Inspectie van het Onderwijs
Ministerie van Onderwijs, Cultuur en Wetenschap

TECHNISCH RAPPORT BESTUREN MET LANGDURIG LAAGPRESTERENDE SCHOLEN

STAAT VAN HET ONDERWIJS 2021

april 2021

Inhoud

1 Inleiding 3

2 Kwantitatieve analyses basisonderwijs 4

2.1 Definities 4

2.2 Fluctuaties in opbrengsten 5

2.3 Laag presteren op schoolniveau 7

2.4 Laag presteren op bestuursniveau 8

2.5 Samenhang school- en bestuurskenmerken en laag presteren 10

2.6 Samenhang school- en bestuurskenmerken en hoog presteren 11

3 Kwantitatieve analyses VO 13

3.1 Definities 13

3.2 Laag presteren op afdelingsniveau 13

3.3 Laag presteren op bestuursniveau 15

3.4 Samenhang afdelings- en bestuurskenmerken en laag presteren 17

3.5 Samenhang afdelings- en bestuurskenmerken en hoog presteren 17

4 Kwantitatieve analyses MBO 19

4.1 Definities 19

4.2 Laag presteren op opleidingsniveau 19

4.3 Laag presteren op instellingsniveau 20

4.4 Samenhang opleidings- en instellingskenmerken en laag presteren 22

4.5 Samenhang opleidings- en instellingskenmerken en hoog presteren 23

5 Uitkomsten Focusgroep 24

5.1 Onderzoeksopzet 24

5.2 Resultaten 24

6 Referenties 28

1 Inleiding

Doel van dit onderzoek:

Dit onderzoek beantwoordt drie vragen voor het funderend onderwijs en mbo:

1. *Hoeveel scholen/afdelingen/instelling zijn er die langere tijd of vaker laag presteren?*

2. *Hoeveel besturen/instellingen zijn er met scholen/afdelingen/opleidingen die relatief vaak laag presteren?*

3. *Welke kenmerken in bestuurlijk handelen, op schoolniveau of in de context gaan samen met laag presteren?*

Dit onderzoek is tweeledig:

Ten eerste zijn er kwantitatieve analyses uitgevoerd:
Met behulp van beschrijvende analyses op basis van registerdata van DUO zijn bovenstaande vragen beantwoord. Voor het basisonderwijs kunnen we spreken van *langdurig* laag presteren omdat we een dataset gebruiken met daarin onderwijsresultaten voor 10 achtereenvolgende schooljaren. Voor het vo en het mbo hebben we de beschikking over een tijdsreeks van respectievelijk 5 en 4 jaar en is er in mindere mate sprake van de term langdurig.

De definities van laag presteren zijn toegelicht in de betreffende hoofdstukken. (hoofdstuk 2 voor bao, hoofdstuk 3 voor vo en hoofdstuk 4 voor mbo).

Focusgroep
De kwantitatieve analyses hebben geleid tot een lijst van besturen met scholen/opleidingen onder hun hoede die relatief vaak laag presteren, en van een lijst van besturen met scholen/opleidingen die relatief vaak hoog presteren. Bij het opstellen van deze lijsten is rekening gehouden met de leerlingenpopulatie op een school of opleiding. Op basis van deze kwantitatieve analyses zijn inspecteurs uitgenodigd voor een focusgroep. Een focusgroep is een verdiepend interview/gesprek met meerdere respondenten aan wie voornamelijk open vragen worden gesteld door een gespreksleider. Uitgenodigde inspecteurs zijn óf contact-inspecteurs van besturen op bovengenoemde lijsten, óf inspecteurs met een brede sectorkennis. Voorafgaand aan het verdiepend gesprek is er getoetst of de inspecteurs de door ons aangewezen besturen ook als laag- of hoogpresterend erkenden. Vervolgens is er met de uitgenodigde inspecteurs gesproken over wat deze besturen kenmerkt en welke handelingsruimte een bestuur heeft om resultaten van scholen/opleidingen te sturen. De uitkomsten van de focusgroep en de kwantitatieve analyses tezamen hebben geleid tot een antwoord op vraag 3.

2 Kwantitatieve analyses basisonderwijs

Voor de totstandkoming van dit hoofdstuk zijn verschillende databronnen gebruikt:

- ✓ Schoolwegingsbestanden; peildata 1-10-2015 tot en met 1-10-2018; bron: CBS, RA-omgeving en CBS-website.
- ✓ Bestanden van de inspectie met daarin toetsresultaten per school voor toetsjaren 2010 tot en met 2019. Deze zijn gebaseerd op de 1cijferbestanden van DUO.
- ✓ Bestanden van de inspectie met daarin achtergrondinformatie per bestuur.
- ✓ Open-data bestanden van het CBS met daarin kenmerken van de wijk waarin de school zich bevindt.
- ✓ IPTO-data met verzuim en personele gegevens.

Selecties in de data
Voor onderstaande analyses maken we gebruik van een dataset met daarin de gemiddelde eindtoetsscores van basisscholen voor toetsjaren 2010 tot en met 2019. Bij de berekening van gemiddeldes zijn leerlingen met een ontheffingsgrond en leerlingen die in leerjaar 7 of 8 zijn ingestroomd niet meegenomen.
Het bestaan van verschillende eindtoetsen maakt de vergelijking van prestaties ingewikkelder. In de analyses voor het basisonderwijs zijn alleen toetsresultaten op de drie grootste eindtoetsen (CET, IEP en Route8) meegenomen.
Scholen zijn meegenomen in de analyses wanneer er voor minimaal 5 jaren eindtoetsgegevens beschikbaar zijn op basis van minstens 10 leerlingen in leerjaar 8. Gemiddeld genomen is er per school voor 8,8 jaren aan toetsdata beschikbaar. Het totaal aantal geanalyseerde scholen is gelijk aan 5.265. Het aantal besturen is gelijk aan 756.

Oud versus nieuw toezicht
Tot augustus 2020 gold de regeling leerresultaten PO, waarbij de ondergrens voor een school afhankelijk was van het percentage gewichtenleerlingen op een school. Hoe lager het percentage gewichtenleerlingen, hoe hoger de grens. De analyses in dit hoofdstuk hebben betrekking op de oude regeling en zet de eindresultaten van een school af tegen de ondergrenzen die tot 1 augustus 2020 van kracht waren.

2.1 Definities

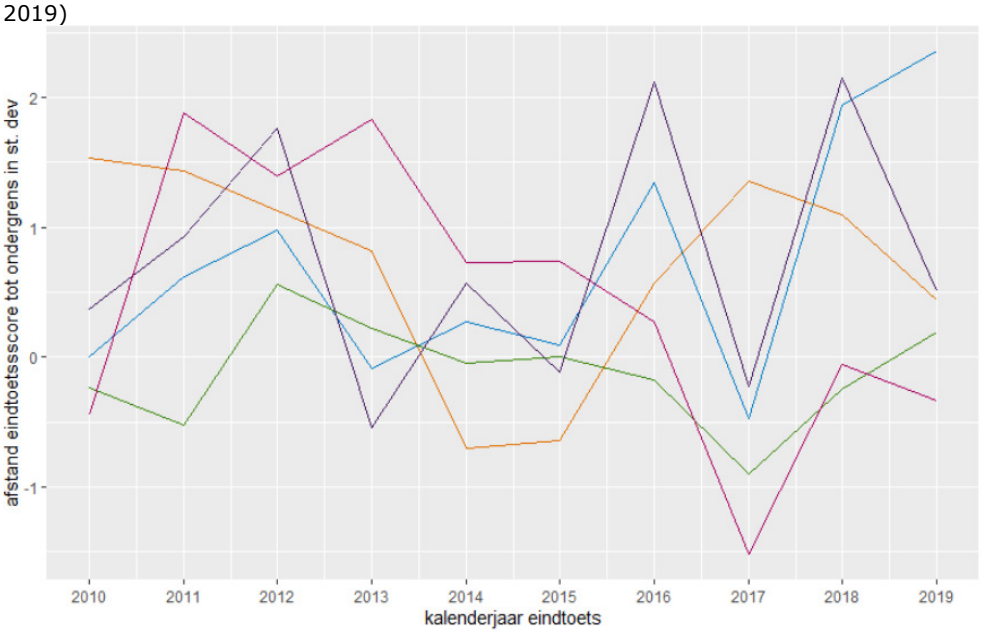
Gemiddelde eindtoetsscore, afstand tot de ondergrens
In dit hoofdstuk berekenen we per school de afstand van de gemiddelde eindtoetsscore in leerjaar 8 tot de ondergrens van de inspectie zoals beschreven in de regeling leerresultaten PO die tot 1 augustus 2020 van kracht was. Leerlingen met een ontheffingsgrond en leerlingen die in leerjaar 7 of 8 zijn ingestroomd zijn uitgesloten van de berekeningen. De afstand tot de ondergrens is gedeeld door de standaarddeviatie, welke is berekend per combinatie toetsjaar – gemaakte toets. We doen dit voor alle toetsjaren van 2010 tot en met 2019 afzonderlijk. Deze berekende maat geeft de afstand van de gemiddelde toetsscore tot de ondergrens weer in aantallen standaarddeviaties. Door dit per toets en per jaar te doen, zijn de verschillende toetsen onderling en over tijd met elkaar te vergelijken.

Registerdata van DUO
Voor de berekening van de gemiddelde eindtoetsscores in leerjaar 8 is gebruik gemaakt van registerdata van DUO. In het oude onderwijsresultatenmodel dat tot augustus 2020 van kracht was, stelde een inspecteur de eindtoetsscore van een school vaker bij. Dit gebeurde bijvoorbeeld als een leerling met een ontheffingsgrond niet als zodanig in de registerdata bekend was. Het uitsluiten van de leerling met een ontheffingsgrond leidde vaak tot een hogere gemiddelde eindtoetsscore. Voor de analyses van dit hoofdstuk hebben wij geen toegang tot de bijgestelde data.

2.2 Fluctuaties in opbrengsten

Als we gemiddelde toetsgegevens van individuele scholen in het Nederlandse basisonderwijs bekijken zien we dat deze fluctueren. Dit is niets nieuws. Het CPB toonde dit eerder ook aan in het rapport over schoolverschillen (CPB, februari 2019).
Het plaatje hieronder maakt de fluctuaties in opbrengsten per individuele school duidelijk. Het toont voor 5 willekeurig geselecteerde basisscholen de afstand van de gemiddelde eindtoetsscore in leerjaar 8 tot de ondergrens voor toetsjaren 2010 tot en met 2019. Om de gemiddelde eindtoetsresultaten over tijd en per toets met elkaar te kunnen vergelijken is de afstand tot de ondergrens uitgedrukt in het aantal standaarddeviaties tot deze ondergrens (zie ook paragraaf 2.1). We zien dat scholen over het algemeen niet continu hoog of continu laag scoren.

