

Schaaleffecten in het basisonderwijs

Eric-Jan Troost



**Onderzoek naar financiële en
onderwijskundige schaaffecten
in het basisonderwijs en de
invloed van schaalgrootte op de
onderwijskwaliteit**

Schaaleffecten in het basisonderwijs

Onderzoek naar financiële en onderwijskundige schaalearfecten in het basisonderwijs en de invloed van schaalgrootte op de onderwijskwaliteit

Auteur

Eric-Jan Troost (358275)

Scriptie ter afronding van de bacheloropleiding Economie & Bedrijfseconomie

Begeleider

Dr. Bas Karreman

Erasmus School of Economics

Capaciteitsgroep Applied Economics

Plaats en tijd

Erasmus School of Economics

Erasmus Universiteit Rotterdam

Juli 2014

Samenvatting

De afgelopen jaren is er een maatschappelijk debat gevoerd over schaalvergroting in het basisonderwijs. De plannen van staatssecretaris Dekker om kleine scholen te laten fuseren of verdwijnen hebben dit debat alleen maar versterkt. Begin 2014 zijn de plannen van de staatssecretaris van tafel gegaan, maar daarmee is de vraag of kleinere basisscholen op een kostenefficiënte manier onderwijs van een goede kwaliteit kunnen leveren, niet beantwoord.

Dit onderzoek wil een beeld geven van de mogelijke relatie in het basisonderwijs tussen schaalgrootte en kostenefficiëntie enerzijds en schaalgrootte en onderwijskwaliteit anderzijds.

Als indicator voor het effect van schaalgrootte op de kostenefficiëntie wordt gebruik gemaakt van de maatstaf gemiddelde vaste kosten per leerling. Het blijkt dat voor scholen en schoolbesturen uit financieel oogpunt een minimale grootte van 150 leerlingen wenselijk is. Tot deze grootte kunnen namelijk grote kostenvoordelen worden behaald. Tot een grootte van ongeveer 500 leerlingen zijn er nog beperkte schaalvoordelen te behalen. Als een school of bestuur meer dan 500 leerlingen aanstuurt, verdwijnt het effect van schaalgrootte en verloopt de kostencurve vlak. Er wordt geen omslagpunt tussen schaalvoor- en schaalnadelen waargenomen.

Gemeten in het aantal scholen per bestuur, is er wel een omslagpunt van schaalvoordelen naar schaalnadelen te bepalen. Een schoolbestuur dat één school aanstuurt, is niet significant duurder of goedkoper dan gemiddeld. Schoolbesturen die twee tot maximaal tien scholen aansturen zijn het meest kostenefficiënt. Schoolbesturen die zestien of meer scholen aansturen behalen significante schaalnadelen.

Uit een analyse op groepsniveau blijkt dat klassen van 19 tot 25 leerlingen het meest kostenefficiënt zijn. Om klassen van deze omvang te bereiken is een bepaalde schoolgrootte van ten minste 150 leerlingen noodzakelijk.

In de analyse naar verschillen tussen denominaties bleek dat reformatorische en protestants-christelijke schoolbesturen per leerling significant lagere kosten behalen dan gemiddeld, terwijl de kosten per leerling bij openbare en met name islamitische schoolbesturen significant hoger zijn dan gemiddeld. Mogelijke verklaringen voor de hogere kosten per leerling op islamitische scholen liggen in de gemiddeld kleinere klassen, het gemiddeld lagere opleidingsniveau van de ouders en de grote hoeveelheid achterstandsleerlingen.

De CITO-score is de indicator voor de onderwijskwaliteit in het bepalen van het effect van schaalgrootte op de kwaliteit van een school. Om het effect van schoolkenmerken te bepalen moet eerst worden gecorrigeerd voor algemene en schoolse leerlingkenmerken. De invloed van schaalgrootte op onderwijskwaliteit is na controle voor deze variabelen klein. Er is niet gebleken dat leerlingen op scholen met 145 leerlingen of minder lagere CITO-scores halen dan gemiddeld. Wel is er in verschillende modellen een effect gevonden dat wijst op schaalnadeel voor grote scholen. Op scholen met meer dan 500 leerlingen is de gemiddelde CITO-score tot 1,5 punt lager.

Inhoudsopgave

	Paginanummer
Titelpagina	1
Samenvatting	2
1. Inleiding	5
2. Theoretisch Kader: Financiële Schaaleffecten	7
2.1 Schaaleffecten	7
2.2 Niveaus van analyse	7
2.3 Schaalvergroting in het basisonderwijs sinds 1985	7
2.4 Schaaleffecten in het onderwijs	8
2.5 Schaalvoordelen	9
2.6 Schaalnadelen	12
2.7 Geen schaaleffect	14
2.8 Hypotheses	15
3. Theoretisch Kader: Onderwijskundige Schaaleffecten	16
3.1 Onderwijskwaliteit	16
3.2 Geneste structuur	16
3.3 Algemene leerling- en gezinskenmerken	17
3.4 Schoolse leerling- en gezinskenmerken	18
3.5 Schoolkenmerken en hypothese	19
4. Data & Methodologie: Financiële Schaaleffecten	22
4.1 Dataselectie	22
4.2 Hypotheses toetsen	23
5. Data & Methodologie: Onderwijskundige Schaaleffecten	24
5.1 Data COOL-cohortonderzoek	24
5.2 Dataselectie leerlinggegevens	25
5.3 Data DUO	25
5.4 Dataselectie schoolgegevens	25
5.5 Methodologisch Model	26
5.6 Hypotheses toetsen	27
6. Resultaten: Financiële Schaaleffecten	29
6.1 Algemene gegevens uitgebreide subgroep	29
6.2 Schaaleffecten op bestuursniveau – aantal leerlingen	29
6.3 Schaaleffecten op schoolniveau	31
6.4 Schaaleffecten op bestuursniveau – aantal scholen	33
6.5 Schaaleffecten op groepsniveau	35
6.6 Schaaleffecten naar denominatie	36

	Paginanummer
7. Resultaten: Onderwijskundige Schaaffecten	41
7.1 Startpositie	41
7.2 Schoolkenmerken	41
7.3 Definitieve modellen	51
8. Conclusie, tekortkomingen en aanbevelingen	53
8.1 Beantwoording hypothesen	53
8.2 Beantwoording probleemstelling	56
8.3 Tekortkomingen	57
8.4 Beleidsaanbevelingen en aanbevelingen tot vervolgonderzoek	58
Dankwoord	60
Literatuurlijst	61
Bijlage 1: Bronnen schaaffecten	66
Bijlage 2: Uitschieter	67
Bijlage 3: Denominaties met minder dan 10 schoolbesturen	68
Bijlage 4: Variabelen op het leerlingniveau	69
Bijlage 5: Informatie over leerlingen CITO-eindtoets I	70
Bijlage 6: Informatie over leerlingen CITO-eindtoets II	72
Bijlage 7: Variabelen op het schoolniveau	74
Bijlage 8: Informatie over scholen CITO-eindtoets I	75
Bijlage 9: Informatie over scholen CITO-eindtoets II	76
Bijlage 10: Correlatietabellen opleidingsniveau en etniciteit	77
Bijlage 11: Modellen algemene leerling- en gezinskenmerken	78
Bijlage 12: Modellen schoolse leerling- en gezinskenmerken	85

1. Inleiding

Na jaren van het stimuleren van schaalvergroting in het basisonderwijs leek het erop dat de Rijksoverheid in 2008 een nieuwe weg was ingeslagen. Destijds was er de nodige bezorgdheid over de te ver doorgeschoten schaalvergroting, waardoor de menselijke maat onder druk zou komen te staan (Rijksoverheid, 2008). Enkele jaren later lijkt het alsof de overheid een grote draai heeft gemaakt in het debat over schaalvergroting in het basisonderwijs. Eind 2013 kwam staatssecretaris Dekker namelijk met schaalvergrotingsplannen om de verwachte krimp in leerlingenaantallen op te kunnen vangen (OC&W, 2013a), gebaseerd op een advies van de Onderwijsraad. Staatssecretaris Dekker wilde de kleinescholentoeslag¹ ‘aanpassen’, waardoor kleine scholen² feitelijk gezien moeten fuseren of verdwijnen. De staatssecretaris lijkt er hierbij vanuit te gaan dat grotere scholen onderwijs van een betere kwaliteit kunnen leveren. Hij stelt namelijk dat samenwerking ervoor zorgt dat *‘met hetzelfde geld meer kwaliteit en diversiteit kan worden geleverd, of in ieder geval kan worden gewaarborgd’* (OC&W, 2013a). Door de afspraken met D66, ChristenUnie en SGP is de afschaffing van de kleinescholentoeslag begin 2014 na maanden van onrust binnen de basisscholen van tafel gegaan (Volkskrant, 2014).

Daarmee is echter de vraag of kleinere basisscholen op een kostenefficiënte manier onderwijs van een goede kwaliteit kunnen leveren niet beantwoord. De Onderwijsraad stelt in haar advies dat de kwaliteit van kleine basisscholen onder druk staat en in de komende jaren verder onder druk komt te staan door dalende leerlingaantallen (Onderwijsraad, 2013). Bepaalde regio’s, voornamelijk op het platteland en in de grensgebieden, zijn extra gevoelig voor deze daling in leerlingaantallen. Daarnaast ziet de Raad in de hogere kosten, hogere kwetsbaarheid en lagere onderwijskwaliteit redenen om het aantal kleine scholen zo veel mogelijk te beperken (Onderwijsraad, 2013). In de wetenschappelijke wereld is er overeenstemming over het feit dat kleine (basis)scholen in Nederland tot een bepaald niveau schaalvoordelen kunnen behalen (Blank, 1990; Van de Venne, 2002; Onderwijsraad, 2005; Waterreus, 2009; IPSE, 2012a; Onderwijsraad, 2013). Een precies omslagpunt tussen schaalvoordelen en nadelen wordt niet altijd gevonden; dit hangt onder andere af van de berekeningsmethode.

Over de relatie tussen schaalgrootte en onderwijskwaliteit is minder eenduidigheid. Uit verschillende Nederlandse studies in het voortgezet onderwijs is gebleken dat er geen kwaliteitsverschil kan worden vastgesteld tussen grote en kleine scholen (Regioplan, 2008). In het basisonderwijs blijken kleine scholen vaker zwak of zeer zwak te zijn ten opzichte van grotere scholen (Onderwijsraad, 2013), maar de Inspectie van het Onderwijs constateert dat het verschil tussen kleine en grotere scholen steeds kleiner wordt (IvhO, 2014). De Inspectie stelt zelfs dat er geen optimale grootte van scholen bestaat, omdat deze altijd afhangt van een specifieke lokale context (IvhO, 2014), waarbij aspecten als de samenstelling van de leerlingenpopulatie, ouderbetrokkenheid, denominatie, professionaliteit van het management en tal van andere zaken een rol kunnen spelen.

¹ Scholen met minder dan 145 leerlingen komen in aanmerking voor een kleinescholentoeslag. De toeslag loopt op als het aantal leerlingen kleiner wordt.

² In het voorstel wordt een school als klein gezien bij minder dan 100 leerlingen.

Dit onderzoek wil een beeld geven van de mogelijke relatie tussen schaalgrootte en kostenefficiëntie enerzijds en schaalgrootte en onderwijskwaliteit anderzijds. Met de resultaten van dit onderzoek in de hand kan het maatschappelijk debat over schaalvergroting in het basisonderwijs verrijkt worden en kunnen specifieke beleidsmaatregelen worden genomen om de onderwijskwaliteit en de kostenefficiëntie te verbeteren. Het onderzoek is tevens een update van eerder wetenschappelijk onderzoek, waarbij het bijzonder is dat er zowel vanuit kosten oogpunt als vanuit kwaliteitsoogpunt naar de invloed van schaalgrootte gekeken wordt. De uiteindelijke probleemstelling in het onderzoek is als volgt geformuleerd:

‘In hoeverre zijn er schaalvoordelen te behalen in het basisonderwijs en welke invloed heeft de mogelijke aanwezigheid van schaaleffecten op de kwaliteit van het basisonderwijs?’

De uitwerking van de probleemstelling valt uiteen in twee delen. Allereerst wordt er gekeken naar het financiële deel: zijn er op het kostengebied schaaleffecten waar te nemen in het basisonderwijs? Bij het beantwoorden van deze vraag staan de gemiddelde vaste kosten per leerling steeds centraal. De vraag wordt geanalyseerd op verschillende niveaus, namelijk op het schoolniveau, het bestuursniveau, naar leerling-docent verhouding en naar denominatie. Daarna wordt de analyse breder gemaakt en wordt er gekeken of schaalgrootte effect heeft op de kwaliteit van het basisonderwijs. Het uitgangspunt bij deze vraag is de individuele CITO-score. De CITO-score geldt als afhankelijke variabele en in de modellen die worden opgebouwd worden steeds bepaalde controlevariabelen toegevoegd. Deze controlevariabelen zijn ingedeeld naar niveau: eerst worden variabelen op het algemene leerlingniveau toegevoegd, daarna op het schoolse leerlingniveau en uiteindelijk worden variabelen op het schoolniveau toegevoegd. Door deze werkwijze kan het effect van schaalgrootte op de onderwijskwaliteit, gecontroleerd voor leerling- en schoolachtergrond, worden vastgesteld. Om tot bepaalde uitgangspunten te komen wordt eerst een literatuurstudie gedaan, waarin schaaleffecten vanuit zowel de economische als de onderwijskundige visie behandeld zullen worden. Aan de hand van de literatuurstudie en in het verleden uitgevoerde onderzoeken worden hypothesen opgesteld. In de conclusie wordt bepaald of de hypothesen verworpen moeten worden en wordt er een antwoord gegeven op de probleemstelling. Tevens zullen er tekortkomingen van het onderzoek en mogelijkheden voor vervolgonderzoek behandeld worden.

2. Theoretisch Kader: Financiële Schaaffecten

2.1 Schaaffecten

Schaaffecten kunnen verschillend gedefinieerd worden. In dit onderzoek verwijzen schaaffecten naar de samenhang tussen de grootte van een organisatie en de kosten per eenheid product (IPSE, 2012a). In een productieproces kunnen schaalvoordelen behaald worden als de gemiddelde kosten lager worden over een bepaald bereik van de output. In dit onderzoek worden de gemiddelde kosten gedefinieerd als de gemiddelde vaste kosten per eenheid product. Vaste kosten zijn altijd even hoog, ongeacht het volume van het productieproces (Hilton & al, 2008). Schaalvoordelen kunnen dan behaald worden als het productievolume stijgt, zodat de vaste kosten kunnen worden gespreid over meer eenheden product (Besanko & al, 2010). Voor kleine organisaties gelden hoge gemiddelde kosten per eenheid product, die snel dalen naarmate de organisatie groter wordt (IPSE, 2012a). Bij toenemende schaalgrootte dalen de gemiddelde kosten nog maar langzaam, of ze kunnen zelfs weer langzaam toenemen. In het laatste geval is er sprake van schaalnadeel. In veel productieprocessen is er sprake van een efficiënt productiepunt, waarop de gemiddelde kosten per eenheid product het laagst zijn. Er is dan sprake van een U-vormige kostencurve. Een L-vormige kostencurve komt ook geregeld voor. De gemiddelde kosten dalen dan tot een bepaald punt, waarna deze nagenoeg constant blijven (IPSE, 2011). Ook een omgekeerde L-vorm is mogelijk, waarbij vanaf een bepaald punt de gemiddelde kosten alleen maar blijven stijgen (IPSE, 2011).

2.2 Niveaus van analyse

In het vervolg worden de financiële schaaffecten geanalyseerd op bestuursniveau, schoolniveau en groepsniveau. In een onderzoek naar schaaffecten in het onderwijs wordt dikwijls de juridische entiteit als uitgangspunt genomen (IPSE, 2011). Met deze juridische entiteit wordt de school of de verzameling scholen bedoeld die gezamenlijk onder één bevoegd gezag vallen. De juridische entiteit (het schoolbestuur) draagt ook de financiële en bestuurlijke verantwoordelijkheid (IPSE, 2012a). De analyse op schoolniveau vindt plaats op een lager niveau. Met een school wordt een onderwijsinhoudelijke en organisatorische entiteit bedoeld die aangeduid wordt met een BRIN-nummer (IPSE, 2012a). Eén school kan daarbij uit meerdere fysieke locaties bestaan. De analyse op groepsniveau gaat nog een niveau lager en kijkt naar de gemiddelde grootte van lesgroepen binnen scholen.

2.3 Schaalvergroting in het basisonderwijs sinds 1985

Met de invoering van de WBO³ in 1985 heeft de Nederlandse overheid schaalvergroting in het basisonderwijs actief gestimuleerd. Met de invoering van deze wet werd de opheffingsnorm gerelateerd aan de leerlingendichtheid per gemeente, waardoor schaalvergroting werd afgedwongen (Onderwijsraad, 2008). Met de invoering van de WBO werden het kleuteronderwijs en het lager onderwijs samengevoegd tot het basisonderwijs, waardoor het aantal vestigingen halveerde van ongeveer 16.000 tot ruim 8.000 (IPSE, 2012b). Met de SBK-regeling⁴ stimuleerde de overheid schaalvergroting ook financieel.

³ Wet op het Basisonderwijs (1985)

⁴ Stimuleringsbijdrage Bestuurlijke Krachtenbundeling (1994)

Met deze regeling werd beoogd het beleidsvoerend vermogen van scholen te versterken door bestuurlijke krachtenbundeling. De regeling zorgde ook binnen een bepaald bereik voor een kostenvoordeel. Bij een groter aantal scholen per bestuur kunnen de vaste lasten, die worden verondersteld gelijk te blijven, verdeeld worden over meerdere scholen.

De laatste grote ingreep in de financiering van het basisonderwijs vond plaats in 2006, toen de lumpsumbekostiging werd ingevoerd. Het bevoegd gezag van een school of een scholengroep ontving voortaan één budget voor personele en materiële uitgaven en mocht dit binnen bepaalde grenzen naar eigen inzicht besteden. De invoering van de lumpsumbekostiging heeft geleid tot een verdere autonomievergroting voor scholen (IPSE, 2012b). Verder heeft de invoering van de lumpsumbekostiging de voortgaande fusiebeweging in het primair onderwijs versterkt. Demografische veranderingen hebben daarbij ook invloed op de schaalvergroting in het basisonderwijs. In 2020 zullen er ongeveer 100.000 leerlingen minder zijn op basisscholen in vergelijking met 2011 (Onderwijsraad, 2013). Deze ontwikkelingen geven aanleiding om te verwachten dat de bestuurlijke schaalvergroting in de toekomst nog verder zal toenemen.

Met name sinds het begin van deze eeuw is het aantal besturen in het basisonderwijs sterk teruggelopen. In 2000 waren er nog 1.672 basisschoolbesturen, in 2012 is dit bijna met de helft teruggelopen tot 853 besturen. Het aantal leerlingen dat onder één bestuur valt, is tegelijkertijd gestegen van 589 in 2000 tot 909 in 2012 (OC&W, 2013b). In 1985 (na invoering van de WBO) telde een school gemiddeld 180 leerlingen (IPSE, 2012b), terwijl dat in de jaren daarna langzaam is opgelopen tot 222 in 2012 (Onderwijsraad, 2013).

2.4 Schaaleardeffecten in het onderwijs

In de navolgende paragrafen wordt er gekeken naar de schaaleardeffecten binnen het onderwijs. Hierbij ligt de focus op de Nederlandse situatie en wordt naast onderzoek dat is uitgevoerd in het basisonderwijs soms gebruik gemaakt van onderzoek dat uitgevoerd is in het voortgezet onderwijs. Er wordt in dit stadium nog geen onderscheid gemaakt tussen schaalgrootte op bestuursniveau, schoolniveau en groepsniveau.

Blank en Urlings (IPSE, 2012a) hebben de bedrijfsvoering in het voortgezet onderwijs in Nederland onderzocht. Daaruit bleek dat vestigingsgrootte een positief effect heeft op de kostendoelmatigheid. Duizend leerlingen extra per vestiging leidt tot een verbetering van de kostendoelmatigheid met 2,1 procent. Elke extra vestiging heeft een positief effect van 0,3 procent. Grotere onderwijsinstellingen werken dus doelmatiger dan instellingen die bestaan uit één vestiging. Uit de studie van Blank en Urlings blijkt verder dat de kosten per eenheid product de vorm van een hockeystick hebben, waarbij het omslagpunt van schaalvoordelen naar schaalnadelen in het voortgezet onderwijs ligt bij ongeveer 6.500 leerlingen.

De Onderwijsraad (2008) stelt vast dat de minimale efficiënte schaal voor een basisschool rond de 200 leerlingen zal liggen, met 8 groepen van ongeveer 25 leerlingen. Dit lijkt onder andere gebaseerd op een onderzoek van Blank (1990), waaruit blijkt dat scholen in het basisonderwijs tot een grootte van 200 leerlingen schaalvoordelen kunnen behalen. Scholen die groter zijn kunnen nog geringe schaalvoordelen behalen. Blank constateert daarom dat er geen sterke financiële argumenten zijn voor zeer grote scholen.

In een recente studie (Blank & al, 2007) blijkt dat basisscholen tot een grootte van 125 leerlingen grote schaalvoordelen kunnen behalen. Scholen tot 300 leerlingen kunnen beperkte schaalvoordelen behalen en daarna slaan de schaalvoordelen om in schaalnadelen. Scholen groter dan 550 leerlingen behalen zeker schaalnadelen (Blank & al, 2007). Uit een oudere studie uit 2000 blijkt dat basisscholen tot een grootte van ongeveer 400 leerlingen schaalvoordelen kunnen behalen (IPSE, 2011).

In een rapport van de Onderwijsraad uit 2005 is de Raad voorzichtig in het aanwijzen van een optimale schaal. In dat rapport stelt de Raad *'dat een optimale schaal moeilijk vast te stellen is'*, omdat deze onder andere afhankelijk is van lokale omstandigheden (Onderwijsraad, 2005). De Onderwijsraad (2005) geeft aan dat verschillende beleidsniveaus en belanghebbenden invloed uitoefenen op de optimale schaal. Het is daarom eigenlijk niet eenvoudig mogelijk om een universele optimale schaal vast te stellen.

2.5 Schaalvoordelen

Schaalvoordelen kunnen ontstaan door een betere benutting van arbeid en kapitaal, bijvoorbeeld omdat functies binnen scholen kunnen worden samengevoegd (IPSE, 2012a) of omdat er professionelere bestuurders met meer expertise kunnen worden aangetrokken (Onderwijsraad, 2008). Verder blijkt dat er binnen grote scholen vaker specialisatie plaatsvindt (Van de Venne, 2002) en dat taken binnen grotere scholen eenvoudiger verdeeld kunnen worden over het personeel (Onderwijsraad, 2013).

Financiële risico's kunnen binnen een groep van scholen gedeeld worden en de bedrijfsvoering kan gezamenlijk plaatsvinden, waardoor doelmatiger gewerkt kan worden (Onderwijsraad, 2008). Daarnaast hebben grote scholen meer financiële middelen om een breed onderwijsaanbod neer te zetten, met meer mogelijkheden tot differentiatie en vakspecialisatie. Grote schoolverbanden kunnen ook meer geld vrijmaken om te experimenteren met onderwijskundige innovaties en om te investeren in de onderwijskwaliteit (Onderwijsraad, 2008; Regioplan, 2008). Ten slotte geeft de Onderwijsraad (2008) aan dat er binnen grote scholen meer professionaliseringsmogelijkheden zijn voor het personeel.

Grote scholen blijken minder kwetsbaar te zijn voor onverwachte tegenvallers. Zieke docenten kunnen bijvoorbeeld zwaar drukken op de continuïteit en onderwijskwaliteit van een kleinschalige onderwijsorganisatie (IvhO, 2013). Ook blijkt dat een zwakke docent zwaarder drukt op de onderwijsresultaten van een kleine school dan van een grote school (Onderwijsraad, 2013). Schaalvergroting op bovenschools niveau biedt nog het voordeel van het ontlasten van de directie op schoolniveau, waardoor de schoolleiding zich kan concentreren op de vormgeving van het primaire proces en onderwijskundig beleid (Van de Venne, 2002). Verder noemt van de Venne (2002) nog voordelen van schaalgrootte in de vorm van flexibele inzetmogelijkheden van het personeel en de grotere carrière mogelijkheden binnen schoolverbanden. Verder is door schaalvergroting het aandeel dat een basisschool uitgeeft aan management substantieel afgenomen (Blank & Haelermans, 2008), waardoor er, tot een bepaald punt, meer budget beschikbaar blijft voor het primaire proces.

Uit het rapport van de Onderwijsraad (2013) blijkt verder dat de kosten per leerling op een kleine school relatief hoog zijn; de kosten variëren van minder dan 4.000 euro per leerling op een gemiddelde school⁵ tot bijna 11.000 euro per leerling op een zeer kleine⁶ school. Kleine scholen zijn een stuk duurder dan grotere scholen, terwijl grotere scholen relatief een betere onderwijskwaliteit leveren (Onderwijsraad, 2005). Kleine scholen⁷ zijn nog steeds iets vaker zwak of zeer zwak dan grote scholen, ook al wordt het verschil steeds kleiner (IvhO, 2014). Als mogelijke verklaring voor de zwakkere kleine scholen noemt de Inspectie de vaker voorkomende combinatiegroepen op kleine scholen. In een kleine school moet de leraar zijn aandacht spreiden over twee, drie of soms wel vier verschillende jaarlagen, waardoor hij minder tijd heeft voor de lesstof van een specifieke groep leerlingen (Onderwijsraad, 2005). Deze groepen vragen dan ook meer managementvaardigheden van de docent. Combinatiegroepen bieden echter ook kansen in de samenwerking tussen leerlingen en in de ontwikkeling van de zelfstandigheid van leerlingen (IvhO, 2014). De Inspectie van het Onderwijs constateert verder dat lessen op grote scholen over het algemeen van een betere kwaliteit zijn omdat leraren op grotere scholen vaker meer didactische- en differentiatievaardigheden hebben (IvhO, 2014). Dat lessen op grotere scholen van een betere kwaliteit zijn kan komen doordat leerprocessen vaker worden gestandaardiseerd en geformaliseerd op grote scholen (Van de Venne, 2002).

De Onderwijsraad (2013) wijst erop dat in verband met de sociale interactie een minimale schoolgrootte wenselijk is. Op een grote school is het voor kinderen makkelijker om in contact te komen met leeftijdsgenootjes en om zich op het sociale vlak te ontwikkelen.

In het voorgaande zijn twee categorieën schaalvoordelen naar voren gekomen. De eerste categorie is statistisch kwantificeerbaar doordat er data beschikbaar is, terwijl de tweede categorie niet statistisch kwantificeerbaar is omdat data ontbreekt of niet openbaar is. De tweede categorie schaalvoordelen kan dienen als mogelijke verklaringsbron. Tabel 1 geeft een samenvatting van de potentiële schaalvoordelen die te behalen zijn in het onderwijs, onderverdeeld naar deze twee categorieën.

⁵ 225 leerlingen

⁶ <24 leerlingen

⁷ Kleine scholen zijn volgens de Inspectie van het Onderwijs scholen tot en met 100 leerlingen.

Schaalvoordelen in het onderwijs

Statistisch kwantificeerbaar

- ✓ *Doelmatigere vormgeving van het primaire proces (lagere kosten per leerling);*
- ✓ *Bovenschools bestuur → meer budget beschikbaar voor het primaire proces;*
- ✓ *Betere benutting van arbeid en kapitaal;*
 - *Relatief minder managers;*

Niet statistisch kwantificeerbaar

- ✓ *Betere benutting van arbeid en kapitaal;*
 - *Samenvoegen van functies;*
 - *Specialisatie;*
 - *Werknemers met meer expertise;*
 - *Verdeling van taken;*
- ✓ *Financiële risico's kunnen worden gedeeld;*
 - *Waarborging van de continuïteit;*
- ✓ *Gezamenlijke bedrijfsvoering;*
- ✓ *Breder onderwijsaanbod mogelijk;*
 - *Meer differentiatiemogelijkheden;*
 - *Meer specialisatiemogelijkheden;*
 - *Meer innovatiemogelijkheden;*
 - *Grotere investeringskracht;*
- ✓ *Grotere invloed op besluitvormende organen;*
- ✓ *Betere inzet van personeel;*
 - *Loopbaanmogelijkheden;*
 - *Professionaliseringsmogelijkheden;*
 - *Concentreren op één jaargroep;*
 - *Flexibele inzetmogelijkheden;*
- ✓ *Bovenschools bestuur;*
 - *Ontlasting van de directie op het schoolniveau;*
 - *Schoolleiding kan zich concentreren op vormgeven onderwijskundig beleid;*
- ✓ *Betere lessen;*
- ✓ *Standaardisatie van leerprocessen;*
- ✓ *Formalisatie van leerprocessen;*
- ✓ *Sociale interactie tussen leerlingen.*

Tabel 1: Schaalvoordelen in het onderwijs

2.6 Schaalnadelen

Schaalnadelen kunnen ontstaan door verregaande bureaucrativering, waardoor een gebrek aan sociale cohesie (IPSE, 2012a), ouderbetrokkenheid (Onderwijsraad, 2008) en een toename van anonimiteit (Onderwijsraad, 2005) kan ontstaan. Ook kan door schaalvergroting op bestuurlijk niveau de afstand tussen schoollocaties, leraren en ouders enerzijds en het bestuur anderzijds, groter worden, met als gevolg dat ouders en personeel een verminderde invloed hebben op de besluitvorming (Onderwijsraad, 2008). Een school kan hierdoor zelfs losraken van zijn maatschappelijke context.

De Onderwijsraad (2008) geeft verder aan dat de keuzevrijheid tussen scholen van een verschillende signatuur onder druk komt te staan als kleinere scholen verdwijnen of opgaan in een bestuur van gemengde signatuur. Hierdoor kunnen lokale of regionale monopolies en arbeidmonopsonies ontstaan (Waterreus, 2009; IPSE, 2011). Door beperkte keuzemogelijkheden kunnen de prikkels om te innoveren verdwijnen en kan er een suboptimale verdeling van middelen plaatsvinden (Waterreus, 2009).

Van de Venne (2002) ziet de ontkoppeling van het beheersmatig leiderschap en het onderwijskundig leiderschap als nadeel. Door de grotere afstand tussen de scholen en de algemene directie kan er op het schoolniveau onvoldoende zicht zijn op financiële en beheersmatige prestaties. Bovendien worden de communicatielijnen langer en worden besluiten vaker centraal genomen (Van de Venne, 2002).

De Onderwijsraad (2008) noemt verder nog het risico van doelmatigheidsverlies bij doorgesloten schaalvergroting, waardoor coördinatiekosten toenemen en het afstemmen van primaire processen lastiger wordt. Een meer complexe organisatie vraagt bovendien meer professionaliteit van het management (Van de Venne, 2002). Daarbij blijkt dat grote schoolbesturen vaker dan kleine schoolbesturen moeite hebben om de onderwijskwaliteit en de financiën onder controle te houden. De Inspectie geeft als verklaring de hogere complexiteit in grootschalige onderwijsorganisaties (IvhO, 2014). Ook blijkt dat sommige bestuurders streven naar budgetmaximalisatie, niet omdat dat per se beter is voor het onderwijs, maar om meer aanzien te verwerven en om hun carrière een boost te geven (Onderwijsraad, 2005).

Ook voor leerlingen zijn er negatieve effecten van doorgesloten schaalvergroting waar te nemen. Het algemeen welbevinden van leerlingen in het voortgezet onderwijs blijkt op grote scholen lager te zijn dan op kleine scholen (Onderwijsraad, 2005). Grote scholen scoren slechter op enkele pedagogische en didactische aspecten en er wordt relatief meer gepest. Daarnaast worden de leerling-docent relaties en de leerling-leerling relaties als minder positief ervaren (Regioplan, 2008).

Tabel 2 geeft een samenvatting van de potentiële schaalnadelen in het onderwijs. Ook in deze tabel wordt het onderscheid gemaakt tussen statistisch kwantificeerbare en niet statistisch kwantificeerbare data.

Schaalnadelen in het onderwijs

Statistisch kwantificeerbaar

- ✓ *Doorgeschoten schaalvergroting;*
 - *Doelmatigheidsverlies (hogere kosten per leerling);*
 - *Bureaucratisering;*
 - *Meer managers;*

Niet statistisch kwantificeerbaar

- ✓ *Bureaucratisering;*
 - *Gebrek aan sociale cohesie;*
 - *Gebrek aan ouderbetrokkenheid;*
 - *Toename van anonimiteit;*
 - *Centrale besluitvorming (standaardisatie, formalisatie);*
- ✓ *Verdwijnen van keuzevrijheid;*
 - *Lokaal onderwijsmonopolie van een school(bestuur);*
 - *Lokaal arbeidmonopolie van een school(bestuur);*
 - *Vermindering van innovatieprikkel;*
 - *Suboptimale verdeling van middelen;*
- ✓ *Schaalvergroting op bestuurlijk niveau;*
 - *Grotere afstand tussen leraren en ouders op de schoollocatie en het bestuur op afstand;*
 - *Verminderde invloed van schoollocaties, leraren en ouders op (strategische) besluitvorming;*
 - *Ontkoppeling van beheersmatig en onderwijskundig leiderschap;*
 - *Op schoolniveau onvoldoende zicht op financiële prestaties;*
 - *Communicatielijnen worden langer;*
- ✓ *Doorgeschoten schaalvergroting;*
 - *Toenemende coördinatiekosten en afstemmingsproblemen;*
 - *Controle houden op financiën en onderwijskwaliteit complexer in grotere organisaties;*
 - *Streven naar budgetmaximalisatie;*
- ✓ *Lager algemeen welbevinden van leerlingen;*
 - *Meer pestgevallen;*
 - *Relatie tussen leerlingen en docenten minder positief;*
 - *Relatie tussen leerlingen minder positief.*

Tabel 2: Schaalnadelen in het onderwijs

2.7 Geen schaaleffect

In een aantal studies komen ook aspecten naar voren waarop de schaalgrootte van een onderwijsinstelling geen invloed heeft. Uit het onderzoek van Regioplan (2008) blijkt dat er in Nederland geen bewijs is dat kleine of grote scholen een positief effect hebben op de leerprestaties. Daar moet wel bij worden opgetekend dat er in Nederland weinig onderzoeken gedaan zijn naar het verband tussen onderwijskwaliteit en schaalgrootte. Verder geeft Regioplan (2008) aan dat ouders en leerlingen zowel op kleine als op grote scholen even tevreden zijn over de onderwijskwaliteit. Het blijkt verder dat de grootte van een school geen effect heeft op de arbeidstevredenheid van docenten (Regioplan, 2008).

Een recente studie van Dijkgraaf en De Jong (2009) geeft aan dat er geen eenduidig verband is tussen schaalgrootte en onderwijskwaliteit in het Nederlandse voortgezet onderwijs. Daaruit volgt dat er onvoldoende wetenschappelijk bewijs lijkt te zijn voor overheidsbeleid gericht op schaalvergroting of -verkleining.

