

## **Groepsgrootte en de schoolresultaten van leerlingen**

ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Department of Economics

Begeleider: dr. Y. Adema

Naam: Arjan van Rijn

Studentnummer: 369313

e-mailadres: arjanvanrijn121@live.nl

### **Samenvatting**

Over de invloed van groepsgrootte op toetsscores is in het verleden al aardig wat geschreven. Deze onderzoeken hadden meestal, in tegenstelling tot dit onderzoek, een instrumentele variabele nodig om tot conclusies te komen die vrij zijn van systematische fouten. Dit onderzoek richt zich op basisschoolleerlingen in Nederland en onderzoekt het verband tussen groepsgrootte en Cito-scores. Ook wordt aandacht besteed aan verschillen in achtergronden van leerlingen en aan verschillende onderwijssystemen die in Nederland gehanteerd worden. Verder worden uitkomsten omgerekend naar gestandaardiseerde effecten, zodat resultaten kunnen worden vergeleken met onderzoeken die andere toetsen gebruiken. De conclusie van dit onderzoek is dat er een significant negatief verband bestaat tussen groepsgrootte en Cito-scores. Het effect is echter minimaal en effecten van leerlingen uit achtergestelde milieus is vele malen groter. Als laatste blijkt dat leerlingen op protestants-christelijke scholen beter presteren dan leerlingen op openbare scholen en leerlingen op het algemeen bijzonder onderwijs presteren slechter dan leerlingen op het reguliere onderwijs.

# Inhoudsopgave

1. Introductie.....	3
2. Het Nederlandse basisschool systeem .....	6
2.1 <i>Leerplicht en kwalificatieplicht</i> .....	6
2.2 <i>Basisschool</i> .....	6
2.3 <i>Schooltype</i> .....	7
2.3.1 <i>Montessori</i> .....	7
2.3.2 <i>Dalton</i> .....	8
2.3.3 <i>Jenaplan</i> .....	8
3. Cito-toets .....	9
4. Cool5-18 en RTL nieuws.....	10
4.1 <i>Cool5-18</i> .....	10
4.2 <i>RTL nieuws</i> .....	10
5. Theoretisch raamwerk .....	12
5.1 <i>Inleiding</i> .....	12
5.2 <i>De regel van Maimonides</i> .....	12
5.3 <i>Vergelijken van onderzoeken</i> .....	12
5.4 <i>Angrist &amp; Lavy</i> .....	12
5.5 <i>Lavy</i> .....	14
5.6 <i>Hoxby</i> .....	14
5.7 <i>Boozer &amp; Rouse</i> .....	15
5.8 <i>Akerhielm</i> .....	16
5.9 <i>Dobbelsteen et al.</i> .....	16
7. Data.....	23
8. Resultaten .....	27
8.1 <i>Model 1</i> .....	27
8.2 <i>Model 2</i> .....	27
8.4 <i>Model 4</i> .....	29
8.5 <i>Robuustheid van de resultaten</i> .....	30
8.6 <i>Interpretatie van de resultaten</i> .....	31
Bibliografie.....	36
Appendix.....	38

## 1. Introductie

"Nederland moet de kenniseconomie versterken" deze boodschap kwam naar voren in het regeerakkoord van Rutte II (Rijksoverheid a, 2013). Het versterken van de kenniseconomie kan gebeuren door te investeren in het onderwijs. Dit kan op verschillende manieren. Bijvoorbeeld door meer personeel aan te nemen om de groepsgrootte te laten dalen of door de lestijden te verlengen. Schoolprestaties zijn zeer belangrijk voor basisschoolleerlingen tot universiteitsstudenten. Werkstukken, toetsen en gestandaardiseerde testen horen daar allemaal bij. Nederland is een kenniseconomie en mensen moeten daarom zo hoog mogelijk zijn opgeleid. Voor de basisscholen is de Cito-toets de afsluiting na acht jaar onderwijs. Vanaf het schooljaar 2014-2015 wordt een eindtoets verplicht gesteld door de overheid voor alle basisschoolleerlingen. Scholen mogen echter wel zelf kiezen of ze de Cito-toets, ROUTE 8 of IEP Eindtoets maken (Rijksoverheid b).

De score van de Cito-toets of andere eindtoets mag niet gebruikt worden om aan de poort te selecteren. Door de toets te verplaatsen van februari naar april wordt dit voorkomen, omdat de adviezen voor leerlingen dan al gegeven zijn door de leraren. Alleen in het geval van een hoge score mag het advies naar boven worden aangepast. Verder mogen ook geen aanvullende voorwaarde worden gesteld door middelbare scholen indien de Cito-toets tegenvalt.

De score voor een eindtoets ligt niet alleen aan de intelligentie van de leerling, maar ook aan de groepsgrootte, afkomst en opleidingsniveau van ouders. In Nederland is geen wettelijk maximum gesteld aan de hoeveelheid leerlingen in een klas (ANP, 2014). Uit een brief van de staatssecretaris Dekker van Onderwijs blijkt dat een gemiddelde klas in Nederland uit 23 leerlingen bestaat. Ook blijkt uit de brief dat er daadwerkelijk klassen zijn van 35 tot 39 leerlingen (Rijksoverheid c). Leraren vakbond 'Leraren in Actie' is in 2013 een burger initiatief genaamd 'Stop de overvolle klassen' gestart om de maximale groepsgrootte terug te brengen naar 28 leerlingen (Bogaerdt, 2013). In 2014 is over dit voorstel gesproken in de Tweede Kamer nadat de benodigde handtekeningen waren verzameld (Overheid, 2014). Staatsecretaris Dekker voelt niets voor het stellen van een maximum aan de groepsgrootte en komt dus ook niet met een wetsvoorstel (Bogaerdt, 2014).

Sinds RTL nieuws erin is geslaagd om gegevens over de Cito-score openbaar te maken, door gebruik te maken van de Wet Openbaarheid van Bestuur, kunnen ouders basisscholen rangschikken aan de hand van de gemiddelde scores. Dit kan echter onbedoelde effecten met zich meebrengen. Basisscholen konden voor 2014-2015 zwakke leerlingen uitsluiten van de eindtoets om zo een kunstmatig hoog gemiddelde te krijgen. Ook is het aantal groepen 8 per basisschool niet meegenomen in het onderzoek. Deze scriptie richt zich daarom op de onderzoeksvraag

*'Wat is het effect van de hoeveelheid leerlingen in een klas op de Cito-toets score in Nederland?'*

Om deze vraag te beantwoorden zijn een aantal hypotheses opgesteld om tot een uitgebreider antwoord te komen.

*'Leerlingen van bijzondere basisscholen scoren op de Cito-toets niet significant lager dan leerlingen van openbare basisscholen ceteris paribus'*

De Cito-toets test de kennis en vaardigheden van leerlingen op het gebied van rekenen en taalvaardigheid. Deze vaardigheden zouden niet moeten verschillen tussen seculiere en niet seculiere basisscholen.

*'Leerlingen van algemeen bijzondere basisscholen hebben significant slechtere resultaten dan leerlingen van openbare en bijzondere basisscholen ceteris paribus'*

In Nederland worden verschillende types onderwijs door de Staat gefinancierd. De grote meerderheid van de leerlingen op de basisschool volgt openbaar of bijzonder onderwijs. Algemeen bijzonder onderwijs, zoals Montessori, Dalton en Jenaplan onderwijs verschilt fundamenteel van de andere onderwijstypes. Doordat leerlingen meer vrijheid krijgen in wat ze doen op school kan het zo zijn dat er verschil zit in de Cito-toetsscores van aan de ene kant openbaar en bijzonder onderwijs en aan de andere kant algemeen bijzonder onderwijs.

Deze scriptie is onderverdeeld in verschillende hoofdstukken. Het doel van deze scriptie is onderzoeken of er een significant verband bestaat tussen groepsgrootte en Cito-scores op basisscholen in Nederland. Verder wordt alleen ingegaan op de groepsgrootte en niet op het verlengen van lestijden, zoals beschreven in

verschillende andere onderzoeken. De scriptie begint met een korte uitleg over de leerplicht en het Nederlandse basisschool systeem. Hierna wordt kort ingegaan op de Cito-toets, het Cool5-18 onderzoek en de analyse van RTL nieuws van de openbare Cito-scores per school. De data uit Cool5-18 staan centraal in deze scriptie. Vervolgens zullen verschillende wetenschappelijke papers behandeld worden in het theoretisch raamwerk over het effect van de groepsgrootte. De overeenkomsten en verschillen zullen aan bod komen, evenals de beperkingen. Ook wordt getracht een verklaring te geven voor de verschillen in resultaten en significantie niveaus. Om het effect te kunnen bepalen zal een OLS regressievergelijking worden opgesteld aan de hand van variabelen uit het Cool5-18 onderzoek. De resultaten van deze regressievergelijking zullen worden geïnterpreteerd en kunnen mogelijk de basis zijn voor nieuw beleid op dit gebied. In het hoofdstuk resultaten zal ook verder worden ingegaan op de verschillen tussen bepaalde groepen leerlingen qua achtergrond en school. Verder zullen resultaten teruggekoppeld worden aan de onderzoeksvraag en de hypothesen. Ten slotte zal ingegaan worden op de beperkingen en limitaties van dit onderzoek.

## **2. Het Nederlandse basisschool systeem**

### *2.1 Leerplicht en kwalificatieplicht*

In Nederland geldt een leerplicht voor iedere inwoner tussen de 5 en 18 jaar. Vanaf de eerste dag van de maand nadat het kind 5 jaar is geworden treedt de leerplicht in werking. Deze wet wordt de leerplicht wet 1969 genoemd en is sindsdien een aantal keer aangepast. De wet zorgt ervoor dat alle kinderen in Nederland onderwijs krijgen. Ouders zijn verplicht kinderen vanaf deze leeftijd naar school te sturen en worden door de leerplichtambtenaar aangepakt als dit niet het geval is. De straf voor het ontduiken van de leerplicht is een proces verbaal en een boete voor de ouders. Vanaf 12 jaar is het kind ook persoonlijk aansprakelijk voor spijbelen en krijgt mogelijk ook een proces verbaal en een leerstraf. Een grote uitzondering op de leeftijdsgrens van de leerplicht is de kwalificatieplicht. Een leerling moet altijd tot zijn of haar 18<sup>e</sup> levensjaar naar school gaan tenzij aan specifieke eisen is voldaan. De leerplicht vervalt bij het behalen van een mbo niveau 2, havo of vwo diploma. Met deze kwalificatieplicht tracht de overheid jongeren zo lang mogelijk in het onderwijs systeem te houden en te zorgen dat ze zo hoog mogelijk zijn opgeleid.