Figuur 1. Fluctuaties in eindtoetsscores voor 5 willekeurig gekozen basisscholen (2010-2019)



Bron: Inspectie van het Onderwijs op basis van registerdata DUO

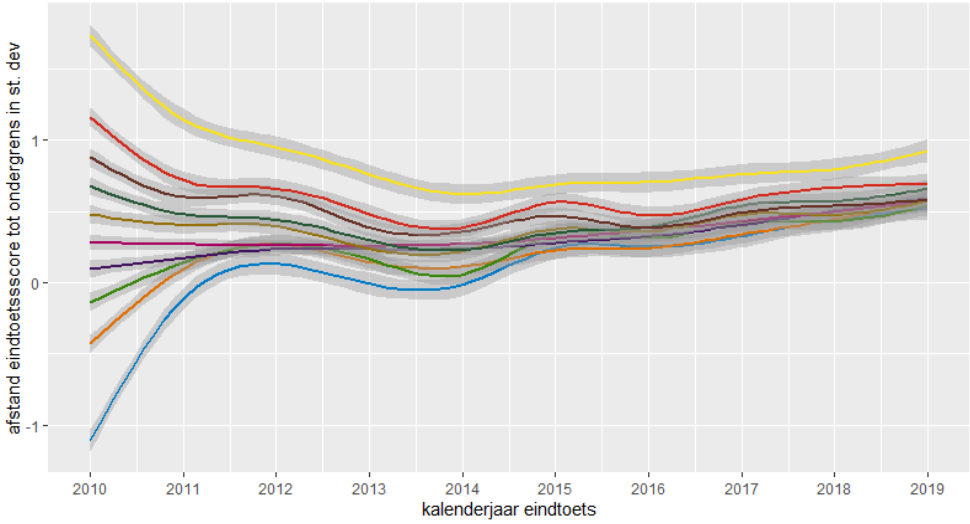
In figuur 2a wijzen we basisscholen toe aan 10 decielgroepen op basis van de afstand van de gemiddelde eindtoetsscore tot de ondergrens in toetsjaar 2010. De figuur toont 10 lijnen, iedere lijn vertegenwoordigt zo'n decielgroep. Scholen die in 2010 tot de top 10 procent behoorden, behaalden gemiddeld genomen een eindtoetsscore gelijk aan 2 standaarddeviaties boven de ondergrens (de bovenste lijn in figuur 2a). De onderste 10 procent scholen behaalde gemiddeld genomen een toetsscore gelijk aan 1 standaarddeviatie onder de ondergrens (de

onderste lijn in figuur 2a). Naarmate we verder naar 2019 gaan worden de verschillen tussen de groepen kleiner, al blijft de score van de hoogpresteerders in 2010 gemiddeld genomen altijd hoger dan die van de laagpresteerders in 2010.

Figuur 2b toont een soortgelijk beeld, maar wijst basisscholen toe aan 10 decielgroepen op basis van prestaties in 2019. Gemiddelde verschillen tussen uiterste groepen waren in de jaren voor 2019 een stuk kleiner, maar verschillen tussen de groepen blijven bestaan.

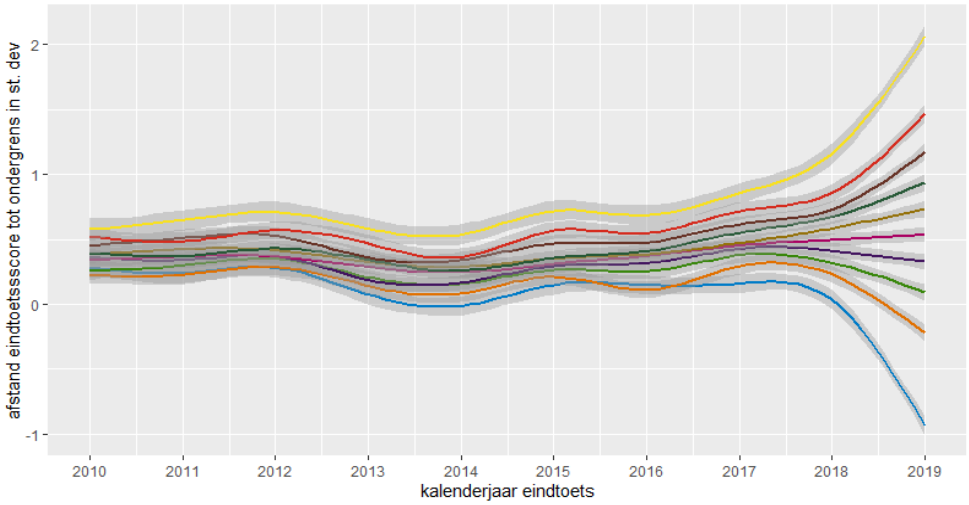
Een gedeeltelijke verklaring voor deze observaties is een statistisch fenomeen, te weten "regressie naar het gemiddelde", waarbij er voor een waarde van een (toevals)variabele die sterk van het gemiddelde afwijkt er een grotere kans bestaat dat de volgende waarneming zich dichter bij het gemiddelde zal bevinden.

Figuur 2a. Gemiddelde afstand eindtoetsscore tot ondergrens per jaar: vergelijking decielgroepen op basis van prestaties in 2010



Bron: Inspectie van het Onderwijs op basis van registerdata DUO

Figuur 2b. Gemiddelde afstand eindtoetsscore tot ondergrens per jaar: vergelijking decielgroepen op basis van prestaties in 2019



Bron: Inspectie van het Onderwijs op basis van registerdata DUO

Er bestaan scholen die langdurig laag presteren. De hierna volgende paragrafen gaan hier dieper op in.

2.3 Laag presteren op schoolniveau

Deze paragraaf beantwoordt de vraag: *Hoeveel scholen zijn er die langere tijd laag presteren?*

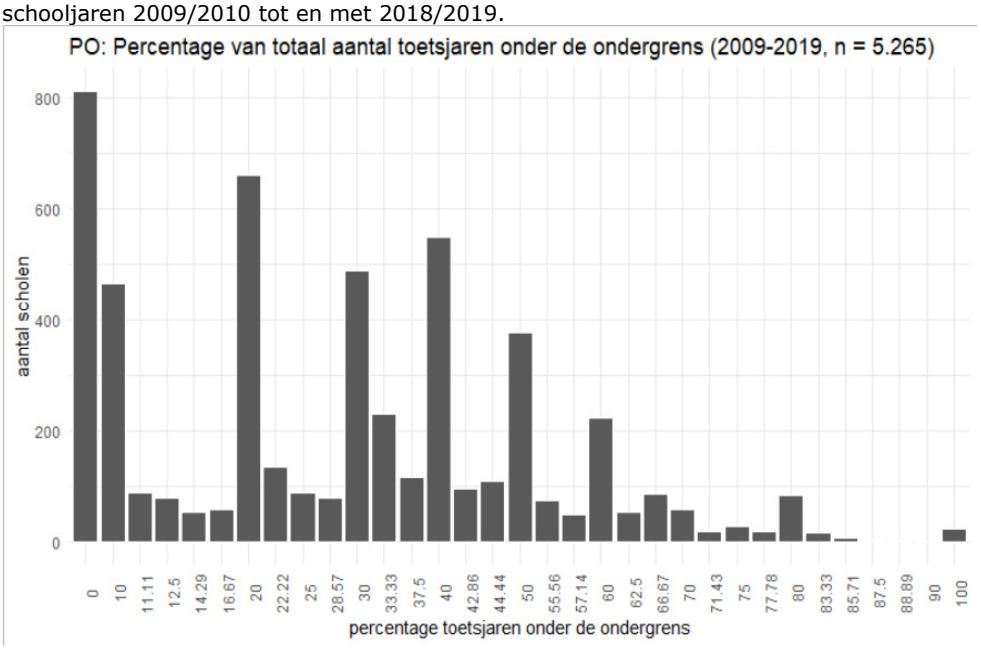
Definitie van laag presteren: onder de ondergrens van de inspectie

Voor het basisonderwijs hebben we voor 5.265 basisscholen en bijbehorende besturen informatie over de toetsresultaten voor 10 achtereenvolgende toetsjaren (2010 tot en met 2019, schooljaren 2009/2010 tot en met 2018/2019). Per school is per jaar bepaald of de gemiddelde toetsuitslag van de school onder de ondergrens van de inspectie ligt. Vervolgens is berekend hoe vaak de school in de afgelopen 10 jaar onder de ondergrens scoorde. Ieder jaar scoort gemiddeld 30 procent van de scholen onder de ondergrens.

Figuur 3 toont de verdeling van het percentage van toetsjaren waarin een school onder de ondergrens presteert. De x-as telt het aantal scholen. De meest rechtse balken vertegenwoordigen de groep scholen die structureel onder de ondergrens presteert. Hoe verder we op de x-as naar links opschuiven hoe minder vaak scholen onder de ondergrens presteren. De meest linkse balk vertegenwoordigt de scholen die structureel boven de ondergrens presteren. Een groot aantal scholen scoort in geen enkel jaar onder de ondergrens (~ 800).

Belangrijk: Voor de berekening van de gemiddelde eindtoetsscores in leerjaar 8 is gebruik gemaakt van registerdata van DUO. In het oude onderwijsresultatenmodel dat tot augustus 2020 van kracht was, stelde een inspecteur de eindtoetsscore van een school vaker bij. Dit gebeurde bijvoorbeeld als een leerling met een ontheffingsgrond niet als zodanig in de registerdata van DUO bekend is. Het uitsluiten van zo'n leerling leidde vaak tot een hogere gemiddelde eindtoetsscore. Het is daarom van belang te noemen dat scholen na het inspecteursoordeel mogelijk vaker boven de ondergrens presteerden dan in figuur 3 te zien is.

Figuur 3. Percentage van toetsjaren waarin een school onder de ondergrens presteert, schooljaren 2009/2010 tot en met 2018/2019.



Bron: Inspectie van het Onderwijs op basis van registerdata DUO

Het aandeel scholen dat in minstens **80% van alle schooljaren** waarvoor er data voor de school beschikbaar is onder de ondergrens presteert is klein. Een groot deel van deze scholen is onder geïntensiveerd toezicht gesteld.

Een voorzichtige schatting leert ons dat er 35.000 leerlingen zitten op scholen die structureel onderpresteren (zie tabel 1). Opvallend in deze tabel is dat er voor deze structureel laag presterende scholen minder toetsjaren aan data beschikbaar is, te weten 6,6 versus 8,8 voor de scholen die minder vaak onder de ondergrens presteren. 8 op de 10 jaar kan dus ook betekenen 4 op 5 jaar, of 7 op 8.