De Inspectie van het Onderwijs (2014) constateert dat het aantal zwakke of zeer zwakke kleine basisscholen al jaren afneemt en dat de verschillen in kwaliteit tussen grote en kleine scholen bijna volledig verdwenen zijn. Desondanks noemt de Onderwijsraad (2013) het een punt van zorg dat het aantal kleine scholen⁸ door dalende geboortecijfers tussen 2010 en 2015 toeneemt met ongeveer 15%.

Tussen lessen in grote en kleine schoolklassen ziet de Inspectie geen kwaliteitsverschillen. Ook tussen groepsgrootte en leerlingenzorg vindt de Inspectie geen verband (IvhO, 2014). De Inspectie van het Onderwijs vindt verder geen aanwijzingen dat er een optimale schaalgrootte van onderwijsinstellingen bestaat (IvhO, 2013). Tabel 3 geeft een samenvatting van de aspecten waarvoor nagenoeg geen schaaleffect is gevonden.

(Nagenoeg) Geen schaaleffect

- ✓ *Onderwijsprestaties op grote en kleine scholen;*
- ✓ *Oudertevredenheid over onderwijskwaliteit op grote en kleine scholen;*
- ✓ *Leerlingtevredenheid over onderwijskwaliteit op grote en kleine scholen;*
- ✓ *Arbeidstevredenheid van docenten op grote en kleine scholen;*
- ✓ *Leskwaliteit grote en kleine schoolklassen;*
- ✓ *Leerlingenzorg grote en kleine schoolklassen.*

Tabel 3: (Nagenoeg) Geen schaaleffect

⁸ Een kleine school is een school met 145 leerlingen of minder (conform de kleinescholentoeslag).

2.8 Hypotheses

Bij het opstellen van de onderstaande hypothesen is gebruik gemaakt van de voorgaande literatuurstudie naar schaalvoor- en schaalnadelen.

Hypothese 1 wordt onderzocht op zowel het bestuursniveau als het schoolniveau. Daarbij is het uitgangspunt dat een heel kleine organisatie hoge kosten per leerling heeft, zodat schaalvoordelen behaald kunnen worden, terwijl een grote organisatie beperkte schaalnadelen ervaart, waardoor er een MES⁹ ontstaat. Op bestuurs- en schoolniveau geldt dat bij toenemende schaalgrootte de vaste kosten over meer leerlingen gespreid kunnen worden, maar ook dat het complexer en dus duurder wordt om de organisatie aan te sturen. Op bestuursniveau wordt er gekeken naar het aantal leerlingen en het aantal scholen dat onder een bestuur valt. Op schoolniveau wordt alleen gekeken naar het aantal leerlingen dat onder de school valt.

Hypothese 1: De gemiddelde kosten per leerling zijn lager als de organisatievorm meer leerlingen heeft, maar er is een omslagpunt tussen schaalvoordelen en schaalnadelen waar te nemen (MES).

Een tweede analyse vindt plaats op groepsniveau. Het is intuïtief aannemelijk dat kleine klassen duurder zijn, omdat de kosten van een docent voor de klas over minder leerlingen te verdelen zijn. Op groepsniveau wordt daarom de leerling-docent verhouding, het aantal leerlingen per voltijd-docent, afgezet tegen de gemiddelde kosten per leerling.

Hypothese 2: Scholen met grotere klassen kennen lagere gemiddelde kosten per leerling.

Een andere interessante analyse kan plaatsvinden op het niveau van denominaties. Binnen het Nederlandse onderwijsbestel neemt het principe van vrije onderwijskeuze een belangrijke plaats in. Verschillende levensbeschouwelijke of religieuze stromingen hebben daarom eigen scholen. De belangrijkste stromingen qua leerlingenaantallen zijn de rooms-katholieke (33,8%), openbare (30,7%), protestants-christelijke (24,0%), algemeen bijzondere (5,2%), reformatorische (2,6%), gereformeerd vrijgemaakte (1,2%), samenwerking PC/RK (1,0%), antroposofische (0,9%) en islamitische (0,7%) (Stamos, 2014). Over eventuele verschillen tussen denominaties in kosten per leerling is weinig bekend. Daarbij lijkt het echter wel aannemelijk dat er verschillen zijn, omdat de leerlingpopulatie op scholen met verschillende levensbeschouwelijke richtingen sterk uiteen loopt. Islamitische scholen hebben relatief veel achterstandsleerlingen (54%) ten opzichte van openbare (15%), rooms-katholieke (13%) en protestants-christelijke (12%) scholen (Gramberg & Jordanov, 2011). Waarschijnlijk zijn de kosten per leerling bij islamitische schoolbesturen daarom hoger.

Hypothese 3: Islamitische schoolbesturen hebben gemiddeld hogere kosten per leerling, terwijl de gemiddelde kosten per leerling bij andere schoolbesturen niet significant van elkaar afwijkend zullen zijn.

⁹ MES = de minimale schaal van productie waarop de laagst mogelijke vaste kosten per eenheid product worden behaald.

3. Theoretisch Kader: Onderwijskundige Schaaffecten

3.1 Onderwijskwaliteit

In veel onderzoeken naar onderwijskwaliteit wordt de CITO-score gebruikt als indicator om de kwaliteit van een school (deels) te bepalen (*bijvoorbeeld in*: Veenstra, 1999; Blank & al, 2007; Noailly & Koning, 2009; Roeleveld & al, 2011; Veerman & al, 2013). De gemiddelde CITO-score vormt, naast andere indicatoren, een belangrijk gegeven voor de Inspectie om de kwaliteit van een school te beoordelen (Bronneman-Helmers, 2009). Daarnaast worden CITO-scores gebruikt om advies te geven over de te volgen opleiding in het voortgezet onderwijs (OC&W, 2013b). Ongeveer 85% van de basisscholen maakt gebruik van de Eindtoets Basisonderwijs van CITO (OC&W, 2012). Verder blijkt dat de CITO-score geen momentopname is, maar stabiel in de tijd en een goede voorspeller voor later succes in het onderwijs en op de arbeidsmarkt (Dronkers, 2013).

Het werken met CITO-scores als kwaliteitsindicator heeft ook zijn beperkingen. Om de toegevoegde waarde van een school goed in beeld te kunnen brengen is het bijvoorbeeld aan te bevelen om de prestaties van leerlingen over hun hele basisschoollooptijd te volgen (Blank & al, 2007). Daarnaast is de onderwijskwaliteit veel breder dan in een enkele testscore uit te drukken is. Sociale vaardigheden, culturele vorming en maatschappelijke ontwikkeling zijn bijvoorbeeld ook van belang. Wel blijkt dat dit soort aspecten van onderwijskwaliteit vaak een sterke samenhang vertonen met leerprestaties (De Moor, 2009).

Een andere veelgebruikte kwaliteitsindicator is de toegevoegde waarde van een school over de basisschoollooptijd van een leerling. De Onderwijsraad ziet de toegevoegde waarde als een te smalle indicator voor het meten van de kwaliteit van scholen, hoewel deze indicator wel kan helpen bij het verbeteren van het onderwijs (Onderwijsraad, 2014).

Het gebruik van de ruwe CITO-scores van de Basistoets Onderwijs (verder *CITO-toets* of *CITO-score*) als afhankelijke variabele lijkt een goede indicator van de onderwijskwaliteit te zijn, mits er gecorrigeerd wordt voor lokale omstandigheden als leerling- en schoolachtergrond.

3.2 Geneste structuur

Veenstra (1999) maakt een zinvol onderscheid tussen bruto- en nettoprestaties van scholen. De brutoprestatie van een school is het gemiddelde niveau dat leerlingen halen, bijvoorbeeld op de CITO-toets. De nettoprestatie van een school is het niveau dat leerlingen gemiddeld halen nadat rekening is gehouden met de samenstelling van de leerlingpopulatie van een school (Veenstra, 1999). Verschillen in brutoprestaties zijn vaak groot, maar verschillen in nettoprestaties zijn veel geringer. Om de onderwijskwaliteit van scholen op een zinvolle manier te vergelijken moeten dus de nettoprestaties van scholen bepaald worden. Een methode om dit te doen is gebruik te maken van een multilevel analyse. In zo'n analyse wordt er rekening gehouden met de geneste structuur van de sociale werkelijkheid (Van Petegem, 2005). Daarmee wordt bedoeld dat metingen niet onafhankelijk van elkaar zijn: leerlingen zitten in één klas, klassen maken deel uit van een school, die weer deel uit kan maken van een overkoepelend schoolbestuur. Om daarom de verschillen tussen scholen en leerlingen in prestaties te verklaren moet er gecorrigeerd worden voor de geneste structuur van de data.

Het gebruik van een multilevel analyse heeft een aantal voordelen ten opzichte van het gebruik van een gewone regressieanalyse (Veenstra, 1999). Door rekening te houden met de geneste structuur van gegevens zijn standaardfouten en significantietoetsen nauwkeuriger. Daarnaast kan er met een multilevel analyse onderscheid gemaakt worden tussen effecten binnen en tussen scholen. Ten slotte maakt een multilevel analyse het ook eenvoudiger om scholen te rangordenen naar nettoprestaties.

In dit onderzoek worden gebruik gemaakt van een geneste structuur van twee niveaus: het leerlingniveau en het schoolniveau. Het blijkt dat de individuele kenmerken van leerlingen de grootste invloed hebben op schoolprestaties (Veenstra, 1999). Leerlingkenmerken kunnen onderverdeeld worden in algemene leerling- en gezinskenmerken en schoolse leerling- en gezinskenmerken. Onder algemene kernmerken vallen de kenmerken die leerlingen en gezinnen buiten school om ook hebben. Onder schoolse kenmerken vallen kenmerken die specifiek zijn voor de onderwijsomgeving.

3.3 Algemene leerling- en gezinskenmerken

Leerlingkenmerken als leeftijd, geslacht, leerlinggewicht en etniciteit spelen de belangrijkste rol bij de te verwachten hoogte van de eindtoetscore (IvhO, 2013). Hoewel de verschillen beperkt zijn halen jongens gemiddeld iets hogere CITO-scores dan meisjes (OC&W, 2013b). Het blijkt dat jongens wat beter zijn in rekenen en meisjes wat beter in taal.

Verder blijkt dat jongens vaker te maken hebben met sociale problemen dan meisjes. Het gaat dan onder andere om spijbelen, schooluitval, leerproblemen, agressief gedrag, pesten en gebrekkige sociale vaardigheden (IvhO, 2013).

Verschillen in thuissituatie hebben ook invloed op de CITO-scores. Opvoedings- en opgroei-omstandigheden zijn minder gunstig voor kinderen van ouders met een lager opleidingsniveau en voor kinderen uit onvolledige gezinnen (IvhO, 2013). In een onderzoek naar correctiefactoren bij opbrengstmaten in het onderwijs bleek dat het opleidingsniveau van de ouders de belangrijkste determinant was (Roeleveld & al, 2011). Het CBS geeft verder aan dat er een sterk verband is tussen de CITO-score en het gezinsinkomen (CBS, 2013). Er zijn verschillende verklaringen voor dit verband. Gezinnen met een hogere sociaaleconomische status hebben meer mogelijkheden om hun kinderen cultureel rijk en zelfstandig op te voeden (Veenstra, 1999). Verder hebben rijke ouders meer middelen beschikbaar voor de ondersteuning van hun kind (CBS, 2013), bijvoorbeeld voor huiswerkhulp.

Er bestaat ook een verband tussen gezinssamenstelling en de CITO-score. Kinderen van gescheiden ouders scoren ruim 2 punten lager op de CITO-toets dan kinderen uit intacte gezinnen (CBS, 2013). Uit eerder onderzoek bleek ook dat opgroeien in een tweoudergezin een positief effect heeft op de wiskundescore (Veenstra, 1999).

Het geboorteland van de vader of moeder is ook van belang. Etniciteit is, na opleidingsniveau van de ouders, de belangrijkste determinant van de CITO-score in het onderzoek van Roeleveld & al (2011). Kinderen van Turkse of Marokkaanse ouders behalen significant slechtere taal- en rekenprestaties.

Verder blijkt dat als de verblijfsduur van ouders in Nederland korter is, de taalprestaties van allochtone kinderen slechter zijn (Oomens, Driessen, & Scheepers, 2003). Gecontroleerd voor sociaaleconomische status heeft etniciteit een extra effect op de prestaties van leerlingen: allochtone leerlingen scoren gemiddeld slechter dan autochtone leerlingen (Veenstra, 1999).

Met betrekking tot de algemene leerling- en gezinskenmerken kunnen we, gebaseerd op eerdere onderzoeken, concluderen dat de CITO-score relatief hoger is als:

- ✓ De leerling van het mannelijk geslacht is;
- ✓ De leerling ouders heeft met een hoog opleidingsniveau;
- ✓ De leerling autochtone ouders heeft;
- ✓ De leerling langer in Nederland verblijft;
- ✓ De leerling leeft in een tweeoudergezin;
- ✓ De leerling geen gedragsproblemen en gezondheidsproblemen heeft.

3.4 Schoolse leerling- en gezinskenmerken

De resultaten van eerder gemaakte toetsen hebben een groot voorspellend effect op de resultaten van nieuwe toetsen (Veenstra, 1999). Scoort een leerling op eerdere toetsen hoog, dan is de kans groot dat hij op nieuwe toetsen ook hoog scoort. Daarnaast blijkt dat leerlingen met een leerlinggewicht, dus leerlingen met een sociaaleconomische achterstand, gemiddeld lagere CITO-scores halen dan leerlingen zonder leerlinggewicht (IvhO, 2013).

Dat leerlingen gemotiveerd worden om hun lestage uit te voeren is belangrijk. Als leerlingen niet uitgedaagd worden, kunnen ze hun motivatie voor schooltaken verliezen (IvhO, 2014). De motivatie om schooltaken uit te voeren is een indicator voor later studiesucces (Veenstra, 1999; IvhO, 2014).

Een veilige leeromgeving voor leerlingen en docent is van groot belang. In een veilige leeromgeving is er een respectvolle omgang met elkaar en een goede relatie tussen leerlingen en volwassenen. In een veilige leeromgeving kunnen leerlingen zelfvertrouwen ontwikkelen (IvhO, 2014). Naarmate het zelfbeeld van een kind positiever is, heeft dit een sterk positief effect op zijn prestaties (Veenstra, 1999). Een positief zelfbeeld draagt bij aan de ontwikkeling van zelfstandigheid, en andersom (CJG, 2014).

Betrokkenheid van de ouders met school heeft een positief effect op de prestaties van hun kinderen (Veenstra, 1999; IvhO, 2013). Ouders die thuis aandacht geven aan school en huiswerk, geven daarmee aan dat school belangrijk voor hen is.

De werkhouding van meisjes wordt door leraren hoger gewaardeerd dan de werkhouding van jongens (OC&W, 2013b). Over het effect van verschillende andere variabelen op de CITO-score is weinig bekend, maar er kunnen wel intuïtief logische veronderstellingen gedaan worden. Als een leerling lekker in zijn vel zit en goed omgaat met leerlingen uit de klas, kan hij zich beter concentreren en tot betere prestaties komen. Het is dus belangrijk dat het algemeen welbevinden van een leerling op school hoog is. Een goede relatie met docenten en medeleerlingen draagt hieraan bij.

Met betrekking tot de schoolse leerling- en gezinskenmerken kunnen we, gebaseerd op eerdere onderzoeken en veronderstellingen, concluderen dat de CITO-score relatief hoger is als:

- ✓ De leerling taakgemotiveerd is;
- ✓ De leerling een positief zelfbeeld en een hoge mate van zelfstandigheid heeft;
- ✓ De leerling een goede werkhouding heeft.
- ✓ De leerling betrokken ouders heeft;
- ✓ De leerling een goede relatie heeft met docenten en medeleerlingen;
- ✓ De leerling op eerdere toetsen hoog scoorde.

3.5 Schoolkenmerken en hypothese

Het tweede niveau in dit onderzoek is het schoolniveau. Uit eerder onderzoek blijkt dat kenmerken op schoolniveau minder verklaringskracht hebben dan kenmerken op leerlingniveau (Veenstra, 1999). Toch worden zoveel mogelijk controlevariabelen op schoolniveau meegenomen. De speciale aandacht gaat uit naar variabelen die met schaalgrootte te maken hebben.

De afgelopen jaren is gebleken dat er verschillen zijn in onderwijskwaliteit tussen de provincies. Deze verschillen zijn de afgelopen jaren wel kleiner geworden. Scholen in Friesland, Groningen en Drenthe hebben hun onderwijskwaliteit de afgelopen jaren verbeterd. Limburg en Zeeland doen het iets beter dan het landelijk gemiddelde, Flevoland doet het minder goed dan gemiddeld (IvhO, 2014).

Volgens de Inspectie is er in het basisonderwijs niet direct een relatie te leggen tussen stedelijkheid en onderwijskwaliteit. De Inspectie (2014) constateert wel dat vo-scholen in verstedelijkte gebieden relatief vaker moeite hebben om de onderwijskwaliteit stabiel te houden, maar het effect is klein. Noailly & Koning (2009) constateren wel een significant negatieve invloed van de mate van verstedelijking op de CITO-scores; leerlingen op scholen in een meer verstedelijkt gebied halen een iets hogere CITO-score dan leerlingen op scholen in minder verstedelijkt gebied.

Scholen met een hoog aantal leerlingen met leerlinggewichten krijgen een hoog schoolgewicht toegekend. Scholen met een hoog schoolgewicht hebben te maken met een moeilijke populatie leerlingen. Het lijkt dus aannemelijk dat scholen met een hoog schoolgewicht lagere gemiddelde CITO-scores halen. Dit beeld wordt versterkt door het feit dat het ministerie van OC&W extra gelden uitkeert aan scholen op basis van schoolgewicht om onderwijsachterstanden weg te werken (OC&W, 2012). Uit ervaringen uit onderwijsland blijkt dat een concentratie van zorgleerlingen op school veel tijd en energie vraagt van het personeel (Visser, 2005). Een hoger aandeel aandachtsleerlingen heeft een significant negatief effect op de gemiddelde CITO-score (Noailly & Koning, 2009).

Het effect van denominatie is, gecontroleerd voor andere variabelen, beperkt. De Inspectie (2014) constateert dat er nauwelijks verschillen zijn tussen denominaties als het gaat om toezichtarrangementen. De islamitische scholen zijn wel vaker zwak.

Vanaf het schooljaar 2008/2009 zijn verschillende scholen aan de slag gegaan met trajecten om hun taal- en/of rekenonderwijs te verbeteren. Uit onderzoek van de Inspectie blijkt dat scholen die verbetertrajecten hebben opgestart in het schooljaar 2008/2009, gemiddeld significant hogere CITO-scores halen in 2010/2011 ten opzichte van 2008/2009. Verder is de stijging in CITO-score op deze scholen in deze periode groter dan op scholen die niet hebben meegedaan aan de verbetertrajecten (IvhO, 2012).

De heterogeniteit van de leerlingenpopulatie op een school speelt ook een rol (IvhO, 2013). Het effect van het percentage allochtone leerlingen op school op de prestaties van leerlingen is klein. Voor het taalonderwijs geldt een negatief effect als er een hoger percentage allochtone leerlingen is, voor het rekenonderwijs is het effect groepsafhankelijk. Verder blijkt dat allochtone leerlingen een groot deel van hun achterstand op autochtone leerlingen inhalen gedurende de basisschooltijd, maar dat er altijd een bepaalde achterstand blijft op autochtone leerlingen. Uit eerder onderzoek bleek ook dat leerlingen op scholen met veel allochtonen slechter scoren dan op scholen met weinig allochtonen (Veenstra, 1999). In een recent onderzoek is dit beeld bevestigd. Een grotere proportie aan allochtone leerlingen in een klas zorgt voor lagere prestaties van de autochtone leerlingen (Veerman, Van de Werfhorst, & Dronkers, 2013). Etnische diversiteit gemeten in het aantal etnische groepen blijkt geen aanvullende invloed te hebben op de prestaties van autochtone leerlingen (Karssen, Van der Veen, & Roeleveld, 2011), hoewel Veerman & al (2013) op enkele deelgebieden zoals het taalonderwijs negatieve effecten waarneemt van etnische diversiteit.

De overige variabelen hebben te maken met schaaffecten. Een uitgebreide bespreking van schaaffecten is te vinden in de paragrafen 2.5 – 2.7. Uit de literatuurstudie rijst een tegenstrijdig beeld op. Er zijn zowel studies waarin kleinere scholen een betere onderwijskwaliteit behalen als studies waarin grotere scholen een betere onderwijskwaliteit behalen. De Inspectie van het Onderwijs geeft aan dat kleine scholen (tot 100 leerlingen) hun achterstand op grotere scholen inlopen, maar dat grotere scholen op eindopbrengsten nog iets beter scoren (IvhO, 2014).

Uit de analyse van Noailly en Koning (2009) blijkt dat scholen die onder een groot schoolbestuur vallen gemiddeld lagere CITO-scores hebben. De Inspectie constateert ook dat grote schoolbesturen vaker een risico lopen op het vlak van onderwijskwaliteit (IvhO, 2014).

Effecten met betrekking tot het personeel worden hier wel besproken maar niet meegenomen in de hypothese. De reden is dat de verwachte effecten op de CITO-score zeer klein zullen zijn. De besproken variabelen worden wel meegenomen in de modellen om mogelijke effecten te ontdekken of uit te sluiten. De gemiddelde leeftijd van het onderwijspersoneel is in de afgelopen jaren toegenomen. Het aandeel 55-plussers is gestegen van 19,1% in 2008 tot 23,1% in 2012, terwijl het aandeel docenten jonger dan 35 gedaald is van 35,1% naar 30,9% (OC&W, 2013b). De Inspectie (2014) neemt waar dat beginnende docenten én docenten met meer dan 20 jaar ervaring gemiddeld minder goede didactische- en differentiatievaardigheden hebben. Beginnende docenten hebben een gebrek aan ervaring, terwijl docenten met veel ervaring soms 'in een sleur' komen.

Leraren met 3 tot 20 jaar ervaring scoren het best op de genoemde vaardigheden. De Inspectie (2014) geeft verder aan dat deeltijdsleraren de algemene didactische en differentiatievaardigheden minder goed beheersen dan voltijdsleraren.

Hoe beter de schoolleider, hoe beter de lessen op een school. Daarbij blijkt dat vrouwen betere schoolleiders zijn dan mannen. Verklaringen kunnen zijn dat de groep vrouwen die schoolleider wordt selecter is en dat vrouwen op een andere manier leidinggeven (IvhO, 2014). In het basisonderwijs wordt geen verschil waargenomen tussen vrouwelijke en mannelijke docenten, maar in het voortgezet onderwijs tonen vrouwelijke docenten vaker dat ze algemene didactische vaardigheden beheersen. Ook blijkt dat vrouwelijke docenten een betere relatie onderhouden met zowel jongens als meisjes in de klas (ScienceGuide, 2012).

Over de relatie tussen het aandeel management, onderwijzend personeel en onderwijs ondersteunend personeel enerzijds en de CITO-scores anderzijds is nagenoeg niets bekend.

Hypothese 4

Met betrekking tot de schoolkenmerken kunnen we, gebaseerd op eerdere onderzoeken, concluderen dat *de CITO-score relatief hoger is als:*

- ✓ *De leerling op een school in Zeeland of Limburg zit;*
- ✓ *De leerling op een school in een meer verstedelijkt gebied zit;*
- ✓ *De leerling op een school zit met een laag schoolgewicht;*
- ✓ *De leerling niet op een islamitische school zit;*
- ✓ *De leerling op een school zit waar taal- en rekenverbetertrajecten zijn ingezet;*
- ✓ *De leerling op een school zit met een homogene leerlingpopulatie;*
- ✓ *De leerling op een school zit groter dan 145 leerlingen;*
- ✓ *De leerling op een school zit die valt onder een klein schoolbestuur.*

4. Data & Methodologie: Financiële Schaafeffecten

Om de hypotheses te toetsen zijn gegevens van verschillende informatiebronnen verzameld. Met betrekking tot het financiële deel van de probleemstelling zijn alle gegevens afkomstig van de Dienst Uitvoering Onderwijs (DUO). Van alle bestanden is gebruik gemaakt van de versie 1 oktober 2012 (*of* schooljaar 2012-2013). Aangezien uitgebreide gegevens over kosten alleen beschikbaar zijn op het bestuursniveau (het niveau van het bevoegd gezag) en helaas niet op het lagere schoolniveau, zijn alle gegevens verzameld op bestuursniveau. De uitspraken op bestuursniveau zijn betrouwbaar, maar uitspraken op schoolniveau en met name groepsniveau zijn gemiddelden en hierdoor minder betrouwbaar als de heterogeniteit binnen een bestuur groot is. In bijlage 1 zijn alle variabelen weergegeven die gebruikt zijn voor dit deel van het onderzoek, met daarachter het bronbestand.

4.1 Dataselectie

Nadat alle beschikbare gegevens waren verzameld, is dataselectie toegepast. Als basis is gekozen voor het bestand met leerlinggegevens (DUO, 2014a). Van 1170 besturen was informatie beschikbaar over de leerlingaantallen op 1 oktober 2012. Uit dit bestand werd informatie verkregen over de denominatie en de leerlingaantallen op het basisonderwijs, het speciaal basisonderwijs, het speciaal onderwijs en het voortgezet onderwijs. Het totaal aantal leerlingen op 1 oktober 2012 is berekend als de som van de hiervoor genoemde categorieën leerlingaantallen. Deze gegevens zijn vervolgens gekoppeld aan het bestand van de personeelsgegevens (DUO, 2014b). Uit dit bestand werd informatie verkregen over het aantal FTE's aan managementfuncties, leraren in opleiding, onderwijzend personeel, onderwijs ondersteunend personeel en het totaal aantal FTE's op 31 december 2012. Van 1163 besturen bleek personeelsinformatie beschikbaar over het jaar 2012. Van sommige besturen bleek wel personeelsinformatie maar geen leerlingeninformatie (1) of andersom (8) beschikbaar. Van 1162 besturen was complete leerling- en personeelsinformatie beschikbaar.

Deze gegevens zijn vervolgens gekoppeld aan het bestand van de lastengegevens (DUO, 2014c). Uit dit bestand werd informatie verkregen over de personeelslasten en de huisvestingslasten. Alle schoolbesturen waarvan deze informatie niet beschikbaar was, zijn verwijderd uit het databestand. Uiteindelijk bleek van 1136 besturen alle benodigde informatie beschikbaar ($\approx 97\%$). Aan de hand van de beschikbare gegevens zijn nog een aantal aanvullende variabelen gecreëerd: *de gemiddelde vaste kosten per leerling, per FTE en per school* en de *leerling-docent verhouding* (zie bijlage 1).

De data is vervolgens geselecteerd op categorie primair onderwijs. In dit onderzoek wordt alleen aandacht besteed aan het (reguliere) basisonderwijs, omdat andere categorieën primair onderwijs kostentechnisch niet te vergelijken zijn met het gewone basisonderwijs. Daarom zijn alleen schoolbesturen geselecteerd die geen leerlingen hebben in het speciaal basisonderwijs, het speciaal onderwijs en het speciaal voortgezet onderwijs. Uiteindelijk bleven 847 schoolbesturen¹⁰ ($\approx 72\%$) over die alleen leerlingen in het regulier basisonderwijs hebben. Dit is de *uitgebreide subgroep*, die gebruikt is voor alle analyses met betrekking tot hypotheses 1 en 2.

¹⁰ Een uitschieter (de Stichting Joods Bijzonder Onderwijs) is buiten beschouwing gelaten: zie bijlage 2.

Voor hypothese 3 is nog een selectie toegepast. Alle denominaties met minder dan 10 schoolbesturen (zie bijlage 3) zijn uit het onderzoek gelaten, zodat de categorieën openbaar, protestants-christelijk, rooms-katholiek, algemeen bijzonder, reformatorisch, antroposofisch, islamitisch en samenwerking PC-RK overbleven. De *bepaalde subgroep* bevat 790 schoolbesturen ($\approx 68\%$).

In het onderzoek naar schaafeffecten gaat het om vaste kosten. Als maat voor de vaste kosten is gekozen voor de categorie *Personeelslasten*. Het is te bepalen dat bijna 95% van het personeel in dienst van de basisscholen in 2012 een contract voor onbepaalde tijd heeft en dus in vaste dienst is (DUO, 2014b). De personeelslasten kunnen dus als vaste kosten worden beschouwd. Daarnaast blijkt uit de jaarrekeningen over 2012 dat ongeveer 82% van alle lasten personeelslasten zijn (DUO, 2014d). Personeelslasten zijn verreweg de belangrijkste categorie lasten voor een school.

Huisvestingslasten zijn in de analyses uiteindelijk buiten beschouwing gelaten. Huisvesting kan namelijk rechtstreeks geregeld worden vanuit gemeentefondsen, maar ook verlopen via de scholen zelf (Blank & al, 2007). Dit leidt tot verschillende uitgangspunten, waardoor een vergelijking niet eerlijk zou verlopen. Huisvestingskosten worden om deze reden vaker niet meegenomen in de analyse, zoals bijvoorbeeld in Blank & al (2007).

4.2 Hypotheses toetsen

Om de eerste hypothese te toetsen op bestuursniveau zijn de gemiddelde vaste kosten per leerling afgezet tegen het aantal leerlingen per schoolbestuur om het effect van bestuursgrootte in leerlingenaantallen te bepalen. Ook zijn de gemiddelde vaste kosten per leerling afgezet tegen het aantal scholen per bestuur om het effect van bestuursgrootte gemeten in het aantal scholen te bepalen.

Om de eerste hypothese te toetsen op schoolniveau zijn alle schoolbesturen geselecteerd die één school besturen. Als gevolg daarvan bleven er 450 schoolbesturen ($\approx 38\%$) over. Vervolgens zijn de gemiddelde vaste kosten per leerling afgezet tegen het aantal leerlingen per school om het effect van schoolgrootte gemeten in het aantal leerlingen te bepalen.

Om de tweede hypothese te toetsen is van elk schoolbestuur de leerling-docentverhouding uitgerekend en vervolgens afgezet tegen de gemiddelde vaste kosten per leerling om het schaafeffect in de vorm van groeps grootte te bepalen.

Om de derde hypothese te toetsen is gebruik gemaakt van de bepaalde subgroep van 790 schoolbesturen. De gemiddelde vaste kosten per leerling zijn afgezet tegen de leerlingaantallen van verschillende denominaties om het effect van levensbeschouwelijke richting te bepalen.

Om de significantie te bepalen, wordt in het onderzoek naar financiële schaafeffecten (tenzij anders vermeld) gebruik gemaakt van een *one sample t-test*, waarin het gemiddelde van de hele steekproef vergeleken wordt met gemiddeldes van verschillende subgroepen. In tabellen wordt met *** aangegeven dat een gegeven significant is op het 1%-niveau, met ** op het 5%-niveau en met * op het 10%-niveau. Geen aanduiding betekent geen significant gegeven.

5. Data & Methodologie: Onderwijskundige Schaaffecten

5.1 Data COOL-cohortonderzoek

Om hypothese 4 te toetsen zijn gegevens van verschillende informatiebronnen verzameld. De belangrijkste bron voor het onderzoek naar onderwijskundige schaaffecten was de studie *Cohort Onderzoek Onderwijsloopbanen*, ook wel afgekort tot COOL. Er is gebruik gemaakt van de gegevens die verzameld zijn op basisscholen. Deze dataverzameling is een samenwerkingsproject van het ITS te Nijmegen en het Kohnstamm Instituut te Amsterdam. Alle variabelen met betrekking tot het leerlingniveau (zie bijlage 4) zijn afkomstig uit het tweede COOL-cohort, afgenomen in het schooljaar 2010/2011 (COOL, 2012). Ook diverse schoolvariabelen zijn afkomstig uit deze cohortstudie (zie bijlage 7).

De technische gegevens over het onderzoek zijn vastgelegd in een achtergrondstudie (Driessen, Mulder, & Roeleveld, 2012). In het vervolg van deze paragraaf wordt een korte beschrijving gegeven van het COOL-onderzoek gebaseerd op dit bestand. Voor een uitgebreide beschrijving van het onderzoek en de variabelen kan het genoemde bestand worden geraadpleegd.

De totale steekproef van COOL bestaat uit 553 scholen, waarvan 406 in de referentiesteekproef en 147 in de aanvullende steekproef. De referentiesteekproef als geheel is een verzameling scholen die representatief is voor Nederland. Scholen in de aanvullende steekproef worden gekenmerkt door een hoge concentratie leerlingen uit lagere sociaaleconomische milieus met zowel allochtone als autochtone achtergrond. Deze scholen zijn aan het onderzoek toegevoegd om ook steekhoudende uitspraken te kunnen doen over scholen met een uiteenlopende sociaal-etnische leerlingenpopulatie. In de totale steekproef zijn ook scholen met onderwijsconcepten opgenomen, maar de spreiding van verschillende onderwijsconcepten in de totale steekproef is niet conform de landelijke verdeling. Zo zijn er relatief veel Daltonscholen terwijl de vrijescholen nagenoeg ontbreken. Dit gebrek zorgt ervoor dat onderwijsconcepten niet worden meegenomen in dit onderzoek.

De COOL-gegevens zijn verzameld over het schooljaar 2010-2011. Scholen zijn benaderd om deel te nemen aan het onderzoek van april 2010 tot september 2010. Achtergrondgegevens van leerlingen zijn verzameld in de periode september 2010 tot december 2010. Toetsafnames en andere dataverzamelingen vonden plaats in de periode januari – april 2011. In de laatste fase van het onderzoek, in mei en juni 2011, zijn gegevens met betrekking tot zorgprofielen, uitstroomgegevens van groep 8 en schoolinformatie verzameld.

Er zijn op de deelnemende scholen gegevens verzameld voor leerlingen in groep 2, 5 en 8. In dit onderzoek worden alleen de gegevens van leerlingen uit groep 8 gebruikt. In totaal zijn er gegevens van 12.538 leerlingen uit groep 8 bekend. Van deze 12.538 leerlingen zijn er van 6.692 leerlingen ($\approx 53\%$) gegevens met betrekking tot de CITO-toets bekend. In dit onderzoek worden alleen de gegevens gebruikt van deze leerlingen. Deze gegevens zijn niet helemaal representatief voor het Nederlandse basisonderwijs, aangezien een groter aandeel allochtone leerlingen de CITO-toets heeft gemaakt. De schoolgegevens die zijn gebruikt uit het COOL-onderzoek zijn te vinden in bijlage 7.

5.2 Dataselectie leerlinggegevens

In bijlage 5 wordt een korte samenvatting van de onderzoeksgegevens gegeven bij aanvang van de analyse, uitgaande van de leerlingen waarvan gegevens met betrekking tot de CITO-toets bekend waren (N=6.692). Uit deze tabel valt op dat er bij elke variabele gegevens missen. Hier diende een beslissing gemaakt te worden over de opzet van de modellen en de toevoeging van de controlevariabelen. Uiteindelijk is ervoor gekozen om geen dataselectie toe te passen. Dat wil zeggen dat alle gegevens gebruikt zijn en dat missende gegevens niet op voorhand uit de dataset verwijderd zijn. Deze keuze heeft als voordeel dat de steekproef waarop de modellen gebaseerd zijn zo groot mogelijk blijft. Anderzijds heeft deze keuze als nadeel dat, als het aantal waarnemingen waarop modellen gebaseerd zijn verandert, onderlinge modellen niet goed vergelijkbaar zijn. Bijlage 6 geeft een samenvatting van leerlinggegevens (N=3.405) op het moment van de definitieve modellen (M₃₃-M₃₆). De weergegeven samenvattingen in bijlage 5 (meeste waarnemingen) en bijlage 6 (minste waarnemingen) zijn dus uitersten.