### *2.2 Basisschool*

Zoals hierboven beschreven valt ieder in Nederland wonend persoon onder de leeftijd van 18 jaar in principe onder de leerplicht wet. De leeftijden tussen 5 en 12 zijn verenigd in de basisschool, ook wel het primaire onderwijs genoemd. De basisschool bestaat uit 8 groepen met leerlingen van ongeveer dezelfde leeftijd. Soms blijven leerlingen zitten of slaan juist een klas over. De basisscholen worden door de overheid gefinancierd op basis van het aantal leerlingen en het leerlinggewicht. Een weegfactor wordt door de overheid toegekend aan kinderen met ouders met weinig opleiding. Er zijn drie weegfactoren waar leerlingen ingedeeld kunnen worden. Leerlingen waarvan beide ouders of de ouder die voor het kind zorgt geen hoger diploma hebben of heeft dan vmbo basis of kader worden ingedeeld in weegfactor 0.3. Als een of beide ouders maximaal een basisschool diploma heeft of hebben dan wordt de leerling in weegfactor 1.2 ingedeeld. Alle andere leerlingen zijn weegfactor 0 (Rijksoverheid d). Vroeger waren er nog meer weegfactoren, zoals 0.25, 0.4, 0.7 en 0.9 maar deze zijn afgeschaft. Scholen krijgen meer geld voor

leerlingen uit achterstandsgezinnen en de achterstand zou daarom idealiter weggewerkt moeten zijn. In de praktijk blijkt dit echter niet het geval te zijn.

In Nederland mag iedereen een basisschool beginnen naar eigen overtuiging. De overheid moet dit financieren en mag geen onderscheid maken tussen verschillende scholen. Nieuwe scholen moeten wel opgenomen worden in het 'Plan van Scholen'. Hier zijn voorwaarden aan verbonden, zoals een minimaal aantal leerlingen. Het minimale aantal leerlingen hangt af van de bevolkingsdichtheid in het gebied van de nieuwe school (Rijksoverheid e).

### *2.3 Schooltype*

Er zijn verschillende soorten basisscholen, namelijk openbare, bijzondere en algemeen bijzondere. Openbare basisscholen zijn scholen zonder levensovertuiging of godsdienst. Bijzondere scholen hebben juist wel een levensovertuiging en dragen dit ook expliciet uit. Waar openbare scholen geen leerlingen of leraren mogen weigeren vanwege geloofsovertuiging mogen bijzondere scholen dit wel doen. Enkele voorbeelden van bijzondere scholen zijn: rooms-katholieke, protestants-christelijke, islamitische, joodse en hindoeïstische scholen. Algemeen bijzondere scholen opereren vanuit hun eigen visie op onderwijs of opvoedkundige overtuiging. Voorbeelden van dit soort scholen zijn Montessori, Dalton en Jenaplan scholen. Combinaties van de drie soorten scholen zijn ook mogelijk. Zo zijn er bijvoorbeeld openbare Jenaplan scholen of rooms-katholieke Dalton scholen. In het vervolg van de scriptie wordt met regulier onderwijs het openbaar en bijzonder onderwijs bedoeld.

#### *2.3.1 Montessori*

Een Montessori basisschool is anders ingedeeld dan een openbare of bijzondere basisschool. In plaats van 8 groepen met ieder een leerjaar zijn er maar 3 groepen. Deze groepen zijn ingedeeld op leeftijd. Kinderen van 4 tot 6, 7 tot 9 en 10 tot 12 zitten samen in een klas. Het Montessori onderwijs draait om zelfontplooiing. Leerlingen mogen zelf beslissen wat ze doen en in welk tempo. De leraar is er om de kinderen te motiveren, maar geeft geen cijfers. Kinderen corrigeren eigen werk en er zijn ook geen tussentijdse rapporten.

### *2.3.2 Dalton*

De belangrijkste punten binnen het Dalton onderwijs zijn vrijheid, zelfstandigheid en samenwerking. Ook bij dit type onderwijs mag de leerling zelf bepalen wat hij of zij gaat doen tijdens de lessen. De docent zorgt ervoor dat de leerlingen wel binnen de grenzen van de lesstof blijven, maar verder zijn ze vrij. Leerlingen moeten zelf inschatten hoelang ze met een taak bezig zullen zijn en na afloop wordt er geëvalueerd. Hoe ouder de leerling wordt hoe meer verantwoordelijkheid bij de leerling komt te liggen. Ook wordt veel samengewerkt in groepen om kinderen voor te bereiden op hun latere leven.

### *2.3.3 Jenaplan*

Jenaplan onderwijs draait om opvoeding. De basisschool neemt een gedeelte van de opvoeding van de kinderen over van de ouders. Kinderen moeten van elkaar leren is de visie en daarom zitten er net als bij het Montessori onderwijs meerdere leerjaren in een groep. Leerlingen zitten in stamgroepen. In stamgroep 1 zitten leerlingen uit groep 1 en 2. Stamgroep 2 bestaat uit groep 3 tot en met 5 en 6 tot en met 8 zitten in stamgroep 3. Een stamgroep is dus eigenlijk een klas. De lesstof sluit aan bij de belevingswereld van de leerlingen en onderzoeken en ontdekken staat centraal.



### 3. Cito-toets

De Cito-toets is een jaarlijkse eindtoets voor basisschoolleerlingen in groep 8 en sinds 2015 verplicht voor alle basisscholen. Voor de toets zijn verschillende momenten tot afname. Scholen mogen kiezen voor een schriftelijke of een digitale toets. In mei is er nog een laatste ronde voor leerlingen die ziek waren of door een andere reden de toets niet of maar gedeeltelijk gemaakt hebben. In 2015 is met het verplichtstellen van de Cito-toets het een en ander veranderd aan het aantal opgaven en onderdelen. In de schooljaren 2007/2008, 2010/2011 en 2013/2014 bestond de toets uit vier verschillende onderdelen namelijk Nederlandse taal, rekenen, studievaardigheden en wereldoriëntatie. Voor de verdeling van de vragen over de categorieën zie Tabel 1 in de appendix. Deze tabel is een samenvoeging van informatie uit een document van CITO<sup>1</sup> en gegevens van het nationaal expertisecentrum leerplanontwikkeling<sup>2</sup>. De gehele toets bestaat uit 290 meerkeuze vragen met vier antwoordmogelijkheden per vraag. Het laatste onderdeel telt niet mee voor de uiteindelijke score en scholen mogen ervoor kiezen om dit onderdeel over te slaan.

De Cito-toets wordt gemeten op een schaal van 501 tot 550 punten. Leerlingen kunnen niet slagen of zakken voor de toets en deze dient alleen voor extra advies bij de schoolkeuze. De Cito-score mag niet gebruikt worden om het advies van de leraar naar beneden bij te stellen als deze zou tegenvallen. Alleen naar boven aanpassen is mogelijk en als er een erg groot verschil tussen zit kan een extra toets worden gemaakt. Tabel 2 en 3 in de appendix bevatten het schooladvies van CITO. Sommige middelbare scholen hebben in de brugklas gemengde klassen, voor deze scholen is Tabel 2 bedoeld. Als de leraar op de basisschool twijfelt over de leerling kan deze een gemengd advies geven. Dit betekent bijvoorbeeld dat een leerling in een havo/vwo klas geplaatst wordt met andere gelijke gevallen. Na afloop van het jaar bekijkt de middelbare school welke richting het beste is voor de leerling aan de hand van cijfers en gesprekken met ouders en de leerling zelf. Niet alle scholen hebben gemengde klassen in de brugklas en Tabel 3 geeft een advies voor deze gevallen.

---

<sup>1</sup> informatie omtrent indeling taal, rekenen en wereldoriëntatie onderdelen (CITO, 2011)

<sup>2</sup> informatie omtrent studievaardigheden (SLO)

## 4. Cool5-18 en RTL nieuws

### 4.1 Cool5-18

Cool5-18 is een cohort onderzoek onder leerlingen uit verschillende leeftijdscategorieën. Dit cohort onderzoek is een vervolg op eerdere cohort onderzoeken onder leerlingen in dezelfde leeftijdscategorieën. Het onderzoek wordt uitgevoerd in opdracht van het Nationaal Regieorgaan Onderwijsonderzoek (NRO) en het Ministerie van Onderwijs, Cultuur en Wetenschappen (OCW).

Aan dit onderzoek doen 550 basisscholen en 50.000 leerlingen mee (Cool5-18 a). Deze leerlingen zitten in groep 2, 5 en 8. Leerlingen worden ook op de middelbare school gevolgd, maar dit onderzoek focust zich alleen op groep 8 van de basisschool. Om de leerlingen door de tijd heen te kunnen volgen wordt het onderzoek in verschillende rondes uitgevoerd. De eerste ronde heeft plaatsgevonden in het schooljaar 2007-2008. Hierbij zijn bij alle leerlingen toetsen afgenomen of opgevraagd in de vorm van een Cito-score. Ook zijn enquêtes door de docenten ingevuld op individueel niveau. Dit betekent dat voor iedere leerling specifieke informatie aanwezig is met bijvoorbeeld huissituatie, afkomst, gedrag, weegfactor enzovoort. Dit biedt unieke mogelijkheden tot analyse. De leerlingen en ouders worden ook betrokken in dit onderzoek en vullen ook een enquête in. In schooljaar 2010-2011 is ronde twee van het onderzoek uitgevoerd. Wederom zijn de groepen 2,5 en 8 uitgebreid onderzocht. De eerste en tweede ronde zijn doormiddel van een uniek leerling-nummer aan elkaar verbonden. Zo kan de vooruitgang van leerlingen uit groep 2 en 5 in ronde 2 worden geanalyseerd (Cool5-18 b). Voor dit onderzoek zijn alleen de gegevens van groep 8 interessant. Ronde drie vond plaats in 2013-2014 en de resultaten zijn pas in mei 2015 gepubliceerd en kunnen daarom nog niet worden meegenomen in deze paper.