Tabel 1. Schatting van aantal leerlingen op structureel laag presterende scholen

Minstens 80% onder grens	Aantal scholen	Gemiddeld aantal leerlingen in 2019	Gemiddeld aantal jaren eindtoetsdata beschikbaar	Schatting aantal leerlingen
Nee	5.122	244,8	8,8	
Ja	143	240,9	6,6	34.453

Minder jaren aan data beschikbaar kan betekenen dat de school nog geen tien jaar bestaat, dat de leerlingaantallen in groep 8 in bepaalde schooljaren kleiner zijn dan 10 of dat er géén of een andere eindtoets dan CET, IEP of Route8 is afgenomen.

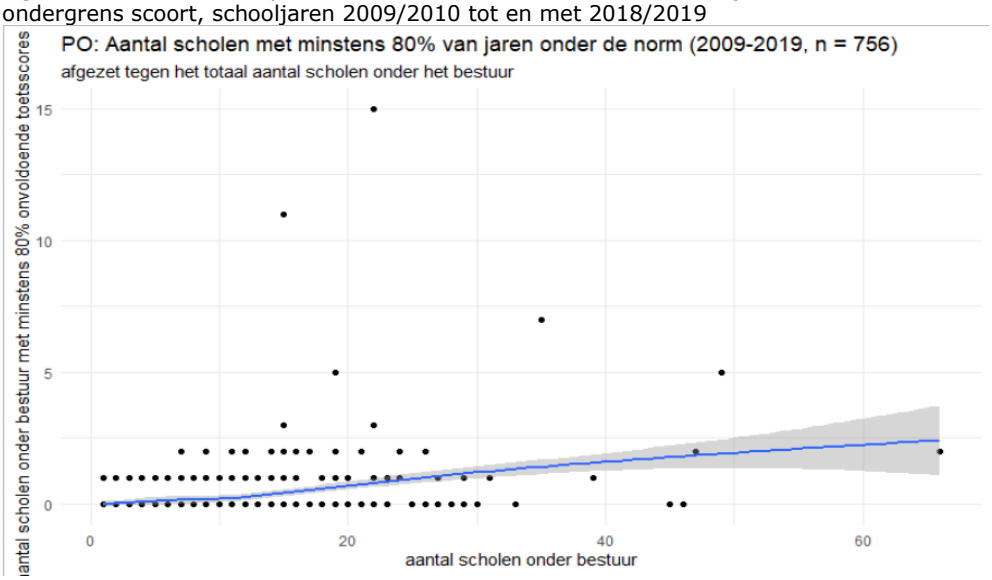
2.4 Laag presteren op bestuursniveau

Deze paragraaf beantwoordt de vraag: *Hoeveel besturen zijn er met scholen die relatief vaak laag presteren?*

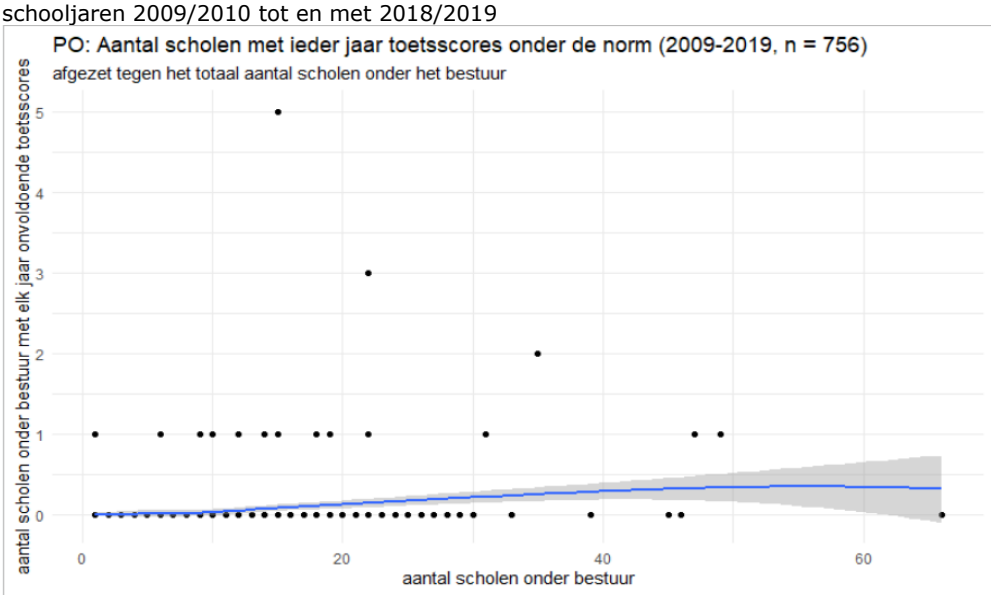
Voor 84 van de 756 besturen geldt dat zij minimaal 1 school onder hun hoede hebben die in minstens **80% van de schooljaren** waarvoor er data voor de

school beschikbaar is onder de ondergrens scoort. Voor 17 besturen geldt dat zij minstens 1 school onder hun hoede hebben die **alle** schooljaren waarvoor er data voor de school beschikbaar is onder de ondergrens scoort.
Grote besturen hebben (logischerwijs) vaker minstens 1 of meer scholen onder hun hoede met structureel onvoldoende opbrengsten (figuur 4, en figuur 5). Iedere stip is een bestuur.

Figuur 4. Aantal scholen per bestuur dat in minimaal 80% van toetsjaren onder de ondergrens scoort, schooljaren 2009/2010 tot en met 2018/2019



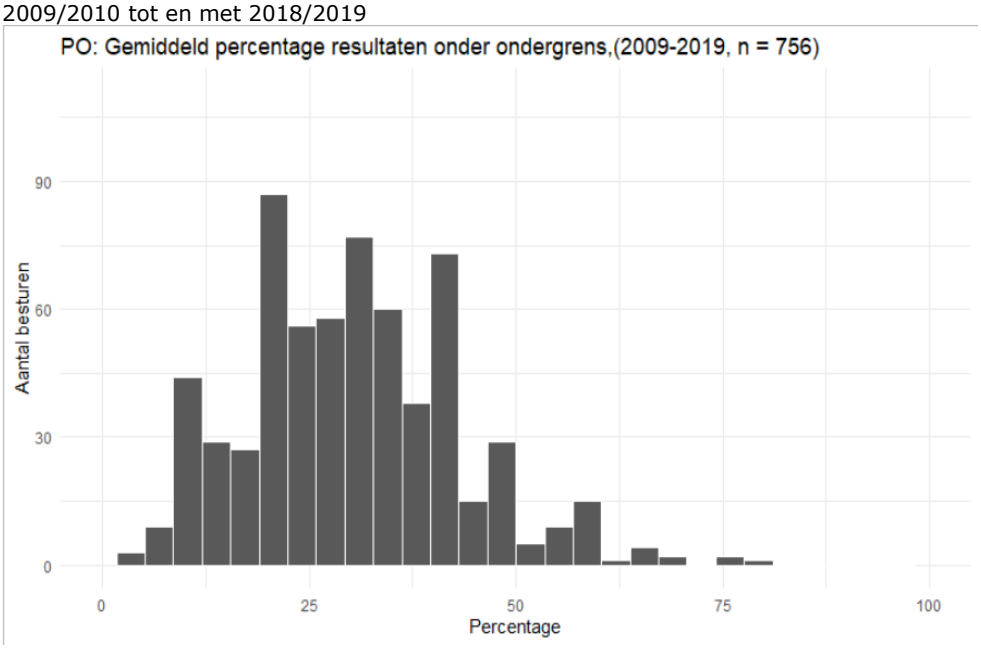
Figuur 5. Aantal scholen per bestuur dat in alle toetsjaren onder de ondergrens scoort, schooljaren 2009/2010 tot en met 2018/2019



Om verschillen tussen besturen op een andere manier te duiden, berekenen we per bestuur het gemiddeld aantal jaren dat scholen onder het betreffende bestuur onder de ondergrens vallen. We drukken dit uit in percentages. Deze berekening levert figuur 6 op. Rechts in de verdeling zijn de besturen met scholen die relatief vaak *onder* de ondergrens presteren, tot in

gemiddeld 75 procent van alle schooljaren. Besturen links in de verdeling hebben scholen onder hun hoede die relatief vaak *boven* de ondergrens presteren.

Figuur 6. Gemiddelde percentage jaren onder de ondergrens per bestuur, schooljaren 2009/2010 tot en met 2018/2019



2.5 Samenhang school- en bestuurskenmerken en laag presteren

Deze paragraaf gaat dieper in op de vraag: *Welke kenmerken in bestuurlijk handelen, op schoolniveau of in de context gaan samen met laag presteren?*

Vervolgens hebben we gekeken naar de samenhang tussen verschillende kenmerken van het bestuur en de school en het percentage jaren dat een school onder de ondergrens presteert. We doen dit middels een regressieanalyse waarbij we rekening houden met de hiërachische structuur van de data (multilevel model). Hieronder bespreken we kort de uitkomsten per onderwerp. We spreken van een samenhang als we een statistisch significante relatie vinden tussen een kenmerk en het percentage jaren dat een school onder de ondergrens presteert.

Belangrijk: In deze analyses kunnen we geen oorzakelijke of causale verbanden vaststellen, alleen correlaties.

Continue verklarende variabelen zijn in gestandardiseerde vorm aan het multilevel model toegevoegd. Hierbij is de volgende formule is toegepast: ((var - gemiddelde (var)) / sd(var).
Regressiecoefficiënten kunnen worden geinterpreteerd als het effect van 1 standaarddeviatie verschil in de verklarende variabele op de te verklaren variabele, in dit geval het percentage van schooljaren dat een school onder de ondergrens presteert.

Kenmerken op schoolniveau
Een hogere schoolweging gaat samen met vaker langdurig onder de ondergrens scoren (B = 6.4, se = 0.3, p < 0.01, sd = 3.7). Ook zien we dat een hoger percentage leerlingen met een ontheffingsgrond samengaat met vaker onder de ondergrens presteren (B = 1.3, se = 0.3, p < 0.01, sd = 0.04), ook al zijn deze leerlingen in deze analyse uitgesloten bij de berekening van de eindtoetsresultaten.