Een aantal variabelen is aangepast ten opzichte van de opzet in het COOL-cohortonderzoek, met name om ervoor te zorgen dat het aantal waarnemingen per categorie groter werd. Zo is van de variabele *gezinsamenstelling* de indeling aangepast van vier categorieën (compleet, woont bij vader, woont bij moeder, overig) naar twee categorieën (compleet, eenoudergezin). Voor de variabelen *geboorteland van de vader* en *geboorteland van de moeder* zijn landen met weinig waarnemingen samengenomen in de categorie overig. Dit resulteerde uiteindelijk in een vijftal categorieën (Nederland, Turkije, Marokko, Suriname en de voormalige Antillen, overig). De variabele *verblijfsduur in Nederland* is aangepast naar twee categorieën (verblijfsduur 5 jaar of langer, verblijfsduur korter dan 5 jaar).

5.3 Data DUO

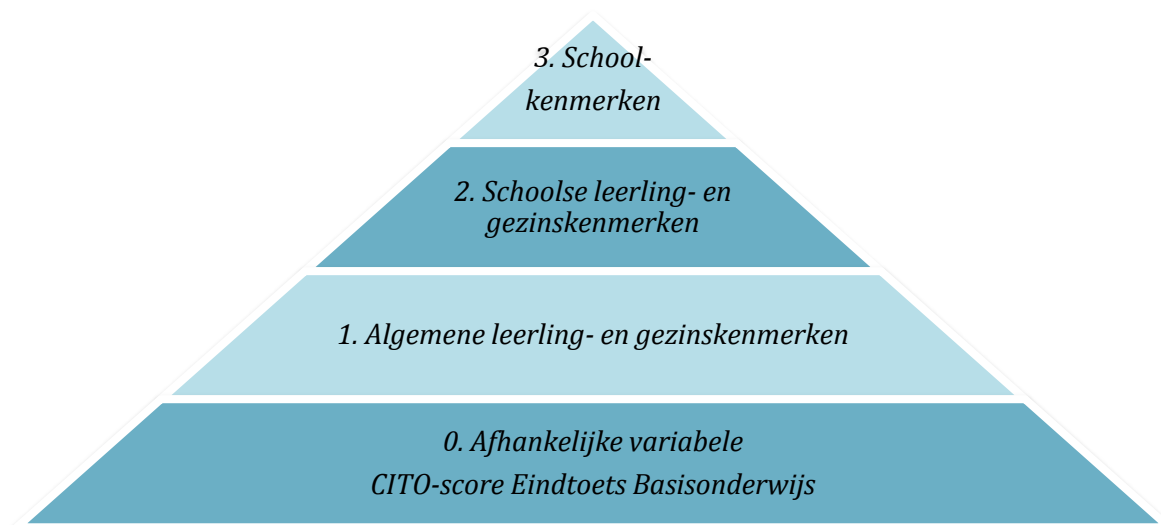
Verschillende schoolgegevens (zie bijlage 7) zijn verkregen via databestanden van het DUO. Van alle bestanden is gebruik gemaakt van de versie 1 oktober 2010 (of schooljaar 2010-2011). De denominatie van de school en het totaal aantal leerlingen (DUO, 2010a) en de gemiddelde leeftijd van het personeel (DUO, 2014e) zijn direct overgenomen. Op basis van andere DUO-gegevens zijn de volgende berekeningen uitgevoerd: de verhouding tussen vast en tijdelijk personeel, het aandeel FTE's aan vrouwen, aan management en aan deeltijdpersoneel (0-0.5FTE) op het totaal aantal FTE (DUO, 2014e) en het aantal scholen per bevoegd gezag (DUO, 2010a). In de bijlagen 8 en 9 wordt een korte samenvatting gegeven van de gegevens van scholen die de CITO-toets afnemen bij aanvang (N=318) respectievelijk bij de definitieve modellen M₃₃-M₃₆ (N=213).

5.4 Dataselectie schoolgegevens

Op basis van de COOL-gegevens zijn nog een tweetal variabelen samengesteld: het *aantal etnische groepen* en het *percentage allochtonen* (in groep 8). Het aantal etnische groepen is berekend aan de hand van het aantal etnische groepen aan vaders en moeders kant. Van één van beide uitkomsten is het hoogste getal genomen, zodat de variabele *aantal etnische groepen* het maximumaantal etnische groepen in de klas aangeeft. De variabele *percentage allochtonen* is berekend aan de hand van het aantal leerlingen met een andere dan Nederlandse afkomst voor vaders en moeders kant. Ook van deze variabele is het hoogste getal genomen, zodat de variabele *percentage allochtonen* het maximumpercentage allochtonen in de klas aangeeft.

5.5 Methodologisch Model

Zoals gezegd vindt het toetsen van de hypothesen plaats door middel van een multilevel analyse. In de analyse worden steeds controlevariabelen toegevoegd om te kijken of de nieuwe gegevens verklaringskracht aan het model toevoegen. De variabelen worden toegevoegd op basis van een methodologisch model gebaseerd op verschillende stappen. De basis van het model wordt gevormd door de afhankelijke variabele, de behaalde *CITO-score*. Tijdens de eerste stap worden algemene leerling- en gezinskenmerken toegevoegd, daarna schoolse leerling- en gezinskenmerken en als laatste stap worden schoolkenmerken toegevoegd. De opbouw van het methodologisch model is schematisch weergegeven in afbeelding 1.



Afbeelding 1: Methodologisch Model

De algemene leerling- en gezinskenmerken die stap voor stap zullen worden toegevoegd aan het model zijn te vinden in de linkerkolom van tabel 4. Te denken valt aan het geslacht, het opleidingsniveau en het geboorteland van beide ouders, de gezinssituatie, de verblijfsduur, de sociaaleconomische status (SES) en het gedrag dat de leerling vertoont. In de rechterkolom van tabel 4 staan de schoolse leerling- en gezinskenmerken, die daarna stap voor stap zullen worden toegevoegd aan het model. Hier valt te denken aan de taakmotivatie, het zelfvertrouwen op school, de werkhouding, de populariteit in de klas, de afhankelijkheid van de docent, de ouderbetrokkenheid met school, het welbevinden met medeleerlingen en de docent en resultaten van een eerder gemaakte taaltoets en rekentoets. De schaalvariabelen *gedrag*, *werkhouding*, *populariteit*, *afhankelijkheid* en *ouderbetrokkenheid* zijn bepaald door vragenlijsten ingevuld door de docent. De schaalvariabelen *taakmotivatie*, *zelfvertrouwen* en *welbevinden* zijn bepaald door vragenlijsten ingevuld door de leerlingen zelf. Een uitgebreide beschrijving van de genoemde variabelen en de manier waarop deze variabelen tot stand gekomen zijn, is te vinden in beschrijving van het COOL-onderzoek (Driessen, Mulder, & Roeleveld, 2012). Een beknopte beschrijving van de variabelen in tabel 4 is te vinden in bijlage 4.

Individuele kenmerken	
Algemene kenmerken	Schoolse kenmerken
<ul style="list-style-type: none"> ✓ Geslacht ✓ Opleidingsniveau vader ✓ Opleidingsniveau moeder ✓ Geboorteland vader ✓ Geboorteland moeder ✓ Gezinsamenstelling ✓ Verblijfsduur in Nederland ✓ Sociaaleconomische status (SES) gezin ✓ Gedrag ✓ Indicatie zorgleerling 	<ul style="list-style-type: none"> ✓ Taakmotivatie ✓ Zelfvertrouwen ✓ Werkhouding ✓ Populariteit in de klas ✓ Afhankelijkheid van de docent ✓ Ouderbetrokkenheid ✓ Welbevinden met medeleerlingen ✓ Welbevinden met docent ✓ Toets Rekenen / Wiskunde LOVS ✓ Toets Taal DMT-2009

Tabel 4: Algemene en schoolse leerling- en gezinskenmerken

Nadat de leerlingkenmerken zijn toegevoegd aan het model worden nog schoolkenmerken toegevoegd. In tabel 5 staan de variabelen op schoolniveau, die meegenomen worden in dit onderzoek. Een beknopte beschrijving van deze variabelen is te vinden in bijlage 7.

Schoolkenmerken
<ul style="list-style-type: none"> ✓ Provincie ✓ Mate van stedelijkheid ✓ Schoolgewichtscore ✓ Denominatie ✓ Percentage allochtonen >50% & >75% ✓ Aantal etnische groepen ✓ Aantal leerlingen <=145 ✓ Aantal leerlingen >=500 ✓ Aantal scholen per schoolbestuur: 6-10, 11-20 & >20 ✓ Aandeel deeltijd-FTE (0-0.5) op totaal aantal FTE ✓ Aandeel vrouwen-FTE op totaal aantal FTE ✓ Aandeel management-FTE op totaal aantal FTE ✓ Gemiddelde leeftijd van het personeel ✓ Verhouding tussen vast en tijdelijk personeel

Tabel 5: Schoolkenmerken

5.6 Hypothese toetsen

Hypothese 4 is opgesteld naar aanleiding van de literatuurstudie naar het effect van schoolkenmerken op de CITO-score van leerlingen. Voordat deze hypothese getoetst kan worden, zullen de variabelen met betrekking tot de algemene leerling- en gezinskenmerken en de schoolse leerling- en gezinskenmerken aan het basismodel worden toegevoegd. Het basismodel wordt dus eerst stap voor stap uitgebreid met de genoemde kenmerken uit tabel 4 en vervolgens met de kenmerken uit tabel 5. Als alle kenmerken van een bepaalde stap uit het methodologisch model zijn toegevoegd wordt gekeken of het model tot dan toe voldoet aan de gestelde verwachtingen. Bij het toetsen van hypothese 4 wordt uiteindelijk gekeken welke waarde de schoolkenmerken voor het model hebben, gecontroleerd voor de algemene en schoolse leerlingkenmerken.

Alle modellen zijn geschat volgens de *Maximum Likelihood* methode, zodat de *Log Likelihood* scores van verschillende modellen onderling vergelijkbaar zijn. Deze scores zijn onderling vergelijkbaar als het model hetzelfde aantal waarnemingen bevat en als model Y voortbouwt op ten minste dezelfde variabelen als model X. Doordat het aantal waarnemingen (met name in de beginmodellen) fluctueert, zijn de *Log Likelihood* scores niet altijd vergelijkbaar. In het algemeen geldt dat hoe groter de *Log Likelihood* score is (hoe dichter de score bij nul ligt), des te meer variantie er door het model verklaard wordt. De vergelijkingsmaat *AIC* is tevens bij elk model weergegeven. Ook deze maat vereist een constant aantal waarnemingen. De *AIC* geeft aan hoe goed een bepaald model bij de data past, waarbij geldt dat een kleiner getal (dichter bij nul) een beter passend model aangeeft. Daarbij corrigeert de *AIC* voor het aantal variabelen dat wordt toegevoegd.

Elke tabel met modellen is opgebouwd volgens hetzelfde systeem. De kolom *Constante* geeft aan wat de constante is van het model en symboliseert dus de gemiddelde CITO-score, gecorrigeerd voor bepaalde *fixed* en *random* effecten. In het geval van een *fixed* effect wordt verondersteld dat het gevonden effect voor de hele steekproef hetzelfde is en dat de gevonden coëfficiënt daarom voor alle gevallen in de dataset geldt. In het geval van een *random* effect kan de waarde van de parameter variëren tussen verschillende groepen (bijvoorbeeld tussen verschillende scholen) (Field, 2009). De kolom *var(Residual)* geeft de unieke individuele variatie in CITO-score aan die niet verklaard kan worden door de onafhankelijke variabelen of door variabelen van een hoger niveau (zoals het schoolniveau) (Verboon, 2012). Hoe kleiner de term *var(Residual)* is hoe beter, want dat betekent dat een groter deel van de leerlingvariantie verklaard kan worden. In latere modellen wordt een random-effect op het schoolniveau toegevoegd, weergegeven met de term *var(School)*. Deze term geeft aan hoeveel variantie op het schoolniveau verklaard wordt, dat wil zeggen, hoeveel scholen onderling verschillen. Met *** respectievelijk ** en * wordt aangegeven dat een gegeven significant is op het 1%-, 5%- en 10%-niveau. Geen aanduiding betekent geen significant gegeven. Onder elke coëfficiënt is tussen haakjes en cursief ook de *standard error* vermeld.

Om te testen of een model met meer variabelen de onafhankelijke variabele beter verklaart, kun je gebruik maken van een *likelihood ratio test* (LR-test). Deze test vergelijkt de verschillende *log likelihoods* van een model en bekijkt of deze significant van elkaar verschillen (Torres-Reyna, 2014). De LR-test vereist een stabiel aantal waarnemingen bij verschillende modellen, dus het doen van deze test is, gegeven het soms veranderend aantal waarnemingen, niet altijd mogelijk. Gedurende de analyse wordt de LR-test af en toe uitgevoerd om te kijken of het nieuwe model significant beter is dan eerdere modellen.

Om te bepalen of het toevoegen van een random effect aan een model zinvol is, kan de *intraclass correlation coëfficiënt* (IC) worden bepaald (Torres-Reyna, 2014). De IC geeft in dit onderzoek aan wat het schooleffect is in het verklaren van de variantie, oftewel, wat de tussenschoolse variantie is. De IC-maatstaf varieert tussen 0 en 1. Als de maatstaf richting 0 gaat is het toevoegen van de random effecten zinloos. De IC wordt als volgt berekend:

$$\text{Intraclass Correlation} = \frac{\text{var}(\text{School})}{\text{var}(\text{School}) + \text{var}(\text{Residual})}$$

6. Resultaten: Financiële Schaafeffecten

6.1 Algemene gegevens uitgebreide subgroep

In tabel 6 staan enkele algemene gegevens over de uitgebreide subgroep (N=847) met betrekking tot het onderzoek naar financiële schaafeffecten. Hieruit blijkt onder andere dat er gemiddeld 906 leerlingen vallen onder één schoolbestuur. De schoolbesturen in de uitgebreide subgroep zijn dus bijna even groot als het landelijk gemiddelde van 909 leerlingen (OC&W, 2013b). Het kleinste schoolbestuur telt slechts 9 leerlingen, terwijl het grootste schoolbestuur de leiding heeft over 7.282 leerlingen. Een gemiddeld schoolbestuur telt voor 58.8 FTE's aan personeelsleden. Hiervan bestaat 5.2 FTE (8,8%) uit managementfuncties, 47.6 FTE (81,0%) uit docerende functies en 5.8 FTE (9,9%) uit onderwijs ondersteunende functies. Gemiddeld heeft een schoolbestuur per 19.3 leerlingen 1 leraar in voltijddienst, maar dit loopt uiteen van 9.1 tot 30 leerlingen per leraar. Gemiddeld bestaat een schoolbestuur uit 4.4 scholen. De gemiddelde vaste kosten per leerling lopen uiteen van € 2.726 tot € 10.539. Het gemiddelde ligt op € 4.238. Dit komt redelijk overeen met de gemiddelde personeelskosten van € 4.350 per leerling die het ministerie over 2012 heeft berekend¹¹ (OC&W, 2013b). De maatstaf *kosten per leerling* kent in vergelijking met de maatstaven *kosten per FTE* en *kosten per school* de minste spreiding en is waarschijnlijk de meest betrouwbare vergelijkingsmaat.

Variabele	Gemiddeld	Minimum	Maximum
Aantal leerlingen	906	9	7282
Totaal aantal FTE's	58,8	1,3	523
FTE's Managers	5,2	0	41,7
FTE's Leraren in opleiding	0,04	0	5,5
FTE's Leraren	47,6	0	417,3
FTE's Onderwijs ondersteunend	5,8	0	77,9
Leerlingen per FTE Leraren	19,3	9,1	30
Scholen per bevoegd gezag	4,4	1	30
Kosten per leerling	€ 4.238	€ 2.726	€ 10.539
Kosten per FTE	€ 63.879	€ 31.442	€ 146.013
Kosten per school	€ 904.429	€ 40.875	€ 4.306.115

Tabel 6: Algemene gegevens uitgebreide subgroep (N=847)

6.2 Schaafeffecten op bestuursniveau – aantal leerlingen

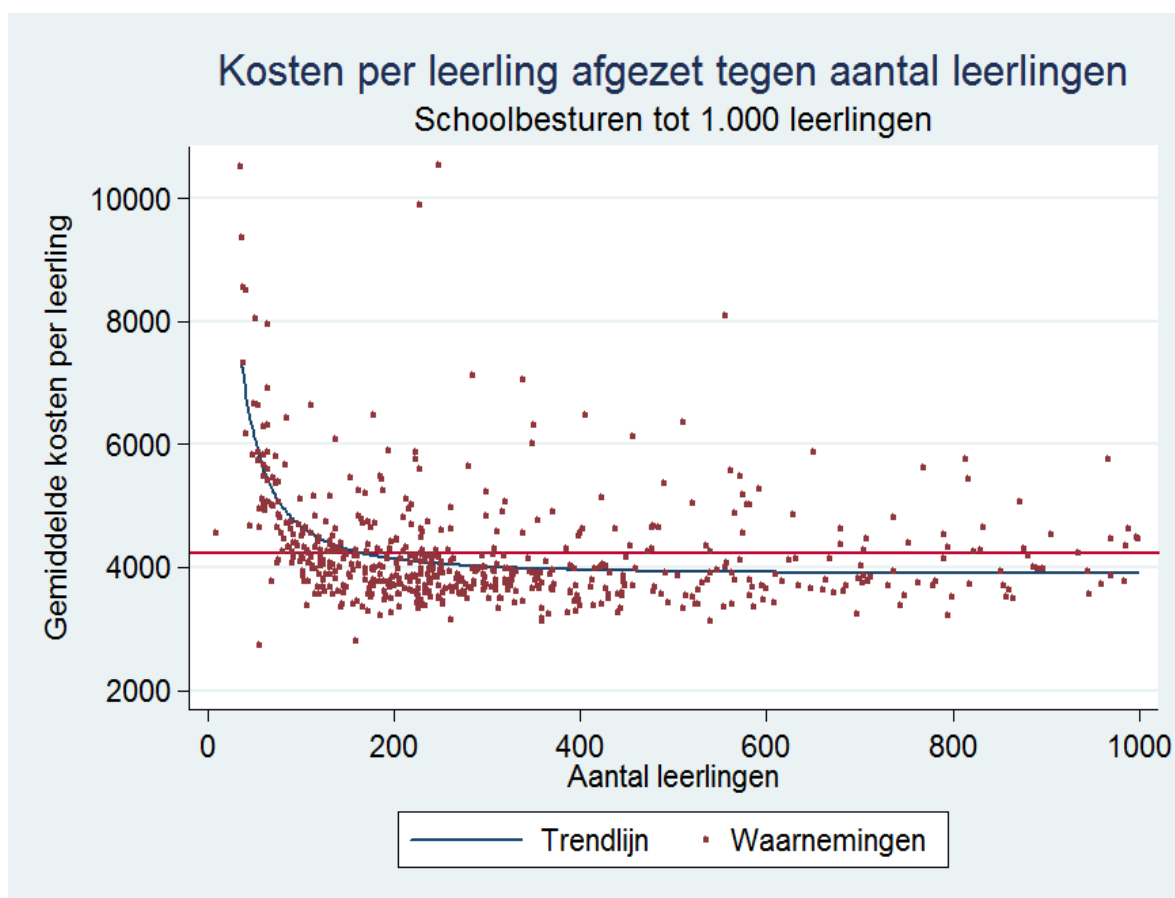
Hypothese 1: De gemiddelde kosten per leerling zijn lager als de organisatievorm meer leerlingen heeft, maar er is een omslagpunt tussen schaalvoordelen en schaalnadelen waar te nemen (MES).

Hypothese 1 wordt eerst getest op het bestuursniveau gemeten in het aantal leerlingen per bestuur. De gemiddelde vaste kosten per leerling zijn afgezet tegen het aantal leerlingen per schoolbestuur. Daarvan zijn twee grafieken gemaakt.

¹¹ $\frac{8.312}{9.555} * 5.000 = 4.350$ euro

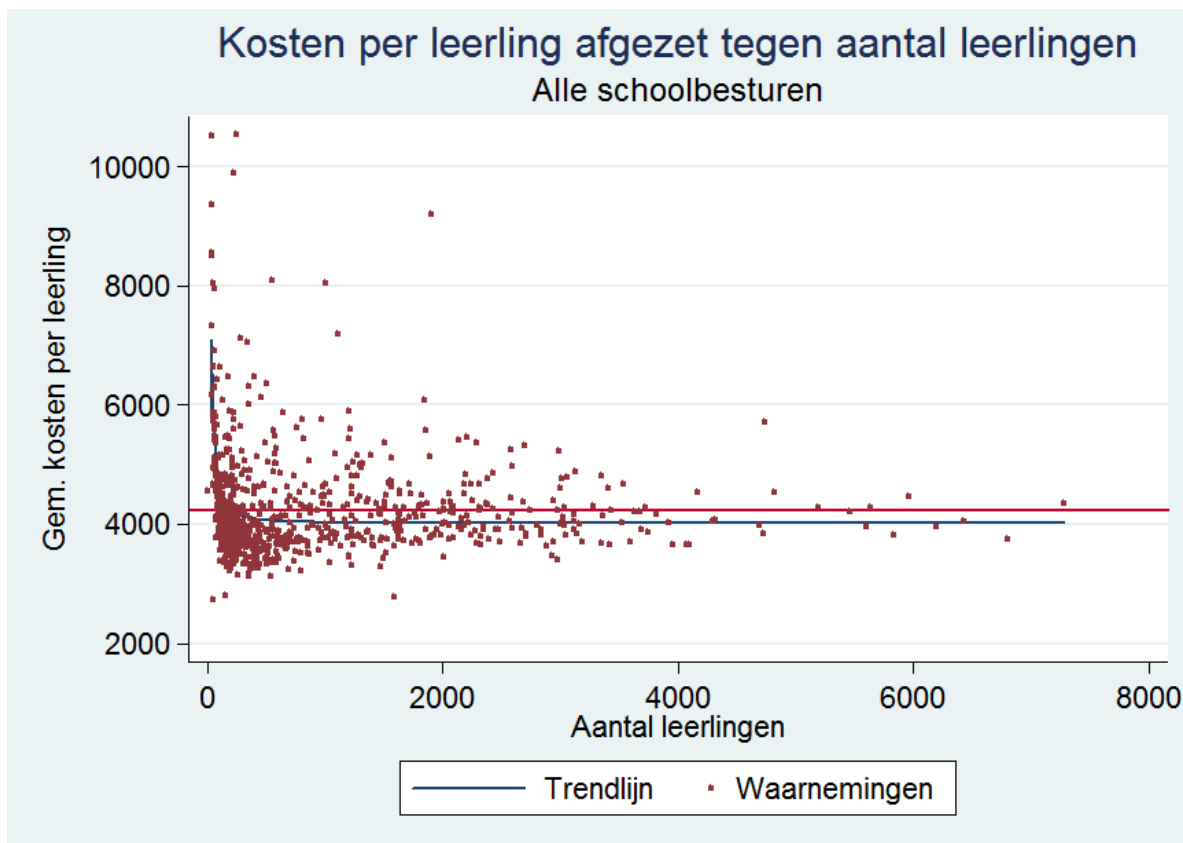
Grafiek 1 geeft een detailopname van de schoolbesturen met maximaal 1.000 leerlingen (N=587). Grafiek 2 geeft een totaalbeeld van alle schoolbesturen. In beide grafieken staat op de x-as het aantal leerlingen en op de y-as staan de gemiddelde vaste kosten per leerling. De rode lijn in beide grafieken geeft aan wat de gemiddelde vaste kosten per leerling zijn (€4.238). De blauwe lijnen zijn trendlijnen op basis van de waarnemingen.

Uit grafiek 1 blijkt dat schoolbesturen tot een grootte van ongeveer 150 leerlingen grote schaalvoordelen kunnen behalen. Als snel loopt de trendlijn min of meer vlak. Uit de grafiek blijkt dat schoolbesturen nog beperkte schaalvoordelen kunnen behalen tot een grootte van ongeveer 400 tot 500 leerlingen. Daarna heeft schaalgrootte op bestuursniveau weinig tot geen invloed meer op de gemiddelde kosten per leerling.



Grafiek 1: Kosten per leerling afgezet tegen het aantal leerlingen voor schoolbesturen tot 1000 leerlingen

Uit grafiek 2 blijkt geen ander beeld. Voor kleine schoolbesturen zijn er grote schaalvoordelen te behalen, maar snel daarna zijn er geen schaalvoordelen meer te behalen en loopt de kostentrendlijn vlak. Uit beide grafieken blijkt ook dat er geen omslagpunt van schaalvoordelen naar schaalnadelen waar te nemen is. Naarmate schoolbesturen meer leerlingen aansturen, stijgen de gemiddelde vaste kosten per leerling vanaf een bepaald punt dus niet. Daaruit volgt dat schoolbesturen wel schaalvoordelen kunnen behalen, maar geen schaalnadelen.



Grafiek 2: Kosten per leerling afgezet tegen het aantal leerlingen voor alle schoolbesturen

Hypothese 1 op bestuursniveau wordt deels verworpen, want er blijkt geen omslag plaats te vinden tussen schaalvoordelen en schaalnadelen. Uit de analyse blijkt dat als een schoolbestuur meer leerlingen aanstuurt, er schaalvoordelen behaald kunnen worden, maar tot een beperkt niveau. Een minimale bestuursgrootte van ongeveer 150 leerlingen is aan te bevelen, omdat tot dat niveau aanzienlijke schaalvoordelen te behalen zijn. Tot een grootte van ongeveer 400 tot 500 leerlingen per schoolbestuur zijn beperkte schaalvoordelen te behalen. Bij grotere schoolbesturen heeft schaalgrootte geen effect meer op de kosten per leerling.

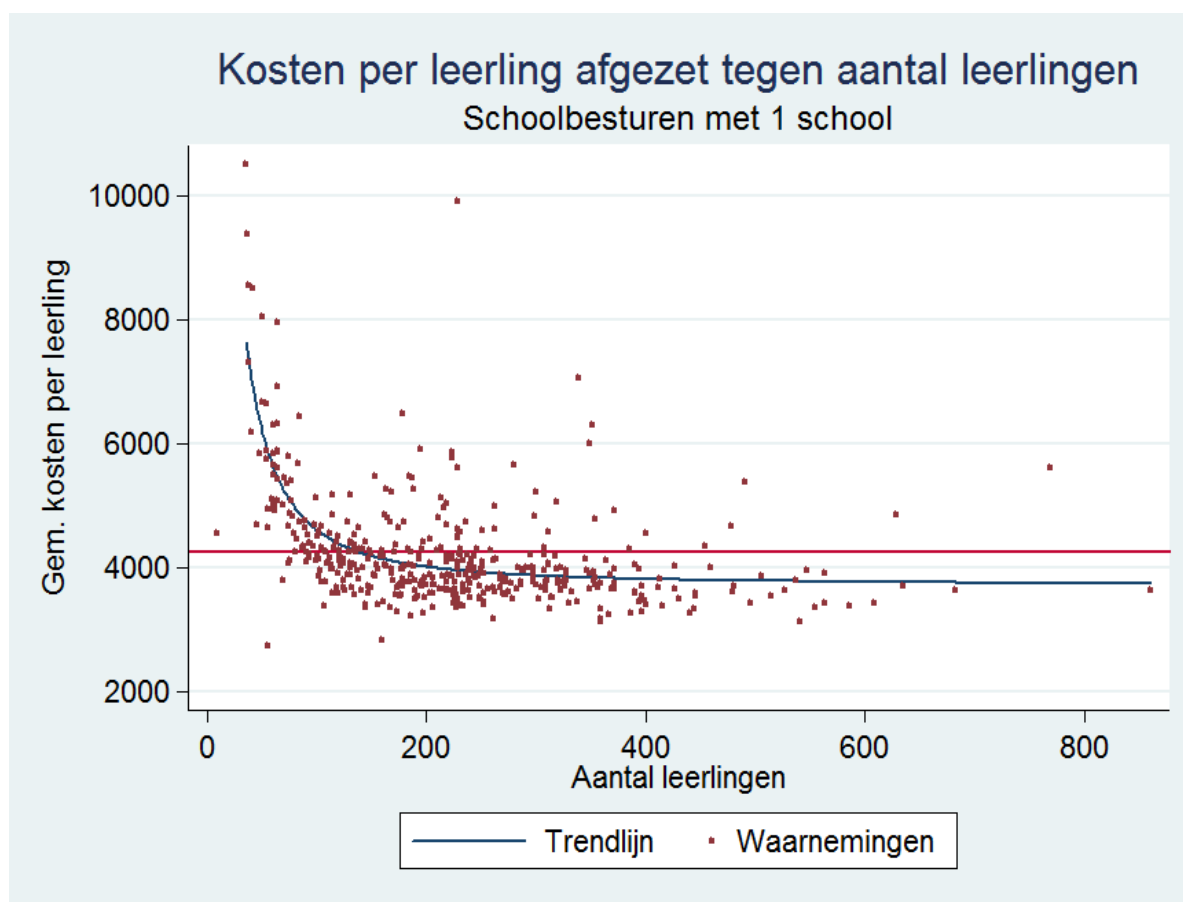
6.3 Schaaleardeffecten op schoolniveau

Hypothese 1 kan naast bestuursgrootte ook getest worden op schoolniveau gemeten in het aantal leerlingen op de school. Om hypothese 1 te toetsen op schoolniveau zijn alle schoolbesturen geselecteerd die één school besturen. Als gevolg daarvan bleven er van de 847 schoolbesturen 450 over. De gemiddelde kosten per leerling zijn in deze subgroep € 4.225. Dit wijkt niet significant af van het gemiddelde van de uitgebreide steekproef.

In de subgroep blijken openbare scholen, katholieke scholen en samenwerkingscholen PC/RK ondervertegenwoordigd en algemeen bijzondere scholen oververtegenwoordigd. Openbare schoolbesturen bestaan nauwelijks uit één school, terwijl algemeen bijzondere schoolbesturen in 85% van de gevallen uit één school bestaan. Dit beperkt de algemene toepasbaarheid van de gegevens.

Het gemiddelde leerlingaantal van 224 per school komt goed overeen met de gemiddelde schoolgrootte van 222 leerlingen die het OCW meldt over 2012 (OC&W, 2013b). Hier verkregen resultaten kunnen dus een goede indicatie zijn voor de situatie op een gemiddelde school.

In grafiek 3 staat op de x-as het aantal leerlingen en op de y-as staan de gemiddelde kosten per leerling. De rode lijn geeft de gemiddelde kosten per leerling uit de hele steekproef aan (€4.238). De blauwe lijn is de trendlijn op basis van de waarnemingen. Uit deze grafiek blijkt dat grote scholen gemiddeld lagere kosten per leerling hebben. Grotere scholen zijn dus goedkoper dan kleine scholen. We zien hier min of meer hetzelfde beeld als eerder in grafiek 1. Scholen kunnen grote schaalvoordelen behalen tot een grootte van ongeveer 150 leerlingen en beperkte voordelen tot een grootte van ongeveer 400 tot 500 leerlingen. Daarna heeft schaalgrootte geen effect meer op de gemiddelde kosten per leerling en loopt de trendlijn vlak. Dit komt overeen met het onderzoek van Blank & al (2007), waarin basisscholen tot ongeveer 125 leerlingen grote schaalvoordelen kunnen behalen. Een verschil is dat in het onderzoek van Blank & al (2007) scholen groter dan 550 leerlingen schaalnadelen behalen, iets wat uit deze analyse niet blijkt. Het hier gevonden schaal-effect komt overeen met een ouder onderzoek uit 2000, wat vindt dat basisscholen tot ongeveer 400 leerlingen schaalvoordelen kunnen behalen (IPSE, 2011).



Grafiek 3: Kosten per leerling afgezet tegen het aantal leerlingen voor schoolbesturen met 1 school

Als we de schaaleffecten op bestuursniveau vergelijken met de schaaleffecten op schoolniveau, dan zien we hetzelfde patroon. Zowel op bestuurs- als op schoolniveau zijn er tot een grootte van ongeveer 150 leerlingen grote schaalvoordelen te behalen. Tot een grootte van ongeveer 400 tot 500 leerlingen zijn er beperkte schaalvoordelen te behalen, terwijl daarna de kostencurve vlak verloopt.

Hypothese 1 op schoolniveau wordt deels verworpen, want er blijkt geen omslag plaats te vinden tussen schaalvoordelen en schaalnadelen. Uit de analyse blijkt wel dat als een school meer leerlingen aanstuurt, er schaalvoordelen behaald kunnen worden, maar dit slechts tot een beperkt niveau. Grotere scholen kennen gemiddeld lagere kosten per leerling, maar naarmate de school groter wordt neemt het effect van schaal op de gemiddelde kosten af.

6.4 Schaaleffecten op bestuursniveau – aantal scholen

Hypothese 1 kan naast bestuurs- en schoolniveau gemeten in aantal leerlingen ook getest worden op bestuursniveau gemeten in aantal scholen per bestuur. Om dit te toetsen zijn alle scholen in de steekproef onderverdeeld naar het aantal scholen per schoolbestuur. 53% van de schoolbesturen stuurt 1 school aan, het grootste bestuur in de steekproef bestuurt 30 scholen.