### 4.2 RTL nieuws

In 2013 heeft RTL nieuws voor elkaar gekregen dat de gemiddelde Cito-scores openbaar gemaakt moesten worden. De rechter vond dat deze gegevens van maatschappelijke belang waren (Rechtspraak, 2013). Wel moesten de gegevens niet herleidbaar zijn naar individuen. Daarom is ook alleen informatie beschikbaar in deze database op klasniveau en vanaf vijf leerlingen per klas. Bij minder dan vijf leerlingen

is er geen sprake meer van privacy, omdat deze gegevens makkelijker te herleiden zijn naar de personen in kwestie.

Met de gegevens heeft RTL nieuws samen met professor Jaap Dronkers van de Universiteit van Maastricht een model gemaakt waarmee alle basisscholen in Nederland met elkaar vergeleken konden worden (RTL nieuws, 2013). Dit gebeurde door de scholen een cijfer tussen de 1.0 en de 10.0 te geven. In dit onderzoek is rekening gehouden met de verschillende achtergronden van leerlingen. Scholen met veel achterstandsleerlingen, de weegfactor, worden in de eindscore gecompenseerd voor het mee laten doen van zwakke leerlingen aan de Cito-toets.

## 5. Theoretisch raamwerk

### 5.1 Inleiding

In dit hoofdstuk komen verschillende relevante academische papers aan bod. Deze papers onderzoeken allemaal de invloed van de groepsgrootte op de resultaten van leerlingen. De onderzoeksmethode verschilt per paper en ook de resultaten zijn niet overal hetzelfde. De resultaten uit deze papers kunnen gebruikt worden om de uitkomsten van deze scriptie te vergelijken.

### 5.2 De regel van Maimonides

Een van de eerste die nadacht over het verband tussen de groepsgrootte en de resultaten van studenten was Moses Maimonides. Hij pleitte al in de 12<sup>e</sup> eeuw voor een maximale grootte van schoolklassen. Maimonides vond dat klassen maximaal uit 40 leerlingen mochten bestaan. Als de 41<sup>ste</sup> leerling zich aanmeldde dan moest de klas worden gesplitst in een groep van 21 en een groep van 20. Het splitsen van klassen bij het bereiken van het maximum aantal leerlingen wordt de regel van Maimonides genoemd.

### 5.3 Vergelijken van onderzoeken

Om onderzoeken onderling te kunnen vergelijken moet er een universele maatstaf zijn om de resultaten te kunnen vergelijken. De coëfficiënten kunnen niet zonder meer met elkaar vergeleken worden omdat elk onderzoek een andere gestandaardiseerde test gebruikt. Om toch de uitkomsten te kunnen vergelijken worden in veel onderzoeken de coëfficiënten omgerekend naar een gestandaardiseerd effect. Dit gestandaardiseerde effect is onafhankelijk van de soort toets en kan dus wel worden vergeleken met andere onderzoeken. Het effect wordt berekend door de coëfficiënt van de regressieanalyse te delen door de standaarddeviatie van de gestandaardiseerde toets. Hieruit volgt een percentage dat geïnterpreteerd moet worden als de procentuele verandering, in termen van toets standaarddeviatie, die optreedt bij het verhogen van het aantal leerlingen in een klas.

### 5.4 Angrist & Lavy

De regel van Maimonides is in de moderne tijd geïntroduceerd door Angrist en Lavy (1999). In deze paper wordt het wettelijk maximum voor het aantal leerlingen in een klas in Israël gebruikt om het verband te onderzoeken tussen groepsgrootte en

schoolresultaten. De regel van Maimonides geeft een zeer betrouwbare exogene variatie. Doordat er een wettelijk maximum is gesteld mag je ervan uitgaan dat dit ook wordt nageleefd. Hierdoor hoeft er minder rekening gehouden te worden tussen verschillen in scholen. Scholen zijn immers verplicht om klassen te splitsen of ze dit nu willen of niet. Een groot voordeel voor dit onderzoek was dat de leerlingen willekeurig over de scholen verspreid zijn. In Israël mogen kinderen alleen naar scholen in de eigen buurt en moeten ook worden aangenomen door deze school. Voor ouders is het onmogelijk om hun kinderen naar andere scholen te sturen omdat de kwaliteit daar beter zou zijn. Dit kan alleen als ouders ook daadwerkelijk verhuizen naar de buurt waar de school zich bevindt. Het is onwaarschijnlijk dat ouders verhuizen puur en alleen om hun kinderen op een specifieke school te krijgen. Individuele data van leerlingen was helaas niet aanwezig en moest daarom geschat worden door middel van een model.

Angrist en Lavy hebben in hun regressie gebruik gemaakt van een variabele die het percentage achtergestelde leerlingen bevat (*PD*). Deze variabele kan vergeleken worden met de eerder genoemde weegfactor die in Nederland gebruikt wordt. Door het toevoegen van deze variabele wordt een poging gedaan om te corrigeren voor het onder brengen van zwakke leerlingen in kunstmatig kleine klassen. Dit leidde ertoe dat voor leesvaardigheid inderdaad een negatieve coëfficiënt gevonden werd (-.031). Echter voor wiskunde was de coëfficiënt nog steeds positief (0.076). De coëfficiënten moeten geïnterpreteerd worden op leerling niveau. Voor elke extra leerling in de klas gaat het klassengemiddelde voor leesvaardigheid met 0.031 punt omlaag en voor wiskunde juist met 0.076 punt omhoog.

Een groot probleem in dit soort onderzoeken is dat er tegenstrijdige resultaten uit kunnen komen. Dit komt grotendeels door het feit dat basisscholen verschillende groepen van hetzelfde leerjaar hebben, maar deze klassen zijn natuurlijk nooit identiek. Ze verschillen in samenstelling en mogelijkheden van individuele leerlingen. De groepsgrootte kan gemanipuleerd worden door alle leerlingen die moeite hebben met leren in een kleine klas te plaatsen en de goede leerlingen in een grote klas. Hierdoor wordt de coëfficiënt van groepsgrootte in regressieanalyse positief. Dit betekent dat grotere klassen betere resultaten behalen. De groepsgrootte heeft in dit geval invloed op de score, maar omgekeerd geldt dit ook. Dit heet omgekeerde causaliteit. Een methode om deze systematische fout op te lossen is door gebruik te

maken van zogenaamde instrumentele variabele. De groepsgrootte in het onderzoek is niet willekeurig en daarom is het waarschijnlijk gecorreleerd met de foutterm van de vergelijking. Hierdoor is geen causale interpretatie van de coëfficiënten mogelijk. Zoals eerder beschreven is geen informatie op leerling niveau aanwezig, maar alleen op schoolniveau. Hierdoor moet een instrumentale variabele gebruikt worden in het model. De regel van Maimonides is een exogene variabele die buiten het model worden verklaard. De hoeveelheid leerlingen in een klas wordt door de overheid bepaald en kan niet worden aangepast in het model. Deze heeft echter wel invloed op de groepsgrootte. Het verband tussen groepsgrootte en resultaten kan nu wel gelegd worden zonder dat er omgekeerde causaliteit plaatsvindt. Het onderzoek concludeert dat er een negatief verband bestaat tussen groepsgrootte en schoolresultaten met leerling aanmeldingen en leerling achtergrond als controle variabele. Voor wiskunde is de afname 0.033 punt per leerling en voor lezen 0.074 per leerling (Angrist & Lavy, 1999). Uit de 2SLS analyse volgt een afname van 0.275 punt voor lezen per extra leerling in de klas en voor wiskunde een afname van 0.230. Het verschil zit hem in het wel of niet controleren voor de trend in de aanmeldingsaantallen van leerlingen aan het begin van het jaar.

### *5.5 Lavy*

Een ander onderzoek van Lavy maakt gebruik van dezelfde dataset als de paper samen met Angrist maar focust zich niet alleen op groepsgrootte, maar ook op het aantal uren (Lavy, 2001). Lavy volgt dezelfde stappen als in de eerdere paper, maar voegt een extra variabele toe aan de analyse. De extra variabele is het aantal uren per week. Schooluren heeft een positieve invloed op de cijfers van leerlingen voor lezen van 0.464 punt per extra uur les in de week. Op wiskunde heeft het extra uur les een invloed van 0.558 punt. De coëfficiënten voor de groepsgrootte zijn echter wederom negatief, -0.266 voor lezen en -0.283 voor wiskunde. Ook hier is gecontroleerd op aanmeldingsaantallen aan het begin van het jaar. Het kan een aantal jaar duren voordat de school kan inspelen op een stijgend aantal aanmeldingen. Het toevoegen van vertraagde variabelen van het aantal aanmeldingen veranderen de resultaten echter nauwelijks.

### *5.6 Hoxby*

Onderzoek onder 649 basisscholen in de Verenigde Staten heeft geen significant statistisch verband opgeleverd tussen groepsgrootte en toetsscores (Hoxby, 2000).

De studie is een longitudinale studie over 24 aaneengesloten jaren. Een voordeel van deze studie is dat er geen sprake is van het zogenaamde Hawthorne effect. Dit effect houdt in dat personen beter gaan presteren vanwege het feit dat ze weten dat ze onderzocht worden. De resultaten worden hierdoor beïnvloed en dit komt de validiteit niet ten goede. Bij het onderzoeken van beleid moet altijd rekening gehouden worden met dit effect om niet tot verkeerde resultaten te komen. In het onderzoek van Hoxby weten scholen niet dat ze onderzocht worden en is er dus ook geen probleem omtrent de interpretatie van de resultaten. De sleutel variabelen zijn de variatie in populatie en het aantal klassen. Een afname van 10% van het gemiddelde aantal leerlingen uit een klas leidt tot een toename van de wiskunde, lezen en schrijven scores met respectievelijk 1.1, 2.6 en 2.0 procent van de toets standaarddeviatie. Als ook nog demografische controle variabele worden toegevoegd veranderen de resultaten naar 1.3, 1.6 en 3.0 procent van de toets standaarddeviatie. Vijf van de zes resultaten zijn significant verschillend van 0 met een 5 procent niveau. Als gebruik wordt gemaakt van het verschil in variatie in populatie dan verdwijnt het significante effect van groepsgrootte op toetsscores. Ook na het toevoegen van variabele met betrekking tot de afkomst en inkomensniveau is er geen significant verband waar te nemen. De interpretatie van dit onderzoek is alleen bedoeld voor klassen met tussen de 10 en 30 leerlingen. Extrapoleren van de resultaten geeft geen betrouwbare resultaten. Klassen met minder dan 10 leerlingen zijn in de regel echter zeer kostbaar en de winst in toename van toetsscores is zeer beperkt (Krueger, 1999).