Scholen buiten de G4 scoren vaker onder de ondergrens (B = 3.4, se = 1.5, p < 0.01). Een hogere omgevingsadressendichtheid (het aantal adressen binnen een cirkel van één kilometer rond dat adres) rondom de school gaat samen met vaker boven de ondergrens presteren (B = -1.9, se = 0.4, p < 0.01, sd = 1504). Hetzelfde geldt voor het percentage van schooljaren dat de CET-toets is afgenomen (B = -2.0, se = 0.3, p < 0.01, sd = 0.2) en een hoger aantal leerlingen in leerjaar 8 (B = -2.4, se = 0.3, p < 0.01, sd = 14.8) Dit betekent overigens niet perse dat grote klassen het beter doen. In onze data is niet bekend in welke klas een leerling zit. Tot slot maakt de denominatie van een school uit. Overig bijzondere scholen doen het relatief beter. Zo zien we bijvoorbeeld dat islamitische scholen, ten opzichte van niet-islamitische scholen met eenzelfde schoolweging, relatief vaak boven de ondergrens scoren (B = -2.1, se = 0.3, p < 0.01)

Verzuim
Een hoger verzuimpercentage op scholen gaat samen met vaker onder de ondergrens presteren (B = 1.6 , se = 0.5, p < 0.01, sd = 2.7).

Achtergrondkenmerken bestuur
Eénpitters doen het beter (ook als we rekening houden met het gegeven dat uitschieters vaker (toevallig) voorkomen bij besturen met minder scholen onder zich (B = -5.5, se = 1.4, p < 0.01). Hetzelfde geldt voor besturen in het zuiden van het land (referentiecategorie is noord, zuiden: B = -6.2, se = 1.2, p < 0.01, midden: B = -1.9, se = 1.0, p < 0.1)

Oordelen op bestuursniveau
We vinden een samenhang tussen een onvoldoende op de standaard Kwaliteitszorg (KA1) en vaker onder de ondergrens presteren (B = 3.3, se = 1.5, p < 0.05). Voor de andere standaarden uit het kwaliteitsgebied Kwaliteitszorg en ambitie (KA2, KA3,) en de standaarden rond financieel beheer (FB1, FB2 en FB3) vinden we geen samenhang, mede ook omdat er op FB1, FB2 en FB3 nauwelijks een onvoldoende oordeel wordt gegeven en er dus weinig variatie in de data is.

2.6 Samenhang school- en bestuurskenmerken en hoog presteren

Deze paragraaf gaat dieper in op de vraag: *Welke kenmerken in bestuurlijk handelen, op schoolniveau of in de context gaan samen met hoog presteren?*

Ook is er gekeken naar scholen die het erg **goed** doen op onderwijsresultaten. In deze analyse is berekend **hoe vaak** de school in de afgelopen 10 jaar tot de top 10 procent hoogst presterende scholen behoorde. Middels een regressieanalyse, waarbij rekening is gehouden met de hiërachische structuur van de data, is de samenhang tussen langdurig hoog presteren van een school en diverse kenmerken van de school en het bestuur bekeken. Hieronder bespreken we kort de uitkomsten per onderwerp.

Kenmerken op schoolniveau

Een hogere schoolweging gaat samen met minder vaak hoog presteren (B = -3.3, se = 0.3, p < 0.01, sd = 3.7) Hetzelfde geldt voor een hoger aantal leerlingen in leerjaar 8 (B = -2.4, se = 0.2, p < 0.01, sd = 14.8). Een hogere omgevingsadressendichtheid (het aantal adressen binnen een cirkel van één kilometer rond dat adres) rondom de school gaat samen met vaker hoog presteren (B = 2.8, se = 0.3, p < 0.01, sd = 1504). Hetzelfde geldt voor het percentage van schooljaren dat de CET-toets is afgenomen (B = 1.4, se = 0.2, p < 0.01, sd = 0.2).

Personele kenmerken
Een hoger percentage onderwijsgevend personeel (t.o.v. directieleden) hangt samen met minder vaak langdurig hoog presteren (B = -1.6, se = 0.3, p < 0.01, sd = 0.04). Ook zien we dat een hoger percentage onderwijsgevend personeel op scholen onder het bestuur dat ouder is dan 45 vaker samengaat met langdurig hoog presteren (B = 5.3, se = 1.3, p < 0.01).

Achtergrondkenmerken bestuur
Eenpitters (besturen met 1 school) doen het gemiddeld genomen goed. We zien dit terug in de regressieanalyse (B = 8.4, se = 1.2, p < 0.01). Maar ook als we besturen ranken op basis van langdurig hoog presteren van scholen onder het bestuur (van vaak naar minder vaak) zien we dat de eenpitters domineren (met of zonder rekening te houden met kenmerken van de leerlingen). Het gaat hier dan veelal om schoolverenigingen in gebieden als het Gooi en de grote steden.

Oordelen op bestuursniveau
We vinden geen samenhang tussen de oordelen op standaarden KA1, KA2, KA3 en FB1, FB2 en FB3 en langdurig hoog presteren.

3 Kwantitatieve analyses VO

Voor de totstandkoming van dit hoofdstuk zijn verschillende databronnen gebruikt:

- ✓ Bestanden van de inspectie met daarin onderwijsresultaten per afdelingen voor de schooljaren 2014-2015 tot en met 2018-2019. Deze zijn gebaseerd op de 1cijferbestanden van DUO.
- ✓ Bestanden van de inspectie met daarin achtergrondinformatie per bestuur.
- ✓ Open-data bestanden van het CBS met daarin kenmerken van de wijk waarin de afdeling zich bevindt.
- ✓ IPTO-data met verzuim en personele gegevens.

Selecties in de data
Voor onderstaande analyses maken we gebruik van een dataset met daarin de driejaarsgemiddelden van de onderwijsresultaten voor schooljaren 2014/2015 tot en met 2018/2019. Afdelingen zijn meegenomen in de analyses wanneer er voor minimaal 3 jaren onderwijsresultaten beschikbaar zijn. Daarnaast zijn alleen afdelingen meegenomen waarvoor een berekend eindoordeel beschikbaar is. Gemiddeld genomen is er per afdeling voor 4,9 jaren aan data beschikbaar. Het totaal aantal geanalyseerde afdelingen is gelijk aan 2.691. Het aantal besturen is gelijk aan 261.

3.1 Definities

Onvoldoende berekend oordeel
De inspectie beoordeelt de onderwijsresultaten in het voortgezet onderwijs aan de hand van 4 indicatoren:

1. positie in leerjaar 3 ten opzichte van het advies van de afdeling,
2. percentage onvertraagde studievoortgang in leerjaar 1 en 2,
3. percentage onvertraagde studievoortgang vanaf leerjaar 3 per afdeling
4. gemiddeld cijfer Centraal Examen van alle vakken per afdeling.

Voor iedere indicator hanteert de inspectie een ondergrens welke afhankelijk is van de leerlingenpopulatie op een afdeling. Voor een afdeling met een meer uitdagende leerlingenpopulatie ligt de ondergrens lager. Bij twee of meer indicatoren met de kwalificatie 'onder de norm' is het berekende oordeel een 'onvoldoende'.

Belangrijk: Voor de berekening van onderwijsresultaten maken we gebruik van registerdata van DUO. Het berekend oordeel is nog niet hetzelfde als een totaaloordeel over de onderwijsresultaten. Dit laatste wordt enkel gegeven door een inspecteur op basis van een onderzoek. Bij een dergelijk onderzoek betrekt de inspecteur bijzondere omstandigheden van de school, het zogenaamde verhaal achter de cijfers.

3.2 Laag presteren op afdelingsniveau

Deze paragraaf beantwoordt vraag: *Hoeveel afdelingen zijn er die relatief vaak laag presteren?*

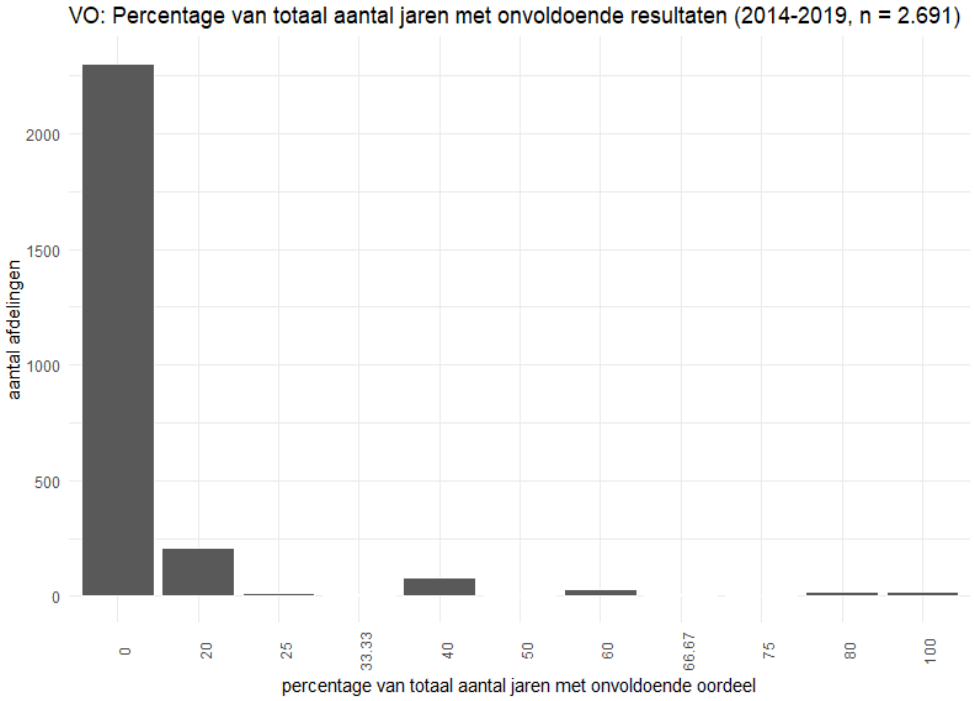
Definitie van laag presteren: een onvoldoende berekend oordeel op basis van de volgende indicatoren:

1. *positie in leerjaar 3 ten opzichte van het advies van de afdeling,*
2. *percentage onvertraagde studievoortgang in leerjaar 1 en 2,*
3. *percentage onvertraagde studievoortgang vanaf leerjaar 3 per afdeling*
4. *gemiddeld cijfer Centraal Examen van alle vakken per afdeling.*

Voor het voortgezet onderwijs hebben we voor 2.691 afdelingen en bijbehorende besturen informatie over de onderwijsresultaten voor 5 achtereenvolgende schooljaren (2014/2015 tot en met 2018/2019). Per afdeling is bepaald of het berekende oordeel een onvoldoende is (zie paragraaf 3.1 voor de definitie van berekend oordeel).

Figuur 7 toont de verdeling van het percentage jaren waarin een afdeling het berekend oordeel 'onvoldoende' heeft. De y-as telt het aantal afdelingen. De meest rechtse balk vertegenwoordigt de groep afdelingen die in alle jaren waarvoor er data voor de afdeling beschikbaar is een onvoldoende scoort. Afdelingen meer links op de x-as hebben minder vaak het berekend oordeel 'onvoldoende'. De meest linkse balk vertegenwoordigt de afdelingen die structureel boven de normen presteert.