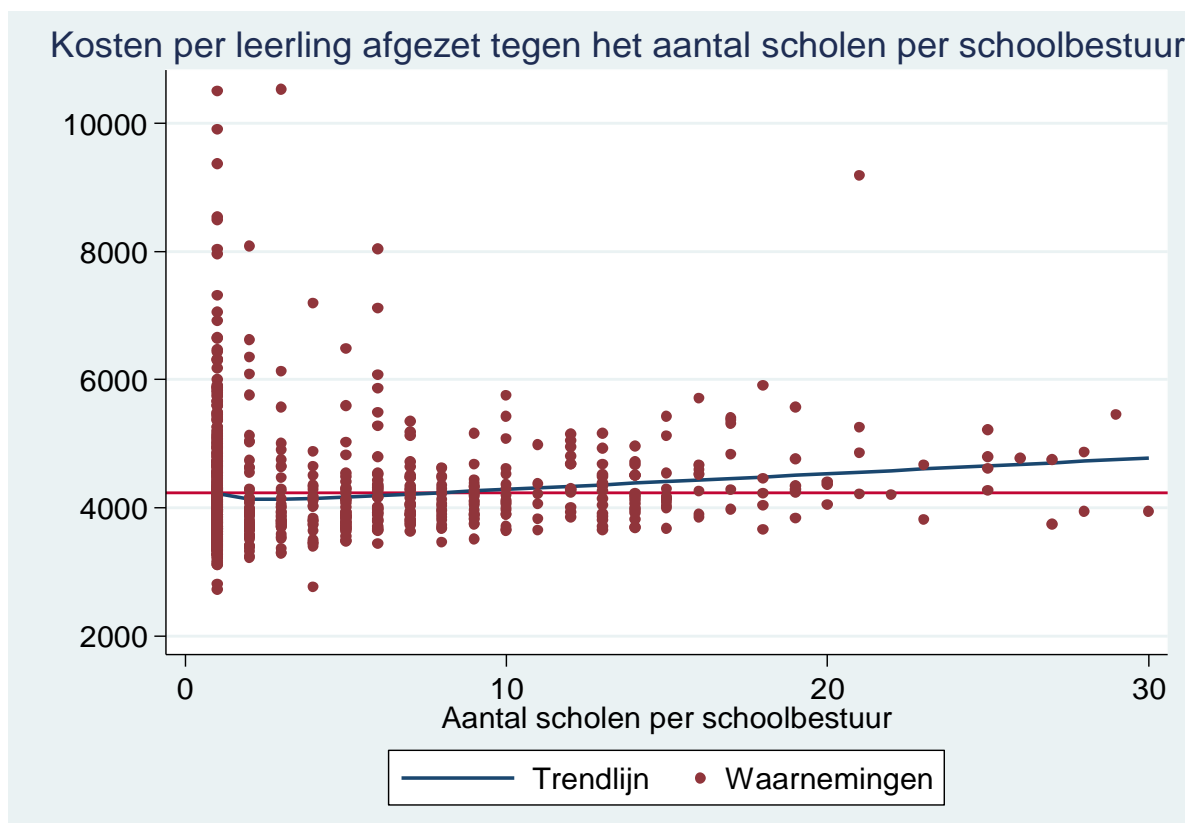
In tabel 7 zijn de gemiddelde kosten per leerling weergegeven voor de verschillende schoolbesturen. Schoolbesturen die weinig voorkomen zijn gegroepeerd, om zo een minimaal aantal observaties van 15 te krijgen. Schoolbesturen met 11-13, 14-15, 16-20 en 21-30 scholen zijn daarom gegroepeerd. Schoolbesturen met 2 tot 10 scholen worden zowel individueel als gegroepeerd weergegeven. In de twee rechtse kolommen staat aangegeven of de kosten per leerling significant lager of hoger zijn dan het gemiddelde van € 4.238.

Aantal scholen	Observaties	Gem. kosten per leerling	Significant lager dan gemiddelde	Significant hoger dan gemiddelde
1 school	450	€ 4.225		
2 scholen	60	€ 4.108		
3 scholen	40	€ 4.204		
4 scholen	27	€ 4.046		
5 scholen	34	€ 4.100		
2-5 scholen	161	€ 4.120	**	
6 scholen	33	€ 4.469		
7 scholen	25	€ 4.248		
8 scholen	20	€ 4.075	**	
9 scholen	27	€ 4.129	**	
10 scholen	20	€ 4.249		
6-10 scholen	125	€ 4.253		
11-13 scholen	35	€ 4.328		
14-15 scholen	29	€ 4.258		
16-20 scholen	29	€ 4.550		***
21-30 scholen	18	€ 4.808		**

Tabel 7: Gemiddelde kosten per leerling afgezet tegen het aantal scholen per schoolbestuur

Uit tabel 7 blijkt dat besturen die 1 school aansturen niet significant goedkoper of duurder zijn dan het gemiddelde schoolbestuur. Wel is er, zoals grafiek 4 laat zien, veel meer spreiding in de gemiddelde kosten per leerling. Schoolbesturen die 2 tot 5 (*gegroepeerd*), 8 of 9 scholen aansturen zijn significant goedkoper, terwijl besturen die 16 of meer scholen aansturen significant duurder zijn. Schoolbesturen die 1 tot ongeveer 10 scholen besturen lijken daarom het meest kostenefficiënt. Een mogelijke verklaring voor de oplopende kosten bij grote schoolbesturen wordt gegeven door Van de Venne (2002), die aangeeft dat bij schaalvergroting de besturing van de organisatie complexer wordt. De Inspectie (2014) signaleert ook dat grote schoolbesturen vaker dan kleine schoolbesturen moeite hebben om de onderwijskwaliteit en de financiën onder controle te houden. Een andere mogelijke verklaring ligt in de afnemende gemiddelde schoolgrootte als het aantal scholen per bestuur toeneemt. De gemiddelde schoolgrootte daalt van 224 leerlingen bij een schoolbestuur met 1 tot 5 scholen tot 192 bij besturen met 16 tot 20 scholen en 154 bij besturen met 21 tot 30 scholen. Kleine scholen zoeken sneller de samenwerking en dit effect wordt vooral zichtbaar bij grote schoolbesturen. In het gedeelte hiervoor is gebleken dat kleine scholen hogere kosten per leerling kennen, dus is het begrijpelijk dat grote schoolbesturen met gemiddeld kleinere scholen duurder zijn.

Uit grafiek 4 blijkt dat er wat grootte van een schoolbestuur betreft een minimum efficiënte schaal is. Op de x-as staat het aantal scholen per schoolbestuur en op de y-as staan de gemiddelde kosten per leerling. De rode lijn geeft de gemiddelde kosten per leerling aan (€4.238). De blauwe trendlijn vertoont een dip in kosten per leerling tussen ongeveer 2 tot 6 scholen per schoolbestuur. Daarna lopen de kosten afgezet tegen het aantal scholen per schoolbestuur op.



Grafiek 4: Kosten per leerling afgezet tegen het aantal scholen per schoolbestuur

Hypothese 1 op bestuursniveau gemeten in aantal scholen per schoolbestuur kan niet worden verworpen. Een school die onderdeel is van een overkoepelend schoolbestuur kent lagere gemiddelde vaste kosten per leerling. Dit blijkt echter alleen te gelden voor schoolbesturen die maximaal ongeveer 10 scholen aansturen. Schoolbesturen die 16 of meer scholen aansturen behalen significante schaalnadelen.

6.5 Schaaleffecten op groepsniveau

Hypothese 2: Scholen met grotere klassen kennen lagere gemiddelde vaste kosten per leerling.

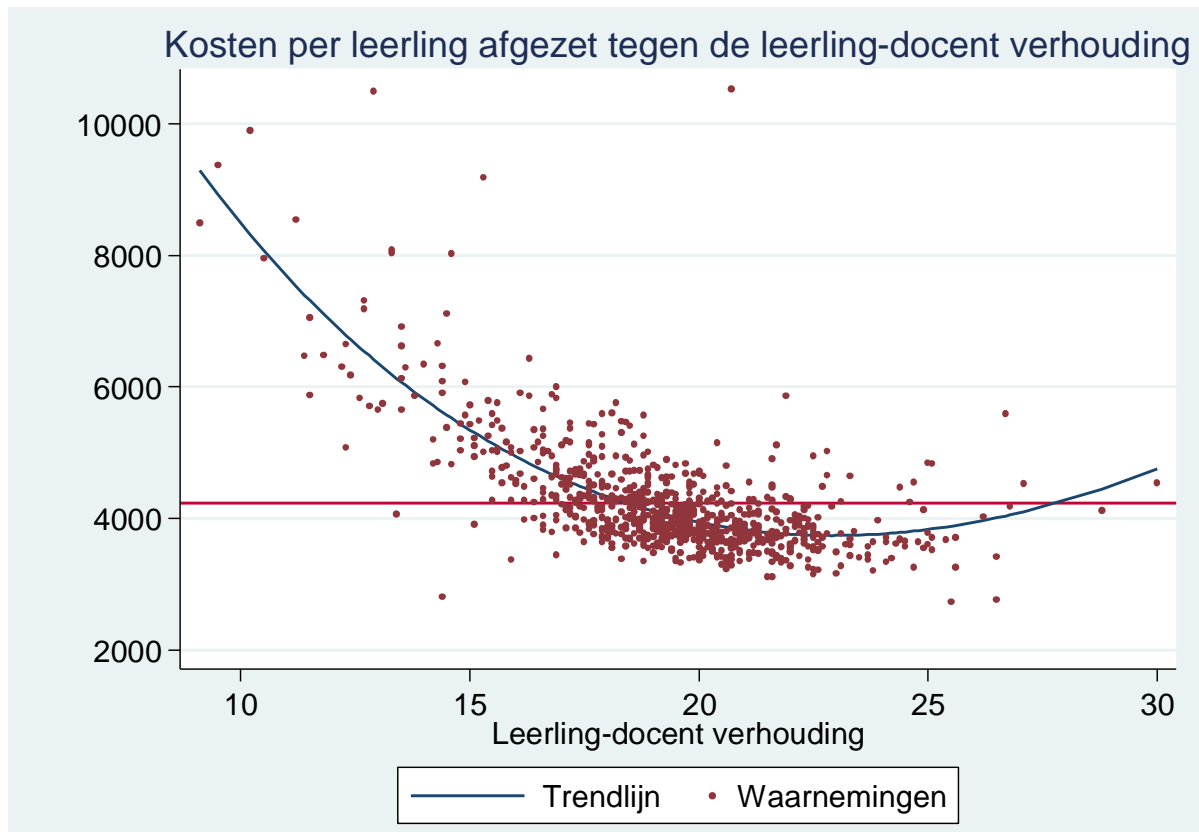
Om de tweede hypothese te toetsen is van alle schoolbesturen de verhouding tussen het aantal leerlingen en de FTE's onderwijzend personeel uitgerekend. Deze leerling-docent verhouding is een indicator voor het aantal leerlingen dat een (voltijd)docent gemiddeld in de klas heeft. Doordat er sprake is van een indicator op bestuursniveau en niet op schoolniveau, kan er vertekening optreden. Scholen binnen hetzelfde bestuur kunnen er immers voor kiezen om klassen klein of groot vorm te geven. Een grafiek van het patroon van de leerling-docent verhouding op schoolbesturen met 1 school wijkt echter niet af van het patroon in grafiek 5.

In tabel 8 zijn de gemiddelde kosten afgezet tegen de leerling-docent verhouding. Leerling-docent verhoudingen die weinig voorkomen zijn gegroepeerd. Uit de tabel blijkt dat besturen met een leerling-docent verhouding kleiner dan 18 significant hogere kosten per leerling hebben dan gemiddeld. Kleine klassen, dus klassen tot 18 leerlingen per docent, zijn duurder dan grotere klassen. Klassen met 19 tot 25 leerlingen per docent zijn significant goedkoper dan het gemiddelde. Klassen groter dan 25 leerlingen per docent worden weer duurder, maar zijn nog wel significant goedkoper dan het gemiddelde. Een verklaring voor de stijgende kosten per leerling bij grote klassen met meer dan 25 leerlingen is moeilijk te geven. Een verklaring zou kunnen zijn dat er in een grote klas meer onderwijs ondersteunend personeel aan het werk is, maar dit verband is niet uit een data-analyse af te leiden.

Leerling-docent verhouding	Observaties	Gemiddelde kosten per leerling	Significant lager dan gemiddelde	Significant hoger dan gemiddelde
≤15	49	€ 6.384		***
15-16	31	€ 5.064		***
16-17	50	€ 4.736		***
17-18	93	€ 4.368		***
18-19	141	€ 4.216		
19-20	184	€ 3.990	***	
20-21	116	€ 3.871	***	
21-22	81	€ 3.835	***	
22-23	52	€ 3.824	***	
23-25	35	€ 3.744	***	
≥25	15	€ 3.907	*	

Tabel 8: Gemiddelde kosten per leerling afgezet tegen de leerling-docent verhouding

In grafiek 5 staat op de x-as de leerling-docent verhouding en op de y-as staan de gemiddelde kosten per leerling. De rode lijn geeft de gemiddelde kosten per leerling weer (€4.238). De blauwe lijn is de trendlijn op basis van de waarnemingen. Uit de grafiek blijkt dat schoolbesturen met een leerling-docent verhouding kleiner dan 18 hogere kosten per leerling hebben dan gemiddeld. Als de verhouding groter wordt dan 24 loopt de trendlijn weer iets op.



Grafiek 5: Kosten per leerling afgezet tegen de leerling-docent verhouding

Hypothese 2 kan niet verworpen worden. Een schoolbestuur met een hogere leerling-docent verhouding kent lagere gemiddelde vaste kosten dan een schoolbestuur met een lagere leerling-docent verhouding. De kostentrendlijn loopt uiteindelijk weer iets op. Scholen met klassen tussen ongeveer 19 en 25 leerlingen zijn het meest kostenefficiënt.

6.6 Schaaffecten naar denominatie

Hypothese 3: Islamitische schoolbesturen hebben gemiddeld hogere kosten per leerling, terwijl de gemiddelde kosten per leerling bij andere schoolbesturen niet significant van elkaar afwijkend zullen zijn.

Deze hypothese is getoetst aan de hand van de beperkte subgroep van 790 besturen. In tabel 9 is een overzicht gegeven van een aantal belangrijke gegevens. Er zijn een aantal opvallende verschillen met de uitgebreide subgroep. Zo is in de beperkte subgroep het gemiddeld aantal leerlingen per schoolbestuur 40 lager en het totaal aan FTE's is 2,8 lager. De gemiddelde kosten per leerling zijn € 31 lager. De minima en maxima veranderen niet.

Variabele	Gemiddeld	Minimum	Maximum
Leerlingen	866	9	7282
Fte's	56	1,3	523
Managers	4,9	0	41,7
Leraren in opleiding	0,04	0	5,5
Leraren	45,4	0	417,3
Onderwijs ondersteunend pers.	5,5	0	77,9
Leerlingen per FTE Leraren	19,3	9,1	30
Scholen per bevoegd gezag	4,19	1	30
Kosten per leerling	€ 4.207	€ 2.726	€ 10.539
Kosten per FTE	€ 63.612	€ 31.442	€ 146.013
Kosten per school	€ 900.862	€ 40.875	€ 4.306.115

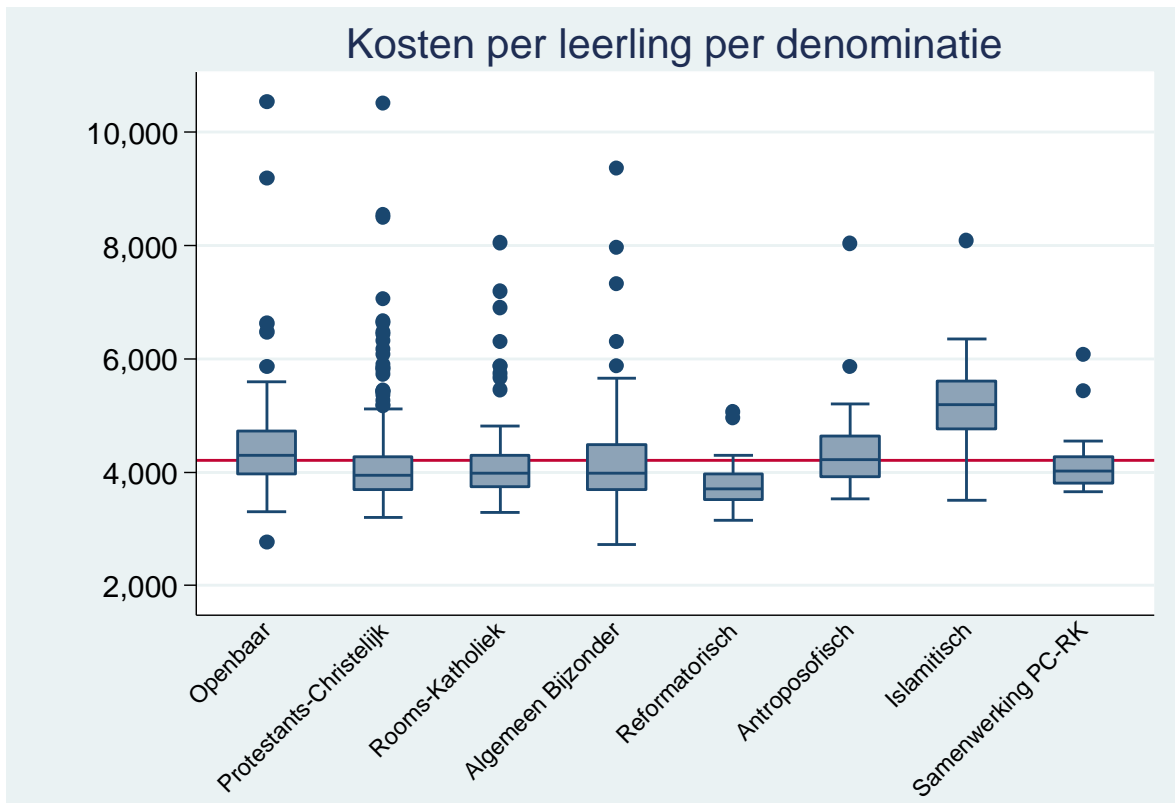
Tabel 9: Algemene gegevens beperkte subgroep (N=790)

Van elke denominatie zijn de gemiddelde kosten per leerling uitgerekend, die vervolgens vergeleken zijn met het gemiddelde van de beperkte subgroep (€4.207). Daaruit blijkt dat reformatorische en protestants-christelijke schoolbesturen significant lagere kosten dan gemiddeld behalen. Verder blijkt dat openbare en islamitische schoolbesturen significant hogere kosten dan gemiddeld behalen. Islamitische schoolbesturen zijn per leerling bijna € 1.100 duurder dan gemiddeld (zie tabel 10).

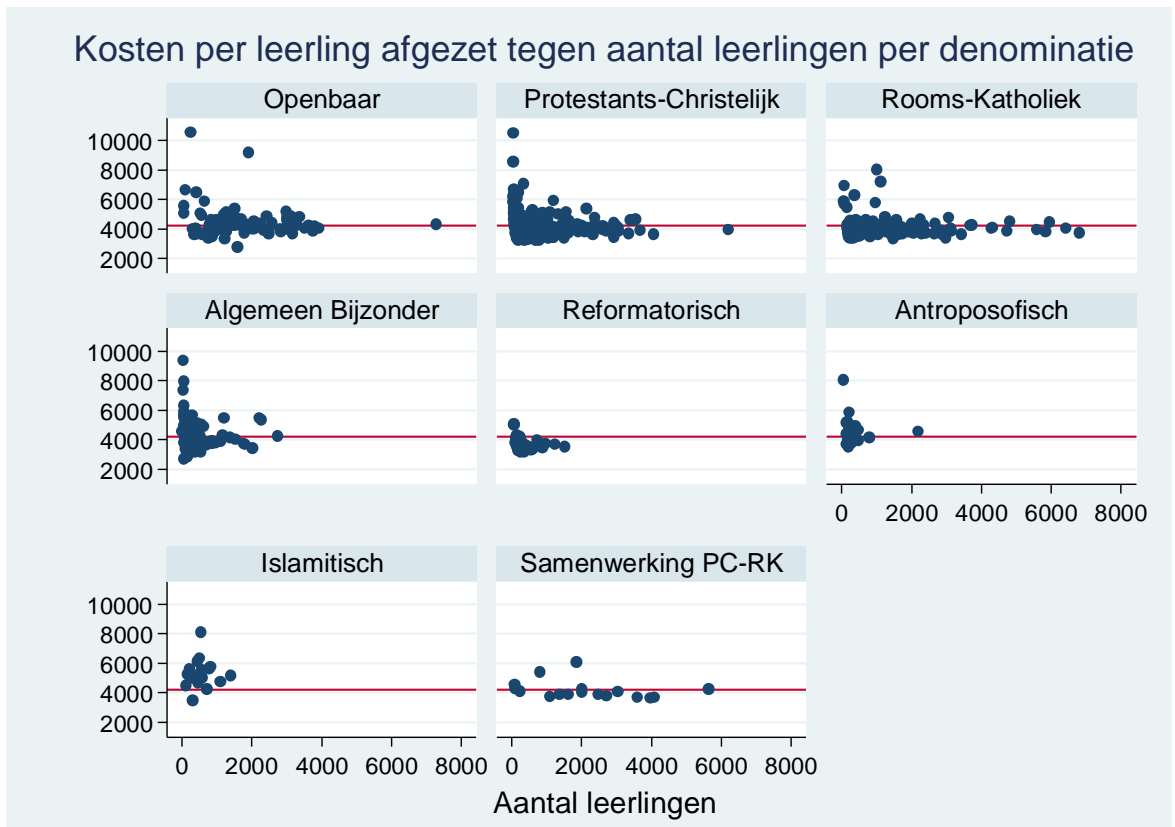
Denominatie	Observaties	Gemiddelde kosten per leerling	Significant lager dan gemiddelde	Significant hoger dan gemiddelde
Openbaar	82	€ 4.494		***
Protestants-Christelijk	325	€ 4.134	*	
Rooms-Katholiek	119	€ 4.162		
Algemeen Bijzonder	161	€ 4.186		
Reformatorisch	38	€ 3.764	***	
Antroposofisch	30	€ 4.423		
Islamitisch	18	€ 5.304		***
Samenwerking PC-RK	17	€ 4.194		

Tabel 10: Gemiddelde kosten per leerling afgezet tegen denominatie

Ook uit een boxplot (grafiek 6) blijkt dat islamitische schoolbesturen per leerling duurder zijn dan gemiddeld. Op de x-as staan de verschillende denominaties, op de y-as de gemiddelde kosten per leerling. De rode lijn geeft de gemiddelde kosten per leerling aan (€4.207). Meer dan driekwart van de islamitische schoolbesturen en meer dan de helft van de openbare schoolbesturen kent hogere kosten dan gemiddeld. Protestants-christelijke en rooms-katholieke schoolbesturen zijn in bijna driekwart van de gevallen goedkoper dan gemiddeld. Bijna alle reformatorische schoolbesturen behalen lagere kosten per leerling dan gemiddeld. Wat verder opvalt, is dat onder openbare, algemeen bijzondere en islamitische scholen de spreiding in leerlingkosten groot is.



Grafiek 6: Kosten per leerling naar denominatie



Grafiek 7: Kosten per leerling afgezet tegen het aantal leerlingen naar denominatie

In grafiek 7 zijn per denominatie de kosten per leerling afgezet tegen het aantal leerlingen. Op de x-as staat het aantal leerlingen en op de y-as staan de gemiddelde kosten per leerling. De rode lijn geeft de gemiddelde kosten per leerling aan (€4.207). In de grafiek per denominatie wordt duidelijk dat kleinere scholen vaker boven de rode lijn liggen, zoals ook het algemene beeld is. Verder is het beeld uit tabel 7 en de boxplot terug te vinden in de grafiek. Reformatorische schoolbesturen liggen bijna allemaal onder de rode lijn, terwijl islamitische schoolbesturen bijna allemaal boven de rode lijn te vinden zijn.

Tabel 11 geeft de correlatie weer tussen het aantal leerlingen dat valt onder een schoolbestuur en de gemiddelde kosten per leerling. De correlatie in de gehele subgroep is zwak negatief (-0.0702), wat aangeeft dat de gemiddelde kosten per leerling dalen als het aantal leerlingen per schoolbestuur toeneemt. Tussen de verschillende denominaties zijn opvallende verschillen waar te nemen. Zo is de correlatie bij reformatorische schoolbesturen veel sterker negatief dan gemiddeld, terwijl de correlatie bij islamitische schoolbesturen zelfs zwak positief is. Islamitische schoolbesturen hebben dus baat bij een kleinschalige vormgeving, omdat de kosten per leerling stijgen als de leerlingaantallen stijgen.

Denominatie	Correlatie leerlingaantal en kosten/leerling	Leerling-docent verhouding
Gehele subgroep	-0.0702	19,34
Openbaar	-0.2188	18,54 (** lager)
Protestants-Christelijk	-0.0988	19,67 (** hoger)
Rooms-Katholiek	-0.1330	19,02 (** lager)
Algemeen Bijzonder	-0.0689	19,38
Reformatorisch	-0.3320	20,67 (** hoger)
Antroposofisch	-0.0753	19,00
Islamitisch	0.0942	16,3 (**lager)
Samenwerking PC-RK	-0.1274	19,66

Tabel 11: Correlatiematrix en leerling-docent verhouding naar denominatie

Verder is in tabel 11 de leerling-docent verhouding per denominatie weergegeven. Gemiddeld heeft een schoolbestuur voor elke 19,34 leerlingen een voltijd docent in dienst. Protestants-christelijke en reformatorische schoolbesturen hebben een significant hoger aantal leerlingen per voltijd docent, terwijl openbare, rooms-katholieke en islamitische schoolbesturen een significant lager aantal leerlingen per voltijd docent hebben. Vooral islamitische schoolbesturen kiezen voor kleine klassen van gemiddeld 16,3 leerlingen per voltijd docent. Dit kan een verklaring zijn voor de hogere kosten per leerling (zie tabel 10).

Een andere verklaring voor de hogere kosten per leerling is de sociaaleconomische status (SES) van de leerlingen op een school. Het blijkt dat er sociaaleconomische verschillen zijn tussen scholen (Herweijer, 2008). De vrije schoolkeuze die in Nederland gehanteerd wordt, kan leiden tot segregatie tussen bepaalde groepen leerlingen (OC&W, 2014a). De denominatie van de onderwijsinstelling lijkt invloed te hebben op die segregatie (Lafeber, 2006).

Islamitische scholen worden vrijwel alleen bevolkt door leerlingen met een niet-westerse allochtone afkomst. Hun ouders zijn gemiddeld lager opgeleid dan de ouders van autochtone leerlingen. Scholen met een hoge concentratie niet-westerse allochtone leerlingen zijn dus bijna per definitie ook scholen met een hoge concentratie leerlingen uit de lagere sociaaleconomische milieus (Herweijer, 2008). Islamitische scholen hebben dus om te gaan met een zwakke leerlingpopulatie, waardoor het begrijpelijk is dat er gekozen wordt voor kleinere schoolklassen. Deze verklaring wordt versterkt door het feit dat islamitische (54%) en openbare scholen (15%) relatief veel achterstandsleerlingen hebben ten opzichte van rooms-katholieke (13%) en protestants-christelijke (12%) scholen (Gramberg & Jordanov, 2011). Uit het onderzoek van Driessen is af te leiden dat leerlingen op reformatorische basisscholen ouders hebben met een relatief hoog opleidingsniveau (Driessen, 2007), waardoor juist grotere klassen mogelijk zijn.

Hypothese 3 wordt verworpen. Er blijken significante verschillen te zijn in de gemiddelde kosten per leerling als deze uitgesplitst worden naar denominatie. Reformatorische en protestants-christelijke schoolbesturen zijn per leerling significant goedkoper dan gemiddeld, terwijl openbare en met name islamitische schoolbesturen per leerling significant duurder zijn dan gemiddeld. Mogelijke verklaringen voor deze verschillen kunnen liggen in de groepsgrootte, het opleidingsniveau van de ouders en de hoeveelheid achterstandsleerlingen.

7. Resultaten: Onderwijskundige Schaalears

7.1 Startpositie

Voordat er gestart wordt met het opbouwen van modellen, wordt het startmodel M_0 bepaald (tabel 12). Dit is het model met als afhankelijke variabele de CITO-score, zonder dat er verklarende variabelen aan zijn toegevoegd (*laag 0 in het methodologisch model*). Er zijn twee startmodellen bepaald. Model M_{0A} is het model waar vanuit de analyse gestart wordt met 6.692 waarnemingen, model M_{0B} is het definitieve model zonder verklarende variabelen met 3.405 waarnemingen. Ten opzichte van M_{0A} zijn de *Log Likelihood* en *AIC* van M_{0B} ongeveer de helft, wat logisch is omdat het aantal waarnemingen van M_{0B} ongeveer de helft is. De gemiddelde CITO-score en onverklaarde variantie ontlopen elkaar in beide modellen niet veel.

	Model M0-A	Model M0-B
<i>Log Likelihood</i>	- 24.484	- 12.423
<i>Aantal observaties</i>	6.672	3.405
<i>AIC</i>	48.973	24.849
<i>Constante</i>	534,66	534,67
<i>var(Residual)</i>	90,16	86,38

Tabel 12: Startmodellen M_{0A} en M_{0B}

Voordat schoolkenmerken aan het model kunnen worden toegevoegd en hypothese 4 beantwoord kan worden moet er worden gecontroleerd voor de algemene en schoolse leerling- en gezinskenmerken (*laag 1 en 2 in het methodologisch model*). Omdat deze kenmerken niet centraal staan in de probleemstelling van dit onderzoek, zijn deze modellen niet in de hoofdttekst verwerkt. Het toewerken naar model M_{24} , het eerste model waaraan schoolkenmerken worden toegevoegd, is weergegeven in twee aparte bijlages. In bijlage 11 zijn de modellen M_1 tot M_{11} te vinden, waarin steeds algemene leerling- en gezinskenmerken worden toegevoegd. In bijlage 12 zijn de modellen M_{12} tot M_{23} te vinden, waarin steeds schoolse leerling- en gezinskenmerken worden toegevoegd. In deze bijlages wordt ook een toelichting gegeven op het tot stand komen van deze modellen. De hoofdttekst vervolgt de analyse van het onderwijskundige deel van de probleemstelling met model M_{24} . In dat model en de daaropvolgende modellen worden steeds variabelen op het schoolniveau toegevoegd om te kijken of deze variabelen nog extra variantie verklaren (*laag 3 in het methodologisch model*).

7.2 Schoolkenmerken

In hypothese 4 werd gesteld dat *de CITO-score relatief hoger is als:*

- ✓ *De leerling op een school in Zeeland of Limburg zit;*
- ✓ *De leerling op een school in een meer verstedelijkt gebied zit;*
- ✓ *De leerling op een school zit met een laag schoolgewicht;*
- ✓ *De leerling niet op een islamitische school zit;*
- ✓ *De leerling op een school zit waar taal- en rekenverbetertrajecten zijn ingezet;*
- ✓ *De leerling op een school zit met een homogene leerlingpopulatie;*
- ✓ *De leerling op een school zit groter dan 145 leerlingen;*
- ✓ *De leerling op een school zit die valt onder een klein schoolbestuur.*

In model M₂₄ wordt de variabele *provincie* toegevoegd met als referentiecategorie de provincie Zeeland, een van de provincies waarin leerlingen beter zouden scoren. In welke provincie de school staat blijkt in de meeste gevallen geen extra verklarend effect te hebben ten opzichte van de al aanwezige variabelen in het model; alle provincie-effecten behalve dat van Limburg zijn niet significant. Limburg scoort ten opzichte van Zeeland en de rest van Nederland significant beter. Waar de Inspectie (2014) constateert dat Limburg en Zeeland het iets beter doen dan het landelijk gemiddelde, blijkt dit dus deels te kloppen. In het geval van Limburg is het verschil zo groot dat het effect heeft op de CITO-scores. In de volgende modellen wordt de variabele *provincie* echter buiten beschouwing gelaten, omdat de AIC van model M₂₄ verslechtert ten opzichte van model M₂₃. De variabele voegt dus geen verklaringskracht toe aan het model, ook al wordt er een significant effect van de provincie Limburg gevonden. De variabele *geboorteland vader is Suriname / Antillen* wordt ook buiten beschouwing gelaten.

In model M₂₅ wordt de variabele *mate van stedelijkheid* toegevoegd, met als referentiecategorie zeer sterk stedelijk gebied. Bij het bepalen van de stedelijkheid is uitgegaan van de omgevingsadressendichtheid (OAD) zoals vastgesteld door het CBS. Een gebied wordt als zeer sterk, sterk, matig, weinig of niet stedelijk gezien bij meer dan 2500 respectievelijk 1500-2500, 1000-1500, 500-1000 en minder dan 500 adressen per vierkante kilometer. Het effect van stedelijkheid is alleen significant in sterk stedelijk gebied. Onder het sterk stedelijk gebied vallen bijvoorbeeld steden als Nijmegen, Breda, Zwolle of Leeuwarden. De coëfficiënten van matig en weinig stedelijk gebied zijn positief maar niet significant. Ten opzichte van het zeer sterk stedelijk gebied scoren leerlingen in weinig tot sterk stedelijk gebied beter, terwijl leerlingen in plattelandsgebieden slechter scoren. Dit is in overeenstemming met het onderzoek van Noailly & Koning (2009) die vinden dat leerlingen in een meer verstedelijkt gebied een hogere CITO-score halen dan leerlingen in een minder verstedelijkt gebied.

In model M₂₆ wordt de variabele *schoolscore* toegevoegd. De schoolscore geeft aan wat de sociaal-etnische samenstelling van de leerlingenpopulatie van de school is, waarbij de score loopt van 100 tot 333. Daarbij geldt dat hoe hoger de score is, hoe groter de sociaal-etnische achterstand van de school. De referentiecategorie is de groep scholen met een schoolscore van 100-104. Dit zijn scholen met weinig tot geen achterstandsleerlingen. Door het toevoegen van de variabele *schoolscore* verandert het effect van de stedelijkheid. Zo zijn scholen die in sterk stedelijk gebied vallen niet meer beter dan scholen in zeer sterk stedelijkheid gebied. Wel zijn de CITO-scores op scholen in plattelandsgebieden significant slechter dan in andere gebieden. Verder scoren scholen met een schoolscore van 105-109 en 110-119 niet slechter dan de referentiecategorie. Leerlingen op scholen met een grote leerlingenpopulatie met sociaal-etnische achterstand scoren significant 2 tot 3,5 punten slechter. Het gevonden effect is in overeenstemming met Noailly & Koning (2009), die stelden dat een hoger aandeel aandachtsleerlingen een significant negatief effect op de CITO-score heeft. Dat het effect van stedelijkheid verandert door de toevoeging van de variabele *schoolscore* is niet verrassend, omdat scholen in (zeer) sterk stedelijk gebied vaker een zware leerlingpopulatie hebben. Scholen op het platteland hebben relatief juist een eenvoudige leerlingenpopulatie, waardoor het effect van stedelijkheid op die scholen wel aanwezig blijft. Wat verder opvalt is dat de onverklaarde variatie op schoolniveau flink afneemt. De variabele *schoolscore* verklaart dus, gecontroleerd voor leerlingkenmerken, een aanzienlijk deel van de verschillen tussen scholen.