Het verschil in variatie in populatie zit hem erin dat er een datum is waarvoor een kind nog aan het lopende schooljaar kan beginnen. Kinderen die later geboren worden moeten het jaar erop naar school. Dit wordt in de volksmond ook wel een jaar tegen genoemd. Het veranderen van het aantal klassen binnen een school heeft ook geen significante invloed op de toetsscores van de klas.

### *5.7 Boozer & Rouse*

Waar andere onderzoeken het leerling-leraar ratio gebruiken of een model schatten om de groepsgrootte als variabele te kunnen onderzoeken gebruiken Boozer en Rouse de echte groepsgroottes in hun onderzoek (Boozer & Rouse, 2001). Zonder een instrumentele variabele komen ze net als Angrist en Lavy uit op een positief verband tussen groepsgrootte en toetsscores. Met de gemiddelde leerling-leraar ratio

als instrumentele variabele is ook hier een negatief verband waar te nemen van 0.4 punt per toegevoegde leerling. In termen van toets standaarddeviatie is dat een afname van 7%.

### *5.8 Akerhielm*

De paper van Akerhielm over de relatie van groepsgrootte en resultaten heeft geprobeerd om het probleem van het niet willekeurig plaatsen van leerlingen in klassen op te lossen. Het onderzoek gebruikt data van de National Education Longitudinal Study (NELS) van het Amerikaanse ministerie van Onderwijs. Ook hier worden de echte groepsgroottes gebruikt en geen benaderingen met modellen en regressie vergelijkingen (Akerhielm, 1995). Het gebruik van de leerling- leraar ratio door andere onderzoeken wordt door Akerhielm een van de redenen genoemd voor de inconsistente resultaten. De leerling-leraar ratio zegt niets over de werkelijke groepsgrootte en zorgt daarom voor een systematische fout. Een ander punt van kritiek is dat voorgaande onderzoeken gebruikmaken van databases die eigenlijk niet bedoeld waren om onderzoek te doen naar het effect van groepsgrootte op toetsscores. Als instrument om het niet willekeurig plaatsen van leerlingen op te vangen wordt gebruik gemaakt van het gemiddelde aantal leerlingen per klas en de instroom van leerlingen in de achtste klas op de school in kwestie. In eerste instantie is de OLS regressie significant positief voor groepsgrootte, maar na het toevoegen van de instrumentele variabele is dat veranderd naar significant negatief. Als de leerlingen en scholen met een hoge weegfactor buiten de analyse worden gelaten dan is de OLS regressie ook significant negatief voor groepsgrootte. De conclusie van Akerhielm is dat politici zich niet op oude onderzoeken moeten baseren, maar juist op onderzoeken met meetinstrumenten zonder systematische fouten.

### *5.9 Dobbelaar et al.*

Onderzoeken op het gebied van educatie vinden niet alleen in de Verenigde Staten of Israël plaats, maar ook gewoon in Nederland. Zo ook op het gebied van groepsgrootte. Het onderzoek uit Nederland lijkt veel op dat van Angrist en Lavy. Ook hier wordt gebruik gemaakt van regelgeving om een instrumentele variabele te creëren. De analyse vindt hier plaats op studentniveau in plaats van op klasniveau. De afhankelijke variabele is de percentiel score van de leerlingen. Een belangrijke uitkomst van het onderzoek is dat de OLS en IV analyse van groep 8 beide geen endogeniteitsprobleem hebben volgens de Hausman test. OLS geeft een positief



verband weer tussen groepsgrootte en percentiel score voor zowel wiskunde als lezen, maar is niet significant verschillend van 0. IV geeft wel positief significante resultaten. Voor wiskunde stijgt het percentiel van een leerlingen met 0.779 als er een extra leerling in de klas zit. Het percentiel van lezen neemt met 0.466 toe. Hieruit kan geconcludeerd worden dat leerlingen beter af zijn in grotere klassen (Dobbelsteen, Levin, & Oosterbeek, 2002). De auteurs zijn van mening dat het toenmalige beleid om het aantal leerlingen in klassen te reduceren niet het beoogde effect zal hebben. Met het dalen van het leerling-leraar ratio zorgt deze maatregel er ook voor dat het verwachte aantal leerlingen met dezelfde intelligentie in een klas daalt. Hierdoor worden de slimme kinderen minder uitgedaagd door mede leerlingen. Dit zorgt ervoor dat de slimme leerlingen luier kunnen worden omdat ze toch al de hoogste cijfers van de klas halen.

## 6. Methodologie

Om de onderzoeksvraag en de hypotheses te kunnen beantwoorden wordt gebruik gemaakt van OLS regressieanalyse. In het begin wordt gebruik gemaakt van enkelvoudige regressie. Deze vergelijking wordt gebruikt als uitgangspunt en wordt als volgt weergegeven:

$$Y = \alpha + \beta A + \varepsilon \quad (1)$$

Hierin is  $Y$  de afhankelijke variabele met de Cito-scores. De  $\alpha$  is de intercept van de vergelijking. Variabele  $A$  is de groeps-grootte variabele met het aantal leerlingen in een klas en coëfficiënt  $\beta$ . Als laatste is ook nog een foutterm  $\varepsilon$  aanwezig. De data is aanwezig op leerling-niveau. Dat wil zeggen dat gegevens beschikbaar zijn voor elke leerling, op een specifieke school in een jaar.

De hierboven beschreven vergelijking geeft natuurlijk geen geloofwaardig beeld van de werkelijkheid. Veel variabele die significant bleken te zijn in andere onderzoeken zijn buiten beschouwing gelaten en ook is er geen onderscheid gemaakt voor het mogelijke verschil in een specifiek jaar. Door de simpelste vergelijking te schatten is er echter wel een basis aanwezig waarvandaan andere vergelijkingen geschat kunnen worden. De resultaten van deze vergelijking kunnen vergeleken worden met die van Angrist en Lavy (1999). Hier zou ook een positieve coëfficiënt uit kunnen komen.

De volgende vergelijking introduceert dummy's voor het schooljaar, de schoolscore en de weefactor van de leerling. De vergelijking wordt als volgt weergegeven:

$$Y = \alpha + \beta A + \gamma_i B_i + \delta_j C_j + \theta_t D_t + \varepsilon_{ij} \quad (2)$$

Hierin zijn  $\gamma$ ,  $\delta$  en  $\theta$  coëfficiënten van de dummy's.  $B$  is een dummy variabele voor de schoolscore. Schoolscore is een variabele die een score toekent aan iedere school door gebruik te maken van de weefactor van iedere leerling. De schoolscore is het gewogen aantal leerlingen van de school gedeeld door het werkelijke aantal leerlingen keer 100. Het subscript  $i$  staat voor de verschillende categorieën waar een school in kan vallen. Deze zijn 1: 100-109, 2: 110-119, 3:120-139, 4:140-159,

5:>159. Categorie 1 zal als referentie dienen en de andere coëfficiënten worden op deze manier geïnterpreteerd als een verandering ten opzichte van een school met geen of bijna geen leerlingen met een weegfactor. Hierdoor kan een school effect worden bepaald. Variabele  $C$  is de weegfactor van een individuele leerling. Zoals eerder beschreven bestaat deze variabele uit de waarden 0, 0.3 of 1.2 en de extra weegfactoren 0.25, 0.4, 0.7 en 0.9 die voortkomen uit oude regelgeving. Het subscript  $j$  staat deze weegfactoren. Leerlingen zonder weegfactor zullen worden gebruikt als referentie en de andere worden dus ook in dit geval geïnterpreteerd door te vergelijken met deze groep. Variabele  $D$  is de jaardummy met subscript  $t$ . Waarde 1 is voor leerjaar 2007/2008, waarde 2 als de leerling in het leerjaar 2010/2011 zat en waarde 3 voor het leerjaar 2013/2014. Het schooljaar 2007/2008 is het referentie jaar. CITO tracht de Cito-scores vergelijkbaar te maken over de jaren. Zo zou idealiter een score van bijvoorbeeld 536 in elk jaar de zelfde betekenis moeten hebben. Om dit te testen is de jaardummy in de vergelijking geïntroduceerd. Als de dummy significant is kan dit op een paar dingen duiden. Het kan zijn dat de leerlingen met alle variabelen hetzelfde in beide jaren een andere intelligentie niveau hebben, dat er juist meer geoefend wordt voor de Cito-toets of dat de Cito-toets toch iets in moeilijkheid verschilt per jaar.

Deze vergelijking geeft al een veel representatiever beeld van de werkelijkheid. Het staat verschil in scholen, leerlingen en jaren toe. In een samenleving heb je altijd goede en minder goede of zelfs slechte scholen. Leerlingen hebben ook niet allemaal dezelfde achtergrond. Sommige komen uit goede gezinnen en andere komen juist uit een achtergesteld milieu. Deze leerlingen worden ondersteund door extra geld, maar dit hoeft nog niet te betekenen dat de achterstand ook is weggewerkt aan het eind van de basisschool.

