55 jaar	1,073	1,255	-0,549	3,022
% medewerkers < 0,2 fte	-5,137*	-3,252	3,169	-8,435
% medewerkers 0,2-0,4 fte	-0,897	0,681	4,069	1,030
% medewerkers 0,4-0,6 fte	1,733	2,065	7,530*	2,791
% medewerkers 0,6-0,8 fte	-2,247	-0,755	-2,887	6,760
Verandering van personeelsbestand				
% nieuwe medewerkers	4,507**	1,575	6,982*	2,471
% vertrekkende medewerkers	-2,677	-1,003	-7,969**	-4,597
Bijdrage van leraren aan eigenschappen (ten opzichte van bijdrage directieleden)				
% leraren van de medewerkers	0,987	-0,824	14,630	-3,887
% leraren van de vrouwen	1,534	1,566	3,809	1,428
omvang baan leraren/omvang baan directie	-0,080	-0,152	0,094	2,247
% leraren van de nieuwe medewerkers	-0,251	-0,167	-0,919	1,426
% leraren van de vertrekkende medewerkers	0,182	0,203	-0,655	-0,066
Aantal observaties				
Aantal observaties	2653	4705	465	339
Aantal scholen				
Aantal scholen	335	509	63	37
R² (within)				
R ² (within)	0,03	0,01	0,16	0,06

Notities: gelijk aan de notities bij Tabel 7.

Tabel B5: Schattingsresultaten voor de binaire individuele maatstaven

	(I) Binaire cijfers centraal eindexamen Specificatie 1		(III) Binair onderbouwrendement Specificatie 1		(V) Binair bovenbouwrendement Specificatie 1		(VI)	
	Slechtst	Best	Slechtst	Best	Slechtst	Best	Slechtst	Best
Algemene eigenschappen								
aantal leerlingen (per 100)	0,959	1,015	1,089	0,793	1,052	0,808		
aantal scholen in gemeente	1,074	1,249	0,944	0,925	0,896	0,950		
% leerlingen in samengestelde brugklas	0,621	0,435	0,881	0,181*	0,699	0,373		
aantal leerlingen per medewerker	0,929	0,811	1,097	1,058	1,315	0,910		
Eigenschappen van medewerkers								
% vrouwelijke medewerkers	0,026	5,698	975,631	2,179	0,109	0,913		
% medewerkers < 25 jaar	0,000	0,000	>1000	0,003	3,322	0,000		
% medewerkers 25-35 jaar	74,839	0,494	2,175	6,344	2,806	7,031		
% medewerkers 35-45 jaar	7,344	0,086	0,360	0,093	22,253	0,454		
% medewerkers > 55 jaar	1,692	6,799	37,653	0,028	0,011	31,454		
% medewerkers < 0,2 fte	549,484	290,114	0,004	0,000*	113,796	0,065		
% medewerkers 0,2-0,4 fte	0,000	20,908	7,680	0,901	0,255	39,481		
% medewerkers 0,4-0,6 fte	0,001	>1000	0,192	0,368	11,674	158,462		
% medewerkers 0,6-0,8 fte	0,041	15,032	90,177	0,006	0,050	111,523		
Verandering van personeelsbestand								
% nieuwe medewerkers	27,663	134,694	0,007	60,411	45,437	70,113		
% vertrekkende medewerkers	0,035	3,874	>1000	0,046	2,625	0,005		
Bijdrage van leraren aan eigenschappen (ten opzichte van bijdrage directieleden)								
% leraren van de medewerkers	2,142	>1000	>1000	33,448	716,197	>1000		
% leraren van de vrouwen	0,000*	0,020	23,447	0,000	0,001	2,803		
omvang baan leraren/omvang baan directie	0,001*	6,830	0,093	0,000*	1,465	0,386		
% leraren van de nieuwe medewerkers	0,582	0,583	4,222	1,203	0,960	0,338		
% leraren van de vertrekkende medewerkers	0,141	1,208	0,248	12,042*	1,640	0,753		
Aantal observaties								
Aantal observaties	922	904	888	1229	923	1024		
Aantal scholen	100	104	114	165	102	115		
Pseudo-R ²	0,08	0,04	0,04	0,08	0,06	0,04		

Notities: gelijk aan de notities bij Tabel 8.

Tabel B6: Robuustheid van de schattingsresultaten voor de binaire combinatiemaatstaf

	(I) Binaire ongecorrigeerde combinatiemaatstaf Specificatie 1 Scholen met 1 vestiging Slechtst	(II) Best	(III) Binaire combinatiemaatstaf Specificatie 1 Alle scholen Slechtst	(IV) Best
Algemene eigenschappen				
aantal leerlingen (per 100)	1,044	0,814	0,995	0,946
aantal scholen in gemeente	0,898	0,946	0,953	1,023
% leerlingen in samengestelde brugklas	1,751	0,728	0,822	0,812
aantal leerlingen per medewerker	1,439**	0,922	1,111	0,991
Eigenschappen van medewerkers				
% vrouwelijke medewerkers	0,003	26,075	5,630	39,789
% medewerkers < 25 jaar	>1000***	174,869	808,200	0,042
% medewerkers 25-35 jaar	>1000*	0,242	0,621	0,278
% medewerkers 35-45 jaar	939,750**	0,126	8,630	0,027
% medewerkers > 55 jaar	0,043	33,799	0,185	3,839
% medewerkers < 0,2 fte	889,404	0,000*	14,813	0,030
% medewerkers 0,2-0,4 fte	180,752	0,000**	0,066	0,145
% medewerkers 0,4-0,6 fte	0,062	0,404	0,012	12,107
% medewerkers 0,6-0,8 fte	4,029	0,003	2,214	0,066
Verandering van personeelsbestand				
% nieuwe medewerkers	0,000***	7,328	0,114	17,853
% vertrekkende medewerkers	308,505	18,427	6,489	0,430
Bijdrage van leraren aan eigenschappen (ten opzichte van bijdrage directieleden)				
% leraren van de medewerkers	0,024	>1000	0,890	0,379
% leraren van de vrouwen	17,709	0,127	0,050	1,430
omvang baan leraren/omvang baan directie	0,612	0,011	1,217	0,185
% leraren van de nieuwe medewerkers	4,222	2,724	1,697	0,849
% leraren van de vertrekkende medewerkers	0,675	0,484	1,039	1,046
Aantal observaties				
Aantal observaties	1038	884	2047	1988
Aantal scholen				
Aantal scholen	118	100	218	213
Pseudo-R²				
Pseudo-R ²	0,10	0,05	0,02	0,02

Notities: gelijk aan de notities bij Tabel 8.



Dit is een uitgave van:

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

info@cpb.nl | www.cpb.nl

Februari 2012