Figuur 7. Percentage van schooljaren waarin een afdeling een onvoldoende berekend oordeel heeft, schooljaren 2014/2015 tot en met 2018/2019



Het plaatje voor vo ziet er anders uit dan voor het basisonderwijs of het mbo, maar dat wil niet zeggen dat vo-afdelingen het beter doen. Dit heeft te maken met de manier waarop laag presteren is gedefinieerd. **De vergelijking tussen verschillende sectoren is niet te maken.**

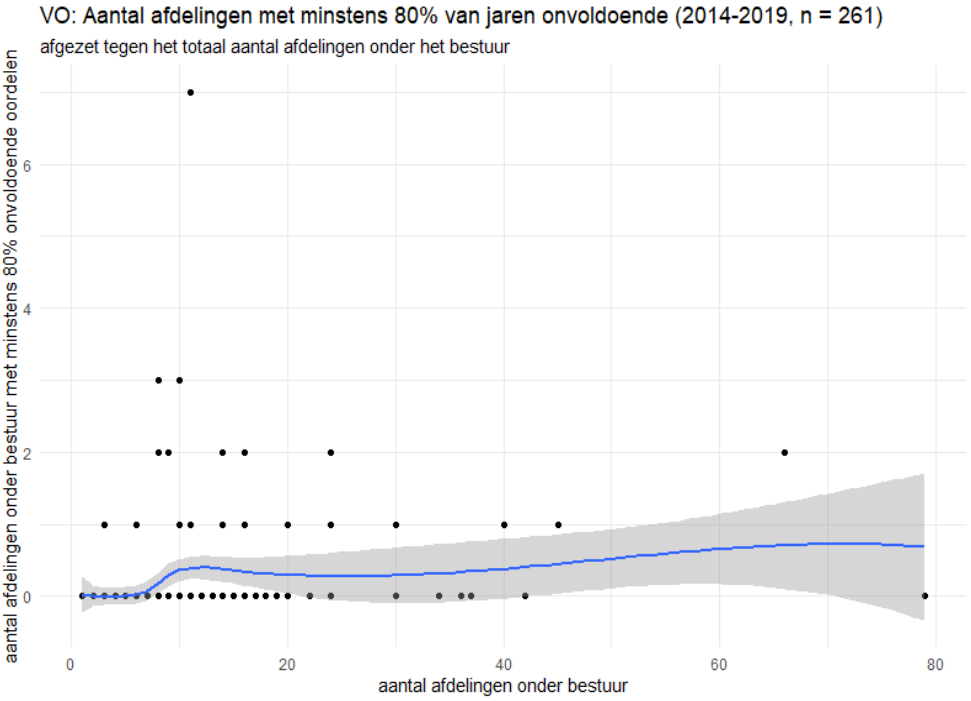
3.3 Laag presteren op bestuursniveau

Deze paragraaf beantwoordt de vraag: *Hoeveel besturen zijn er met afdelingen die relatief vaak laag presteren?*

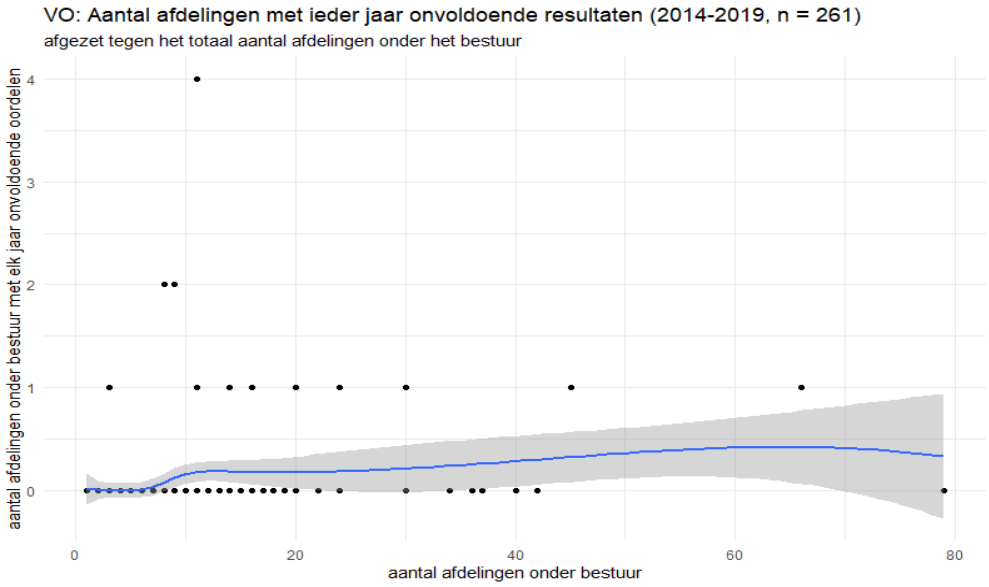
Voor 22 van de 261 besturen geldt dat zij minimaal 1 afdeling onder hun hoede hebben die in minstens **80 procent van de schooljaren** waarvoor er data voor de afdeling beschikbaar is een onvoldoende berekend oordeel scoort. 80 procent van schooljaren betekent in deze analyses 4 op 5 jaar, 4 op 4 jaar of 3 op 3 jaar. Gemiddeld genomen is er voor afdelingen 4,5 jaar aan data beschikbaar. Hierin zien we geen verschillen tussen hoog- en laagpresteerders. Voor 15 besturen geldt dat zij minstens 1 afdeling onder hun hoede hebben die elk jaar een onvoldoende berekend oordeel scoort.

Grotere besturen hebben (logischerwijs) vaker minstens 1 of meer afdelingen onder hun hoede met structureel onvoldoende resultaten (figuur 8, en figuur 9). Iedere stip is een bestuur.

Figuur 8. Aantal afdelingen per bestuur dat in minimaal 80% van schooljaren een onvoldoende berekend oordeel heeft, schooljaren (2014/2015 tot en met 2018/2019)

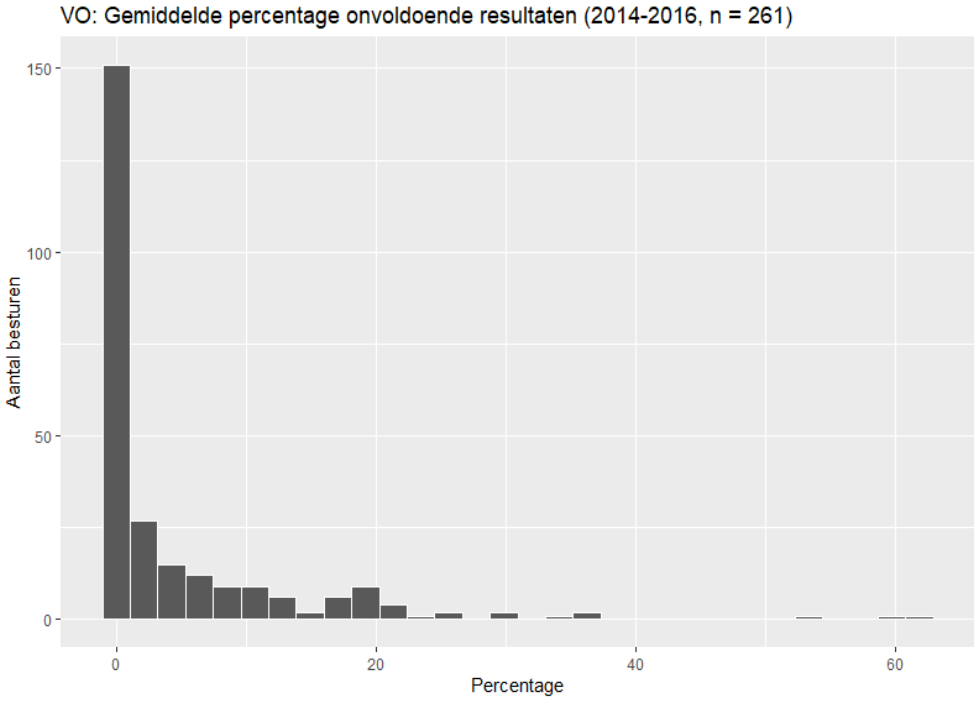


Figuur 9. Aantal afdelingen per bestuur dat in alle schooljaren een onvoldoende berekend oordeel heeft, schooljaren (2014/2015 tot en met 2018/2019)



Om verschillen tussen besturen op een andere manier in kaart te brengen berekenen we per bestuur het gemiddelde van het aantal keren dat afdelingen onder het bestuur een berekende onvoldoende scoren. We drukken dit uit in percentages. Deze berekening levert figuur 10 op. Rechts in de verdeling zijn de besturen met afdelingen die relatief vaak onvoldoende presteren. Besturen links in de verdeling hebben afdelingen onder hun hoede die relatief vaak tot altijd voldoende presteren.

Figuur 10. Gemiddelde percentage onvoldoende berekend oordeel per bestuur, 2014/2015 tot en met 2018/2019



3.4 Samenhang afdelings- en bestuurskenmerken en laag presteren

Deze paragraaf gaat dieper in op de vraag: *Welke kenmerken in bestuurlijk handelen, op afdelingsniveau of in de context gaan samen met laag presteren?*

> **Belangrijk:** In deze analyses kunnen we geen oorzakelijke of causale verbanden vaststellen, alleen correlaties.

Vervolgens hebben we gekeken naar de samenhang tussen verschillende kenmerken van het bestuur en de afdeling en het onder de norm presteren van afdelingen. In deze analyse is rekening gehouden met de hiërarchische structuur van de data (multilevel model). Hieronder bespreken we kort de uitkomsten per onderwerp.

> Continue verklarende variabelen zijn in gestandardiseerde vorm aan het multilevel model toegevoegd. Hierbij is de volgende formule is toegepast: ((var - gemiddelde (var)) / sd(var).
> Regressiecoëfficienten kunnen worden geinterpreteerd als het effect van 1 standaarddeviatie verschil in de verklarende variabele op de te verklaren variabele. De te verklaren variabele in deze analyse is het percentage indicatoren dat in vijf schooljaren onder de norm valt. Dit verschilt van de definitie die we hanteren in paragraaf 3.2 en 3.3 om gebruik te kunnen maken van alle variatie in de verschillende indicatoren die ten grondslag liggen aan het berekend oordeel.

Kenmerken op schoolniveau

Afdelingen met een hoger percentage leerlingen uit een apcg-gebied (B = 1.4, se = 0.4, $p < 0.01$, sd = 0.2) hebben vaker onvoldoende resultaten.