	Model M24	Model M25	Model M26
<i>Log Likelihood</i>	- 10.210	-10.215	- 10.205
<i>Aantal observaties</i>	3.405	3.405	3.405
<i>AIC</i>	20.474	20.467	20.458
<i>Constante</i>	469,18	467,64	469,53
Fixed			
Opleiding Vader HBO-WO	0,89 *** (0.215)	0,91 *** (0.215)	0,88 *** (0.215)
Opleiding Moeder MBO	1,11 *** (0.207)	1,10 *** (0.207)	1,04 *** (0.207)
Opleiding Moeder HBO-WO	2,00 *** (0.275)	2,00 *** (0.275)	1,92 *** (0.276)
Geboorteland Vader Turkije	- 1,00 *** (0.389)	- 1,02 *** (0.390)	- 0,90 ** (0.391)
Geboorteland Vader Suriname of Antillen	- 0,75 (0.620)		
Zorgleerling	- 1,90 *** (0.238)	- 1,91 *** (0.238)	- 1,90 *** (0.237)
Zelfvertrouwen	1,30 *** (0.147)	1,30 *** (0.147)	1,33 *** (0.147)
Werkhouding	0,82 *** (0.108)	0,82 *** (0.108)	0,82 *** (0.108)
Afhankelijkheid	- 0,56 *** (0.125)	- 0,55 *** (0.125)	- 0,55 *** (0.125)
Ouderbetrokkenheid	0,72 *** (0.114)	0,75 *** (0.114)	0,74 *** (0.114)
Welbevinden met medeleerlingen	- 0,54 *** (0.132)	- 0,53 *** (0.132)	- 0,54 *** (0.132)
Toets Rekenen / Wiskunde	0,46 *** (0.009)	0,46 *** (0.009)	0,46 *** (0.009)
Toets Taal	0,08 *** (0.007)	0,08 *** (0.007)	0,08 *** (0.007)
Groningen	0,62 (2.025)		
Friesland	- 1,41 (1.328)		
Drenthe	- 0,87 (1.541)		
Overijssel	- 1,29 (1.257)		
Flevoland	- 0,38 (1.585)		
Gelderland	0,39 (1.194)		
Utrecht	0,47 (1.683)		
Noord-Holland	- 0,61 (1.174)		
Zuid-Holland	0,03 (1.240)		

<i>Vervolg</i>	Model M24	Model M25	Model M26
Noord-Brabant	- 0,35 (1.123)		
Limburg	2,75 ** (1.311)		
Sterk stedelijk		1,56 ** (0.792)	0,67 (0.807)
Matig stedelijk		0,78 (0.797)	- 0,51 (0.831)
Weinig stedelijk		0,83 (0.777)	- 0,89 (0.863)
Niet stedelijk		- 0,69 (0.829)	- 2,48 *** (0.922)
Schoolscore 105-109			- 0,27 (0.625)
Schoolscore 110-119			0,32 (0.660)
Schoolscore 120-139			- 2,21 *** (0.764)
Schoolscore 140-159			- 2,05 ** (0.900)
Schoolscore > 159			- 3,61 *** (1.119)
Random			
var(Residual)	21,0	21,0	21,0
var(School)	8,3	8,7	7,8

Tabel 13: Geschatte Modellen M₂₄ tot M₂₆

De modellen M₂₇ tot M₂₉ zijn te vinden in tabel 14. In model M₂₇ wordt de variabele *denominatie* toegevoegd, die is ingedeeld in de categorieën openbaar (*referentiecategorie*), protestants-christelijk, rooms-katholiek, algemeen bijzonder, islamitisch en overig bijzonder. De groep overig bijzonder is een restgroep waarin alle denominaties, die slechts een aantal keren voorkwamen, zijn opgenomen. Dit betreft onder andere de reformatorische, gereformeerd vrijgemaakte en hindoeïstische scholen. Leerlingen op protestants-christelijke scholen en rooms-katholieke scholen scoren 1 respectievelijk 1,23 punt hoger op de CITO-score. Van de andere denominaties is geen effect waar te nemen, mogelijk ook door het kleinere aantal scholen dat in deze groepen valt. Driessen (2007) stelt dat de verschillen tussen verschillende denominaties, gecorrigeerd voor ouderlijk opleidingsniveau, klein zijn. Dat beeld blijkt ook uit dit model. Driessen (2007) en de Inspectie (2014) stellen wel dat islamitische scholen tot de zwakst presterende basisscholen behoren. Dat blijkt niet uit dit model, maar gezien het klein aantal islamitische scholen in deze steekproef (6) kunnen er aan de prestaties van deze scholen in het algemeen geen conclusies verbonden worden.

In model M₂₈ worden de variabelen *sterk, matig en weinig stedelijk* en *schoolscore 105-109 en 110-119* buiten beschouwing gelaten. Tevens worden de niet significante effecten van de *denominatie* uit het model verwijderd. In model M₂₈ worden de variabelen *taalverbetertraject* en *rekenverbetertraject* toegevoegd. Deze variabelen geven aan of een school deelneemt aan onderwijstrajecten om het reken- en/of taalniveau van de leerlingen te verbeteren. Beide onderwijsverbetertrajecten hebben geen significante invloed op de CITO-score van de leerlingen. De Inspectie (2012) concludeerde dat scholen waarop taal- en rekenverbetertrajecten werden gestart, hun CITO-score in het schooljaar 2010/2011 significant hadden verbeterd ten opzichte van het schooljaar 2008/2009. Dit is echter een vergelijking tussen verschillende schooljaren op dezelfde scholen in plaats van een vergelijking tussen verschillende scholen in hetzelfde schooljaar, zoals in dit model. De conclusie uit M₂₈ is dat leerlingen op scholen die taal- en rekenajecten hebben opgezet niet significant slechter (maar ook niet beter) scoren dan leerlingen op andere scholen. Dit is een gunstige ontwikkeling met betrekking tot de gemiddelde onderwijskwaliteit, omdat leerlingen op voorheen taalzwakke en rekenzwakke scholen nu dezelfde CITO-scores halen als leerlingen op niet taal- en rekenzwakke scholen.

In model M₂₉ worden variabelen met betrekking tot de diversiteit op school toegevoegd. De eerste variabele meet hoeveel *etnische groepen* er in groep 8 zijn. De tweede variabele en derde variabele zijn afgeleiden van de variabele *percentage allochtonen* in groep 8. Er zijn hier afgeleiden toegevoegd, omdat de correlatie tussen het aantal etnische groepen en het percentage allochtonen hoog was ($r=0.6979$). De afgeleide variabelen meten of het percentage allochtonen in groep 8 groter is dan 50 respectievelijk 75 procent. De correlatie van deze variabelen met het aantal etnische groepen is lager ($r=0.5414$ respectievelijk $r=0.3594$), zodat het collineariteitsprobleem minder speelt. De toevoeging van de variabelen heeft een groot effect op het model. Zo worden de schoolscores 140-159 en >159 niet significant. Dit betekent dat het percentage allochtonen op deze scholen groot is, waardoor de variabelen voor een (groot) deel dezelfde variantie verklaren. Verder is het effect van protestants-christelijke en rooms-katholieke scholen ten opzichte van andere scholen niet meer significant. In volgende modellen wordt de variabele *allochtonen >75%* behouden, de twee andere toegevoegde variabelen worden niet meegenomen. Het beeld uit model M₂₉ geeft aan, in overeenstemming met het jaarrapport van de Inspectie (2013), dat de heterogeniteit een rol speelt op een school. Leerlingen op scholen met veel allochtonen scoren slechter (zie ook Veenstra, 1999; Veerman & al, 2013). Etnische diversiteit, gemeten in het aantal etnische groepen, blijkt geen effect te hebben en dit is in overeenstemming met het onderzoek van Roeleveld & al (2011).

In de volgende modellen worden variabelen met betrekking tot schaal toegevoegd. In model M₃₀ worden variabelen gekoppeld aan leerlingaantallen toegevoegd. De eerste variabele geeft aan of een school *145 leerlingen of minder* heeft¹² (zodat de school in aanmerking komt voor de kleinescholentoeslag) en de tweede variabele geeft aan of een school *500 leerlingen of meer* heeft. Deze laatste maat is gekozen, omdat eerder is gebleken dat er vanaf de maat van ongeveer 500 leerlingen geen financiële schaalvoordelen te behalen zijn.

¹² De maat kleiner stellen, bijvoorbeeld op 100 leerlingen, leverde een klein aantal scholen op.

	Model M27	Model M28	Model M29
<i>Log Likelihood</i>	-10.202	- 10.201	- 10.201
<i>Aantal observaties</i>	3.405	3.405	3.405
<i>AIC</i>	20.461	20.457	20.450
Fixed			
<i>Constante</i>	468,84	468,35	468,20
Fixed			
Opleiding Vader HBO-WO	0,88 *** (0.215)	0,89 *** (0.215)	0,90 *** (0.215)
Opleiding Moeder MBO	1,04 *** (0.207)	1,02 *** (0.207)	1,01 *** (0.207)
Opleiding Moeder HBO-WO	1,92 *** (0.276)	1,92 *** (0.276)	1,90 *** (0.276)
Geboorteland Vader Turkije	- 0,93 ** (0.392)	- 0,90 ** (0.391)	- 0,80 ** (0.393)
Zorgleerling	- 1,90 *** (0.237)	- 1,90 *** (0.237)	- 1,92 *** (0.237)
Zelfvertrouwen	1,32 *** (0.147)	1,33 *** (0.147)	1,33 *** (0.147)
Werkhouding	0,82 *** (0.108)	0,83 *** (0.108)	0,83 *** (0.108)
Afhankelijkheid	- 0,55 *** (0.125)	- 0,54 *** (0.125)	- 0,54 *** (0.125)
Ouderbetrokkenheid	0,74 *** (0.114)	0,74 *** (0.114)	0,74 *** (0.114)
Welbevinden medeleerlingen	- 0,54 *** (0.132)	- 0,54 *** (0.132)	- 0,54 *** (0.132)
Toets Rekenen / Wiskunde	0,46 *** (0.009)	0,46 *** (0.009)	0,46 *** (0.009)
Toets Taal	0,08 *** (0.007)	0,08 *** (0.007)	0,08 *** (0.007)
Sterk stedelijk	0,51 (0.801)		
Matig stedelijk	- 0,74 (0.836)		
Weinig stedelijk	- 1,07 (0.859)		
Niet stedelijk	- 2,37 *** (0.917)	- 1,97 *** (0.595)	- 1,84 *** (0.597)
Schoolscore 105-109	- 0,17 (0.621)		
Schoolscore 110-119	0,58 (0.665)		
Schoolscore 120-139	- 2,07 *** (0.775)	- 1,69 ** (0.674)	- 1,58 ** (0.748)
Schoolscore 140-159	- 1,65 * (0.929)	- 1,11 (0.802)	- 0,35 (1.106)
Schoolscore > 159	- 3,51 *** (1.119)	- 3,12 *** (0.974)	- 1,47 (1.261)

<i>Vervolg</i>	Model M27	Model M28	Model M29
Protestants Christelijk	1,00 * (0.607)	1,09 * (0.596)	0,83 (0.584)
Rooms Katholiek	1,23 ** (0.532)	1,13 ** (0.511)	0,81 (0.498)
Algemeen Bijzonder	0,51 (2.168)		
Islamitisch	0,65 (1.390)		
Overig Bijzonder	- 0,46 (1.296)		
Taalverbetertraject		- 0,36 (0.470)	
Rekenverbetertraject		0,60 (0.615)	
Percentage allochtonen >= 50%			- 0,50 (1.121)
Percentage allochtonen >= 75%			- 2,27 ** (1.050)
Aantal etnische groepen			0,22 (0.137)
Random			
var(Residual)	21,0	21,0	21,0
var(School)	7,5	7,7	7,5

Tabel 14: Geschatte Modellen M₂₇ tot M₂₉

Het toevoegen van beide variabelen zorgt ervoor dat het effect van de protestants christelijke denominatie nog minder sterk wordt. Het heeft geen significant effect op de CITO-score van leerlingen of een school 145 leerlingen of minder heeft. Het effect van een grote school is wel significant. Leerlingen op scholen met 500 leerlingen of meer scoren 1,55 punt lager op hun CITO-toets. Wetenschappelijke onderzoeken zijn tegenstrijdig op dit punt, maar het lijkt er dus op dat leerlingen op kleinere scholen, gecorrigeerd voor leerlingkenmerken, geen slechtere gemiddelde CITO-scores halen, terwijl er van grote scholen juist wel een negatief schaafeffect uitgaat.

De niet significante variabelen *schoolscore 140-159*, *aantal leerlingen <=145* en *protestants-christelijke denominatie* uit model M₃₀ worden vanaf nu buiten beschouwing gelaten. In model M₃₁ wordt de variabele *aantal scholen per schoolbestuur* toegevoegd, onderverdeeld naar 6 tot 10 scholen, 11 tot 20 scholen en meer dan 20 scholen. Deze variabele blijkt geen enkel effect te hebben op de CITO-scores van leerlingen. Dat de schaalgrootte van het bestuur geen effect heeft, is in overeenstemming met een onderzoek van Dijkgraaf & al (2009), dat geen effect vond van bestuursschaalgrootte in het middelbaar onderwijs op leeropbrengsten.

		Model M30	Model M31	Model M32
<i>Log Likelihood</i>		- 10.200	- 10.202	- 10.204
<i>Aantal observaties</i>		3.405	3.405	3.405
<i>AIC</i>		20.450	20.451	20.445
<i>Constante</i>		468,94	469,45	469,39
Fixed	Opleiding Vader HBO-WO	0,91 *** (0.215)	0,91 *** (0.215)	0,92 *** (0.215)
	Opleiding Moeder MBO	1,01 *** (0.207)	1,02 *** (0.207)	1,03 *** (0.207)
	Opleiding Moeder HBO-WO	1,90 *** (0.276)	1,91 *** (0.275)	1,92 *** (0.275)
	Geboorteland Vader Turkije	- 0,77 ** (0.393)	- 0,77 ** (0.392)	- 0,76 * (0.392)
	Zorgleerling	- 1,93 *** (0.237)	- 1,93 *** (0.237)	- 1,93 *** (0.237)
	Zelfvertrouwen	1,33 *** (0.147)	1,32 *** (0.147)	1,32 *** (0.147)
	Werkhouding	0,83 *** (0.108)	0,83 *** (0.108)	0,83 *** (0.108)
	Afhankelijkheid	- 0,54 *** (0.125)	- 0,55 *** (0.125)	- 0,54 *** (0.125)
	Ouderbetrokkenheid	0,73 *** (0.114)	0,73 *** (0.114)	0,74 *** (0.114)
	Welbevinden medeleerlingen	- 0,53 *** (0.132)	- 0,53 *** (0.132)	- 0,53 *** (0.132)
	Toets Rekenen / Wiskunde	0,46 *** (0.009)	0,46 *** (0.009)	0,459 *** (0.009)
	Toets Taal	0,08 *** (0.007)	0,08 *** (0.007)	0,077 *** (0.007)
	Niet stedelijk	- 2,10 *** (0.602)	- 2,17 *** (0.580)	- 2,22 *** (0.573)
	Schoolscore 120-139	- 1,25 * (0.682)	- 1,27 * (0.665)	- 1,10 * (0.654)
	Schoolscore 140-159	- 0,22 (0.857)		
	Schoolscore > 159	- 1,47 (1.160)	- 1,28 (1.099)	
	Protestants Christelijk	0,77 (0.587)		
	Rooms Katholiek	0,87 * (0.509)	0,62 (0.451)	
	Percentage allochtonen >= 75%	- 2,10 ** (0.867)	- 2,41 *** (0.776)	- 2,93 *** (0.674)
	Aantal leerlingen <= 145	- 0,12 (0.506)		
	Aantal leerlingen >= 500	- 1,55 * (0.907)	- 1,39 (0.921)	- 1,38 (0.906)
	6-10 scholen per bestuur		- 0,44 (0.744)	
	11-20 scholen per bestuur		- 0,27 (0.652)	
	>21 scholen per bestuur		- 0,26 (0.693)	
Random	var(Residual)	21,0	21,0	21,0
	var(School)	7,5	7,5	7,6

Tabel 15: Geschatte Modellen M₃₀ tot M₃₂

In model M_{31} verandert een aantal variabelen van significant naar niet significant. De variabelen *aantal scholen per schoolbestuur*, *schoolscore >159* en *rooms-katholieke denominatie* worden in vervolgmogelijken daarom niet meer meegenomen. Verschillende andere variabelen met betrekking tot schoolkenmerken zijn na model M_{31} nog toegevoegd, zoals het aandeel FTE deeltijd (0-0.5 FTE), het aandeel FTE vrouwen en het aandeel FTE management op het totaal aantal FTE, de gemiddelde leeftijd van het personeel en de verhouding tussen vast en tijdelijk personeel. Deze variabelen bleken allemaal geen significante invloed op het model te hebben en om ruimte te besparen worden deze modellen niet weergegeven.

Model M_{32} is het model dat gebruikt wordt om hypothese 4 te beoordelen. In hypothese 4 werd gesteld dat de CITO-score relatief hoger is als:

- ✓ De leerling op een school in Zeeland of Limburg zit;
- ✓ De leerling op een school in een minder verstedelijkt gebied zit;
- ✓ De leerling op een school zit met een laag schoolgewicht;
- ✓ De leerling niet op een islamitische school zit;
- ✓ De leerling op een school zit waar taal- en rekenverbetertrajecten zijn ingezet;
- ✓ De leerling op een school zit met een homogene leerlingpopulatie;
- ✓ De leerling op een school zit groter dan 145 leerlingen;
- ✓ De leerling op een school zit die valt onder een klein schoolbestuur.

Aan de hand van model 32 kan vergelijking [V.5] worden opgesteld:

$$\begin{aligned}
 CITO - score = & [469,39 + \mu_{school}] + \alpha X_{optvader-hbowo} + \alpha X_{optmoeder-mbo} \\
 & + \alpha X_{optmoeder-hbowo} - \alpha X_{gl-turkije} - \alpha X_{zorgleerling} + \alpha X_{zelfvertrouwen} \\
 & + \alpha X_{werkhouding} - \alpha X_{afhankelijkheid} + \alpha X_{ouderbetrokkenheid} \\
 & - \alpha X_{welbevinden-medeleerlingen} + \alpha X_{toets-rekenen} + \alpha X_{toets-taal} \\
 & - \alpha X_{niet-stedelijk} - \alpha X_{schoolscore-120-139} - \alpha X_{allochtoon \geq 75\%} + \varepsilon_{residual}
 \end{aligned}$$

Met behulp van de geschatte parameters uit model M_{32} kunnen de coëfficiënten α worden ingevuld, waardoor vergelijking [V.6] ontstaat:

$$\begin{aligned}
 CITO - score = & [469,39 + \mu_{7,6}] + 0,92 * X_{optvader-hbowo} + 1,03 * X_{optmoeder-mbo} \\
 & + 1,92 * X_{optmoeder-hbowo} - 0,73 * X_{gl-turkije} - 1,93 * X_{zorgleerling} \\
 & + 1,32 * X_{zelfvertrouwen} + 0,83 * X_{werkhouding} - 0,54 * X_{afhankelijkheid} \\
 & + 0,74 * X_{ouderbetrokkenheid} - 0,543 X_{welbevinden-medeleerlingen} \\
 & + 0,459 * X_{toets-rekenen} + 0,077 * X_{toets-taal} - 2,22 * X_{niet-stedelijk} \\
 & - 1,10 * X_{schoolscore-120-139} - 2,93 * X_{allochtoon \geq 75\%} + \varepsilon_{residual}
 \end{aligned}$$

Wat betreft de variabelen op schoolniveau kan uit vergelijking [V.6] worden afgeleid dat de provincie waarin de school staat geen invloed heeft op de CITO-score. De mate van stedelijkheid heeft wel een impact; een leerling op een school in een niet stedelijk gebied scoort ruim 2 punten lager op de CITO-toets.

Wellicht dat scholen op het platteland meer moeite hebben met het aantrekken van voldoende gekwalificeerd personeel. Een andere verklaring kan zijn dat scholen in niet stedelijk gebied minder concurrentie hebben van scholen in de buurt. Soms zijn scholen de enige in het dorp en meestal zijn de scholen de enige van een bepaalde denominatie in het dorp, waardoor scholen minder hoeven te concurreren.

Het schoolgewicht voegt weinig verklarende kracht toe aan het model. Alleen leerlingen op een school met een schoolgewicht tussen de 120 en 139 scoren significant slechter. Dat leerlingen op een school met nog een hoger schoolgewicht niet slechter scoren, kan deels worden verklaard door de toegevoegde variabele *75% of meer allochtonen*. De correlatie van deze variabele met de variabelen *schoolgewicht 140-159* en *>159* is vrij hoog, namelijk ($r=0.4411$) respectievelijk ($r=0.4912$). De correlatie tussen *schoolgewicht 120-139* en de variabele *75% of meer allochtonen* is daarentegen nihil ($r=0.0692$).

Denominatie heeft uiteindelijk geen effect op de CITO-score, hoewel er van protestants-christelijke en rooms-katholieke scholen in een aantal modellen een significant positief effect uitging op de CITO-score van hun leerlingen.

Leerlingen op een school waarop taal- en rekenverbetertrajecten zijn ingevoerd, scoren niet anders op de CITO-score dan leerlingen op scholen, waarop deze trajecten niet zijn ingevoerd. Dat er geen verschil is tussen CITO-scores is eigenlijk goed nieuws, omdat reken- en taalzwakke scholen voorheen significant slechter scoorden dan gemiddeld. In de toekomst worden de effecten van onderwijsverbetertrajecten misschien nog groter. Het kan bijvoorbeeld zo zijn dat leerlingen die nu in groep 8 zitten niet alle vruchten van het programma hebben kunnen plukken, omdat ze maar een deel van het traject hebben gevolgd.

De samenstelling van de leerlingpopulatie is van belang. Scholen met een zware leerlingpopulatie scoren over het algemeen slechter (model M_{27} / M_{28}). In latere modellen wordt dit verband meer verklaard door de variabele *75% allochtonen of meer* op een school. Het aantal etnische groepen op een school heeft geen directe verklarende kracht. Het percentage allochtonen op een school zegt niet direct dat de populatie niet homogeen is (het kan tenslotte gaan om dezelfde bevolkingsgroep), maar een hoog percentage allochtone leerlingen zorgt wel voor slechtere CITO-scores.

De schaal van een school heeft geen significante invloed op de resultaten van leerlingen. Leerlingen op scholen met minder dan 145 leerlingen scoren niet slechter dan gemiddeld. Leerlingen op scholen met meer dan 500 leerlingen scoren wel slechter, gemiddeld 1,38 punt, maar dit effect is niet significant. Of een school onder een groot of klein schoolbestuur valt heeft geen invloed op de prestaties van leerlingen.

Hypothese 4, zoals die is opgesteld, wordt verworpen. Slechts enkele punten van de hypothese bleken overeen te komen met de werkelijkheid. Leerlingen bleken slechter te scoren op scholen in niet stedelijk gebied en op scholen met een hoog percentage allochtonen. De andere variabelen, zoals die genoemd zijn in hypothese 4, hebben geen significant verklarend effect.

7.3 Definitieve modellen

Uit de vorige paragraaf blijkt dat variabelen op schoolniveau weinig verklaringkracht toevoegen aan het model. De AIC van de modellen M_{24} tot M_{32} is, ondanks de toevoeging van variabelen op het schoolniveau, nauwelijks veranderd (M_{24} 20.474; M_{32} 20.445). Het random schooleffect wordt door toevoeging van de variabelen wel kleiner, maar al met al kan gesteld worden dat variabelen op schoolniveau nauwelijks bijdragen aan een beter model.

Een model is beter, als het met zo weinig mogelijk termen zo veel mogelijk variantie verklaart. Daarom wordt in model M_{33} de variabele *schoolscore 120-139* verwijderd, in model M_{34} de variabele *500 leerlingen of meer*, in model M_{35} de variabele *niet stedelijk gebied* en in model M_{36} de variabele *75% of meer allochtonen*. Door deze aanpassingen verandert de AIC van de modellen nauwelijks. Eventuele verdere aanpassingen zouden wel zorgen voor grotere verschuivingen in de AIC. In de modellen M_{33} tot M_{36} neemt de tussenschoolse variantie wel toe; dit is begrijpelijk, omdat verschillende verklarende variabelen op schoolniveau worden verwijderd. De definitieve modellen M_{33} tot M_{36} zijn te vinden in tabel 16.

Nu de definitieve modellen zijn vastgesteld wordt de *intraclass correlation coefficient* (IC) voor het schooleffect nogmaals berekend. De IC voor het schooleffect geeft aan welk deel van de variantie in de data verklaard wordt door de variantie tussen scholen. In de modellen M_{33} tot M_{36} is de variantie tussen leerlingen 21,0 en tussen scholen variërend van 7,8 tot 9,3. Door het verwijderen van variabelen op schoolniveau in de modellen neemt de variantie tussen scholen begrijpelijkerwijs toe. De IC-coëfficiënten voor het schooleffect zijn als volgt:

$$IC_{\text{schooleffect}-M_{33}} = \frac{7,8}{7,8 + 21,0} \approx 0,270$$

$$IC_{\text{schooleffect}-M_{36}} = \frac{9,3}{9,3 + 21,0} \approx 0,307$$

De tussenschoolse variantie komt uit op ongeveer 27% tot 31% van de onverklaarde variantie. De leerlingvariantie bepaalt ongeveer 69% tot 73% van de onverklaarde variantie. Het toevoegen van het random effect op schoolniveau is getuige een LR-test zinvol ($p=0.000$).

De verklaringkracht van de modellen is ten opzichte van het startmodel M_{0B} vrij groot. De R^2 van de modellen M_{33} en M_{36} is als volgt:

$$R_{M_{33}}^2 = \left[1 - \left[\frac{21,0 + 7,8}{86,4} \right] \right] * 100 \approx 66,7\%$$

$$R_{M_{36}}^2 = \left[1 - \left[\frac{21,0 + 9,3}{86,4} \right] \right] * 100 \approx 64,9\%$$

Uiteindelijk wordt ongeveer 66% van de variantie in CITO-scores verklaard door de definitieve modellen.

	Model M33	Model M34	Model M35	Model M36
<i>Log Likelihood</i>	- 10.205	- 10.206	- 10.212	- 10.220
<i>Aantal observaties</i>	3.405	3.405	3.405	3.405
<i>AIC</i>	20.446	20.446	20.456	20.470
<i>Constante</i>	469,25	469,13	468,76	468,27
Fixed				
Opleiding Vader HBO-WO	0,92 *** (0.215)	0,92 *** (0.215)	0,92 *** (0.215)	0,92 *** (0.215)
Opleiding Moeder MBO	1,03 *** (0.207)	1,03 *** (0.207)	1,02 *** (0.207)	1,09 *** (0.206)
Opleiding Moeder HBO-WO	1,93 *** (0.275)	1,92 *** (0.275)	1,91 *** (0.276)	1,98 *** (0.275)
Geboorteland Vader Turkije	- 0,76 * (0.392)	- 0,79 ** (0.392)	- 0,76 * (0.393)	- 0,99 ** (0.389)
Zorgleerling	- 1,94 *** (0.237)	- 1,94 *** (0.237)	- 1,95 *** (0.237)	1,92 *** (0.238)
Zelfvertrouwen	1,32 *** (0.147)	1,32 *** (0.147)	1,33 *** (0.147)	1,31 *** (0.147)
Werkhouding	0,83 *** (0.108)	0,83 *** (0.108)	0,83 *** (0.108)	0,82 *** (0.108)
Afhankelijkheid	- 0,54 *** (0.125)	- 0,54 *** (0.125)	- 0,54 *** (0.125)	- 0,55 *** (0.126)
Ouderbetrokkenheid	0,73 *** (0.114)	0,74 *** (0.114)	0,72 *** (0.114)	0,73 *** (0.114)
Welbev. medeleerlingen	- 0,53 *** (0.132)	- 0,53 *** (0.132)	- 0,53 *** (0.132)	- 0,53 *** (0.132)
Toets Rekenen / Wiskunde	0,459 *** (0.009)	0,459 *** (0.009)	0,459 *** (0.009)	0,461 *** (0.009)
Toets Taal	0,077 *** (0.007)	0,077 *** (0.007)	0,077 *** (0.007)	0,077 *** (0.007)
Percent. alloctonen >= 75%	- 2,99 *** (0.678)	- 3,17 *** (0.674)	- 2,77 *** (0.681)	
Niet stedelijk	- 2,10 *** (0.573)	- 2,01 *** (0.575)		
Aantal leerlingen >= 500	- 1,53 * (0.908)			
Random				
var(Residual)	21,0	21,0	21,0	21,0
var(School)	7,8	7,9	8,4	9,3

Tabel 16: Definitieve Modellen M₃₃ tot M₃₆

8. Conclusie, tekortkomingen en aanbevelingen

8.1 Beantwoording hypothesen

Ter beantwoording van de probleemstelling zullen de belangrijkste resultaten worden herhaald. De probleemstelling is als volgt gedefinieerd:

‘In hoeverre zijn er schaalvoordelen te behalen in het basisonderwijs en welke invloed heeft de mogelijke aanwezigheid van schaaffecten op de kwaliteit van het basisonderwijs?’

Aan de hand van deze probleemstelling zijn vier hypothesen geformuleerd, drie over financiële schaaffecten en één over onderwijskundige schaaffecten.

Hypothese 1

De gemiddelde kosten per leerling zijn lager als de organisatievorm meer leerlingen heeft, maar er is een omslagpunt tussen schaalvoordelen en schaalnadelen waar te nemen (MES).

Hypothese 1 is gemeten op bestuursniveau in het aantal leerlingen, op schoolniveau in het aantal leerlingen en op bestuursniveau in het aantal scholen dat wordt aangestuurd. Uit de analyse blijkt dat een minimale bestuurs- en schoolgrootte van 150 leerlingen vanuit kosten oogpunt wenselijk is. Tot deze grootte kunnen scholen namelijk grote kostenvoordelen behalen. Tot een grootte van ongeveer 400 tot 500 leerlingen zijn er nog beperkte schaalvoordelen te behalen. Als een school of bestuur meer dan 500 leerlingen aanstuurt, verdwijnt het effect van schaalgrootte op de kosten en loopt de kostencurve vlak. Er vindt geen omslag plaats tussen schaalvoordelen en schaalnadelen, wat wil zeggen dat scholen wel schaalvoordelen maar geen schaalnadelen kunnen behalen.

De analyse van hypothese 1 op bestuursniveau gemeten in het aantal scholen laat een ander beeld zien. Hier zien we namelijk een omslagpunt van schaalvoordelen naar schaalnadelen. Schoolbesturen kunnen schaalvoordelen behalen als ze tot maximaal tien scholen aansturen. Schoolbesturen die zestien of meer scholen aansturen behalen significante schaalnadelen.

Hypothese 1 wordt op school- en bestuursniveau gemeten in het aantal leerlingen verworpen, omdat er geen omslag plaatsvindt tussen schaalvoordelen en schaalnadelen. Als echter gekeken wordt naar schaalgrootte gemeten in het aantal scholen per schoolbestuur, nemen we zo'n omslagpunt wel waar. Op bestuursniveau gemeten in het aantal besturen kunnen we hypothese 1 dus niet verwerpen.

Hypothese 2

Scholen met grotere klassen kennen lagere gemiddelde kosten per leerling.

Hypothese 2 is gemeten op basis van de leerling-docent verhouding. Uit de analyse blijkt dat een schoolbestuur met een hogere leerling-docent verhouding lagere kosten kent dan een schoolbestuur met een lage leerling-docent verhouding. De kostentrendlijn loopt uiteindelijk iets op, maar grotere klassen blijven significant goedkoper dan gemiddeld. Klassen van 19 tot 25 leerlingen blijken het meest kostenefficiënt. Hypothese 2 wordt niet verworpen.

Hypothese 3

Islamitische schoolbesturen hebben gemiddeld hogere kosten per leerling, terwijl de gemiddelde kosten per leerling bij andere schoolbesturen niet significant van elkaar afwijkend zullen zijn.

In de analyse naar verschillen tussen denominaties bleek dat reformatorische en protestants-christelijke schoolbesturen per leerling significant lagere kosten halen dan gemiddeld, terwijl de kosten per leerling bij openbare en islamitische schoolbesturen significant hoger zijn dan gemiddeld. Islamitische schoolbesturen zijn per leerling tot wel € 1.100 duurder dan gemiddeld. Dit hoeft niet te betekenen dat islamitische schoolbesturen niet efficiënt met de beschikbare middelen omgaan. Mogelijke verklaringen voor de hogere kosten per leerling op islamitische scholen liggen in de kleinere klassen, het gemiddelde opleidingsniveau van de ouders en de hoeveelheid achterstandsl leerlingen. Hypothese 3 wordt verworpen, want er blijken significante verschillen te zijn tussen de verschillende denominaties.

Hypothese 4

De CITO-score van een leerling is relatief hoger als:

- ✓ *De leerling op een school in Zeeland of Limburg zit;*
- ✓ *De leerling op een school in een meer verstedelijkt gebied zit;*
- ✓ *De leerling op een school zit met een laag schoolgewicht;*
- ✓ *De leerling niet op een islamitische school zit;*
- ✓ *De leerling op een school zit waar taal- en rekenverbetertrajecten zijn ingezet;*
- ✓ *De leerling op een school zit met een homogene leerlingpopulatie;*
- ✓ *De leerling op een school zit groter dan 145 leerlingen;*
- ✓ *De leerling op een school zit die valt onder een klein schoolbestuur.*

Om de effecten van schoolkenmerken te bepalen moet eerst worden gecorrigeerd voor algemene en schoolse leerling- en gezinskenmerken. Aan de hand van literatuuronderzoek zijn een tweetal uitgangspunten opgesteld. Deze uitgangspunten zijn eerst onderzocht, waarbij gepoogd is resultaten van eerdere studies te repliceren. De resultaten met betrekking tot de algemene en schoolse leerlingkenmerken zullen eerst besproken worden, alvorens hypothese 4 beoordeeld zal worden.

Algemene leerling- en gezinskenmerken

Uit eerder onderzoek is gebleken dat de CITO-score van een leerling is relatief hoger als:

- ✓ *De leerling van het mannelijk geslacht is;*
- ✓ *De leerling ouders heeft met een hoog opleidingsniveau;*
- ✓ *De leerling autochtone ouders heeft;*
- ✓ *De leerling langer in Nederland verblijft;*
- ✓ *De leerling leeft in een tweeoudergezin;*
- ✓ *De leerling geen gedragsproblemen en gezondheidsproblemen heeft.*

De algemene leerling- en gezinskenmerken zijn beoordeeld aan de hand van model M₁₁. Uit de analyse blijkt dat de CITO-score van een jongen tot wel 1,7 punt hoger is dan de CITO-score van een meisje. Het opleidingsniveau van beide ouders heeft een groot effect. Leerlingen met ouders die beiden MBO hebben gevolgd scoren tot 4,75 punt hoger, leerlingen met ouders die HBO of WO hebben gevolgd scoren zelfs tot 10 punten hoger.

Een leerling scoort lager als deze een Turkse (-2,78), Marokkaanse (-1,85) of Surinaams Antilliaanse (-3,76) vader heeft. Als een leerling langer dan 5 jaar of altijd in Nederland verblijft, is de score tot 3,3 punt hoger. De gezinssamenstelling heeft ook effect op de CITO-score. Als een kind gescheiden ouders heeft, is de CITO-score 1,5 punt lager. De mate van goed gedrag heeft een effect van maximaal 5,5 punt. Als een leerling is aangemerkt als zorgleerling, is de CITO-score tot 7 punten lager. De gevonden effecten zijn allemaal in lijn met eerdere onderzoeken.

Schoolse leerling- en gezinskenmerken

Uit eerder onderzoek is gebleken dat de CITO-score van een leerling is relatief hoger als:

- ✓ De leerling taakgemotiveerd is;
- ✓ De leerling een positief zelfbeeld en een hoge mate van zelfstandigheid heeft;
- ✓ De leerling een goede werkhouding heeft.
- ✓ De leerling betrokken ouders heeft;
- ✓ De leerling een goede relatie heeft met docenten en medeleerlingen;
- ✓ De leerling op eerdere toetsen hoog scoorde.