Een andere variatie die van invloed kan zijn op de Cito-score is het opleidingsniveau van de ouders. Deze is wel gedeeltelijk meegenomen in de weegfactor van de leerling, maar deze variabele heeft gegevens op leerling niveau. Ouders die hoog zijn opgeleid kunnen hun kinderen helpen met het schoolwerk. De kans is ook aanwezig dat intelligente ouders een intelligent kind krijgen. De vergelijking die hiervoor wordt gebruikt is:

$$Y = \alpha + \beta A + \gamma_i B_i + \delta_j C_j + \theta_t D_t + \vartheta_e E_e + \rho_e F_e + \sigma G + \tau_o H_o + \varphi_u I_u + \varepsilon_{ij} \quad (3)$$

Ook hier zijn  $\vartheta$ ,  $\rho$ ,  $\sigma$ ,  $\tau$  en  $\varphi$  de coëfficiënten. Variabele  $E$  en  $F$  zijn respectievelijk het opleidingsniveau van de vader en de moeder. Het subscript  $e$  staat voor welke opleiding de vader of moeder gedaan heeft. Als referentie wordt categorie 1 gebruikt en hierin vallen alle ouders met maximaal een lagere school diploma. De andere categorieën zijn als volgt opgebouwd 2: maximaal lager beroepsonderwijs, 3: maximaal mbo en 4: hbo of wetenschappelijk onderwijs. Dummy variabele  $G$  bevat het geslacht van de leerling met jongens als referentie categorie. Variabele  $H$  bevat de provincie waar de school staat met als referentie Zuid-Holland. Het subscript  $o$  bevat verder de volgende categorieën 2:Groningen, 3:Drenthe, 4: Friesland, 5: Noord-Holland, 6: Flevoland 7: Utrecht, 8: Gelderland, 9: Overijssel, 10: Zeeland, 11: Noord-Brabant, 12: Limburg. Als laatste is variabele  $I$  toegevoegd met de mate van stedelijkheid in het gebied waar de school staat. Het subscript  $u$  heeft ook hier verschillende categorieën 1: zeer stedelijk, 2: sterk stedelijk; 3: matig stedelijk, 4: weinig stedelijk, 5: niet stedelijk. Hier is categorie 5 gekozen als referentie.

In Nederland is vrijheid van godsdienst en levensovertuiging ook in het onderwijs van toepassing en daarom zijn er veel verschillende soorten scholen. Vergelijking 3 kan daarom uitgebreid worden met verschillende variabele om het verschil tussen scholen in kaart te brengen. Deze kunnen gebruikt worden om de hypothesen te kunnen beantwoorden.

$$Y = \alpha + \beta A + \gamma_i B_i + \delta_j C_j + \theta_t D_t + \vartheta_e E_e + \rho_e F_e + \sigma G + \tau_o H_o + \varphi_u I_u + \omega_a J_a + \mu_n K_n + rL + gM + mN + \varepsilon_{ij} \quad (4)$$

Variabele  $J$  is een dummy variabele en bevat het schooltype van de basisschool met coëfficiënt  $\omega$ . Subscript  $a$  bevat de categorieën 1: openbaar onderwijs, 2: protestants-christelijk onderwijs, 3: rooms-katholiek onderwijs en 4: overig bijzonder onderwijs. Het openbaar onderwijs is de referentie variabele. Bij het overig bijzonder onderwijs moet gedacht worden aan islamitisch onderwijs, joods onderwijs enzovoort. Als laatste van de dummy variabele is variabele  $K$  toegevoegd met coëfficiënt  $\mu$ . Deze variabele bevat alle algemeen bijzondere scholen die niet gemengd zijn. Subscript  $n$

heeft vier categorieën, namelijk 1: nee, 2: Montessori, 3: Dalton en 4: Jenaplan. In de regressie wordt als referentie gebruikt dat er geen sprake is van algemeen bijzonder onderwijs. Verder zijn ook variabele  $L$ ,  $M$  en  $N$  toegevoegd met respectievelijke  $r$ ,  $g$  en  $m$  als coëfficiënten. Deze variabelen staan voor de gemengde algemeen bijzondere scholen. Sommige scholen gebruiken verschillende stromingen door elkaar. Zo zijn er Montessori-Dalton scholen en Dalton-Jenaplan scholen.  $L$  staat voor gemengde Montessori school,  $M$  voor gemengde Dalton school en  $N$  voor gemengde Jenaplan school. Deze variabelen nemen de waarden 1 aan in het geval dat de school een gemengde school is van het specifieke type en een 0 in alle andere situaties. Een opsomming van de betekenis van alle variabelen is weergegeven in Tabel 4.

In al deze regressie vergelijkingen is er mogelijk sprake van een endogeniteitsprobleem. Sommige scholen besteden het extra geld dat ze krijgen per leerling voor kleinere klassen. Hierdoor komen deze leerlingen in een kunstmatig kleine klas. De Cito-scores van deze leerlingen zullen vaak ook lager liggen dan het landelijk gemiddelde en dit leidt tot resultaten met systematische fouten. Variabele  $B$  speelt een sleutelrol in de oplossing voor dit probleem. De schoolscore bevat informatie over de hoeveelheid leerlingen met een achterstand. Als de regressievergelijkingen worden geschat met alle categorieën en vervolgens ook nog een met alleen categorie 1 dan kan gekeken worden of de resultaten veranderen. In categorie 1 zitten de scholen met geen tot zeer weinig achterstand leerlingen en daarvan kan verwacht worden dat ze voor deze groep geen aparte klas inrichten. Het zijn er simpelweg te weinig. Voor de analyse in dit onderzoek zullen significantie niveaus worden gegeven van 1, 5 en 10 procent.

Variabele	Betekenis
<i>A</i>	Groepsgrootte
<i>B</i>	Schoolscore
<i>C</i>	Weegfactor
<i>D</i>	schooljaar
<i>E</i>	Opleiding vader
<i>F</i>	Opleiding moeder
<i>G</i>	geslacht
<i>H</i>	provincie
<i>I</i>	stedelijkheid
<i>J</i>	schooltype
<i>K</i>	Geheel algemeen bijzondere school
<i>L</i>	Gedeeltelijke Montessori
<i>M</i>	Gedeeltelijke Dalton
<i>N</i>	Gedeeltelijke Jenaplan

**Tabel 4: Betekenis van variabelen**

## 7. Data

De gegevens die gebruikt worden in deze scriptie komen van de longitudinale studie Cool5-18. Gegevens van basisschoolleerlingen zijn voor de schooljaren 2007-2008 (Kohnstamm Instituut, 2009), 2010-2011 (Kohnstamm Instituut, 2012) en 2013-2014 (Kohnstamm Instituut, 2015) beschikbaar. Het onderzoek bevat uitsluitend Nederlandse basisscholen.

Alle variabelen die zijn beschreven in de methodologie komen uit deze onderzoeken. De meeste variabele konden zonder meer overgenomen worden, maar enkele variabelen verdiende speciale aandacht. De opleiding van de vader en de moeder moesten gehercodeerd worden om de groepen 8 uit 2010/2011 en 2013/2014 met die van 2007/2008 te kunnen vergelijken. Hierbij zijn de variabelen uit 2010/2011 en 2013/2014 minder gedetailleerd geworden, maar het voordeel is dat ze in één keer in de regressieanalyse kunnen worden ingevoegd. De variabele voor de groeps grootte is afgeleid uit de beschikbare gegevens. Het aantal leerlingen op een school is opgeteld, rekeninghoudend met de subnaam van de klassen. Zo zijn bijvoorbeeld alle 8A en 8B per school afzonderlijk opgeteld. Hierdoor ontstaat een betrouwbaar beeld van de werkelijke groeps grootte per school. De schoolscore moest ook worden gehercodeerd om deze hetzelfde te laten zijn voor het gehele onderzoek. De oorspronkelijke categorieën 1: 100-104 en 2:105-109 zijn samengevoegd tot een enkele categorie. Verder zijn de nieuwe categorieën 3:120-139 en 4: 140-159 ook een samenvoegen van verschillende categorieën. Van de variabele die over het algemeen bijzonder onderwijs gaan zijn 4 nieuwe variabele gemaakt. Allereerst is een variabele gemaakt die beschrijft of een algemeen bijzondere school niet gemengd is. Daarnaast zijn 3 afzonderlijke variabelen gecreëerd die weergeven of een school een gemengde algemeen bijzondere school is. Het kan voorkomen dat een school een gemengde Montessori, Dalton of Jenaplan school is en toch maar een keer voorkomt bij de gemengde variabele. Dit komt doordat er zeer veel algemeen bijzondere scholen zijn en deze kunnen niet allemaal meegenomen worden in het onderzoek. Veelal zijn er te weinig observaties om tot conclusies te komen.

In het jaar 2014 mist de variabele van de provincie van de scholen. Ook stedelijkheid, schoolscore en het soort school zijn niet aanwezig. Het is toch nuttig

om de gegevens van 2014 mee te nemen in dit onderzoek om zodoende meer observaties te krijgen. De variabelen, afgezien van het soort school, zijn niet heel erg belangrijk. Met het toevoegen van deze gegevens wordt de data van de algemeen bijzondere scholen uitgebreid. Dit is een zeer kleine groep scholen en uitbreiden van het aantal observaties kan hier zeker geen kwaad.

De verschillende schooljaren bevatten gezamenlijk 25717 unieke observaties. Doordat de oorspronkelijke database ook de subklas heeft gegeven kunnen de combinatieklassen eruit worden gefilterd. Deze klassen zijn ongunstig voor het onderzoek omdat het onmogelijk is om te weten hoeveel leerlingen er exact in de klas zitten. Een combinatieklas van groep 7 en 8 staat in de data alleen weergegeven voor de leerlingen uit groep 8. Stel dat daar maar 7 leerlingen inzitten met hele hoge Cito-scores dan lijkt het alsof er tussen de groeps grootte en Cito-scores een negatief verband bestaat. De combinatieklas in kwestie bestaat echter uit meer leerlingen. De leerlingen die in groep 7 zitten worden niet meegenomen. Het is daarom verstandig om alle combinatie klassen uit de data te filteren. Na het filteren van de combinatieklassen zijn nog 24608 observaties over. Vervolgens zijn ook de klassen verwijderd die meer dan 40 leerlingen in de klas hadden. Uit de eerder genoemde brief van staatssecretaris Dekker van Onderwijs blijkt dat klassen met meer leerlingen een verwaarloosbaar deel uitmaken van de totale klassen in Nederland. Hierdoor blijven nog maar 23105 observaties over. Als laatste worden ook de leerlingen buiten beschouwing gelaten die om wat voor reden dan ook de toets niet gemaakt hebben en dan blijven er uiteindelijk 18656 observaties over. Door het verwijderen van observaties is categorie 7 van de variabele *C* komen te vervallen doordat geen enkele observatie overgebleven is. Bij sommige variabelen in de database komen onvolledig en de missende waardes voor. Deze worden buiten beschouwing gelaten in de analyse.