Personele kenmerken

Een hoger verzuim op afdelingen gaat samen met vaker resultaten onder de norm (B = 1.4, se = 0.5, $p < 0.01$, sd = 2.3). Een hoger percentage fte's mannen onder het onderwijsgevend personeel gaat samen met vaker boven de norm scoren (B = -1.4, se = 0.5, $p < 0.01$, sd = 0.1).

Achtergrondkenmerken bestuur

Multi-sectorale besturen hebben vaker afdelingen met langdurig onvoldoende resultaten (B = 1.9, se = 0.5, $p < 0.01$) en besturen in de regio Midden hebben minder vaak afdelingen die onder de norm scoren (B = -3.1, se = 1.1, $p < 0.01$) dan besturen in de regio Noord.

Oordelen op bestuursniveau

We hebben de samenhang bekeken tussen het oordeel op de standaard Kwaliteitszorg (KA1) en onderpresteren en vinden een samenhang tussen een onvoldoende op KA1 en vaker onder de norm presteren (B = 6.0, se = 1.4, $p < 0.01$). Een bestuur met een onvoldoende voor kwaliteitszorg heeft vaker meer afdelingen die vaker onder de norm scoren dan een bestuur dat een voldoende op deze standaard scoort.

3.5 Samenhang afdelings- en bestuurskenmerken en hoog presteren

Deze paragraaf gaat dieper in op de vraag: *Welke kenmerken in bestuurlijk handelen, op afdelingsniveau of in de context gaan samen met hoog presteren?*

We hebben ook gekeken naar de samenhang tussen verschillende kenmerken van het bestuur en de afdeling en het percentage jaren dat een afdeling tot de top 10 best presterende scholen van dat jaar behoort. Hieronder bespreken we kort de uitkomsten per onderwerp.

Kenmerken op schoolniveau
Afdelingen met een meer uitdagende leerlingenpopulatie, leerlingen uit een apcg-gebied, scoren vaker boven de P90-grens ($B = 1.5$, $se = 0.1$, $p < .01$, $sd = 0.2$). Net als afdelingen met meer jongens ($B = 0.3$, $se = 0.1$, $p < 0.05$, $sd = 0.1$). Wat hier precies de achterliggende reden voor is vergt verdiepend onderzoek.

Personele kenmerken
Een hoger percentage tijdelijk onderwijsgevend personeel gaat samen met vaker tot de top 10 best presterende afdelingen behoren ($B = 0.6$, $se = 0.2$, $p < 0.01$, $sd = 0.1$).

Achtergrondkenmerken bestuur
Eenpitters (besturen met 1 school) hebben vaker afdelingen die tot de top 10 procent hoogst scorende afdelingen behoren ($B = 0.3$, $se = 0.1$, $p < 0.01$). Ook als we corrigeren voor het aantal afdelingen dat onder verantwoordelijkheid van een bestuur valt.

Oordelen op bestuursniveau
We hebben de samenhang bekeken tussen het oordeel op de standaard KA1 en hoogpresteerders en vinden geen samenhang ($B = -0.2$, $se = 0.5$, $p > 0.05$).

4 Kwantitatieve analyses MBO

Voor de totstandkoming van dit hoofdstuk zijn verschillende databronnen gebruikt:

- ✓ Bestanden van de inspectie met daarin opbrengstgegevens per opleiding (brin*bc*niveau) voor schooljaren 2015/2016 tot en met 2018/2019. Deze bestanden zijn gebaseerd op de 1cijferbestanden van DUO.
- ✓ Bestanden van de inspectie met daarin achtergrondinformatie per bestuur.
- ✓ IPTO-data met personele gegevens.

Selecties in de data
Voor onderstaande analyses maken we gebruik van een dataset met daarin de opbrengstgegevens van opleidingen in het mbo. Niveau 1-opleidingen zijn uitgesloten. Opleidingen zijn meegenomen wanneer er voor minimaal 3 jaren een opbrengstenoordeel beschikbaar is op basis van minstens 30 studenten. Gemiddeld genomen zijn er per opleiding 3,9 jaren aan data beschikbaar. Het totaal aantal geanalyseerde opleidingen is gelijk aan 2.539. Het aantal instellingen in deze analyses is gelijk aan 59.

Door de Invoering Herziene Kwalificatiedossiers in augustus 2015 is het niet mogelijk alle opleidingen voor de herziening te koppeling aan opleiding na de herziening. Om deze reden kunnen we niet verder terug dan schooljaar 2015/2016.

4.1 Definities

Onvoldoende berekend oordeel
De opleiding heeft een onvoldoende berekend oordeel als slechts één of geen van de drie indicatoren: jaarresultaat, diplomaresultaat of startersresultaat aan de inspectienormen voor voldoende voldoet.

Belangrijk: We spreken hier van een **berekend** oordeel, omdat het louter gaat om de vergelijking met de norm. Van een daadwerkelijk oordeel van de onderwijsresultaten is pas sprake wanneer de inspecteur dat oordeel heeft uitgesproken.

4.2 Laag presteren op opleidingsniveau

Deze paragraaf beantwoordt de vraag: *Hoeveel opleidingen (brin x bc x niveau) zijn er die relatief vaak laag presteren?*

Definitie van laag presteren: onvoldoende berekend oordeel op opbrengsten op basis van 3 indicatoren:

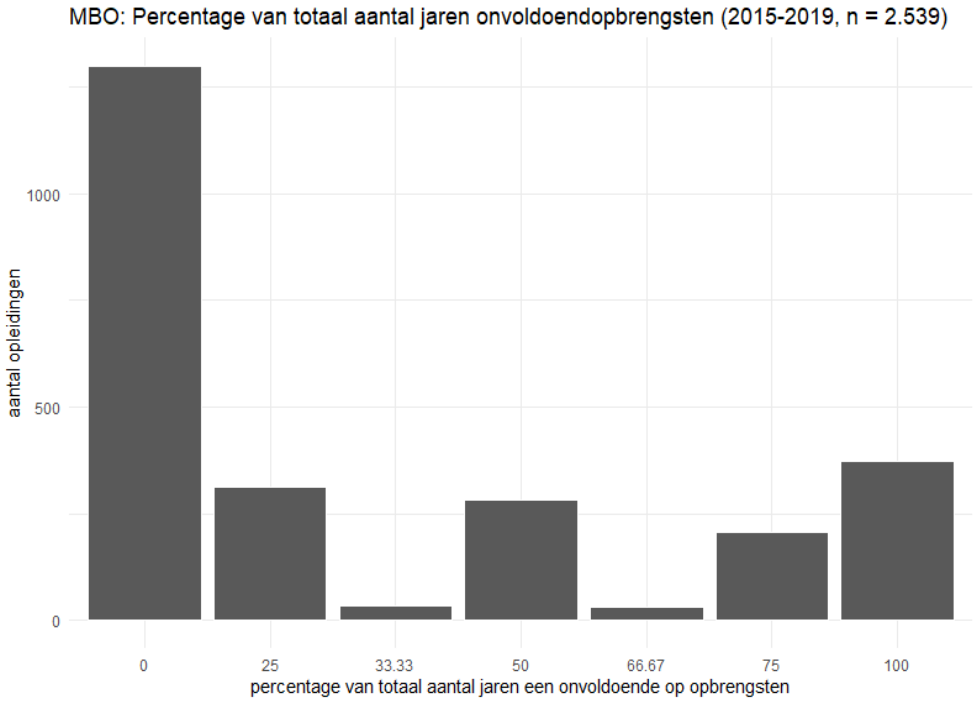
1. *jaarresultaat*
2. *diplomaresultaat*
3. *startersresultaat*

Voor het middelbaar beroepsonderwijs hebben we voor 2.539 opleidingen en bijbehorende instellingen/besturen het oordeel op opbrengsten voor 4

achtereenvolgende schooljaren (2015/2016 tot en met 2018/2019). Per opleiding is berekend hoe vaak de opleiding in 4 jaar een onvoldoende berekend oordeel heeft.

Figuur 11 toont de verdeling van het percentage schooljaren waarin de opleiding een onvoldoende berekend oordeel heeft. De y-as telt het aantal opleidingen. De meest rechtse balk vertegenwoordigt de groep opleidingen die structureel een onvoldoende scoort, namelijk in alle jaren waarin er data voor de opleiding beschikbaar is. De meest linkse balk vertegenwoordigt de groep opleidingen die structureel een voldoende op opbrengsten scoort.

Figuur 11. Percentage van schooljaren waarin een opleiding een onvoldoende berekend oordeel heeft, schooljaren 2015/2016 tot en met 2018/2019

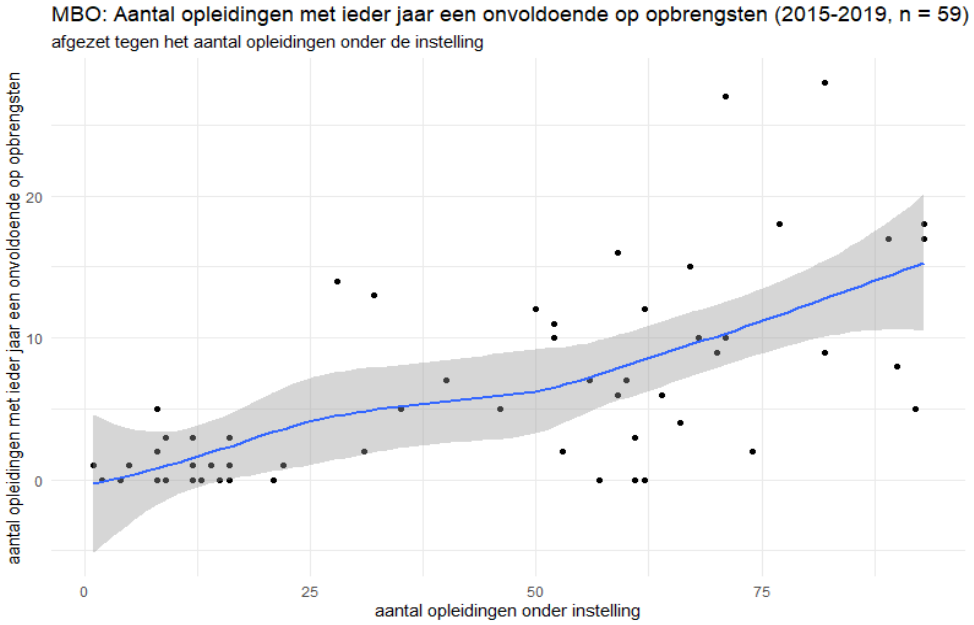


4.3 Laag presteren op instellingsniveau

Deze paragraaf beantwoordt de vraag: *Hoeveel instellingen/besturen zijn er met opleidingen onder hun hoede die relatief vaak laag presteren?*