De schoolse leerling- en gezinskenmerken zijn beoordeeld aan de hand van model M_{23} . Een aantal algemene leerling- en gezinskenmerken hebben in dit model geen significant effect meer op de CITO-score. Het gaat dan om het geslacht, de verblijfsduur, de gezinssamenstelling, of de leerling een Marokkaanse, Surinaamse of Antilliaanse vader heeft, het opleidingsniveau van de vader op LBO en MBO-niveau en het opleidingsniveau van de moeder op LBO-niveau.

Het effect van eerder gemaakte toetsen is erg groot. Als een leerling op een taal- of rekentoets goed gescoord heeft, is de kans erg groot dat de leerling ook een hoge CITO-score haalt. Verder blijkt dat de score van een leerling hoger is als de leerling een goede werkhouding laat zien (wat hoog correleert met het gedrag van de leerling). Als de leerling veel zelfvertrouwen heeft (wat hoog correleert met taakmotivatie), heeft dit ook een gunstig effect op de CITO-score. Leerlingen die een bepaalde zelfstandigheid tonen in de klas en leerlingen met betrokken ouders scoren significant hoger. De relatie met de docent heeft geen significant effect op de hoogte van de CITO-score. Een andere variabele die een afwijkend beeld laat zien, is het welbevinden van de leerling met andere leerlingen. Deze variabele heeft een negatief effect op de CITO-score: als het welbevinden hoger is, is de CITO-score lager. Een verklaring hiervoor kan zijn, dat de variabele niet goed is samengesteld en de variantie tussen leerlingen niet goed verklaart. Een andere verklaring kan zijn, dat het effect wel echt is en dat leerlingen zichzelf en elkaar stoerder vinden als er lage cijfers worden gehaald.

Enkele effecten uit eerdere onderzoeken worden niet gerepliceerd, zoals het effect van welbevinden met de docent en welbevinden met medeleerlingen. De andere schoolse leerling- en gezinskenmerken zijn echter in lijn met wat vanuit de literatuurstudie werd verwacht.

Hypothese 4 is opgesteld naar aanleiding van de schoolkenmerken en beoordeeld aan de hand van model M_{32} . In dit model zijn alle algemene en schoolse leerling- en gezinskenmerken uit model M_{23} overleefd.

De invloed van schoolkenmerken op de CITO-score was kleiner dan verwacht. Leerlingen in de provincie Limburg scoren significant beter dan leerlingen elders in het land, maar verder was er geen effect van provincie op de CITO-score waar te nemen. Het effect van de mate van stedelijkheid viel alleen te ontdekken in niet stedelijk gebied; leerlingen op het platteland scoren tot ruim 2 punten lager op de CITO-toets. Het schoolgewicht heeft invloed op de CITO-score, maar het percentage allochtonen (wat aanzienlijk correleert met het schoolgewicht) heeft meer invloed. Leerlingen op scholen met $\geq 75\%$ allochtonen scoren tot bijna 3 punten lager op de CITO-toets.

Het effect van denominatie is verwaarloosbaar. Ook schaalgrootte lijkt weinig effect te hebben op de CITO-score. Leerlingen op kleine scholen tot 145 leerlingen blijken ook niet slechter te scoren op de CITO-toets. Daarentegen scoren leerlingen op scholen met meer dan 500 leerlingen in verschillende modellen tot 1,5 punt slechter. De grootte van het schoolbestuur heeft geen invloed op de prestaties van leerlingen. Of een school deelneemt aan onderwijsverbetertrajecten heeft ook geen invloed op de CITO-scores van leerlingen.

Hypothese 4 dient verworpen te worden, want weinig van de in hypothese 4 vastgestelde schoolkenmerken hebben effect op de CITO-score. Een aantal schoolkenmerken hebben wel een significant effect op de CITO-score. Leerlingen op scholen in niet stedelijk gebied en leerlingen op scholen met een heterogene leerlingenpopulatie ($\geq 75\%$ of meer allochtoon) scoren slechter. Het effect van schaalgrootte is onzeker, maar wat wel blijkt is dat leerlingen op kleine scholen geen significant slechtere CITO-scores halen. Het tegendeel lijkt juist het geval: leerlingen op grote scholen hebben last van schaalgrootte, waardoor ze tot 1,5 punt lagere CITO-scores halen.

8.2 Beantwoording probleemstelling

'In hoeverre zijn er schaalvoordelen te behalen in het basisonderwijs en welke invloed heeft de mogelijke aanwezigheid van schaaleffecten op de kwaliteit van het basisonderwijs?'

Kleine scholen in het basisonderwijs kunnen significante schaalvoordelen behalen. Op scholen met 50 leerlingen of minder zijn de gemiddelde vaste kosten per leerling € 7.211 en op scholen met 100 leerlingen of minder € 5.406, terwijl het gemiddelde van de hele steekproef uitkomt op € 4.238 per leerling. Tot een grootte van ongeveer 150 leerlingen zijn de te behalen kostenvoordelen groot. Daarna neemt het effect van schaalgrootte op de kosten af. Tot een schaalgrootte van ongeveer 500 leerlingen zijn er nog kleine schaalvoordelen te behalen. Voor scholen groter dan 500 leerlingen zijn er geen schaaleffecten gevonden.

Schoolbesturen die 1 school aansturen zijn niet significant duurder of goedkoper dan gemiddeld. Schoolbesturen die bestaan uit 2 tot maximaal 10 scholen behalen significant lagere kosten, terwijl schoolbesturen die 16 scholen of meer aansturen significant hogere kosten behalen dan gemiddeld. Het is het meest kostenefficiënt om schoolbesturen tot maximaal 10 scholen aan te laten sturen. Grotere schoolbesturen hebben last van de schaalgrootte en behalen significante schaalnadelen.

De analyse op groeps grootte geeft aan dat een groeps grootte van 19 tot 25 leerlingen het meest kostenefficiënt is. Om lesgroepen van deze grootte samen te kunnen stellen is een minimale schoolgrootte van ongeveer 150 leerlingen gewenst. Groepen van 18 leerlingen en kleiner zijn significant duurder dan gemiddeld.

De invloed van schaal op onderwijskwaliteit is klein. Er is niet gebleken dat leerlingen op kleine scholen met 145 leerlingen of minder lagere CITO-scores halen. Wel is er in verschillende modellen een effect gevonden dat wijst op schaalnadeel voor grote scholen. Op scholen met meer dan 500 leerlingen is de gemiddelde CITO-score tot 1,5 punt lager.

8.3 Tekortkomingen

Bij het uitvoeren van het onderzoek zijn een aantal tekortkomingen aan het licht gekomen. Als maatstaf voor de onderwijskwaliteit is uitgegaan van de CITO-scores, maar die scores kunnen niet worden gezien als een volledige afspiegeling van de onderwijskwaliteit. Kwaliteit is zo veel breder dan alleen toetscijfers. Onderwijskwaliteit bestaat ook uit het leren van sociale vaardigheden, normen en waarden, omzien naar elkaar, culturele waarden, persoonlijke ontwikkeling, enzovoort. Het is erg moeilijk om de volledige onderwijskwaliteit op een kwantitatieve manier te vatten, maar het is niet ondenkbaar dat de gevonden effecten anders zouden zijn als een vollediger maat van onderwijskwaliteit gebruikt zou kunnen worden.

In het onderzoek is uitgegaan van enkeljarige CITO-scores, terwijl het misschien beter zou zijn om een langjarig gemiddelde te nemen. Dit geldt voornamelijk voor de kleinere scholen, waar een uitschieter een grote invloed kan hebben op de gemiddelde score. Langjarige gemiddeldes zijn in het COOL-onderzoek echter niet beschikbaar en CITO-scores per leerling zijn er sowieso niet over een langere periode.

Een andere tekortkoming is het ontbreken van bepaalde belangrijke controlevariabelen (*'omitting bias'*). Hoewel geprobeerd is zoveel mogelijk controlevariabelen op zowel het leerling- als het schoolniveau mee te nemen in de modellen, zijn er waarschijnlijk variabelen niet meegenomen die wel een belangrijke invloed op de uitkomsten zouden kunnen hebben gehad. Dit heeft gevolgen voor de nu gevonden effecten; deze kunnen overschat zijn of zelfs volledig verdwijnen als er een variabele wordt toegevoegd die de variantie beter verklaart.

De gebruikte COOL-dataset geeft een rijk beeld van de leerling- en gezinsachtergrond. Het gebruiken van deze dataset brengt echter ook beperkingen met zich mee, die achteraf soms niet op te lossen waren. Zo is het de vraag of de variabele *welbevinden met medeleerlingen* op een bevredigende manier is geconstrueerd. Ook ontbreken er bij veel leerlingen en scholen cruciale gegevens die het aantal subjecten in de steekproef soms fors vermindert.

De onderwijskundige schaaleardeffecten zijn bepaald via modellen die stapje voor stapje opgebouwd zijn. Bij dat opbouwen is ervoor gekozen om de dataset niet op voorhand te beperken door missende gegevens te verwijderen. Deze keuze brengt het nadeel met zich mee dat de modellen onderling moeilijker te vergelijken zijn, omdat er sprake is van een (soms) verschillend aantal waarnemingen.

Doordat het aantal waarnemingen in de uitgebreide modellen nog maar ongeveer de helft is van het aantal waarnemingen in het startmodel, brengt dit de nodige beperkingen met zich mee. Zo is het aantal scholen met een algemeen bijzondere, islamitische en overig bijzondere denominatie te beperkt om algemene uitspraken over te kunnen doen. Ook zijn er soms variabelen waarvan er in bepaalde categorieën weinig waardes zijn, waardoor de effecten in werkelijkheid anders kunnen uitvallen. Als het aantal waarnemingen groter was geweest, zou met meer zekerheid kunnen worden vastgesteld of een effect juist is bepaald.

Een andere tekortkoming is de eenzijdige blik op de kosten in het onderzoek naar financiële schaaffecten. Of een school goedkoop is of duur, gemeten in kosten per leerling, zegt niet alles over het belang van een school op een bepaalde plaats. Hier geldt hetzelfde als bij de CITO-score als maatstaf voor de onderwijskwaliteit; een school is veel breder dan het financiële alleen. Verder zijn de precieze kosten alleen beschikbaar op het bestuursniveau en niet op het individuele schoolniveau, wat de toepasbaarheid van de analyses op het schoolniveau en met name het groepsniveau aanzienlijk beperkt. Bij schoolbesturen met een heterogeen schoolbestand is de zeggingskracht van de analyses klein, omdat het puur draait om gemiddeldes. Een andere beperking is dat de kostenanalyse alleen gebaseerd is op de personeelskosten. Hoewel de personeelslasten ruim 80% van de totale lasten vertegenwoordigen, kan het werkelijke beeld afwijken van het hier gevonden beeld omdat bepaalde kosten niet zijn meegenomen in de analyse.

De dataselectie van schoolbesturen is ook een beperking. In de steekproef bestuurt het grootste schoolbestuur 30 scholen, terwijl er in werkelijkheid ook schoolbesturen zijn die meer dan 70 scholen besturen. Door ontbrekende gegevens zijn verhoudingsgewijs veel grote schoolbesturen uit de analyse gevallen. Het beeld dat de Inspectie geeft van een complexere organisatie lijkt dus te kloppen, want veel grote schoolbesturen hebben geen volledige gegevens aangeleverd bij het DUO. Dit beperkt echter wel de zeggingskracht van de analyse wat betreft grote schoolbesturen. (Anderzijds zijn er van grote schoolbesturen altijd maar enkelen, waardoor deze besturen toch gegroepeerd hadden moeten worden, zodat het feitelijke beeld misschien niet veel afwijkt van het beeld zoals in dit onderzoek geschetst is).

8.4 Beleidsaanbevelingen en aanbevelingen tot vervolgonderzoek

De Rijksoverheid dient verdere schaalvergroting in het basisonderwijs met de nodige zorg tegemoet te treden. Schaalvergroting biedt aanzienlijke voordelen in bijvoorbeeld een betere benutting van arbeid en kapitaal, de verminderde kwetsbaarheid, een breder onderwijsaanbod en grotere personeelsmogelijkheden. Door de schaal van kleine scholen toe te laten nemen kunnen flinke kostenvoordelen worden behaald. Met het oog alleen op de financiën gericht is echter niet alles gezegd. De overheid dient te letten op bredere belangen. Schaalvergroting brengt ook de nodige nadelen met zich mee, die niet per se of direct in geld om te zetten zijn. Er kan hierbij gedacht worden aan een verminderde leefbaarheid van (kleine) dorpen, een gebrek aan (ouder)betrokkenheid en het verdwijnen van keuzevrijheid. Dit zijn aspecten die in beslissingen ook moeten worden meegewogen. Dit onderzoek richt zicht echter alleen op de eenvoudiger meetbare aspecten, zoals de kosten per leerling.

Met deze reserves in het achterhoofd, verdient het vanuit kosten oogpunt de aanbeveling dat er schaalvergroting plaatsvindt. Scholen opereren veel efficiënter als ze ten minste een schaalgrootte van ongeveer 150 leerlingen bereiken. Het onderzoek geeft echter ook grenzen van schaalvergroting aan. Er is namelijk geen enkele financiële ratio om scholen groter te laten worden dan ongeveer 500 leerlingen. Het verdient juist de aanbeveling dat scholen kleiner blijven dan 500 leerlingen omdat er aanwijzingen zijn dat, zelfs gecontroleerd voor diverse leerlingkenmerken, leerlingen slechtere CITO-scores behalen op grote scholen. Een mogelijkheid tot vervolgonderzoek is te bekijken of dit negatieve effect van grote scholen ook gevonden wordt in een grotere steekproef. Nu kan echter gesteld worden dat een school met ongeveer 200 tot 500 leerlingen vanuit zowel het kosten oogpunt als het oogpunt van de onderwijskwaliteit het meest ideaal is.

In de afgelopen 15 jaar heeft een enorme bestuurlijke schaalvergroting plaatsgevonden in het primair onderwijs, waarbij het totaal aantal besturen gehalveerd is. Er kan wel gesteld worden dat deze schaalvergroting in sommige gevallen is doorgeschoten. Uit het onderzoek blijkt dat schoolbesturen die 16 of meer schoolbesturen aansturen significante schaalnadelen behalen. Het verdient de aanbeveling dat de rijksoverheid kijkt naar maximumgrenzen die gesteld kunnen worden aan de bestuurlijke schaal van een organisatie in het onderwijs.

De verschillen tussen denominaties bleken aanzienlijk. Er zijn een aantal verklaringen aangedragen voor deze verschillen, maar in een vervolgonderzoek zou er dieper op deze verschillen ingegaan kunnen worden. Als blijkt dat reformatorische scholen ondanks een aantal controlevariabelen goedkoper opereren, kunnen daar lessen voor andere scholen uit worden getrokken. Aan de andere kant geldt dat, als blijkt dat islamitische scholen ondanks het meenemen van controlevariabelen duurder zijn dan andere scholen, zij lessen kunnen krijgen om de (financiële) organisatie beter op poten te zetten.

Uit het onderzoek naar onderwijskundige schaaffecten bleek dat op scholen die voor 75% of meer bestaan uit leerlingen met een allochtone achtergrond, de CITO-scores flink slechter zijn. Dit is een bekend segregatie-effect (zie bijvoorbeeld *Lafeber, 2006* en *Herweijer, 2008*). Verder onderzoek zou moeten uitwijzen of dit effect consistent is. Als dit effect inderdaad consistent is, zou de politiek (hoewel dit controversieel is) na kunnen denken over bepaalde richtlijnen om zogenaamde zwarte en blanke scholen te voorkomen.

Verder bleken de CITO-scores op scholen in niet stedelijk gebied slechter dan elders. Wellicht hangt het gevonden effect samen met het groter aantal kleine scholen op het platteland, maar over de precieze oorzaak van dit effect is weinig bekend. Het gevonden resultaat biedt een interessante mogelijkheid tot vervolgonderzoek.

Binnen het onderwijs ligt de nadruk op het (bestuurlijk) laten fuseren van scholen. Vervolgonderzoek zou zich kunnen toespitsen op de vraag of samenwerking tussen scholen niet beter is dan bestuurlijk fuseren. Wellicht kan men door bijvoorbeeld gezamenlijke inkoop of schoonmaak wel profiteren van de lusten, maar geen hinder ondervinden van de nadelen schaalvergroting.

Dankwoord

*‘Wat voor wijzen bleef verborgen,
is aan kinderen geopenbaard.’*

Mattheüs 11:25

Het onderwerp van deze scriptie komt voort uit mijn belangstelling voor zowel de economie als het onderwijs. Gedurende mijn bachelor Economie & Bedrijfseconomie heb ik alle kans gehad om mijn economische kennis te verdiepen. Als verbreding op deze studie heb ik in het derde jaar een halfjaar de educatieve minor gevolgd op een middelbare school. In deze periode heb ik ervaren dat een docent heel veel voor zijn of haar leerlingen kan betekenen, maar ook dat je als docent heel veel kunt leren van de wijsheid die kinderen op jonge leeftijd al bezitten. Het lag daarom voor de hand dat mijn scriptie zou gaan over een onderwerp op het snijvlak van economie en onderwijs. Met het onderzoek naar financiële en onderwijskundige schaaffecten hoop ik mijn steentje bij te kunnen dragen aan de verdere verbetering van het Nederlandse basisonderwijs.

Om uiteindelijk tot de scriptie te komen zoals die nu is, heb ik adviezen en tips gehad van een aantal mensen. Allereerst wil ik mijn begeleider, dr. Bas Karreman, bedanken. In de afgelopen maanden heeft hij mij begeleid tijdens het schrijfproces en heeft hij mij met zijn feedback steeds verder geholpen. Zijn prikkelende vragen waren aanleiding om zaken beter te onderzoeken of verder te bestuderen. Dank daarvoor.

Verder wil ik mijn ouders bedanken, die mij alle ruimte en gelegenheid hebben gegeven om te studeren aan de Erasmus. Ik wil hen ook bedanken dat ze de conceptversies van deze scriptie hebben gelezen. De adviezen die hieruit voortgekomen zijn heb ik gebruikt om de scriptie verder te verbeteren.

Daarnaast verdienen drs. Joris Voskuilen van de *Nederlandse Organisatie voor Wetenschappelijk Onderzoek* (NWO) en dr. Jaap Roeleveld van het *Kohnstamm Instituut* een woord van dank. Zonder hun medewerking was het onmogelijk geweest om het onderzoek met de COOL-data en de aanvullende gegevens uit te voeren.

Ten slotte spreek ik de hoop uit dat deze scriptie niet in een stoffige la beland, maar dat de gevonden resultaten gebruikt kunnen worden om vervolgonderzoek op te starten en om eventueel nieuw beleid vorm te geven.

Eric-Jan Troost

Literatuurlijst

- Besanko, D., & al, e. (2010). *Economics of Strategy*. John Wiley & Sons.
- Blank, J. (1990). Schaalvergroting in het basisonderwijs. *Economische Statistische Berichten*, 1156-1159.
- Blank, J., & al, (2007). *Basisonderwijs en bureaucratie; Een empirisch onderzoek naar de allocatie van middelen in basisscholen*. Delft/Rotterdam: Institute for Public Sector Efficiency Studies TU Delft/ECORYS.
- Blank, J., & Haelermans, C. (2008). Trends in onderwijsbureaucratie. *Tijdschrift voor Openbare Financiën*, 58-73.
- Bronneman-Helmers, H. (2009). De kwaliteit van het onderwijs. In W. D. Financiën, *Jaarboek Overheidsfinanciën 2009* (pp. 123-139). Den Haag: Sdu Uitgevers.
- CBS. (2013, april 16). *Kinderen met stiefouder scoren lager op de Citotoets*. Opgeroepen op juni 5, 2014, van CBS in de klas: <http://www.cbs.nl/nl-NL/menu/informatie/onderwijs/actueel/maatschappijleer/2013-3815-wm.htm>
- CJG. (2014). *Zelfvertrouwen*. Opgeroepen op juni 5, 2014, van Centrum voor Jeugd en Gezin: <http://www.cjg.nl/puber/groei-en-ontwikkeling/sociaalemotionele-ontwikkeling/zelfvertrouwen>
- COOL. (2012, April 26). Onderwijsloopbanen van 5-18 jaar - COOL 5-18 - Basisonderwijs 2010/2011; Tweede meting basisonderwijs 2010/2011. *Cohortonderzoek COOL5-18*. DANS Dataset 49841, <https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:49841>: Kohnstamm Instituut, Universiteit van Amsterdam; ITS, Radboud Universiteit Nijmegen.
- De Moor, A. (2009). Concurrentie en kwaliteit in het primair en voortgezet onderwijs. Amsterdam: TPEdigitaal.
- Dijkgraaf, E., & De Jong, M. (2009). Schaaleardeffecten en onderwijskwaliteit. *Economische Statistische Berichten*, 87-89.
- Driessen, G. (2007). *Opbrengsten van islamitische basisscholen*. Nijmegen: ITS, Radboud Universiteit Nijmegen.
- Driessen, G., Mulder, L., & Roeleveld, J. (2012). *Cohortonderzoek COOL5-18; Technisch rapport basisonderwijs, tweede meting 2010/11*. Nijmegen / Amsterdam: ITS / Kohnstamm Instituut.
- Dronkers, J. (2013). De berekening van de toegevoegde waarde van basisscholen., (Kennislunch OCW). Den Haag.
- DUO. (2010a, oktober 1). *01. Leerlingen basisonderwijs naar leerlinggewicht en per vestiging het schoolgewicht en impulsgebied*. Opgeroepen op juni 2014, van DUO Onderwijsdata:

- http://duo.nl/organisatie/open_onderwijsdata/databestanden/po/Leerlingen/Leerlingen/po_leerlingen1.asp
- DUO. (2014a, Januari 19). *07. Leerlingen primair onderwijs per bevoegd gezag naar denominatie en onderwijssoort*. Opgeroepen op April 2014, van DUO Onderwijsdata: http://www.ib-groep.nl/organisatie/open_onderwijsdata/databestanden/po/Leerlingen/Leerlingen/po_leerlingen7.asp
- DUO. (2014b, Maart 24). *02. Onderwijspersoneel in aantal fte*. Opgeroepen op April 2014, van DUO Onderwijsdata: http://www.ib-groep.nl/organisatie/open_onderwijsdata/databestanden/po/Onderwijspersoneel/Personeel/po_personeel_fte.asp
- DUO. (2014c, Januari 7). *12. Lasten*. Opgeroepen op April 2014, van DUO Onderwijsdata: http://www.ib-groep.nl/organisatie/open_onderwijsdata/databestanden/po/Financien/Jaarrekeninggegevens/Lasten.asp
- DUO. (2014d, Januari 7). *02. Staat van baten en lasten*. Opgeroepen op 2014 April, van DUO Onderwijsdata: http://www.ib-groep.nl/organisatie/open_onderwijsdata/databestanden/po/Financien/Jaarrekeninggegevens/Staat_baten_lasten.asp
- DUO. (2014e, maart 24). *02. Onderwijspersoneel in aantal fte*. Opgeroepen op juni 2014, van DUO Onderwijsdata: http://duo.nl/organisatie/open_onderwijsdata/databestanden/po/Onderwijspersoneel/Personeel/po_personeel_fte.asp
- Field, A. (2009). *Discovering statistics using SPSS*. London: SAGE Publications Ltd.
- Goes, L. (2009). *De invloed van intrinsieke- en extrinsieke motivatie op schoolprestaties; Etnische Achtergrond als moderator op de relatie tussen motivatie en schoolprestaties*. Utrecht: Universiteit Utrecht.
- Gramberg, P., & Jordanov, R. (2011). *"Achterstandsleerlingen, denominatie en stedelijkheid"*. Oberon / Bond KBO.
- Herweijer, L. (2008). Segregatie in het basis- en voortgezet onderwijs. In P. Schnabel, R. Bijl, & J. De Hart, *Betrekkelijke betrokkenheid, Studies in sociale cohesie, Sociaal en Cultureel Rapport* (pp. 206-233). Den Haag: Sociaal Cultureel Planbureau.
- Hilton, R., & al, (2008). *Cost Management: Strategies for Business Decisions*. New York: McGraw-Hill.
- IPSE. (2011). *Naar een optimale schaal van publieke voorzieningen; Een quick scan van de literatuur*. Delft: Centrum voor Innovatie en Publieke Sector Efficiëntie Studies, Technische Universiteit Delft.
- IPSE. (2012a). *Benchmark bedrijfsvoering voortgezet onderwijs; Kwantitatief onderzoek naar de bedrijfsvoering van instellingen in het voortgezet onderwijs*. Delft: Centrum voor Innovaties en Publieke Sector Efficiëntie Studies, Technische Universiteit Delft.

- IPSE. (2012b). *Productiviteitstrends in het primair onderwijs; Een empirisch onderzoek naar het effect van regulering op de productiviteitsontwikkeling tussen 1970-2010*. Delft: Centrum voor Innovaties en Publieke Sector Efficiënte Studies, Technische Universiteit Delft.
- IvhO. (2012). *Samenvattend rapport monitor verbetertrajecten taal en rekenen, 2008/2009, 2009/2010 en 2010/2011*. Utrecht: Inspectie van het Onderwijs.
- IvhO. (2013). *De staat van het onderwijs; Onderwijsverslag 2011/2012*. Utrecht: Inspectie van het Onderwijs.
- IvhO. (2014). *De staat van het Onderwijs - Onderwijsverslag 2012/2013*. Utrecht: Inspectie van het Onderwijs.
- JBO. (2014a). *Joodse Identiteit*. Opgeroepen op mei 16, 2014, van Joodse Basisschool Rosj Pina: http://www.jbo.nl/pags_rosjpina/identiteit/24.html
- JBO. (2014b). *Waartoe dient de ouderbijdrage?* Opgeroepen op mei 16, 2014, van Stichting Joodse Scholengemeenschap: http://www.jbo.nl/pags_jbo/ouderbijdrage/index.html
- JMouders. (2010, september 13). *Pesten maakt populair*. Opgeroepen op juni 14, 2014, van J/M voor ouders: Opvoeding: <http://www.jmouders.nl/Themas/Opvoeding/Opvoeden/Pesten-maakt-populair.htm>
- Karssen, M., Van der Veen, I., & Roeleveld, J. (2011). *Effecten van schoolsamenstelling op schoolprestaties in het Nederlandse basisonderwijs*. Amsterdam: Kohnstamm Instituut Universiteit van Amsterdam.
- Koomen, H. M., Spilt, J. L., Roorda, D. L., Oort, F., & Thijs, J. T. (2010). *Leraar-leerlingrelaties, schools leren van leerlingen en welbevinden van leraren*. Den Haag: NWO.
- Lafeber, J. (2006). *Segregatie op basisscholen; Een onderzoek naar de invloed van individuele preferenties van ouders op het segregatieniveau van basisscholen*. Groningen: Rijksuniversiteit Groningen.
- Menheere, A., & Hooge, E. (2010). *Ouderbetrokkenheid in het onderwijs; Een literatuurstudie naar de betekenis van ouderbetrokkenheid voor de schoolse ontwikkeling van kinderen*. Amsterdam: Kenniscentrum Onderwijs en Opvoeding, Hogeschool van Amsterdam.
- Noailly, J., & Koning, P. (2009). Schoolkeuze, concurrentie en kwaliteit in het basisonderwijs. *Economische Statistische Berichten*, 118-120.
- NOS. (2013, februari). *'Beperk aantal kleine scholen'*. Opgeroepen op mei 15, 2014, van NOS Binnenland: <http://nos.nl/artikel/474011-beperk-aantal-kleine-scholen.html>
- OC&W. (2012). *Overdrachtsdossier OCW*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschap.
- OC&W. (2013a). *31293 nr. 167*. Den Haag: Officiële Bekendmakingen Tweede Kamer.

- OC&W. (2013b). *Kerncijfers 2008-2012*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschap.
- OC&W. (2014a). *Onderwijsdeelname naar sociaaleconomisch milieu*. Opgeroepen op mei 15, 2014, van Monitor Trends in Beeld: http://www.trendsinbeeld.minocw.nl/vervolg.php?h_id=1&s_id=1&v_id=5
- OC&W. (2014b). *Onderwijsdeelname naar etniciteit*. Opgeroepen op mei 15, 2014, van Monitor Trends in Beeld: http://www.trendsinbeeld.minocw.nl/vervolg.php?h_id=1&s_id=1&v_id=6&titel=Onderwijsdeelname_naar_etniciteit
- Onderwijsraad. (2005). *Variëteit in schaal; Keuzevrijheid, sociale samenhang en draagvlak bij grote onderwijsorganisaties*. Den Haag: Onderwijsraad.
- Onderwijsraad. (2008). *De bestuurlijke ontwikkeling van het Nederlandse onderwijs; waarborgen voor keuzevrijheid en legitimatie*. Den Haag: Onderwijsraad.
- Onderwijsraad. (2013). *Grenzen aan kleine scholen; Sterk en pluriform onderwijs in tijden van krimp*. Den Haag: Onderwijsraad.
- Onderwijsraad. (2014, april 7). *Toegevoegde waarde: een instrument voor onderwijsverbetering - niet voor beoordeling*. Opgeroepen op juni 4, 2014, van Onderwijsraad: <http://www.onderwijsraad.nl/publicaties/2014/toegevoegde-waarde/volledig/item7108>
- Oomens, S., Driessen, G., & Scheepers, P. (2003). De integratie van allochtoonse ouders en onderwijsprestaties van hun kinderen; enkele allochtone groepen vergeleken. *Tijdschrift voor Sociologie*, 289-313.
- Regioplan. (2008). *De menselijke maat - Effecten van schaalgrootte in het voortgezet onderwijs*. Amsterdam: Regioplan Beleidsonderzoek.
- Rijksoverheid. (2008, November 28). *Andere kijk op schaalgrootte in het onderwijs. Persberichten Rijksoverheid*. Opgeroepen van: <http://www.rijksoverheid.nl/documenten-en-publicaties/persberichten/2008/11/28/andere-kijk-op-schaalgrootte-in-het-onderwijs.html>.
- Roeleveld, J., & al, (2011). *Correctiefactoren bij opbrengstmaten in het primair onderwijs*. Amsterdam: Kohnstamm Instituut.
- ScienceGuide. (2012, juli 4). *Leraarschap: Zijn extra mannen voor de klas wel nodig?* Opgeroepen op juni 5, 2014, van ScienceGuide: <http://www.scienceguide.nl/201207/zijn-extra-mannen-voor-de-klas-wel-nodig.aspx>
- ScienceGuide. (2013, oktober 10). *Grote school drukt kosten*. Opgeroepen op mei 15, 2014, van ScienceGuide: <http://www.scienceguide.nl/201310/grote-school-drukt-kosten.aspx>

- Stamos. (2014, maart). *Leerlingen/studentenaantallen naar denominatie (uitgebreid)*, in *Basisonderwijs*. Opgeroepen op mei 15, 2014, van STAMOS:
<http://www.stamos.nl/index.rfx?verb=showitem&item=3.27.8&view=table>
- Torres-Reyna, O. (2014). *Multilevel Analysis* (ver. 1.0). New Jersey: Princeton University.
- Trouw. (2001, juni 12). Op universiteit volgen generaties elkaar op. *Trouw*,
<http://www.trouw.nl/tr/nl/5009/Archief/archief/article/detail/2483928/2001/06/12/Op-universiteit-volgen-generaties-elkaar-op.dhtml>.
- Urlings, T., & Blank, J. (2012). *Benchmark bedrijfsvoering voortgezet onderwijs*. Delft: Centrum voor Innovaties en Publieke Sector Efficiënte Studies.
- Van Damme, E. (2003). *Moeder en kind*. Tilburg: Universiteit van Tilburg.
- Van de Venne, L. (2002). *Schaal in het primair onderwijs : een studie naar de relatie tussen schaal en organisatie-effectiviteit*. Amsterdam: UVA.
- Van Petegem, P. (2005). *Vormgeven aan schoolbeleid; Effectieve-scholenonderzoek als inspiratiebron voor de zelfevaluatie van scholen*. Leuven: Uitgeverij Acco.
- Veenstra, R. (1999). *Leerlingen – klassen – scholen : Prestaties en vorderingen van leerlingen in het voortgezet onderwijs*. Groningen: Rijksuniversiteit Groningen.
- Veerman, G., Van de Werfhorst, H., & Dronkers, J. (2013). Ethnic composition of the class and educational performance in primary education in The Netherlands. *Educational Research and Evaluation: An International Journal on Theory and Practice*, 370-401.
- Verboon, P. (2012). *Multilevel Analyse (versie 2.2)*. Open Universiteit.
- Visser, S. (2005). *Kwaliteit mag geen toeval zijn; Zeven basisscholen voor het voetlicht*. Groningen: Gronings Instituut voor Onderzoek van het Onderwijs.
- Volkscrant. (2014, Februari 1). Kleine scholen blijven open door behoud toeslag. *Volkscrant*,
<http://www.volkscrant.nl/vk/nl/2686/Binnenland/article/detail/3588301/2014/02/01/Kleine-scholen-blijven-open-door-behoud-toeslag.dhtml>.
- Waterreus, I. (2009, Oktober). Is er een economische reden voor een fusietoets in het onderwijs? *TPE Digitaal*, pp. 80-102.
- Afbeelding voorblad*: RTV Oost. (2013, november 19). *Hardenberg scoort slechtst in Cito-toets*. Opgeroepen op juni 27, 2014, van RTV Oost Nieuws:
<http://www.rtvooost.nl/nieuws/default.aspx?nid=176291#prettyPhoto>

Bijlage 1: Bronnen schaaleffecten

Voor alle hieronder genoemde variabelen geldt dat zij gemeten zijn op 1 oktober of over 2012 en gelden per bevoegd gezag (tenzij anders vermeld).

Verklaring variabele	Referentie
Nummer	(DUO, 2014a)
Naam	(DUO, 2014a)
Denominatie	(DUO, 2014a)
Aantal leerlingen dat regulier basisonderwijs volgt	(DUO, 2014a)
Aantal FTE onderwijzend personeel (docenten)	(DUO, 2014b)
Totaal aantal FTE	(DUO, 2014b)
Leerling-docent verhouding	Berekening: Aantal leerlingen dat regulier basisonderwijs volgt / aantal FTE docenten
Aantal scholen per bevoegd gezag	Berekening a.d.h.v. BRIN-nummers (DUO, 2014b)
Totale personeelslasten	(DUO, 2014c)
Gemiddelde vaste kosten per leerling	Berekening: personeelslasten / aantal leerlingen dat regulier basisonderwijs volgt
Gemiddelde vaste kosten per FTE	Berekening: personeelslasten / totaal aantal FTE
Gemiddelde vaste kosten per school	Berekening: personeelslasten / aantal scholen per bevoegd gezag

Bijlage 2: Uitschieter

De stichting Joods Bijzonder Onderwijs (bevoegd gezag nummer 21464) is in de analyse buiten beschouwing gelaten. De gemiddelde vaste kosten per leerling voor dit schoolbestuur zijn € 15.887, dat is bijna 375% van het gemiddelde. Het daaropvolgende hoogste bedrag per leerling in de steekproef is € 10.539. De Joodse school heeft een grote invloed op o.a. de gemiddelde vaste kosten per leerling, die € 13 dalen als de Joodse school niet in beschouwing wordt genomen.