In Tabel 5 in de appendix zijn de correlaties weergegeven tussen variabelen die worden gebruikt in de regressieanalyse. Hierbij zitten enkele opvallende significante correlaties. Zo zijn variabele *B* en *C* sterk positief gecorreleerd (0.627). Het verband dat wordt waargenomen tussen de schoolscore en de weegfactor is niet heel vreemd. De variabele van de schoolscore is opgebouwd uit de weegfactoren van leerlingen, dus hoe hoger de weegfactor hoe hoger de schoolscore. Dit verklaart de positieve relatie tussen beide variabelen. Variabele *C* is ook sterk negatief



gecorrleerd met variabelen *D*, *E* en *F*. De correlaties zijn respectievelijk -0.510, -0.510 en -0.575. Het verband tussen de weegfactor (*C*) en de jaardummy (*D*) is op het eerste gezicht vreemd. Hierbij moet echter wel rekening gehouden worden met het veranderen van regelgeving voor de weegfactoren. In schooljaar 2007/2008 waren er nog zeven categorieën van weegfactoren, namelijk 0, 0.3, 1.2, 0.25, 0.4, 0.7 en 0.9. In deze volgorde zijn ze ook gecodeerd. Voor de schooljaren 2010/2011 en 2013/2014 zijn de laatste vier categorieën echter komen te vervallen. Leerlingen die eerst in de 0.4 categorie vielen moeten in latere jaren in een andere categorie worden ingedeeld. Het zijn weliswaar niet dezelfde leerlingen, maar de achtergronden van deze leerlingen kunnen wel gelijk zijn. Deze leerlingen worden niet meer in dezelfde categorie ingedeeld, maar zeer waarschijnlijk in de 0.3 categorie. Dit zou een verklaring kunnen zijn voor de negatieve correlatie tussen de weegfactor en de jaardummy. De relatie tussen de weegfactor en het opleidingsniveau van de vader (*E*) en de moeder (*F*) is ook geen verrassing. De weegfactor wordt mede beïnvloed door het opleidingsniveau van de ouders. Hoe hoger de opleiding, hoe lager de weegfactor. Ook het opleidingsniveau van de vader en de moeder zijn sterk positief gecorrleerd (0.606). In veel relaties of huwelijken hebben de man en vrouw een vergelijkbaar opleidingsniveau en dit is ook terug te zien in deze dataset. Ten slotte zijn ook de schoolscore en de mate van stedelijkheid (*I*) sterk negatief gecorrleerd (-0.588). De variabele stedelijkheid is gecodeerd op een schaal die loopt van zeer stedelijk naar niet stedelijk. Hoe minder stedelijk een gebied is waar de school staat, hoe lager de schoolscore. In minder stedelijk gebied wonen minder mensen en zijn ook minder kinderen die naar de basisschool gaan. Veel achterstand gezinnen wonen echter in de grote steden. In de grote steden zullen relatief meer categorie 1.2 leerlingen naar de basisschool gaan dan in minder stedelijk gebied. Dit resulteert in een lagere schoolscore voor basisscholen in minder stedelijk gebied.

Tabel 6 bevat beschrijvende statistieken van de twee interval variabelen groeps grootte (*A*) en afhankelijke variabele Cito-score (*Y*). Het gemiddelde van 22.40 leerlingen in een groep verschilt niet veel van het gemiddelde van 23 dat staatssecretaris Dekker in zijn brief aan de Tweede Kamer noemt. Een ander punt uit de brief dat naar voren komt is de maximale groeps grootte van 35-39 leerlingen in

een klas. Deze groepsgroottes liggen maar twee standaarddeviaties van het gemiddelde af en zijn daarom zeer goed mogelijk.

Variabele	Minimum	Maximum	Gemiddelde	Standaarddeviatie
Groepsgrootte	1	40	22,40	7,275
Cito-score	501	550	533,58	10,186

**Tabel 6: Beschrijvende statistieken interval variabelen**

De gemiddelde score van de Cito-toets in de dataset is afgerond 534. Dit is iets lager dan het landelijk gemiddelde, maar de afwijking is niet meer dan één tot anderhalve punt. De standaarddeviatie van de Cito-scores is 10,186. Deze standaarddeviatie zal gebruikt worden om het gestandaardiseerd effect van de variabelen te berekenen.

Tabel 7 in de appendix laat zien dat de meeste categorieën genoeg observaties hebben om er zinnige uitspraken over te kunnen doen. In de variabele weegfactor (C) bevatten de categorieën 0.4 en 0.7 zeer weinig observaties. Hier moet rekening mee gehouden worden bij het interpreteren van de coëfficiënten in de regressieanalyse. Verder hebben Groningen, Drenthe en Flevoland wel heel weinig observaties. In het gedeelte van de algemeen bijzondere scholen heeft de categorie van een gemengde Montessori school maar 57 observaties. Mogelijk moet deze worden samengevoegd met Dalton en Jenaplan.

## 8. Resultaten

In dit hoofdstuk zullen de regressievergelijken 1 tot en met 4 die zijn beschreven in de methodologie worden uitgevoerd. Regressievergelijking 1 komt overeen met model 1, regressievergelijking 2 met model 2 enzovoort. Ook worden van de significante variabelen de gestandaardiseerde effecten berekend.

### 8.1 Model 1

In Tabel 8 is model 1 weergegeven. In dit model zijn zowel de constante als de groeps-grootte positief significant. De coëfficiënt van de groeps-grootte is 0,028. De interpretatie van deze coëfficiënt is dat een stijging van de groeps-grootte met één leerling de gemiddelde Cito-score in de klas doet stijgen met 0,028 punt of 0,275% van de standaarddeviatie van de toets<sup>3</sup>. De verklarende kracht van dit model is nihil en hierdoor niet geschikt om conclusies op te baseren.

Variabele	Coëfficiënt
Constante	532,948*** (0,241)
Groeps-grootte	0,028*** (0,010)
R <sup>2</sup>	0,000

Tabel 8: Model 1 met Cito-score als afhankelijke variabele

\*\*\* significant op 1% niveau

\*\* significant op 5% niveau

\* significant op 10% niveau

### 8.2 Model 2

Model 2 is weergegeven in Tabel 9. Hierbij zijn variabele voor schoolscore, weegfactor en schooljaar toegevoegd. De belangrijkste verandering in model 2 ten opzichte van model 1 is dat variabele groeps-grootte niet significant is geworden. De invloed van de groeps-grootte op de Cito-score is daarom niet significant verschillend van 0. Alle andere toegevoegde variabele zijn zelfs significant bij een 1 procent niveau, behalve categorie 2013/2014 van de variabele schooljaar. De coëfficiënten van de categorieën ten opzichte van de referentie categorie zijn negatief, behalve bij schooljaar 2010/2011 (1,900). De verklarende kracht van het model is niet hoog met

<sup>3</sup> Het gestandaardiseerd effect wordt op de volgende manier berekend:  $0,028/10,186 \cdot 100\% = 0,2748\%$ .  
Alle andere gestandaardiseerde effecten volgen dezelfde berekening.

een  $R^2$  van 0,074. Verder zijn de gestandaardiseerde effecten van de significante<sup>4</sup> categorieën in procenten van de standaarddeviatie van de toets weergegeven. Het effect van schoolscore wordt steeds negatiever naarmate de schoolscore stijgt. Het effect van de weegfactoren die in alle drie de schooljaren voorkomen zijn min of meer constant op -56 en -57 procent van de standaarddeviatie van de toets. De laatste drie categorieën uit 2007/2008 volgen het patroon van de schoolscores. Schooljaar 2010/2011 heeft als enige een positieve invloed ten opzichte van de referentie categorie van ruim 18% van de standaarddeviatie.

Variabele	Categorie	Coëfficiënt	Gestandaardiseerd effect
Constante		535,139*** (0,265)	
Groepsgrootte		-0,002 (0,010)	
Schoolscore	110-119	-1,730*** (0,250)	-16,984
	120-139	-2,652*** (0,274)	-26,036
	140-159	-3,217*** (0,325)	-31,583
	>159	-3,472*** (0,356)	-34,086
Weegfactor	0,3	-5,782*** (0,307)	-56,764
	1,2	-5,711*** (0,322)	-56,067
	0,25	-3,020*** (0,308)	-29,649
	0,4	-6,342*** (1,487)	-62,262
	0,7	-7,952*** (1,588)	-78,068
schooljaar	2010/2011	1,900*** (0,196)	18,653
	2013/2014	-0,259 (0,202)	
$R^2$	0,074		

Tabel 9: Model 2 met Cito-score als afhankelijke variabele

\*\*\* significant op 1% niveau

\*\* significant op 5% niveau

\* significant op 10% niveau

<sup>4</sup> Alles tot 10% wordt hier en in het vervolg van de resultaten aangemerkt als significant

### *8.3 Model 3*

Tabel 10 in de appendix bevat model 3. Met het uitbreiden van model 2 met een aantal variabelen zijn wat verschuivingen waar te nemen. De belangrijkste worden hier genoemd. De groepsgrootte is in dit model wel significant en dat was in model 2 niet zo. Ook is het verband tussen groepsgrootte en Cito-scores een negatief verband geworden in plaats van een positief verband in model 1. De weegfactor van 0.4 is hier ook significant geworden zij het op een 5 procent niveau. Andere variabelen die voorkomen in zowel model 2 als 3 verschillen niet van teken. Iets wat verder opvalt is dat het verschil in Cito-scores tussen jongens en meisjes volgens dit model niet significant verschilt van 0. Alleen de provincies Groningen, Noord-Brabant en Limburg zijn significant met de eerste twee op een 5 procent niveau en Limburg zelf op een 1 procent niveau. Hierbij moet opgemerkt worden dat Groningen weinig observaties had. Het gestandaardiseerde effect van de groepsgrootte is  $-0.255\%$  van de toets standaarddeviatie. Dit betekent dat voor elke extra leerling in een klas het gemiddelde van de Cito-score in de klas daalt met  $0.255\%$  van de toets standaarddeviatie.