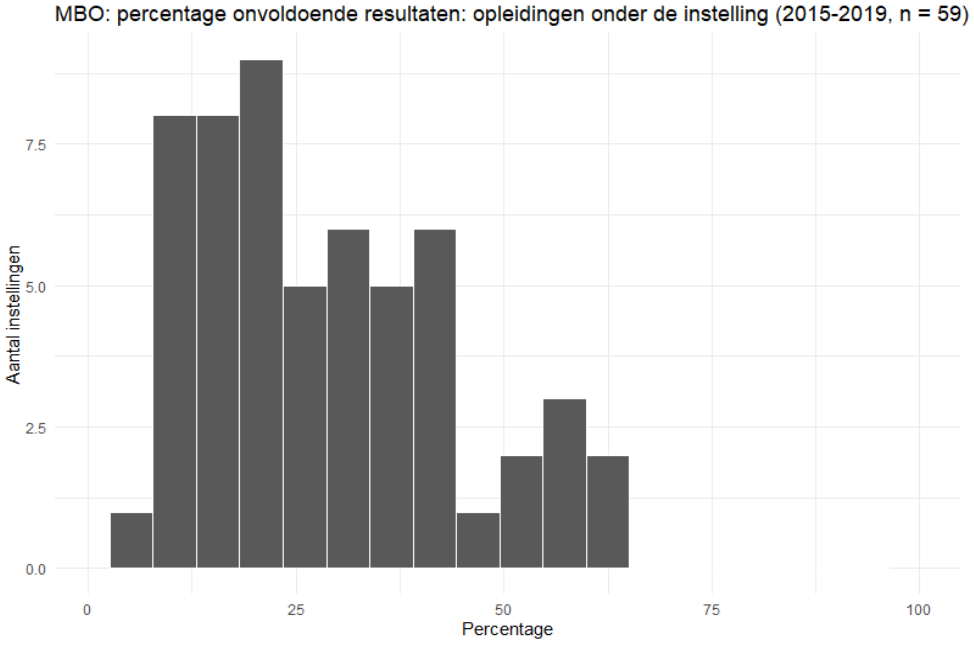
Voor 45 van de 59 instellingen geldt dat zij minimaal 1 opleiding onder hun hoede hebben die in **alle jaren waarvoor er data voor de opleiding beschikbaar** is een onvoldoende berekend oordeel scoren. We zien geen verschillen in het aantal jaren aan data beschikbaar tussen hoog- en laagpresteerders. Grote instellingen hebben (logischerwijs) vaker minstens 1 of meer opleidingen onder hun hoede met structureel onvoldoende opleidingen (figuur 12). Iedere stip is een instelling.

Figuur 12. Aantal opleidingen per instelling dat in alle schooljaren een onvoldoende berekend oordeel heeft, schooljaren 2015/2016 tot en met 2018/2019



Om verschillen tussen instellingen op een andere manier te duiden berekenen we per instelling het gemiddelde van het aantal keren dat opleidingen onder de instelling een onvoldoende scoren. We drukken dit uit in percentages. Deze berekening levert figuur 13 op. Rechts in de verdeling zijn de instellingen met opleidingen die relatief vaak onvoldoende presteren. Instellingen links in de verdeling hebben opleidingen onder hun hoede die relatief vaak voldoende presteren.

Figuur 13. Gemiddelde percentage jaren een onvoldoende op opbrengsten voor opleidingen onder de instelling, schooljaren 2015/2016 tot en met 2018/2019



4.4 Samenhang opleidings- en instellingskenmerken en laag presteren

Deze paragraaf gaat dieper in op de vraag: *Welke kenmerken in bestuurlijk handelen, op opleidingsniveau of in de context gaan samen met laag presteren?*

Vervolgens hebben we gekeken naar de samenhang tussen verschillende kenmerken van het bestuur en de opleiding en onvoldoende resultaten. In deze analyse is rekening gehouden met de hiërarchische structuur van de data (multilevel model). Hieronder bespreken we kort de uitkomsten per onderwerp.

Belangrijk: In deze analyses kunnen we geen oorzakelijke of causale verbanden vaststellen, alleen correlaties.

Continue verklarende variabelen zijn in gestandardiseerde vorm aan het multilevel model toegevoegd. Hierbij is de volgende formule is toegepast: ((var - gemiddelde (var)) / sd(var).
Regressiecoefficienten kunnen worden geinterpreteerd als het effect van 1 standaarddeviatie verschil in de verklarende variabele op de te verklaren variabele, in dit geval het percentage van schooljaren dat een opleiding een onvoldoende berekend oordeel heeft.

Kenmerken op opleidingsniveau

Opleidingen met een hoger percentage leerlingen uit apcg-gebieden scoren vaker een onvoldoende ($B = 8.5$, $se = 1.3$, $p < .001$, $sd = 0.2$). Hetzelfde geldt voor opleidingen in de G4. Ook zien we verschillen tussen niveaus, waarbij niveau 3 opleidingen vaker een onvoldoende scoren dan opleidingen niveau 2 en niveau 4 (referentiecategorie is niveau 3, niveau 2: $B = -7.8$, $se = 1.9$ $p < 0.01$, niveau 4: $B = -11.6$, $se = 1.8$, $p < 0.01$).
Ook bestaan er verschillen tussen de domeinen. Opleidingen in de domeinen Voedsel, natuur en leefomgeving ($B = -17.9$, $se = 5.4$, $p < 0.01$), Techniek en procesindustrie ($B = -15.9$, $se = 4.1$, $p < 0.01$) doen het relatief goed. In tegenstelling tot opleidingen in de domeinen Ambacht, laboratorium en gezondheidstechniek ($B = 17.0$, $se = 6.1$, $p < 0.01$), Media en vormgeving ($B = 9.5$, $se = 4.9$, $p < 0.01$), die vaker een onvoldoende scoren (referentiecategorie is Afbouw, hout en onderhoud). Bij de verschillen tussen opleidingen en niveaus is er in deze analyses rekening gehouden met verschillen in leerlingenpopulaties. Opleidingen die vaker onderpresteren krijgen een lager rapportcijfer ($B = -3.1$, $se = 0.8$, $p < 0.01$, $sd = 0.6$) en vaker een onvoldoende ($B = 3.2$, $se = 0.7$, $p < 0.01$) van studenten die aan de JOB-monitor hebben meegewerkt.

Verzuim

Voor het mbo hebben we geen beschikking over verzuim per opleiding. Verder is er geen samenhang gevonden met personele kenmerken.

Achtergrondkenmerken bestuur

Een hoger percentage opleidingen (brin x bc_id x niveau) met een onvoldoende rapportcijfer (tevredenheidcijfers vanuit de JOB, bron: bestuursmonitor) gaat samen met vaker onder de norm scoren ($r = 3.7$, $se = 1.3$, $p < 0.001$). We zien geen samenhang met het aantal studenten per instelling.

Oordelen op bestuursniveau

We zien een samenhang met een voldoende op de standaard Verantwoording en dialoog (KA3) en vaker een voldoende op resultaten (referentiecategorie is voldoende oordeel op KA3: B = 14.1, se = 6.5, $p < 0.05$).

4.5 Samenhang opleidings- en instellingskenmerken en hoog presteren

Deze paragraaf gaat dieper in op de vraag: *Welke kenmerken in bestuurlijk handelen, op opleidingsniveau of in de context gaan samen met laag presteren?*

Vervolgens hebben we gekeken naar de samenhang tussen verschillende kenmerken van het bestuur en de opleiding en (vaker) hoog presteren. Hieronder bespreken we kort de uitkomsten per onderwerp.

Kenmerken op opleidingsniveau
Opleidingen met een hoger percentage leerlingen uit apcg-gebieden behoren minder vaak tot de top 10 best presterende opleidingen (B = -2.8, se = 0.8, $p < 0.01$, sd = 0.2). Niveau 3 lijkt het ook hier minder goed te doen. Al zijn de verschillen tussen niveaus hier minder overtuigend dan wanneer we kijken naar onder de norm presteren. Opleidingen in de domeinen Voedsel, natuur en leefomgeving (B = 10.3, se = 3.5, $p < 0.01$), Techniek en procesindustrie (B = 5.9, se = 2.3, $p < 0.01$), Uiterlijke verzorging (B = 7.3, se = 2.8 , $p < 0.01$), en Bouw en Infra (B = 10.6, se = 3.0, $p < 0.01$) presteren relatief vaak hoger boven de normen. We zien geen verschillen tussen grote steden.

Verzuim
Voor het mbo hebben we geen beschikking over verzuim per opleiding.

Achtergrondkenmerken bestuur
We zien geen samenhang met de tevredenheidscijfers over de opleidingen onder de instelling zoals we dat wel zagen bij onder de norm presteren. Ook zien we geen verschillen tussen regio's en tussen instellingen met verschillende groottes.

Oordelen op bestuursniveau
We zien geen samenhang met oordelen op bestuursniveau en vaker hoger boven de normen presteren.

5 Uitkomsten Focusgroep

Om de uitkomsten van het kwantitatieve deel van dit onderzoek duiding te geven zijn we aan de hand van de kwantitatieve uitkomsten in gesprek gegaan met inspecteurs uit de sectoren primair onderwijs, voortgezet onderwijs, (voortgezet) speciaal onderwijs en middelbaar beroepsonderwijs. Hieronder beschrijven we de onderzoeksopzet en inzichten.

5.1 Onderzoeksopzet

De kwantitatieve analyses hebben geleid tot een lijst van besturen met scholen/opleidingen onder hun hoede die relatief vaak laag presteren, en een lijst van besturen met scholen/opleidingen die relatief vaak hoog presteren. Bij het opstellen van deze lijsten is rekening gehouden met de leerlingenpopulatie op een school of opleiding. Uitgenodigde inspecteurs zijn óf contactinspecteurs van betreffende besturen, óf inspecteurs met specifieke sectorkennis. Voorafgaand aan de focusgroep is er getoetst of zij de door ons aangewezen besturen ook als laag- of hoogpresterend erkenden.

In totaal vonden er twee focusgroepen plaats. Eén voor funderend onderwijs (po, vo en (v)so) met 9 deelnemende inspecteurs, en één voor het middelbaar beroepsonderwijs, waar 6 inspecteurs met elkaar in gesprek gingen. Eén inspecteur die niet bij de focusgroep mbo aanwezig kon zijn is later nog één-op-één geïnterviewd. Dezelfde vragen zijn voorgelegd.