De hoge kosten per leerling kunnen onder andere worden verklaard door de kleine klassen op deze school (14,5 leerlingen per voltijd docent). Verder heeft de Joodse school relatief veel onderwijs ondersteunend personeel in dienst (12,1 FTE), terwijl een school met 250 tot 350 leerlingen gemiddeld 2,3 FTE aan ondersteunend personeel in dienst heeft. Een andere reden waarom de kosten per leerling zo hoog zijn, is dat de school joodse identiteitsvakken aanbiedt (JBO, 2014a). Er wordt ook een ouderbijdrage verlangd, o.a. om de beveiliging van de school te bekostigen (JBO, 2014b). Gezien deze bijzondere omstandigheden is de stichting Joods Bijzonder Onderwijs in het onderzoek buiten beschouwing gelaten.

Bijlage 3: Denominaties met minder dan 10 schoolbesturen

Denominatie	Aantal schoolbesturen
Evangelisch	7
Evangelische broedergemeenschap	1
Gereformeerd vrijgemaakt	7
Hindoeïstisch	2
Interconfessioneel	2
Joods	1
Samenwerking Opb., PC	1
Samenwerking Opb., PC, RK	2
Samenwerking Opb., RK	5
Samenwerking PC, RK, Alg. Bijz	2
Geen opgaaf	27

Bijlage 4: Variabelen op het leerlingniveau

Bron: (COOL, 2012)

Variabele	Waarden				
Leerlingnummer	Identificatienummer				
Geslacht	0 = jongen	1 = meisje			
Verblijfsduur in Nederland	0 = langer dan 5 jaar / altijd	1 = korter dan 5 jaar			
Gezinssamenstelling	0 = volledig	1 = eenoudergezin (alleen vader of moeder)			
Geboorteland vader	0 = Nederland	1 = Turkije	2 = Marokko	3 = Suriname of Antillen	4 = Overig
Geboorteland moeder	0 = Nederland	1 = Turkije	2 = Marokko	3 = Suriname of Antillen	4 = Overig
Opleidingsniveau vader	1 = max LO	2 = max LBO	3 = max MBO	4 = HBO / WO	
Opleidingsniveau moeder	1 = max LO	2 = max LBO	3 = max MBO	4 = HBO / WO	
Sociaaleconomische status gezin (SES)	0 = onbekend	1 = max LBO allochtoon	2 = max LBO autochtoon	3 = max MBO allochtoon	
	4 = max MBO autochtoon	5 = HBW/WO allochtoon	6 = HBO/WO autochtoon		
Indicatie zorgleerling	0 = geen zorgleerling	1 = zorgleerling			
Schaal leerlingprofiel: onderpresteren	1 = beslist onwaar	2 = onwaar	3 = niet waar/onwaar	4 = waar	5 = beslist waar
Schaal leerlingprofiel: gedrag leerling	1 = beslist onwaar	2 = onwaar	3 = niet waar/onwaar	4 = waar	5 = beslist waar
Schaal leerlingprofiel: werkhouding	1 = beslist onwaar	2 = onwaar	3 = niet waar/onwaar	4 = waar	5 = beslist waar
Schaal leerlingprofiel: populariteit in klas	1 = beslist onwaar	2 = onwaar	3 = niet waar/onwaar	4 = waar	5 = beslist waar
Schaal relatie leerkracht-leerling: afhankelijkheid	1 = beslist onwaar	2 = onwaar	3 = niet waar/onwaar	4 = waar	5 = beslist waar
Schaal leerlingprofiel: ouderbetrokkenheid	1 = beslist onwaar	2 = onwaar	3 = niet waar/onwaar	4 = waar	5 = beslist waar
Schaal leerlinglijst: welbevinden leerkracht	1 = klopt helemaal niet	2 = klopt niet	3 = klopt soms wel / soms niet	4 = klopt	5 = klopt precies
Schaal leerlinglijst: welbevinden medeleerlingen	1 = klopt helemaal niet	2 = klopt niet	3 = klopt soms wel / soms niet	4 = klopt	5 = klopt precies
Schaal leerlinglijst: zelfvertrouwen op school	1 = klopt helemaal niet	2 = klopt niet	3 = klopt soms wel / soms niet	4 = klopt	5 = klopt precies
Schaal leerlinglijst: taakmotivatie	1 = klopt helemaal niet	2 = klopt niet	3 = klopt soms wel / soms niet	4 = klopt	5 = klopt precies
Toets Rekenen/Wiskunde LOVS	24-168				
Toets Taal DrieMinutenToets 2009	18-124				
Deelname aan Cito Eindtoets	1 = nee	2 = ja			
Score Cito Eindtoets totaal	501-550				

Bijlage 5: Informatie over leerlingen CITO-eindtoets I

Informatie over de leerlingen in de steekproef, waarvan gegevens bekend waren met betrekking tot de CITO-eindtoets. De tabellen geven een overzicht van de aantallen per variabele vanuit de aanvangssituatie (N=6.692). De eerste tabel geeft een overzicht van de aantallen. De tweede tabel geeft een overzicht van gemiddelden van schaalvariabelen.

Variabele	Aantal
Aantal leerlingen	6.692
Aantal jongens ¹³	3.223
Aantal meisjes ¹³	3.321
Verblijfstatus Nederland korter dan 5 jaar ¹⁴	64
Verblijfstatus Nederland langer dan 5 jaar of altijd ¹⁴	5.507
Compleet (tweeouder) gezin ¹⁵	5.692
Eenoudergezin (vader óf moeder) ¹⁵	619
Geboorteland vader: Nederland ¹⁶	4.880
Geboorteland vader: Turkije ¹⁶	367
Geboorteland vader: Marokko ¹⁶	412
Geboorteland vader: Suriname of Nederlandse Antillen ¹⁶	135
Geboorteland vader: overig ¹⁶	425
Geboorteland moeder: Nederland ¹⁷	5.015
Geboorteland moeder: Turkije ¹⁷	350
Geboorteland moeder: Marokko ¹⁷	424
Geboorteland moeder: Suriname of Nederlandse Antillen ¹⁷	202
Geboorteland moeder: overig ¹⁷	488
Opleidingsniveau vader: lager onderwijs ¹⁸	637
Opleidingsniveau vader: lager basisonderwijs ¹⁸	1.650
Opleidingsniveau vader: MBO ¹⁸	2.175
Opleidingsniveau vader: HBO – WO ¹⁸	1.633
Opleidingsniveau moeder: lager onderwijs ¹⁹	764
Opleidingsniveau moeder: lager basisonderwijs ¹⁹	1.392
Opleidingsniveau moeder: MBO ¹⁹	2.791
Opleidingsniveau moeder: HBO – WO ¹⁹	1.434
Geen indicatie zorgleerling ²⁰	4.954
Wel indicatie zorgleerling ²⁰	1.472

¹³ Van 148 leerlingen was het geslacht onbekend.

¹⁴ Van 1.121 leerlingen was de verblijfstatus onbekend.

¹⁵ Van 381 leerlingen was de gezinsstatus onbekend.

¹⁶ Van 473 leerlingen was het geboorteland van de vader onbekend.

¹⁷ Van 213 leerlingen was het geboorteland van de moeder onbekend.

¹⁸ Van 597 leerlingen was het opleidingsniveau van de vader onbekend.

¹⁹ Van 311 leerlingen was het opleidingsniveau van de moeder onbekend.

²⁰ Van 266 leerlingen was de zorgindicatie onbekend.

Schaalvariabele	Gemiddeld	Eerste Kwartiel	Derde Kwartiel
Gedrag leerling	3,75	3,25	4,50
Werkhouding	3,51	3,00	4,00
Populariteit	3,67	3,00	4,00
Afhankelijkheid	1,96	1,20	2,40
Ouderbetrokkenheid	3,63	3,00	4,00
Welbevinden docent	3,73	3,43	4,14
Welbevinden medeleerlingen	4,20	3,83	4,67
Zelfvertrouwen	3,71	3,33	4,00
Taakmotivatie	3,97	3,60	4,40
Toets Rekenen (LOVS)	110,02	102	118
Toets Taal (DMT2009)	98,41	91	107

Bijlage 6: Informatie over leerlingen CITO-eindtoets II

Informatie over de leerlingen in de steekproef, waarvan gegevens bekend waren met betrekking tot de CITO-eindtoets. De tabellen geven een overzicht van de aantallen per variabele vanuit de definitieve modellen M_{33} - M_{36} ($N=3.405$). De eerste tabel geeft een overzicht van de aantallen. De tweede tabel geeft een overzicht van gemiddelden van schaalvariabelen.

Variabele	Aantal
Aantal leerlingen	3.405
Aantal jongens ²¹	1.665
Aantal meisjes ²¹	1.710
Verblijfstatus Nederland korter dan 5 jaar ²²	29
Verblijfstatus Nederland langer dan 5 jaar of altijd ²²	2.845
Compleet (tweeouder) gezin ²³	3.082
Eenoudergezin (vader óf moeder) ²³	160
Geboorteland vader: Nederland	2.692
Geboorteland vader: Turkije	205
Geboorteland vader: Marokko	224
Geboorteland vader: Suriname of Nederlandse Antillen	78
Geboorteland vader: overig	206
Geboorteland moeder: Nederland ²⁴	2.676
Geboorteland moeder: Turkije ²⁴	187
Geboorteland moeder: Marokko ²⁴	221
Geboorteland moeder: Suriname of Nederlandse Antillen ²⁴	84
Geboorteland moeder: overig ²⁴	234
Opleidingsniveau vader: lager onderwijs	355
Opleidingsniveau vader: lager basisonderwijs	928
Opleidingsniveau vader: MBO	1.240
Opleidingsniveau vader: HBO – WO	882
Opleidingsniveau moeder: lager onderwijs	382
Opleidingsniveau moeder: lager basisonderwijs	759
Opleidingsniveau moeder: MBO	1.539
Opleidingsniveau moeder: HBO – WO	725
Geen indicatie zorgleerling	2.667
Wel indicatie zorgleerling	738

²¹ Van 30 leerlingen was het geslacht onbekend.

²² Van 531 leerlingen was de verblijfstatus onbekend.

²³ Van 163 leerlingen was de gezinsstatus onbekend.

²⁴ Van 3 leerlingen was het geboorteland van de moeder onbekend.

Schaalvariabele	Gemiddeld	Eerste Kwartiel	Derde Kwartiel
Gedrag leerling	3,78	3,25	4,50
Werkhouding	3,53	3,00	4,00
Populariteit	3,69	3,33	4,00
Afhankelijkheid	1,94	1,00	2,40
Ouderbetrokkenheid	3,67	3,00	4,00
Welbevinden docent	3,75	3,43	4,17
Welbevinden medeleerlingen	4,22	3,83	4,67
Zelfvertrouwen	3,72	3,33	4,17
Taakmotivatie	3,99	3,60	4,40
Toets Rekenen (LOVS)	110,37	103	118
Toets Taal (DMT2009)	98,48	92	107

Bijlage 7: Variabelen op het schoolniveau

TOELICHTING	WAARDEN				BRONBESTAND
Schoolnummer	Identificatienummer				(COOL, 2012)
Deel van de steekproef	0 = aanvullende steekproef	1 = referentiesteekproef			
Provincie	20 = Groningen	21 = Friesland	22 = Drenthe	23 = Overijssel	
	24 = Flevoland	25 = Gelderland	26 = Utrecht	27 = Noord-Holland	
	28 = Zuid-Holland	29 = Zeeland	30 = Noord-Brabant	31 = Limburg	
Mate van stedelijkheid	1 = zeer sterk stedelijk	2 = sterk stedelijk	3 = matig stedelijk		
	4 = weinig stedelijk	5 = niet stedelijk			
Schoolgewichtscore	1 = 100-104	2 = 105-109	3 = 110-119		
	4 = 120-139	5 = 140-159	6 = 160-220		
Denominatie	1 = Openbaar	2 = Protestants-Christelijk	3 = Rooms-Katholiek		(DUO, 2010a)
	4 = Algemeen Bijzonder	5 = Islamitisch	6 = Overig Bijzonder		
Taalverbetertraject	1 = nee	2 = ja	3 = in voorbereiding		(COOL, 2012)
Rekenverbetertraject	1 = nee	2 = ja	3 = in voorbereiding		
Percentage allochtonen >50%	0 = nee	1 = ja			<i>Berekening</i>
Percentage allochtonen >75%	0 = nee	1 = ja			<i>Berekening</i>
Aantal etnische groepen	Tussen 1 en 11				<i>Berekening</i>
Aantal leerlingen <145	0 = nee	1 = ja			<i>Berekening op basis van (DUO, 2010a)</i>
Aantal leerlingen >500	0 = nee	1 = ja			
Aantal scholen per bestuur: 6-10	0 = nee	1 = ja			
Aantal scholen per bestuur: 11-20	0 = nee	1 = ja			
Aantal scholen per bestuur: >20	0 = nee	1 = ja			
Aandeel FTE 0-0.5	Percentage van totaal aantal FTE				<i>Berekening op basis van (DUO, 2014e)</i>
Aandeel FTE vrouwen	Percentage van totaal aantal FTE				
Aandeel management	Percentage van totaal aantal FTE				
Gemiddelde leeftijd personeel	Getal tot 55,7 jaar				
Verhouding vast en tijdelijk personeel	Percentage van totaal aantal FTE				

Bijlage 8: Informatie over scholen CITO-eindtoets I

Informatie over de scholen waarop leerlingen hebben meegedaan aan de CITO-eindtoets. De tabel geeft een overzicht van de aantallen per variabele vanuit de aanvangssituatie, waarbij het aantal leerlingen N=6.692 en het aantal scholen N=318 was.

Variabele	Aantal
Aantal scholen	318
Scholen in de referentiesteekproef	254
Scholen in de aanvullende steekproef	64
Scholen in zeer sterk stedelijk gebied	42
Scholen in sterk stedelijk gebied	65
Scholen in matig stedelijk gebied	66
Scholen in weinig stedelijk gebied	80
Scholen in niet stedelijk gebied	65
Openbare scholen	110
Protestants-christelijke scholen	70
Rooms-Katholieke scholen	114
Algemeen Bijzondere scholen	6
Islamitische scholen	10
Overig Bijzondere scholen	8
Scholen die gebruik maken van taalverbetertraject	115
Scholen die gebruik maken van rekenverbetertraject	48
Schoolscore 100-104	112
Schoolscore 105-109	63
Schoolscore 110-119	57
Schoolscore 120-139	43
Schoolscore 140-159	27
Schoolscore >159	16
Gemiddeld aantal FTE per school	15,23
Gemiddeld aantal scholen per bestuur	16,9
Gemiddeld percentage allochtone leerlingen (groep 8)	22,18
Gemiddeld aantal etnische groepen (groep 8)	2,99
Gemiddeld aantal leerlingen per voltijd-docent	16,47

Bijlage 9: Informatie over scholen CITO-eindtoets II

Informatie over de scholen waarop leerlingen hebben meegedaan aan de CITO-eindtoets. De tabel geeft een overzicht van de aantallen per variabele vanuit de definitieve modellen (M₃₃-M₃₆), waarbij het aantal leerlingen N=3.405 en het aantal scholen N=213 was.

Variabele	Aantal
Aantal Scholen	213
Scholen in de referentiesteekproef	168
Scholen in de aanvullende steekproef	45
Scholen in zeer sterk stedelijk gebied	28
Scholen in sterk stedelijk gebied	46
Scholen in matig stedelijk gebied	46
Scholen in weinig stedelijk gebied	52
Scholen in niet stedelijk gebied	41
Openbare scholen	72
Protestants-christelijke scholen	49
Rooms-Katholieke scholen	77
Algemeen Bijzondere scholen	3
Islamitische scholen	6
Overig Bijzondere scholen	6
Scholen die gebruik maken van taalverbetertraject	78
Scholen die gebruik maken van rekenverbetertraject	33
Schoolscore 100-104	75
Schoolscore 105-109	41
Schoolscore 110-119	38
Schoolscore 120-139	28
Schoolscore 140-159	21
Schoolscore >159	11
Gemiddeld aantal FTE per school	15,11
Gemiddeld aantal scholen per bestuur	16,67
Gemiddeld percentage allochtone leerlingen (groep 8)	23,24
Gemiddeld aantal etnische groepen (groep 8)	3,02
Gemiddeld aantal leerlingen per voltijd-docent	16,47

Bijlage 10: Correlatietabellen opleidingsniveau en etniciteit

Voor de onderstaande correlatietabellen geldt dat alleen gegevens zijn meegenomen van leerlingen die de CITO-toets hebben gemaakt.

Correlatietabel 10.A: Opleidingsniveau vader en moeder (N=6.007)								
	Vader LO	Vader LBO	Vader MBO	Vader HBO-WO	Moeder LO	Moeder LBO	Moeder MBO	Moeder HBO-WO
Vader LO	1.000							
Vader LBO	-0.2072	1.000						
Vader MBO	-0.2535	-0.4530	1.000					
Vader HBO-WO	-0.2067	-0.3694	-0.4520	1.000				
Moeder LO	0.6190	-0.0521	-0.1789	-0.1804	1.000			
Moeder LBO	-0.0451	0.3248	-0.0697	-0.2188	-0.1924	1.000		
Moeder MBO	-0.2249	-0.0655	0.2850	-0.0874	-0.3223	-0.4627	1.000	
Moeder HBO-WO	-0.1659	-0.2010	-0.1307	0.4565	-0.1994	-0.2862	-0.4795	1.000

Correlatietabel 10.B: Geboorteland vader en opleidingsniveau vader (N=5.931)				
	Vader LO	Vader LBO	Vader MBO	Vader HBO-WO
Geboorteland Nederland	-0,4890	0,0479	0,1328	0,1462
Geboorteland Turkije	0,2842	-0,0055	-0,0761	-0,1087
Geboorteland Marokko	0,3848	-0,0538	-0,1006	-0,1029
Geboortel. Suriname / Antillen	0,0101	0,0318	0,0114	-0,0516
Geboorteland overig	0,1405	-0,0379	-0,0514	-0,0032

Correlatietabel 10.C: Geboorteland moeder en opleidingsniveau moeder (N=6.249)				
	Moeder LO	Moeder LBO	Moeder MBO	Moeder HBO-WO
Geboorteland Nederland	-0,5387	0,0435	0,1897	0,1535
Geboorteland Turkije	0,3449	-0,0153	-0,1314	-0,0989
Geboorteland Marokko	0,3845	-0,0328	-0,1346	-0,1088
Geboortel. Suriname / Antillen	-0,0087	0,0168	0,0362	-0,0537
Geboorteland overig	0,2005	-0,0362	-0,0848	-0,0198

Bijlage 11: Modellen algemene leerling- en gezinskenmerken

Aan de eerste geschatte modellen worden steeds algemene leerling- en gezinskenmerken toegevoegd (*laag 1 in het methodologisch model*). Als verwezen wordt naar correlatietabellen wordt verwezen naar bijlage 10, waarin deze tabellen zijn opgenomen. In tabel 1 zijn de eerste vijf geschatte modellen M_1 tot M_5 weergegeven. In model M_1 wordt de variabele *geslacht* toegevoegd aan het model. Geslacht is genoteerd als '0' voor jongens en '1' voor meisjes. Dit betekent dat de gemiddelde CITO-score voor jongens gelijk is aan de constante (535,25) en voor meisjes gelijk is aan de constante min de coëfficiënt voor het geslacht (534,09). Dit is in overeenstemming met onderzoek van het ministerie, dat vindt dat jongens gemiddeld iets hogere CITO-scores halen dan meisjes (OC&W, 2013b).

In model M_2 wordt het *opleidingsniveau van de vader* toegevoegd. Het opleidingsniveau is onderverdeeld in de categorieën maximaal lager onderwijs (1, *referentiecategorie*), maximaal lager basisonderwijs (2), maximaal MBO (3) en HBO/WO (4). Het effect van het opleidingsniveau van de vader is sterk. Het verschil in CITO-score tussen een leerling waarvan de vader alleen lager onderwijs heeft gevolgd en een leerling waarvan de vader gestudeerd heeft aan een hogeschool of universiteit is ruim 10 punten. Het effect van geslacht wordt ten opzichte van model M_1 iets kleiner. Dat opleidingsniveau een belangrijke determinant is van de behaalde CITO-score, is in overeenstemming met het onderzoek van Roeleveld & al (2011).

In model M_3 wordt het *opleidingsniveau van de moeder* toegevoegd aan het model. Dit zorgt voor een verschuiving in effecten: het opleidingsniveau van de moeder heeft (behalve op het LBO-niveau) meer invloed op de CITO-score dan het opleidingsniveau van de vader. De effecten van het opleidingsniveau van beide ouders zijn samengenomen nog groter dan het enkele effect van het opleidingsniveau van een van de ouders. Een leerling die ouders heeft die beiden op de hogeschool of universiteit hebben gestudeerd, heeft een CITO-score die bijna 13 punten hoger is dan een leerling met ouders die lager onderwijs hebben gevolgd. Uit correlatietabel 10.A blijkt dat bepaalde opleidingsniveaus van vader en moeder sterk gecorreleerd zijn. Uit de tabel blijkt dat zowel de laagste opleidingsniveaus ($r=0.6190$) als de hoogste opleidingsniveaus ($r=0.4565$) van vader en moeder een sterke correlatie kennen. Dit verband werd in een ouder onderzoek ook al door het CBS gevonden. Uit dat onderzoek bleek dat universiteitsstudenten in 6 van de 10 gevallen ouders hebben die zelf ook aan de universiteit gestudeerd hebben (Trouw, 2001).

In model M_4 wordt het *geboorteland van de vader* toegevoegd. De variabele geboorteland is gedefinieerd als Nederland (0, *referentiecategorie*), Turkije (1), Marokko (2), Suriname en de voormalige Nederlandse Antillen (3) en overig²⁵ (4). Door de toevoeging van het geboorteland van de vader zijn zowel het opleidingsniveau van de vader als dat van de moeder op LBO-niveau niet meer significant. Dit betekent dat het geboorteland van de vader voor een deel dezelfde variantie verklaart als de lagere opleidingsniveaus. Uit correlatietabel 10.B blijkt dit verband inderdaad vooral voor het laagste opleidingsniveau te gelden.

²⁵ Onder 'overig' vallen de landen Joegoslavië, (voormalige) Sovjetunie, Polen, China, Irak, Afghanistan, Somalië, de Molukken en overig Westers. Deze landen zijn voor de overzichtelijkheid van het model samengenomen. Tevens was het aantal waarnemingen voor deze landen beperkt.

	Model M1	Model M2	Model M3	Model M4	Model M5
<i>Log Likelihood</i>	- 23.932	- 21.487	- 21.010	- 20.415	- 20.563
<i>Aantal observaties</i>	6.525	5.972	5.886	5.725	5.762
<i>AIC</i>	47.871	42.985	42.038	40.856	41.151
<i>Constante</i>	535,25	529,90	529,05	530,61	530,23
Fixed					
Geslacht	- 1,16 *** (0.167)	- 1,09 *** (0.229)	- 1,04 *** (0.224)	- 1,00 *** (0.226)	- 1,01 *** (0.226)
Opleiding Vader LBO		2,67 *** (0.418)	0,92 * (0.492)	0,31 (0.513)	0,53 (0.509)
Opleiding Vader MBO		5,51 *** (0.403)	2,61 *** (0.500)	1,95 *** (0.520)	2,21 *** (0.516)
Opleiding Vader HBO-WO		10,21 *** (0.417)	5,60 *** (0.530)	4,90 *** (0.550)	5,14 *** (0.547)
Opleiding Moeder LBO			0,83 * (0.480)	0,17 (0.513)	0,25 (0.521)
Opleiding Moeder MBO			3,97 *** (0.468)	3,17 *** (0.508)	3,25 *** (0.517)
Opleiding Moeder HBO-WO			7,22 *** (0.517)	6,33 *** (0.559)	6,42 *** (0.567)
Geboorteland Vader Turkije				- 2,76 *** (0.537)	
Geboorteland Vader Marokko				- 1,59 *** (0.529)	
Geboorteland Vader Suriname of Antillen				- 2,80 *** (0.786)	
Geboorteland Vader Overig				0,15 (0.485)	
Geboorteland Moeder Turkije					-2,25 *** (0.558)
Geboorteland Moeder Marokko					- 1,32 ** (0.535)
Geboorteland Moeder Suriname of Antillen					- 2,57 *** (0.752)
Geboorteland Moeder Overig					0,25 (0.476)
Random					
var(Residual)	89,8	78,1	73,8	73,3	73,6

Tabel 1: Geschatte Modellen M₁ tot M₅

Verder valt op dat de andere effecten van het opleidingsniveau afnemen (alle coëfficiënten worden kleiner). De categorie 'geboorteland overig' is niet significant, waarschijnlijk omdat dit een restgroep is (waardoor het resultaat moeilijk interpreteerbaar is). Na toevoeging van de variabele valt op dat leerlingen met een Turkse, Marokkaanse, Surinaamse of Antilliaanse vader significant lagere CITO-scores behalen. Gecontroleerd voor het opleidingsniveau van beide ouders blijkt etniciteit, in overeenstemming met Veenstra (1999), een extra effect te hebben op de prestaties van leerlingen.

In model M_5 nemen we alleen het *geboorteland van de moeder* mee. Hier ontstaat hetzelfde patroon als in model M_4 : opleidingsniveau op LBO-niveau is voor beide ouders niet meer significant, evenals de categorie overig bij het geboorteland van de moeder. De correlatie tussen het laagste opleidingsniveau en het geboorteland van de moeder is weer sterk (zie correlatietabel 10.C). Het valt op dat de coëfficiënten voor het geboorteland van de moeder in alle categorieën iets kleiner zijn dan de coëfficiënten voor het geboorteland van de vader, hoewel het effect nog steeds fors is. Leerlingen met een Turkse, Marokkaanse of Surinaamse moeder behalen ten opzichte van leerlingen met een Nederlandse moeder een CITO-score die 2,25 respectievelijk 1,32 en 2,57 punt lager is.

In tabel 2 staan de volgende geschatte modellen M_6 tot M_8 . In model M_6 wordt zowel het *geboorteland van de moeder* als dat van de *vader* meegenomen. Dit zorgt voor tegenstrijdige effecten. Alle categorieën van het geboorteland van de moeder zijn nu insignificant. Wat verder opvalt is dat de verschillende categorieën bij vader en moeder soms tegen elkaar in lijken te werken. De correlatie tussen het geboorteland van de vader en van de moeder is ook erg hoog ($r = 0.7002$), waardoor het collineariteitsprobleem opspeelt. Dat de correlatie tussen beide variabelen hoog is blijkt ook uit de correlatietabellen 10.B en 10.C, die beiden min of meer hetzelfde patroon weergegeven. Beide variabelen verklaren voor een groot deel dezelfde variantie. Dat dit probleem opspeelt, is bijvoorbeeld te zien als Turkije het geboorteland is. De coëfficiënt is bij de vader sterker negatief geworden (ten opzichte van model M_4), terwijl die van de moeder matig positief is. Deze coëfficiënten werken tegen elkaar in en verklaren samen een negatief effect. Door het collineariteitsprobleem is het beter één van beide variabelen uit het model te laten. In het vervolg wordt daarom de variabele *geboorteland van de moeder* uit het model gelaten en verder gewerkt met de variabele *geboorteland van de vader*.

In model M_7 wordt de variabele *gezinssamenstelling* toegevoegd. De gezinssamenstelling is een dummy-variabele die gedefinieerd is als een volledig (tweeouder)gezin (0, *referentie-categorie*) of als eenoudergezin (1). Het effect van gescheiden ouders op de CITO-score is flink. Als een leerling gescheiden ouders heeft, haalt deze gemiddeld een score die 2,28 punt lager is dan van een leerling uit een tweeoudergezin. Er blijkt geen verband te zijn tussen opleidingsniveau van de ouders en gezinssamenstelling: gescheiden ouders zijn niet vaker van een lager opleidingsniveau. In de andere verklarende effecten treden geen noemenswaardige verschillen op. Het gevonden effect is in lijn met onderzoek van het CBS naar de gevolgen van echtscheiding, waaruit bleek dat kinderen met gescheiden ouders ruim 2 punten lager scoorden op de CITO-toets (CBS, 2013). Kinderen die in gebroken gezinnen opgroeien zijn minder gelukkig en uiteindelijk minder succesvol (Van Damme, 2003).

	Model M6	Model M7	Model M8
<i>Log Likelihood</i>	- 20.387	- 19.393	- 16.690
<i>Aantal observaties</i>	5.718	5.439	4.694
<i>AIC</i>	40.808	38.814	33.411
<i>Constante</i>	530,55	530,56	530,72
Fixed			
Geslacht	- 1,00 *** (0.227)	- 0,97 *** (0.232)	- 1,09 *** (0.248)
Opleiding Vader LBO	0,35 (0.515)	0,34 (0.531)	0,50 (0.600)
Opleiding Vader MBO	1,98 *** (0.521)	2,00 *** (0.537)	2,24 *** (0.612)
Opleiding Vader HBO-WO	4,93 *** (0.552)	4,86 *** (0.570)	5,04 *** (0.641)
Opleiding Moeder LBO	0,21 (0.524)	0,28 (0.531)	0,23 (0.593)
Opleiding Moeder MBO	3,19 *** (0.520)	3,33 *** (0.526)	2,98 *** (0.595)
Opleiding Moeder HBO-WO	6,33 *** (0.569)	6,46 *** (0.579)	6,06 *** (0.643)
Geboorteland Vader Turkije	- 3,94 *** (1.100)	- 2,81 *** (0.547)	- 3,23 *** (0.623)
Geboorteland Vader Marokko	- 0,68 (1.194)	- 1,63 *** (0.548)	- 2,26 *** (0.644)
Geboorteland Vader Suriname of Antillen	- 2,28 ** (1.088)	- 2,85 *** (0.795)	- 3,61 *** (1.018)
Geboorteland Vader Overig	0,14 (0.623)	0,15 (0.492)	0,56 (0.535)
Geboorteland Moeder Turkije	1,32 (1.145)		
Geboorteland Moeder Marokko	- 0,92 (1.207)		
Geboorteland Moeder Suriname of Antillen	- 0,83 (1.062)		
Geboorteland Moeder Overig	0,09 (0.612)		
Gezinssamenstelling		- 2,28 *** (0.549)	- 2,39 *** (0.566)
Verblijfsduur in Nederland			- 5,30 *** (1.293)
Random			
var(Residual)	73,1	73,6	71,7

Tabel 2: Geschatte Modellen M₆ tot M₈

In model M_8 wordt de variabele *verblijfsduur in Nederland* toegevoegd. De variabele verblijfsduur in Nederland is ingedeeld in 5 jaar of langer (0, *referentiecategorie*) en korter dan 5 jaar (1). De verblijfsduur van een leerling in Nederland heeft grote invloed op de CITO-score. Als een leerling 5 jaar of korter in Nederland verblijft, is de CITO-score gemiddeld ruim 5 punten lager. Het gevonden effect is in overstemming met het onderzoek van Oomen & al (2003), dat stelt dat taalprestaties van allochtone kinderen slechter zijn als de verblijfsduur van de ouders in Nederland korter is. Dit effect op de CITO-score is begrijpelijk, want naast taalproblemen die kunnen optreden zijn deze kinderen het Nederlandse onderwijssysteem niet gewend, waardoor er problemen kunnen optreden. Er vinden verder een aantal kleine verschuivingen in het model plaats. Zo wordt het opleidingsniveau van de vader relatief wat belangrijker ten opzichte van het opleidingsniveau van de moeder. Ook worden de negatieve effecten van de etniciteit van de vader nog groter.

In de tabel 3 staan de volgende drie geschatte modellen M_9 - M_{11} . In model M_9 wordt de verzamelvariabele *Sociaal Economische Status* (SES) toegevoegd. Deze variabele is een combinatie van opleidingsniveau en afkomst, ingedeeld in maximaal lager basis onderwijs allochtoon (1, *referentiecategorie*) en autochtoon (2), maximaal MBO allochtoon (3) en autochtoon (4) en HBO/WO allochtoon (5) en HBO/WO autochtoon (6). De toevoeging van de variabele *SES* geeft een diffuus beeld. De categorieën zijn (op een na) niet significant. Alle coëfficiënten die te maken hebben met opleidingsniveau stijgen. De coëfficiënten die te maken hebben met het geboorteland van de vader worden sterker negatief. Ook hier blijkt het probleem van collineariteit te spelen. De correlatie tussen de variabelen *opleidingsniveau van de vader* en *SES* is erg hoog ($r = 0.8139$). Beide variabelen verklaren voor een groot deel dezelfde variantie. Het is daarom redelijk om één variabele uit het model te laten. In dit geval wordt de variabele *SES* buiten het model gelaten, omdat dit een samengestelde variabele is. Het is beter om de variabelen *geboorteland van de vader* en *opleidingsniveau van vader* en *moeder* afzonderlijk op te nemen.

In model M_{10} is een maatstaf voor het *gedrag* van de leerling toegevoegd. Dit is een schaal op basis van beoordelingen van docenten over het gedrag van hun leerlingen. Een hoge score impliceert goed gedrag. De schaal loopt van 1 tot 5, waarbij een hoge score (5) zeer goed gedrag en een lage score (1) zeer slecht gedrag impliceert. De toevoeging van gedrag aan het model zorgt voor verschuivingen. De invloed van geslacht op de CITO-score wordt bijvoorbeeld een halve punt groter. Dit is voor een groot deel te verklaren uit het feit dat het gedrag van meisjes gemiddeld hoger (3,89) gewaardeerd wordt dan het gedrag van jongens (3,57), zodat zodoende meisjes gemiddeld een hogere gedragsscore halen. Het model corrigeert voor de invloed van geslacht door de coëfficiënt van de variabele *geslacht* te vergroten. Deze vinding is in overeenstemming met het oordeel van de Inspectie (2013), dat jongens vaker agressief gedrag en gebrekkige sociale vaardigheden vertonen. Verder valt op dat de invloed van gedrag op de CITO-score groot is: het verschil in score tussen een leerling die zeer slecht gedrag vertoont en een leerling die zeer goed verdrag vertoont is ruim 6,5 punten.