### *8.4 Model 4*

In deze vergelijking is model 3 uitgebreid met variabelen voor de verschillende soorten onderwijs en geloofsovertuigingen. De resultaten voor model 4 zijn te vinden in Tabel 11 in de appendix. Door toevoeging van nieuwe variabelen veranderen de coëfficiënten wederom niet van teken. De significantie van categorieën matig en weinig stedelijk van variabele stedelijkheid verdwijnt. Hierbij moet wel opgemerkt worden dat matig stedelijk pas significant was in model 3 op een 10 procent niveau en weinig stedelijk nog net significant was op een 5 procent niveau. De nieuw toegevoegde variabelen zijn van belang voor de twee hypothesen die zijn opgesteld. Het schooltype is alleen significant voor protestants-christelijk onderwijs met een coëfficiënt van 0.513. Protestants-christelijke basisscholen presteren volgens dit model  $5.036\%$  beter dan openbare basisscholen. Van de geheel algemeen bijzondere scholen presteert alleen het Daltononderwijs significant verschillend aan het reguliere onderwijs. Leerlingen op het Daltononderwijs presteren  $6.479\%$  minder dan leerlingen op het reguliere onderwijs. Gedeeltelijke Daltonscholen presteren met  $-10.514\%$  zelfs nog slechter dan gehele Daltonscholen. Als alle geheel of gedeeltelijke algemeen bijzondere scholen bij elkaar worden gevoegd in een enkele

dummyvariabele dan ontstaat er een significant negatief verband tussen algemeen bijzonder onderwijs en Cito-scores<sup>5</sup>. De nieuwe variabele heeft een coëfficiënt van -1,305 met standaardfout 0,284 en is significant op een 1 procent niveau. Het gestandaardiseerde effect is bijna -13%. Dit betekent dat leerlingen op een school die gekwalificeerd wordt als algemeen bijzondere school bijna 13% slechter presteren dan leerlingen op reguliere basisscholen. Alle andere variabelen komen qua coëfficiënt en significantie niveaus overeen met model 4.

### *8.5 Robuustheid van de resultaten*

Scholen kunnen leerlingen met een hoge weefactor in kustmatig kleine klassen plaatsen en zo de resultaten beïnvloeden. Dit probleem kan opgelost worden door een nieuw model te schatten met alleen categorie 1 van de schoolscore, zoals beschreven in Akerhielm (1995). Deze scholen hebben geen tot zeer weinig leerlingen met een weefactor en zijn hierdoor ook minder geneigd om achterstandsleerlingen in kleinere klassen te plaatsen. Model 3 dient als basis voor deze nieuwe regressieanalyse. Van de variabelen over het schooltype, geheel bijzondere school, gemengde Montessori, gemengde Dalton en gemengde Jenaplan blijven te weinig observaties over. Hierdoor voldoet model 4 niet. Uit model 3 wordt de jaardummy van 2013/2014 verwijderd vanwege het feit dat in dit jaar de schoolscore ontbreekt.

In Tabel 12 in de appendix is model 5 weergegeven. De groepsgrootte blijft significant en negatief. Alleen de coëfficiënt wordt iets negatiever van -0.026 in model 3 naar -0.076 in model 5. Veranderingen in coëfficiënten komen ook naar voren in de het percentage van de standaarddeviatie. Voor groepsgrootte was deze -0,255% in model 3 en wordt -0,746% in model 5.

Een andere robuustheidcontrole komt uit de paper van Krueger (1999). Hierin wordt betoogd dat klassen van één tot en met tien leerlingen veel te duur zijn voor de school. Met deze klassen moet dus wel iets aan de hand zijn als ze toch bestaan. Daarom worden ze uit de analyse gelaten en wordt gekeken of de resultaten veranderen. De basis voor deze analyse is model 4. In Tabel 13 in de appendix is te zien dat de groepsgrootte in model 6 nog steeds significant is, maar alleen op een

---

<sup>5</sup> Dit model is niet opgenomen in de scriptie

veel hoger significantie niveau van 10%. Ook de coëfficiënt stijgt iets van -0.027 naar -0.020.

### *8.6 Interpretatie van de resultaten*

De uitkomsten van model 1 zijn te vergelijken met die uit de paper van Angrist en Lavy (1999). De groepsmaat heeft in model 1 een significant positief effect op de Cito-score van 0.028 punt per extra leerling in de klas of 0.275% van de toets standaarddeviatie. Angrist en Lavy hadden de mogelijkheid om de effecten van wiskunde (0.076) en leesvaardigheid (-0.031) afzonderlijk op te nemen in het onderzoek. Iets wat voor deze paper niet mogelijk was. Aangezien de Cito-toets een combinatie is van wiskunde en leesvaardigheid kunnen de modellen wel vergeleken worden. Over de coëfficiënten kan alleen gezegd worden dat ze in de zelfde orde van grootte liggen. De toets standaarddeviatie is niet bekend en daarom zijn de getallen niet te vergelijken. Verder komen de resultaten van model 1 ook overeen met die van Booser & Rouse (2001). Ook hier wordt een significant positief verband gevonden tussen groepsmaat en toetscores.

Model 2 komt al iets meer in de buurt van een geloofwaardig model. Door toevoeging van dummy's voor schoolscore, weegfactor en schooljaar gaat de verklarende kracht van het model omhoog naar 0.074. Een ander zeer belangrijk punt in dit model is dat de coëfficiënt van groepsmaat niet significant meer is. De groepsmaat heeft geen invloed meer op Cito-score van de klas en het maakt dan dus niet meer uit of een klas uit 1 of 40 leerlingen bestaat. De weegfactor heeft echter wel een significant negatief verband met de afhankelijke variabele. Voor deze variabele zijn de gestandaardiseerde effecten zeer groot en liggen tussen de 29 en 79 procent van de toets standaarddeviatie.

Als vervolgens in model 3 extra variabelen worden toegevoegd met onder andere het opleidingsniveau van de ouders dan verandert de groepsmaat wederom. De coëfficiënt geeft nu een significant negatief verband van -0.026 weer. In termen van toets standaarddeviatie is dat -0.255 procent. Deze verandering is zeer belangrijk. In model 1 waren grote klassen wenselijk, maar in model 3 zijn juist kleine klassen goed voor de resultaten van basisschoolleerlingen. Een verklaring hiervoor kan zijn dat er sprake is van een endogeniteitsprobleem in model 1. In model 1 wordt geen onderscheid gemaakt tussen scholen met veel en weinig achterstandsléerlingen.

Hiermee samen hangt verschil in opleidingsniveau van de ouders dat niet is meegenomen in dit model. Scholen kunnen achterstandsleerlingen in kunstmatig kleine klassen plaatsen. Hierdoor lijkt het alsof kleine klassen slechter presteren dan grote klassen. Dit kan resulteren in het positief verband tussen groepsgrootte en Cito-score in model 1. Model 3 voegt extra variabelen toe waardoor het endogeniteitsprobleem verkleind, zo niet opgelost wordt. Het opleidingsniveau van de ouders heeft ook een zeer groot effect op de Cito-score. Voor de categorie hoger beroepsonderwijs en wetenschappelijk onderwijs kan dit oplopen tot ruim 43% voor vader en bijna 49% voor moeders. Moeders met maximaal lager beroepsonderwijs hebben een negatieve invloed, terwijl dit voor vaders niet het geval is. Dit zou mogelijk verklaard kunnen worden door de traditionele gezinnen. In deze gezinnen is de man de kostwinner en de vrouw blijft thuis bij de kinderen. Kinderen brengen in deze gezinnen het meeste tijd door met hun moeder en dit beïnvloed ook de ontwikkeling van het kind. Het verschil tussen jongens en meisjes is niet significant verschillend aan 0. De verschillen tussen provincies zijn bijna allemaal niet significant. Groningen en Noord-Brabant zijn nog maar net significant op een 5 procent niveau. Limburg is de enige provincie die ook bij 1 procent significant is. Leerlingen uit deze provincies presteren bijna 15% beter in termen van standaarddeviatie dan leerlingen uit Zuid-Holland. Mate van stedelijkheid heeft een relatief kleine invloed van 5 tot bijna 11 procent.

Om de hypothesen te kunnen beantwoorden is model 4 geschat. Hierin zijn variabelen opgenomen van het schooltype en of er mogelijk sprake is van een gedeeltelijke Montessori, gedeeltelijke Dalton of gedeeltelijke Jenaplan school. Hier blijft het negatieve verband van de groepsgrootte met de Cito-score bestaan. Leerlingen op Rooms-katholieke en overige bijzondere basisscholen verschillen niet significant met leerlingen op openbare basisscholen. Op protestants-christelijke basisscholen is er echter wel een significant verschil waar te nemen. Al is de coëfficiënt van 0.513 maar net significant op een 5 procent niveau. Leerlingen op protestants-christelijke basisscholen presteren juist iets meer dan 5% beter dan leerlingen op openbare basisscholen. Dit is toch wel verrassend, omdat er eigenlijk geen verschil mag bestaan tussen deze scholen. Het enige verschil is de geloofsovertuiging en die zou geen invloed op de Cito-toets moeten hebben. De andere hypothese spitste zich toe op de algemeen bijzonder scholen. Hierbij kan



geconcludeerd worden op basis van Tabel 12 in de appendix dat leerlingen op scholen die geheel Daltononderwijs aanbieden bijna 6,5% slechter presteren dan leerlingen in het reguliere onderwijs. De uitkomst is wel alleen significant bij een 10 procent niveau. Voor Montessori en Jenaplan onderwijs is geen significant verband gevonden, maar hierbij moet opgemerkt worden dat deze twee onderwijs systemen zeer weinig observaties hebben. Gedeeltelijke Daltonschole hebben wel een zeer significant negatief verband met de Cito-score. Het gaat hierbij om een significantie niveau van 1 procent en een effect van -10.5%. Gedeeltelijke Montessori en Jenaplan scholen hebben ook hier te weinig observaties om uitspraken over te kunnen doen. Het lijkt er dus op dat leerlingen op algemeen bijzondere scholen slechter presteren dan leerlingen op reguliere scholen. Dit kan komen door de grote verschillen in de manier waarop les wordt gegeven op scholen. In het regulier onderwijs bepaalt de leraar wat er moet gebeuren en hoelang de klas met een opdracht bezig is en op bij het algemeen bijzonder onderwijs mag de leerling dit zelf beslissen. Als alle algemeen bijzondere scholen worden samengevoegd in een enkele variabele dan ontstaat een gestandaardiseerd effect van bijna -13%. Leerlingen op algemeen bijzondere scholen presteren volgens dit model significant slechter dan leerlingen op het reguliere onderwijs.