Tijdens de focusgroep zijn eerst de besturen besproken met scholen/opleidingen die relatief vaak tot de laagpresteerders behoorden. Er is dieper ingegaan op wat deze besturen kenmerkt in termen van o.a. laag presteren, visie, onderwijsleiderschap, organisatorische kenmerken, historie en continuïteit. Daarna is tijdens het gesprek gesproken over welke handelingsruimte besturen hebben om de resultaten van hun scholen/opleidingen te verbeteren. Ditzelfde is herhaald voor besturen met goed presterende opleidingen. Ook is ingegaan op de vraag wat besturen met systematisch zwak presterende scholen/opleidingen kunnen leren van besturen met goed presterende scholen/opleidingen.

De focusgroepen vonden digitaal plaats. De gespreksleider leidde de discussie aan de hand van een semi-gestructureerde gespreksleidraad. Van beide focusgroepen is een verslag gemaakt dat ter verificatie naar de deelnemers is gestuurd en is aangepast op basis van feedback van de deelnemers.

Voorafgaand aan het onderzoek zijn alle deelnemers op de hoogte gesteld van het doel van de focusgroep en de vragen die er gesteld zouden gaan worden.

Twee onderzoekers werkten samen tijdens de dataverzameling en data-analyse. Op verschillende momenten controleerden de onderzoekers elkaars werk. Een deel van de analyse bestond uit het herkennen van gemeenschappelijke factoren uit beide focusgroepen. Dit is gedaan door het coderen van de twee verslagen. Relevante fragmenten zijn gecodeerd, daarna zijn de codes in centrale categorieën geïntegreerd en zijn er verbanden gelegd tussen deze categorieën.

5.2 Resultaten

Uit beide focusgroepen kwam naar voren dat de relatie tussen de prestaties op leerlingniveau (onderwijsresultaten) en bestuurlijk handelen complex is. Zo is deze relatie niet altijd objectiveerbaar en is een breed scala aan (vaak niet meetbare) factoren van invloed.

Onderstaand staan de gemeenschappelijke (bestuurlijke) kenmerken uit beide focusgroepen genoemd die volgens de bevraagde inspecteurs van invloed zijn op de hoogte van de onderwijsresultaten van scholen/opleidingen onder het bestuur. Eerst beschrijven we factoren uit de focusgroep mbo, daarna uit de focusgroep funderend onderwijs (po, vo, (v)so).

5.2.1 *Middelbaar beroepsonderwijs*

Tijdens de focusgroep voor de sector mbo komt de relatie tussen bestuurlijk handelen en onderwijsresultaten van opleidingen minder duidelijk naar voren dan tijdens de focusgroep funderend onderwijs. Een mogelijke verklaring hiervoor is dat een bestuurder in het mbo over het algemeen verder van zijn opleidingen afstaat dan een bestuurder in andere onderwijssectoren. Ook zijn de verschillen tussen opleidingen, die in verschillende domeinen en op verschillende niveaus worden aangeboden, veel groter dan bijvoorbeeld tussen afdelingen en scholen in het po en vo.

Contextfactoren

Voor opleidingen met een meer uitdagende leerlingenpopulatie is het over het algemeen moeilijker om hoge resultaten te behalen. Maar ook de regio waarin een opleiding wordt aangeboden is van invloed. In sommige regio's is het moeilijker adequaat personeel aan te trekken en zijn sociale problemen groter dan in andere regio's. Toch zijn verschillen in resultaten niet altijd toe te kennen aan context. Een voorbeeld hiervan zijn verschillende roc's die opereren in een vergelijkbare context maar toch verschillende resultaten behalen.

Analyseren en monitoren onderwijsresultaten

Goede besturen monitoren de onderwijsresultaten en sturen tijdig en adequaat bij. Zij stellen doelen, evalueren deze doelen en zijn in staat resultaten planmatig te verbeteren. Een uitkomst van de evaluatie kan zijn dat doelen worden bijgesteld.
De deelnemers noemen het negatieve effect van analyses met te weinig diepgang. Een bestuur moet bijvoorbeeld weten met wat voor leerlingen hij te maken heeft. En als er maatregelen worden ingezet om onderwijsresultaten te verbeteren, moet men weten wat het effect van deze maatregelen is geweest. Als studenten om input wordt gevraagd is het belangrijk ook daadwerkelijk met deze input aan de slag te gaan. Ook is het van belang verder te kijken dan de eigen organisatie. Wat verklaart bijvoorbeeld verschillen in resultaten met andere instellingen ondanks het bestaan van (ogenschijnlijk) gelijke contextfactoren?

Personeel

Een hogere kwaliteit van het personeel, zowel leidinggevenden als docenten, leidt tot betere resultaten. Kwaliteit kan worden verbeterd door middel van professionalisering van zittend personeel maar ook door het aantrekken van nieuw personeel. Soms functioneren personeelsleden niet goed maar blijven zij dezelfde positie bekleden. Als onvoldoende functionerende professionals te lang op dezelfde posities blijven kan dit enorme frustraties opleveren en zet dit resultaten onder druk. Ook worden een hoge werkdruk, veel wisselingen van personeel en een (relatief) hoog ziekteverzuim genoemd als belemmerende factoren.

Bestuurskracht

De deelnemers in de focusgroepen noemen het negatieve effect van gebrek aan bestuurskracht, maar ook het positieve effect van het bestaan ervan. Goede besturen hebben het vermogen om met kennis van zaken uitdagingen succesvol aan te gaan. Bestuurders laten zich zien op opleidingen en zijn betrokken. Ze

hebben een sterke onderwijskundige visie en expertise en weten vanuit die positie de juiste gesprekken te voeren en beweging in gang te zetten. In deze visie speelt toekomstgerichtheid een belangrijke rol.
Dit sterke onderwijskundige leiderschap werd ook gelinkt aan de kwaliteitscultuur op opleidingen en binnen het bestuur. Het bestuur stimuleert het leren van en met elkaar, door bijvoorbeeld opleidingen best-practices met elkaar te laten delen. Het neemt zelf een voorbeeldfunctie in door het stimuleren van een open dialoog.
Het bestuur voelt een gemeenschappelijke verantwoordelijkheid en neemt de stafdienst en de opleidingsprofessionals mee in zijn beslissingen. Het bestuur stelt kaders en laat tegelijkertijd opleidingen en professionals voldoende autonoom.

Financiën
Tijdens de focusgroep werd er kort ingegaan op de financiële situatie van een bestuur. Als een bestuur een slechtere financiële positie heeft, kan er weinig tot niets geïnvesteerd worden in de kwaliteitsagenda. Dit zet resultaten onder druk.

5.2.2 *Funderend onderwijs (po, vo, (v)so)*

Contextfactoren
Ook tijdens de focusgroep funderend onderwijs werd de essentiële rol van contextfactoren benadrukt. Eén van de eerst genoemde contextfactoren met een uitwerking op resultaten is een meer uitdagende leerlingenpopulatie op een school. Scholen met een meer uitdagende leerlingenpopulatie hebben over het algemeen meer moeite om de inspectiegrenzen te behalen. Ook de regio waarin het bestuur opereert speelt een rol. In sommige regio's zijn sociale problematieken groter en verhuizen leerlingen regelmatig naar buurten met hogere sociaal-economische status. Daarnaast heerst er in sommige regio's een gebrek aan gekwalificeerd personeel. Dit alles zet resultaten onder druk.

Zicht hebben, monitoren en (tijdig) bijsturen
De deelnemers in de focusgroepen noemen het negatieve effect van een gebrek aan monitoren, doelen stellen en tijdig bijstellen van doelen, maar ook het positieve effect van het bestaan hiervan.
Goede besturen monitoren de onderwijsresultaten en sturen tijdig en adequaat bij. Zij stellen doelen die passen bij de leerlingenpopulatie op hun scholen. Ze evalueren deze doelen en zijn in staat resultaten planmatig te verbeteren. Maar ze vergelijken bijvoorbeeld ook de eigen resultaten met de resultaten van andere scholen/besturen. Bij het verzamelen en analyseren van gegevens wordt zowel gedoeld op kwantitatieve als kwalitatieve gegevens. Men noemt ook het 'gevaar' van zwakke zelfevaluaties. De gegevens moeten op een deugdelijke manier verzameld zijn en een reëel beeld geven van de daadwerkelijke situatie op een school.
Uiteindelijk moeten doelen evalueren worden. Het bijstellen van doelen kan een conclusie zijn van deze evaluatie.

Kwaliteit personeel
De kwaliteit van het personeel, van zowel leerkrachten als schoolleiders, hangt sterk samen met het functioneren van een school. Voor sommige scholen is het moeilijk voldoende gekwalificeerd personeel te vinden. Soms functioneren bestaande personeelsleden onvoldoende, maar blijven zij op dezelfde positie omdat er geen betere leraren kunnen worden aangetrokken. Hier verschilt de handelingsruimte van het bestuur nog weleens. In sommige regio's zijn personeelsproblemen nijpender dan in andere.

Bestuurskracht

De deelnemers in de focusgroepen noemen het negatieve effect van een gebrek aan bestuurskracht, maar ook het positieve effect van het bestaan ervan. Besturen met voldoende bestuurskracht weten met kennis van zaken uitdagingen structureel aan te gaan. Zij hebben zicht op de problematieken die spelen op hun scholen en weten doortastende maatregelen te treffen. De bestuurder gaat regelmatig in gesprek met schoolleiders en werkt samen aan duurzame verbetering. Ook legt hij blinde vlekken bloot. Hij is bijvoorbeeld kritisch of het onderwijsconcept op een school wel past bij de leerlingenpopulatie op die school.
Ook is het belangrijk dat scholen een duidelijk beeld hebben van wat er van hen verwacht wordt en dat zij hierin worden begeleid/aangestuurd door het bestuur. Hier wordt de gulden middenweg genoemd. Een te grote autonomie van scholen leidt vaak niet tot uitvoering van deze visie in de praktijk. Te weinig ruimte voor eigen beslissingen leidt vaak ook niet tot het gewenste resultaat.
Overigens werd ook meerdere malen externe attributie genoemd als een belemmerende factor. Met externe attributie wordt bedoeld dat oorzaken van bijvoorbeeld lagere resultaten worden gezien als liggend buiten de beïnvloedingsmogelijkheid van de school of het bestuur.

Randstedelijke problematieken
Wat tijdens de focusgroep ook naar voren komt, zijn de besturen in kleinere (rand)gemeenten met grootstedelijke issues. Men doelde hier op regio's waar er relatief veel sociale problemen zijn, waar leerlingenpopulaties uitdagender zijn, maar waar dit niet als dusdanig erkend wordt.

6 Referenties

CPB (2019). *Verschillen in leerresultaten tussen basisscholen*. Den Haag: Centraal Planbureau.