		Model M9	Model M10	Model M11
<i>Log Likelihood</i>		- 16.685	- 15.976	- 15.677
<i>Aantal observaties</i>		4.694	4.510	4.510
<i>AIC</i>		33.411	31.984	31.388
<i>Constante</i>		531,47	524,73	527,72
Fixed				
	Geslacht	- 1,08 *** (0.247)	- 1,55 *** (0.255)	- 1,71 *** (0.239)
	Opleiding Vader LBO	0,91 (0.631)	0,66 (0.611)	0,68 (0.572)
	Opleiding Vader MBO	2,40 *** (0.709)	2,44 *** (0.624)	2,20 *** (0.584)
	Opleiding Vader HBO-WO	5,19 *** (0.833)	5,08 *** (0.653)	4,61 *** (0.611)
	Opleiding Moeder LBO	0,70 (0.624)	0,00 (0.602)	0,24 (0.563)
	Opleiding Moeder MBO	3,08 *** (0.705)	2,82 *** (0.605)	2,54 *** (0.567)
	Opleiding Moeder HBO-WO	6,07 *** (0.810)	5,88 *** (0.653)	5,42 *** (0.611)
	Geboorteland Vader Turkije	- 3,81 *** (0.757)	- 2,89 *** (0.633)	- 2,78 *** (0.592)
	Geboorteland Vader Marokko	- 2,75 *** (0.788)	- 1,91 *** (0.654)	- 1,85 *** (0.612)
	Geboorteland Vader Suriname of Antillen	- 4,03 *** (1.097)	- 3,49 *** (1.027)	- 3,76 *** (0.961)
	Geboorteland Vader Overig	0,38 (0.598)	0,71 (0.539)	0,30 (0.505)
	Gezinssamenstelling	- 2,31 *** (0.567)	- 1,99 *** (0.564)	- 1,54 *** (0.528)
	Verblijfsduur in Nederland	- 5,34 *** (1.293)	- 4,75 *** (1.305)	- 3,32 *** (1.222)
	SES basis autochtoon	- 1,99 ** (0.828)		
	SES MBO allochtoon	- 0,63 (0.937)		
	SES MBO autochtoon	- 1,10 (0.945)		
	SES HBO-WO allochtoon	- 1,81 (1.177)		
	SES HBO-WO autochtoon	- 0,87 (1.108)		
	Gedrag (schaal 1 tot 5)		1,63 *** (0.152)	1,39 *** (0.142)
	Zorgleerling			- 7,07 *** (0.280)
Random	var(Residual)	71,6	69,9	61,3

Tabel 3: Geschatte Modellen M₉ tot M₁₁

In model M₁₁ wordt de variabele *zorgleerling* toegevoegd. Een leerling is als zorgleerling gekwalificeerd als op een of meerdere van de volgende vragen een bevestigend antwoord volgt: of er voor de leerling een rugzakje beschikbaar is, of er sprake is van ambulante begeleiding, of er een procedure is gestart voor plaatsing op het speciaal onderwijs, of er een zorghandelingsplan voor de leerling is opgesteld en of het kind een eigen leerlijn volgt. De variabele *zorgleerling* is een dummy variabele: 0 (*referentiecategorie*) voor een gewone leerling en 1 voor een zorgleerling. Een zorgleerling behaalt gemiddeld een CITO-score die ruim 7 punten lager ligt dan de CITO-score van een gewone leerling. Door de toevoeging van de variabele *zorgleerling* vinden een aantal kleine veranderingen plaats. Een opvallende verandering is het effect van de variabele *verblijfsduur*, dat fors afneemt. Daaruit kan worden geconcludeerd dat zorgleerlingen relatief vaak leerlingen zijn die minder lang in Nederland verblijven. Aangezien de variabele *zorgleerling* onder andere gebaseerd is op een apart leerplan voor een leerling is dit logisch: leerlingen die oorspronkelijk uit een ander land komen, zullen waarschijnlijk een ander leerplan volgen dan de reguliere leerling.

Aan de hand van model 11 kunnen we de gevonden effecten tussentijds toetsen aan de verwachtingen die er op voorhand waren met betrekking tot de algemene leerlingkenmerken. Er was vastgesteld dat de CITO-score relatief hoger is als:

- ✓ De leerling van het mannelijk geslacht is;
- ✓ De leerling ouders heeft met een hoog opleidingsniveau;
- ✓ De leerling autochtone ouders heeft;
- ✓ De leerling langer in Nederland verblijft;
- ✓ De leerling leeft in een tweeoudergezin;
- ✓ De leerling geen gedragsproblemen en gezondheidsproblemen heeft.

Aan de hand van deze verwachting kan vergelijking [V.1] worden opgesteld:

$$\begin{aligned} CITO - score = & 527,72 + \alpha X_{\text{geslacht}} + \alpha X_{\text{oplvader-mbo}} + \alpha X_{\text{oplvader-hbowo}} \\ & + \alpha X_{\text{oplmoeeder-mbo}} + \alpha X_{\text{oplmoeeder-hbowo}} + \alpha X_{\text{gl-turkije}} + \alpha X_{\text{gl-marokko}} \\ & + \alpha X_{\text{gl-suriname}} + \alpha X_{\text{gezin}} + \alpha X_{\text{verblijfsduur-NL}} + \alpha X_{\text{gedrag}} + \alpha X_{\text{zorgleerling}} \\ & + \varepsilon_{\text{residual}} \end{aligned}$$

Met behulp van de geschatte parameters uit model M₁₁ kunnen de coëfficiënten α worden ingevuld, waardoor vergelijking [V.2] ontstaat:

$$\begin{aligned} CITO - score = & 527,72 - 1,71 * X_{\text{geslacht}} + 2,20 * X_{\text{oplvader-mbo}} + 4,61 * X_{\text{oplvader-hbowo}} \\ & + 2,54 * X_{\text{oplmoeeder-mbo}} + 5,42 * X_{\text{oplmoeeder-hbowo}} - 2,78 * X_{\text{gl-turkije}} \\ & - 1,85 * X_{\text{gl-marokko}} - 3,76 * X_{\text{gl-suriname}} - 1,54 * X_{\text{gezin}} \\ & - 3,32 * X_{\text{verblijfsduur-NL}} + 1,39 * X_{\text{gedrag}} - 7,07 * X_{\text{zorgleerling}} + \varepsilon_{\text{residual}} \end{aligned}$$

Aan de hand van vergelijking [V.2] kan worden geconcludeerd dat de CITO-score hoger is als de leerling mannelijk is, als de ouders een opleiding hebben gevolgd op HBO-WO niveau, als de leerling geen Turkse, Marokkaanse, Surinaamse of Antilliaanse ouders heeft, als de leerling langer dan 5 jaar of altijd in Nederland verblijft, als de leerling leeft in een tweeoudergezin en als de leerling goed gedrag vertoont en geen zorgleerling is. De tot nu toe gevonden effecten wijken niet af van de vastgestelde effecten uit eerder onderzoek.

Bijlage 12: Modellen schoolse leerling- en gezinskenmerken

Voordat er nieuwe onafhankelijke variabelen op het schoolse leerlingniveau worden toegevoegd wordt model M_{11} uitgebreid met een *random-effect op het schoolniveau* (als constante). Dit betekent dat de geschatte parameters in een model niet hetzelfde hoeven te zijn voor elke leerling, maar dat ze kunnen verschillen op schoolniveau. In tabel 1 op de volgende bladzijde staan de modellen M_{12} tot M_{14} genoteerd. Daarbij valt op dat de effecten van het geboorteland van de vader en het opleidingsniveau van de vader op de CITO-score afnemen. De variantie op schoolniveau verklaart dus voor een deel hetzelfde als verschillen in geboorteland en opleidingsniveau. Met de analyse van Herweijer (2008) over zwarte en witte scholen in het achterhoofd is dit begrijpelijk. Scholen blijken namelijk vaak een min of meer homogene leerlingenpopulatie te hebben, bestaande uit leerlingen met eenzelfde afkomst of sociaaleconomische status. Een deel van de variatie in CITO-scores kan daarom verklaard worden op schoolniveau.

Een LR-test geeft aan dat de toevoeging van het random schooleffect het model verbeterd ($p=0.000$). Of de toevoeging van het random effect zinvol is kan ook worden bepaald door de IC-maatstaf (Torres-Reyna, 2014). In model M_{12} is de berekening van de IC als volgt:

$$IC_{schooleffect} = \frac{6,9}{6,9 + 54,7} \approx 0,11$$

De IC- en de LR-test geven voldoende aanleiding om het random-schooleffect te behouden.

In de volgende modellen worden steeds schoolse leerling- en gezinskenmerken toegevoegd (*laag 2 in het methodologisch model*). In model M_{13} wordt de variabele *taakmotivatie* toegevoegd. Deze variabele geeft aan in hoeverre leerlingen hun vaardigheden willen ontwikkelen en of ze willen leren van hun schoolwerk. Taakmotivatie is een schaalvariabele, gemeten van helemaal niet gemotiveerd (1) tot zeer gemotiveerd (5). De motivatie van een leerling heeft een flink effect op zijn CITO-score. Het verschil tussen een leerling die helemaal niet gemotiveerd is en een leerling die zeer gemotiveerd is, is ruim 7 punten. Zoals de Inspectie (2014) concludeert is het heel belangrijk dat leerlingen gemotiveerd worden om hun lestaken uit te voeren. Dit blijkt ook uit het hier gevonden effect. Dit effect geeft ook aan dat taakmotivatie niet alleen een indicator is voor later studiesucces, zoals Veenstra (1999) en de Inspectie (2014) beweren, maar ook voor direct studiesucces. Verder valt op dat de variantie verklaart door interschoolse verschillen toeneemt. Dit duidt erop dat er verschillen zijn tussen scholen in de mate waarop ze hun leerlingen kunnen motiveren voor de lestaken.

In model M_{14} wordt de variabele *zelfvertrouwen* toegevoegd. Het gaat om een schaalvariabele die meet of de leerling vertrouwen heeft in het feit dat hij schooltaken aankan, waarbij 1 aangeeft dat de leerling helemaal geen zelfvertrouwen heeft en 5 aangeeft dat de leerling heel veel zelfvertrouwen heeft. Het toevoegen van de variabele *zelfvertrouwen* heeft een groot effect op het model. Het verschil in CITO-score tussen een leerling met helemaal geen zelfvertrouwen en heel veel zelfvertrouwen is bijna 17 punten. Het gevonden effect van zelfvertrouwen op prestaties is in overeenstemming met Veenstra (1999), die stelt dat naarmate het zelfbeeld van een kind positiever is het kind beter presteert.

	Model M12	Model M13	Model M14
<i>Log Likelihood</i>	- 15.565	- 15.294	- 15.083
<i>Aantal observaties</i>	4.510	4.446	4.442
<i>AIC</i>	31.165	30.626	30.205
<i>Constante</i>	528,21	521,68	513,77
Fixed			
Geslacht	- 1,77 *** (0.230)	- 1,78 *** (0.230)	- 1,12 *** (0.222)
Opleiding Vader LBO	0,25 (0.555)	0,21 (0.551)	0,22 (0.527)
Opleiding Vader MBO	1,75 *** (0.567)	1,64 *** (0.563)	1,62 *** (0.538)
Opleiding Vader HBO-WO	3,89 *** (0.596)	3,73 *** (0.592)	3,38 *** (0.566)
Opleiding Moeder LBO	0,44 (0.548)	0,41 (0.543)	0,27 (0.519)
Opleiding Moeder MBO	2,78 *** (0.552)	2,84 *** (0.548)	2,66 *** (0.524)
Opleiding Moeder HBO-WO	5,57 *** (0.596)	5,68 *** (0.591)	5,26 *** (0.565)
Geboorteland Vader Turkije	- 1,81 *** (0.617)	- 2,24 *** (0.619)	- 2,53 *** (0.593)
Geboorteland Vader Marokko	- 1,38 ** (0.642)	- 1,86 *** (0.641)	- 2,29 *** (0.615)
Geboorteland Vader Suriname of Antillen	- 3,31 *** (0.970)	- 3,51 *** (0.960)	- 3,82 *** (0.919)
Geboorteland Vader Overig	0,42 (0.510)	0,03 (0.510)	0,07 (0.489)
Gezinssamenstelling	- 1,32 ** (0.522)	- 1,03 ** (0.519)	- 1,20 ** (0.497)
Verblijfsduur in Nederland	- 2,79 ** (1.377)	- 2,25 (1.409)	- 2,36 * (1.351)
Gedrag (schaal 1 tot 5)	1,39 *** (0.145)	1,23 *** (0.145)	1,18 *** (0.139)
Zorgleerling	- 7,74 *** (0.289)	- 7,53 *** (0.288)	- 6,74 *** (0.279)
Taakmotivatie		1,82 *** (0.194)	- 0,15 (0.210)
Zelfvertrouwen			4,23 *** (0.212)
Random			
var(Residual)	54,7	53,4	48,7
var(School)	6,9	7,2	7,0

Tabel 1: Geschatte Modellen M₁₂ tot M₁₄

Het effect van de meeste andere variabelen in model M_{14} neemt af. De coëfficiënt van *geslacht* neemt af met ongeveer 0,7 punt, waaruit volgt dat meisjes gemiddeld iets meer zelfvertrouwen hebben dan jongens. Ook valt op dat de coëfficiënten van het hoogste opleidingsniveau van vader en moeder beiden afnemen met ongeveer 0,4 punt. Leerlingen met ouders met een hoog opleidingsniveau hebben zodoende meer zelfvertrouwen. Verder valt op dat de coëfficiënten van het geboorteland van vader sterker negatief worden. Leerlingen met een Turkse (-0,29), Marokkaanse (-0,43) en Surinaams Antilliaanse (-0,31) vader hebben minder zelfvertrouwen dan autochtone leerlingen. Verrassend genoeg is het effect van *taakmotivatie* niet meer significant na de toevoeging van de variabele zelfvertrouwen. Dit duidt erop dat beide variabelen voor een groot deel dezelfde variantie verklaren, waarbij de variabele *zelfvertrouwen* dominant is.

De modellen M_{15} tot M_{17} staan weergegeven in tabel 2. In model M_{15} wordt de variabele *werkhouding* toegevoegd. Deze variabele geeft het oordeel van de docent weer over de werkhouding van de leerling. De werkhouding wordt gewaardeerd op een schaal van 1 tot 5, waarbij 1 staat voor een zeer slechte werkhouding en 5 voor een heel goede werkhouding. Het effect van geslacht neemt nu erg toe. De werkhouding van meisjes wordt dus slechter gewaardeerd door docenten dan de werkhouding van jongens. Dat is in tegenspraak met de gegevens van het OC&W (2013b), waarin juist werd gesteld dat de werkhouding van meisjes door docenten hoger wordt gewaardeerd dan de werkhouding van jongens. Verder valt op dat het effect van *opleidingsniveau* van zowel de vader als de moeder verder afneemt. Het effect van de variabele *zorgleerling* neemt af met ruim een punt, waaruit volgt dat de werkhouding van zorgleerlingen slechter is dan van gewone leerlingen.

In model M_{16} wordt de variabele *populariteit in de klas* toegevoegd. Deze variabele is een oordeel van de docent en geeft onder andere aan of de leerling goed kan opschieten met klasgenoten, of de leerling bij klasgenoten populair is en of de leerling vriend(inn)en in de klas heeft. De variabele is weer gemeten in een schaal van 1 tot 5, waarbij 1 aangeeft dat de leerling weinig populair is en 5 dat de leerling heel populair is. De variabele heeft een positief effect: populaire leerlingen halen tot 0,6 punt hogere CITO-scores, maar het effect is klein en niet significant.

In de modellen M_{15} en M_{16} zijn de variabelen *zelfvertrouwen* en *werkhouding* belangrijke verklarende variabelen voor de te behalen CITO-score. In deze modellen hebben de schalen *gedrag* en *taakmotivatie* een negatief effect op de CITO-score. Dit komt omdat de schalen *zelfvertrouwen* en *taakmotivatie* en de schalen *gedrag* en *werkhouding* vrij hoog correleren, respectievelijk $r=0.5034$ en $r=0.5432$. *Zelfvertrouwen* en *taakmotivatie* verklaren dus voor een groot deel dezelfde variantie, evenals *gedrag* en *werkhouding*. In de volgende modellen worden alleen de schalen *zelfvertrouwen* en *werkhouding* nog meegenomen.

In model M_{17} worden ook de variabelen *verblijfsduur in Nederland* en *populariteit in de klas* uit het model gehaald. Deze variabelen zijn namelijk niet significant. Door te werken met minder variabelen neemt het aantal waarnemingen dat meegenomen wordt in het model toe. Tevens blijkt dat de variantie-niveaus nauwelijks veranderen.

		Model M15	Model M16	Model M17
<i>Log Likelihood</i>		- 14.670	- 14.639	- 16.991
<i>Aantal observaties</i>		4.400	4.390	5.092
<i>AIC</i>		29.383	29.322	34.019
<i>Constante</i>		513,00	512,72	510,83
Fixed				
	Geslacht	- 2,37 *** (0.216)	- 2,35 *** (0.216)	- 2,39 *** (0.199)
	Opleiding Vader LBO	0,01 (0.497)	- 0,02 (0.499)	- 0,09 (0.442)
	Opleiding Vader MBO	1,30 ** (0.508)	1,26 ** (0.510)	1,08 ** (0.448)
	Opleiding Vader HBO-WO	2,72 *** (0.535)	2,65 *** (0.537)	2,57 *** (0.477)
	Opleiding Moeder LBO	0,50 (0.490)	0,50 (0.491)	0,54 (0.439)
	Opleiding Moeder MBO	2,56 *** (0.494)	2,56 *** (0.495)	2,76 *** (0.436)
	Opleiding Moeder HBO-WO	4,93 *** (0.533)	4,92 *** (0.534)	5,09 *** (0.480)
	Geboorteland Vader Turkije	- 2,58 *** (0.561)	- 2,58 *** (0.562)	- 2,60 *** (0.500)
	Geboorteland Vader Marokko	- 2,27 *** (0.584)	- 2,29 *** (0.585)	- 2,23 *** (0.509)
	Geboorteland Vader Suriname of Antillen	- 3,65 *** (0.868)	- 3,63 *** (0.868)	- 3,37 *** (0.712)
	Geboorteland Vader Overig	- 0,06 (0.465)	- 0,06 (0.466)	- 0,58 (0.423)
	Gezinssamenstelling	- 1,10 ** (0.468)	- 1,13 ** (0.470)	- 0,96 ** (0.451)
	Verblijfsduur in Nederland	- 1,54 (1.290)	- 1,52 (1.291)	
	Gedrag (schaal 1 tot 5)	- 0,61 *** (0.151)	- 0,65 *** (0.154)	
	Zorgleerling	- 5,68 *** (0.268)	- 5,68 *** (0.269)	- 6,10 *** (0.251)
	Taakmotivatie	- 0,45 ** (0.198)	- 0,46 *** (0.199)	
	Zelfvertrouwen	3,54 *** (0.201)	3,54 *** (0.202)	3,30 *** (0.167)
	Werkhouding	3,42 *** (0.143)	3,40 *** (0.145)	3,17 *** (0.115)
	Populariteit in de klas		0,15 (0.138)	
Random				
	var(Residual)	42,6	42,6	42,9
	var(School)	8,0	8,1	8,3

Tabel 2: Geschatte Modellen M₁₅ tot M₁₇

De modellen M₁₈ tot M₂₀ staan in tabel 3 op de volgende bladzijde. In model M₁₈ wordt de variabele *afhankelijkheid* toegevoegd. Deze variabele is een oordeel van de docent over de afhankelijk van de leerling van de docent. De schaal loopt van 1 tot 5, waarbij 1 aangeeft dat de leerling niet afhankelijk is (dus een hoge mate van zelfstandigheid toont) en 5 aangeeft dat de leerling zeer afhankelijk is (dus een hoge mate van onzelfstandigheid toont). De afhankelijkheid van de docent heeft een significante invloed op de CITO-score. Een leerling die zeer afhankelijk is van de docent, behaalt een CITO-score die tot ruim 5 punten lager ligt dan een leerling die zeer zelfstandig werkt. Uit andere onderzoeken blijkt ook dat de mate van zelfstandigheid invloed heeft op schoolprestaties (Goes, 2009). Als leerlingen autonoom zijn op school, vergroot dit de intrinsieke motivatie, wat zorgt voor hogere schoolprestaties. Verder valt op dat het effect van de variabele *zorgleerling* afneemt. Dat geeft aan dat zorgleerlingen afhankelijker zijn van de docent.

In model M₁₉ wordt de variabele *ouderbetrokkenheid* toegevoegd. In deze variabele is opgenomen of in dit gezin de ouders actief betrokken zijn bij de school, of leren wordt bevorderd en of de ouders het kind bij het leren tot steun zijn. Ook in dit geval is er sprake van een schaal van 1 tot 5, waarbij 1 staat voor een lage ouderbetrokkenheid en 5 voor een hoge ouderbetrokkenheid. Door de toevoeging van deze variabele verandert het model aanzienlijk. Zo neemt het effect van het opleidingsniveau van de vader op MBO-niveau af en wordt insignificant. Het gevonden effect van ouderbetrokkenheid is in overeenstemming met de literatuurstudie van Menheere & Hooge (2010). Uit een overzicht van negen studies blijkt dat de betrokkenheid van ouders van positieve invloed is op de schoolse ontwikkeling van hun kinderen. Ouderbetrokkenheid zorgt voor socialer gedrag, meer doorzettingsvermogen en een positievere houding ten aanzien van leren. Goed ouderschap biedt een veilige en stabiele thuisomgeving, stimuleert de intellectuele ontwikkeling en kent hoge verwachtingen van het kind (Menheere & Hooge, 2010). Vanuit goed ouderschap is te verklaren waarom de variabele *gezinssamenstelling* niet significant wordt. In een gezin met gescheiden ouders is er een minder stabiele thuisomgeving, wat direct invloed heeft op de prestaties van het kind.

In model M₂₀ wordt de variabele *welbevinden met medeleerlingen* toegevoegd. Deze variabele is samengesteld door leerlingen een vragenlijst in te laten vullen. In de vragenlijst kwam onder andere ter sprake of leerlingen veel contact hebben met klasgenoten, of de leerling goed overweg kan met klasgenoten, of de leerling het leuk vindt om om te gaan met klasgenoten en of de leerling de klas waarin hij zit leuk vindt. De variabele is samengesteld door een factoranalyse. De (1) staat voor een laag en de (5) voor een hoog welbevinden met medeleerlingen. Verrassend genoeg heeft de variabele een significant negatief effect op de CITO-score. Een mogelijke oorzaak ligt in de samenstelling van deze variabele, waardoor de variantie tussen leerlingen niet goed verklaard wordt. Deze aanname lijkt ondersteund te worden door het COOL-rapport, dat aangeeft dat de verklaarde variantie²⁶ slechts 48% is. Daarbij is ook de Cronbach's Alpha met 0.76 aan de lage kant. Het kan zijn dat de verschillende omgekeerde vragen (zoals 'ik voel me soms alleen in de klas') verkeerd zijn ingevuld door de leerlingen.

²⁶ Door de factoren welbevinden met docent, welbevinden met medeleerlingen, zelfvertrouwen en taakmotivatie gezamenlijk wordt 48% van de variantie verklaard.

	Model M18	Model M19	Model M20
<i>Log Likelihood</i>	- 16.928	- 16.842	- 16.824
<i>Aantal observaties</i>	5.087	5.078	5.077
<i>AIC</i>	33.895	33.725	33.689
<i>Constante</i>	514,97	512,48	515,50
Fixed			
Geslacht	- 2,23 *** (0.198)	- 2,13 *** (0.196)	- 2,15 *** (0.196)
Opleiding Vader LBO	- 0,12 (0.438)	- 0,38 (0.434)	- 0,40 (0.433)
Opleiding Vader MBO	1,00 ** (0.444)	0,51 (0.441)	0,42 (0.440)
Opleiding Vader HBO-WO	2,52 *** (0.473)	1,92 *** (0.470)	1,80 *** (0.469)
Opleiding Moeder LBO	0,52 (0.435)	0,22 (0.431)	0,24 (0.430)
Opleiding Moeder MBO	2,75 *** (0.433)	2,07 *** (0.432)	2,09 *** (0.431)
Opleiding Moeder HBO-WO	5,01 *** (0.476)	4,12 *** (0.478)	4,12 *** (0.476)
Geboorteland Vader Turkije	- 2,50 *** (0.496)	- 2,04 *** (0.493)	- 1,97 *** (0.492)
Geboorteland Vader Marokko	- 2,16 *** (0.505)	- 1,88 *** (0.501)	- 1,79 *** (0.500)
Geboorteland Vader Suriname of Antillen	- 3,29 *** (0.707)	- 2,45 *** (0.703)	- 2,40 *** (0.702)
Geboorteland Vader Overig	- 0,48 (0.421)	- 0,15 (0.418)	- 0,17 (0.417)
Gezinssamenstelling	- 0,93 ** (0.447)	-0,49 (0.444)	- 0,54 (0.442)
Zorgleerling	- 5,79 *** (0.251)	- 5,77 *** (0.248)	- 5,77 *** (0.247)
Zelfvertrouwen	3,19 *** (0.165)	3,05 *** (0.164)	3,21 *** (0.166)
Werkhouding	2,83 *** (0.120)	2,43 *** (0.124)	2,42 *** (0.124)
Afhankelijkheid	- 1,33 *** (0.137)	- 1,31 *** (0.135)	- 1,40 *** (0.136)
Ouderbetrokkenheid		1,41 *** (0.130)	1,44 *** (0.130)
Welbevinden met medeleerlingen			- 0,82 *** (0.149)
Random			
var(Residual)	42,0	41,1	40,8
var(School)	8,6	8,5	8,6

Tabel 3: Geschatte Modellen M₁₈ tot M₂₀

Aan de andere kant kan het gevonden effect ook een geldige weergave van de werkelijkheid zijn. Slechte cijfers kunnen statusverhogend werken en de leerling populairder maken (JMouders, 2010). Een andere verklaring kan zijn, dat leerlingen die veel contact hebben met andere leerlingen minder geconcentreerd zijn op hun schoolwerk (of huiswerk) en dat daardoor de invloed van welbevinden op de CITO-score negatief is.

De modellen M_{21} tot M_{23} zijn te vinden in tabel 4. In model M_{21} wordt de *variabele welbevinden met de docent* toegevoegd. Ook deze variabele is samengesteld uit een door leerlingen ingevulde vragenlijst. In deze vragenlijst kwam onder andere aan de orde of de docent meestal weet hoe de leerling zich voelt, of de leerling met de docent kan praten over problemen, of de leerling zich op het gemak voelt bij de docent en of de leerling een goed contact heeft met de docent. De schaalvariabele loopt van 1 tot 5, waar (1) staat voor een laag en (5) voor een hoog welbevinden met de docent. De coëfficiënt van de variabele *welbevinden met de docent* is negatief. Ook in dit geval is de verklaarde variantie door de factoren laag²⁶, waarbij de Cronbach's Alpha 0.74 is. Het gevonden effect is in tegenstelling met resultaten uit eerdere onderzoeken. In een omvangrijk onderzoek van het NWO blijkt juist dat leerlingen op school beter presteren als ze een positieve relatie hebben met hun leraar en meer betrokken zijn bij de lessen (Koomen, Spilt, Roorda, Oort, & Thijs, 2010).

Het is bekend dat resultaten van eerdere toetsen grote voorspellende waarde hebben. Daarom worden de toetsen die door de meeste leerlingen zijn gemaakt toegevoegd. Voor taal is dit de *Drie Minuten Toets (DMT)* en voor rekenen is dit de toets *Rekenen / Wiskunde LOVS*. Door de toevoeging van deze toetsen in model M_{22} wordt het aantal leerlingen dat meegenomen wordt in het model beperkt tot 3.210. De score voor de *DMT* loopt uiteen van 18 tot 124 gescoorde punten, de score voor de *LOVS* loopt uiteen van 24 tot 169 gescoorde punten. De toetsen hebben een grote invloed op het model. De variabelen *geslacht*, *gezinssamenstelling* en *welbevinden met de docent* zijn niet meer significant. Aangezien het effect van welbevinden met de docent ook niet in overeenstemming was met de theoretische verwachting, biedt dit een basis om deze variabele uit het model te verwijderen.

In model M_{22} verandert ook het effect van het *opleidingsniveau* flink, zowel bij de vader als de moeder. Bij de vader is alleen het opleidingsniveau HBO-WO nog een significante voorspeller voor de CITO-score, bij de moeder worden de effecten van haar opleiding op het MBO en HBO-WO een stuk kleiner. De resultaten van eerder gemaakte toetsen verklaren dus voor een deel dezelfde variantie als de lagere opleidingsniveaus. Hogere opleidingsniveaus van ouders blijven echter een significant positief effect uitoefenen op de CITO-score. Opvallend is dat leerlingen met een Marokkaanse vader niet meer significant slechter scoren, in tegenstelling tot leerlingen met een Turkse, Surinaamse of Antilliaanse vader.

In model 23 worden alle variabelen die in model 21 of 22 niet significant waren verwijderd, namelijk *geslacht*, *opleiding vader LBO*, *opleiding vader MBO*, *opleiding moeder LBO*, *geboorteland vader Marokko*, *geboorteland vader overig*, *gezinssamenstelling* en *welbevinden met de docent*.

		Model M21	Model M22	Model M23
<i>Log Likelihood</i>		- 16.806	- 9.624	- 10.219
<i>Aantal observaties</i>		5.073	3.210	3.405
<i>AIC</i>		33.666	19.295	20.471
<i>Constante</i>		516,16	468,05	468,36
Fixed				
	Geslacht	- 2,13 *** (0.196)	0,14 (0.182)	
	Opleiding Vader LBO	- 0,38 (0.433)	- 0,05 (0.388)	
	Opleiding Vader MBO	0,43 (0.441)	0,43 (0.397)	
	Opleiding Vader HBO-WO	1,80 *** (0.470)	1,16 *** (0.425)	0,91 *** (0.215)
	Opleiding Moeder LBO	0,25 (0.430)	0,64 * (0.382)	
	Opleiding Moeder MBO	2,07 *** (0.431)	1,50 *** (0.383)	1,11 *** (0.207)
	Opleiding Moeder HBO-WO	4,09 *** (0.476)	2,50 *** (0.428)	1,99 *** (0.275)
	Geboorteland Vader Turkije	- 1,96 *** (0.492)	- 1,14 ** (0.454)	- 1,02 *** (0.389)
	Geboorteland Vader Marokko	- 1,84 *** (0.499)	- 0,37 (0.458)	
	Geboorteland Vader Suriname of Antillen	- 2,43 *** (0.701)	- 1,13 * (0.646)	- 0,77 (0.615)
	Geboorteland Vader Overig	- 0,20 (0.417)	0,13 (0.401)	
	Gezinssamenstelling	- 0,53 (0.442)	- 0,32 (0.399)	
	Zorgleerling	- 5,75 *** (0.247)	- 1,84 *** (0.246)	- 1,93 *** (0.238)
	Zelfvertrouwen	3,31 *** (0.169)	1,19 *** (0.155)	1,31 *** (0.147)
	Werkhouding	2,46 *** (0.125)	0,84 *** (0.121)	0,83 *** (0.108)
	Afhankelijkheid	- 1,37 *** (0.136)	- 0,62 *** (0.130)	- 0,55 *** (0.126)
	Ouderbetrokkenheid	1,47 *** (0.130)	0,64 *** (0.122)	0,72 *** (0.115)
	Welbevinden met medeleerlingen	- 0,72 *** (0.153)	- 0,53 *** (0.141)	- 0,53 *** (0.132)
	Welbevinden met docent	- 0,48 *** (0.164)	0,10 (0.150)	
	Toets Rekenen / Wiskunde		0,46 *** (0.009)	0,46 *** (0.009)
	Toets Taal		0,07 *** (0.007)	0,08 *** (0.007)
Random				
	var(Residual)	40,7	20,9	21,0
	var(School)	8,6	9,2	9,2

Tabel 4: Geschatte Modellen M₂₁ tot M₂₃

Aan de hand van model 23 kunnen we de gevonden effecten tussentijds toetsen aan de verwachtingen die er op voorhand waren met betrekking tot de schoolse leerlingkenmerken. Alleen effecten die in model M₂₃ significant waren worden meegenomen in de beoordeling. Er was vastgesteld dat de CITO-score relatief hoger is als:

- ✓ De leerling taakgemotiveerd is;
- ✓ De leerling een positief zelfbeeld en een hoge mate van zelfstandigheid heeft;
- ✓ De leerling een goede werkhouding heeft.
- ✓ De leerling betrokken ouders heeft;
- ✓ De leerling een goede relatie heeft met docenten en medeleerlingen;
- ✓ De leerling op eerdere toetsen hoog scoorde.

Aan de hand van model 23 kan vergelijking [V.3] worden opgesteld:

$$\begin{aligned}
 CITO - score = & [468,36 + \mu_{school}] + \alpha X_{oplvader-hbowo} + \alpha X_{oplmoeder-mbo} \\
 & + \alpha X_{oplmoeder-hbowo} - \alpha X_{gl-turkije} - \alpha X_{zorgleerling} + \alpha X_{zelfvertrouwen} \\
 & + \alpha X_{werkhouding} - \alpha X_{afhankelijkheid} + \alpha X_{ouderbetrokkenheid} \\
 & - \alpha X_{welbevinden-medeleerlingen} + \alpha X_{toets-rekenen} + \alpha X_{toets-taal} + \varepsilon_{residual}
 \end{aligned}$$

Met behulp van de geschatte parameters uit model M₂₃ kunnen de coëfficiënten α worden ingevuld, waardoor vergelijking [V.4] ontstaat:

$$\begin{aligned}
 CITO - score = & [468,36 + \mu_{9,2}] + 0,91 * X_{oplvader-hbowo} + 1,11 * X_{oplmoeder-mbo} \\
 & + 1,99 * X_{oplmoeder-hbowo} - 1,02 * X_{gl-turkije} - 1,93 * X_{zorgleerling} \\
 & + 1,31 * X_{zelfvertrouwen} + 0,83 * X_{werkhouding} - 0,55 * X_{afhankelijkheid} \\
 & + 0,72 * X_{ouderbetrokkenheid} - 0,53 * X_{welbevinden-medeleerlingen} \\
 & + 0,46 * X_{toets-rekenen} + 0,08 * X_{toets-taal} + \varepsilon_{residual}
 \end{aligned}$$

Aan de hand van vergelijking [V.4] kan worden geconcludeerd dat het voorspellend effect van eerder gemaakte toetsen erg groot is. Als een leerling op eerdere toetsen hoog scoorde, is het waarschijnlijk dat de leerling ook een hoge CITO-score haalt. Tevens is de score hoger als de leerling een goede werkhouding laat zien (wat hoog correleert met *gedrag*) en als de leerling veel zelfvertrouwen heeft (wat hoog correleert met *taakmotivatie*). Leerlingen die zelfstandig zijn, scoren ook hoger dan leerlingen die onzelfstandig zijn. Daarnaast scoren leerlingen die betrokken ouders hebben, significant hoger. De variabele *welbevinden met de docent* is niet significant en heeft dus geen invloed op de CITO-score. De andere variabele die een afwijkend beeld laat zien ten opzichte van de hypothese, is de variabele *welbevinden met medeleerlingen*. Deze variabele laat een negatieve relatie zien; de CITO-score is hoger als het welbevinden met medeleerlingen lager is. Een verklaring hiervoor kan zijn dat de variabele geen goed beeld geeft van de relatie tussen leerlingen. Anderzijds kan het effect ook werkelijk bestaan en aangeven dat een deel van de populariteit van leerlingen ontstaat door het halen van slechte cijfers, waardoor leerlingen 'stoer' gevonden worden. De tot nu toe gevonden effecten wijken, afgezien van de variabelen *welbevinden met medeleerlingen* en *met docenten*, niet af van de vastgestelde effecten uit eerder onderzoek.