## 9. Conclusie, limitaties en beperkingen

In de meeste modellen komt een negatief verband tussen groepsgrootte en Cito-score naar voren. Grote klassen hebben dus een negatieve invloed op de gemiddelde Cito-scores van leerlingen. Het effect in toets standaarddeviaties is echter niet heel groot. Dit effect is vele malen groter bij de weegfactor van leerlingen en scholen. De achterstand die leerlingen oplopen doordat ze opgroeien in een achtergesteld milieu is niet weggewerkt. Hiermee samen hangt het opleidingsniveau van de ouders. De overheid zal dus doorleren nog meer moeten stimuleren, vooral onder vrouwen. Verder presteren leerlingen op protestants-christelijke basisscholen 5% beter dan leerlingen op openbare basisscholen. Ook presteren leerlingen op Daltonscholen 6,5% slechter dan leerlingen op het regulier onderwijs. Voor gedeeltelijke Daltonscholen is dit zelfs -10,5%. Als alle algemeen bijzondere scholen worden samengevoegd in een enkele variabele dan is het gezamenlijke effect bijna -13%.

De robuustheidcontrole met alleen de schoolscore van categorie 1 verandert niets aan het teken en de significantie van de groepsgrootte. De groepsgrootte blijft een significant negatief effect hebben op de Cito-score van leerlingen. Het weglaten van de klassen met een tot tien leerlingen zorgt wel voor een verschuiving in het significantie niveau van de groepsgrootte, maar niet van het teken. Ook het effect van de protestants-christelijke scholen blijft bestaan. Bij geheel algemeen bijzondere scholen vindt wel een verschuiving plaats. Nu is alleen Jenaplan onderwijs significant op een 5 procent niveau. De conclusie voor gedeeltelijk Dalton scholen blijft wel overeind. Resultaten voor geheel algemeen bijzondere scholen fluctueren dus bij een robuustheidcontrole. De conclusie voor hypothese 2 moet dus worden afgezwakt. Het gedeelte over de gedeeltelijk algemeen bijzondere Dalton scholen blijft wel overeind. Als alle algeheel bijzondere scholen worden samengevoegd tot een enkele variabele dan is een significant negatief verband waar te nemen van bijna 13%. Hierdoor kan geconcludeerd worden dat algemeen bijzonder onderwijs als geheel slechter presteert dan het reguliere onderwijs.

Beperkingen aan dit onderzoek zijn er zeker ook. Allereerst is het niet mogelijk om het intelligentieniveau van de leerling mee te nemen in het onderzoek. Een IQ score zou van zeer veel toegevoegde waarde zijn. Hiermee kan afgeleid worden wat de

potentie is van de leerling. Ook de omgeving van de school kon niet goed genoeg worden meegenomen in dit onderzoek. De variabele die beschrijft in welke mate de basisschool in stedelijk gebied staat probeert dit echter wel. Het maakt natuurlijk uit of een school in de Schilderswijk in Den Haag staat of in Wassenaar. Een specifiekere variabele kan de resultaten verfijnen. Een andere variabele die hierbij aansluit is het wijkinkomen. Deze zou ook toegevoegd kunnen worden. In dit onderzoek was het echter onmogelijk doordat de scholen die meedoen aan het Cool5-18 onderzoek volledig anoniem zijn. Verder zou de ervaring van de leraar van invloed kunnen zijn op de Cito-scores.

Het argument van leraren dat een daling van de groepsgrootte de resultaten positief zou beïnvloeden is waar. Zij het maar zeer minimaal. Uit dit onderzoek blijkt dat een daling van de groepsgrootte bijna geen invloed heeft. Het argument is daarom niet erg overtuigend. Het echte argument is het verlagen van de werkdruk onder leraren in het primair onderwijs. Het beleid van de overheid in de toekomst moet zich niet richten op het laten dalen van de groepsgrootte, maar juist op het aanpakken van achterstandsleerlingen en het stimuleren van doorleren. Hier is de echte winst te halen voor het versterken van de kenniseconomie.

## Bibliografie

- Akerhielm, K. (1995). Does class size matter. *Economics of Education Review*, 14 (issue 3), 229-241.
- Angrist, J., & Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on student achievement. *The Quarterly Journal of Economics*, volume 114 (No. 2), 533-575.
- ANP. (2014, april 2). *Groepsgrootte op scholen*. Opgeroepen op mei 29, 2015, van Website van nu.nl: bron <http://www.nu.nl/politiek/3742044/geen-landelijke-grens-klassengrootte-scholen.html>
- Bogaerd, D. M. (2013, oktober 28). *Burgerinitiatief 'stop de overvolle klassen'*. Opgeroepen op mei 29, 2015, van Website van vosabb: <http://www.vosabb.nl/burgerinitiatief-stop-de-overvolle-klassen/>
- Bogaerd, D. M. (2014, april 2). *Informatie over uitkomst burgerinitiatief*. Opgeroepen op mei 29, 2015, van Website van vosabb: <http://www.vosabb.nl/tag/burgerinitiatief/>
- Boozer, M., & Rouse, C. (2001). Intraschool Variation in Class Size: Patterns and Implications. *Journal of Urban Economics*, 50 (issue 1), 163-189.
- CITO. (2011). *Informatie cito-toets 2011*. Opgeroepen op juni 15, 2015, van Website van CITO: [www.cito.nl/~media/cito\\_nl/.../cito\\_meting\\_taal\\_rekenen\\_2011.ashx](http://www.cito.nl/~media/cito_nl/.../cito_meting_taal_rekenen_2011.ashx)
- Cool5-18 a. (sd). *Informatie aantal scholen en leerlingen*. Opgeroepen op juni 17, 2015, van Website van Cool5-18: <http://www.cool5-18.nl/>
- Cool5-18 b. (sd). *Informatie opzet Cool5-18 onderzoek*. Opgeroepen op juni 17, 2015, van Website van Cool5-18: <http://www.cool5-18.nl/doelenopzetbo/>
- Dobbelsteen, S., Levin, J., & Oosterbeek, H. (2002). The causal effect of class size on scholastic achievement: distinguishing the pure class size effect from the effect of changes in class composition. *Oxford bulletin of economics and statistics*, volume 64 (issue 1), 17 - 38.
- Hoxby, C. M. (2000). The Effects of Class Size and Composition on Student Achievement: New Evidence from Natural Population Variation. *Quarterly Journal of Economics*, 115, 1239-1285.
- Kohnstamm Instituut. (2009, juli 14). *Cool data 2007-2008*. Opgeroepen op juni 18, 2015, van Website van narcis: <http://www.narcis.nl/dataset/RecordID/oai%3Aeasy.dans.knaw.nl%3Aeasy-dataset%3A44290/uquery/cool5-18/id/1/Language/NL>
- Kohnstamm Instituut. (2012, mei 26). *Cool data 2010-2011*. Opgeroepen op juni 19, 2015, van Website van narcis: <http://www.narcis.nl/dataset/RecordID/oai%3Aeasy.dans.knaw.nl%3Aeasy-dataset%3A49841/uquery/cool5-18/id/6/Language/NL>
- Kohnstamm Instituut. (2015, mei 28). *Cool data 2013-2014*. Opgeroepen op juni 19, 2015, van Website van narcis: <http://www.narcis.nl/dataset/RecordID/oai%3Aeasy.dans.knaw.nl%3Aeasy-dataset%3A61290/uquery/cool5-18/id/7/Language/NL>
- Krueger, A. B. (1999). Experimental Estimates of Education Production Functions. *Quarterly Journal of Economics*, 114 (2), 497-532.

Lavy, V. (2001). Estimating the Effect of School Time of Instruction on Student Achievements. *Social Science Research Centre , Working Paper No. 01-4.*

Overheid. (2014, mei 16). *Kamerdebat over burgerinitiatief*. Opgeroepen op mei 29, 2015, van Website van officiële bekendmakingen: <https://zoek.officielebekendmakingen.nl/h-tk-20132014-70-3.html>

Rechtspraak. (2013, september 4). *Uitspraak WOB verzoek RTL nieuws*. Opgeroepen op juni 17, 2015, van Website van rechtspraak: <http://www.rechtspraak.nl/Organisatie/Rechtbanken/Midden-Nederland/Nieuws/Pages/Cito-scoresscholenmogenopenbaargemaaktworden.aspx>

Rijksoverheid a. (2013, Februari 11). *Regeerakkoord 2013*. Opgeroepen op Juni 12, 2015, van Website van Rijksoverheid: <http://www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2013/02/11/uitwerking-regeerakkoord-voor-versterking-kenniseconomie.html>

Rijksoverheid b. (sd). *Informatie over eindtoets basisschool*. Opgeroepen op mei 29, 2015, van Website van Rijksoverheid: <http://www.rijksoverheid.nl/onderwerpen/toelating-middelbare-school/verplichte-eindtoets-basisonderwijs>

Rijksoverheid c. (sd). *Kamerbrief over klassengrootte*. Opgeroepen op mei 29, 2015, van Website van Rijksoverheid: <http://www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2015/02/02/kamerbrief-over-groepsgrootte-in-het-basisonderwijs.html>

Rijksoverheid d. (sd). *Informatie over leerlinggewicht*. Opgeroepen op juni 3, 2015, van Website van Rijksoverheid: <http://www.rijksoverheid.nl/onderwerpen/taalachterstand/vraag-en-antwoord/wat-is-de-gewichtenregeling-in-het-basisonderwijs.html>

Rijksoverheid e. (sd). *Informatie over school oprichting*. Opgeroepen op juni 3, 2015, van Website van Rijksoverheid: <http://www.rijksoverheid.nl/onderwerpen/vrijheid-van-onderwijs/zelf-eeen-school-oprichten>

RTL nieuws. (2013, September 20). *Vergelijkingstool basisscholen*. Opgeroepen op juni 17, 2015, van Website van RTL nieuws: <http://www.rtlnieuws.nl/nieuws/special/publicatie-cito-scores-kijk-hoe-jouw-school-scoort>

SLO. (sd). *Informatie over indeling cito toets*. Opgeroepen op juni 15, 2015, van Website van SLO: <http://www.slo.nl/primair/leergebieden/ned/taalsite/lexicon/00815/>

## Appendix

Onderdeel	Totale aantal opgave	Subonderdeel	Aantal opgaven
Taal	100	Schrijven van teksten	30
		Spelling	30
		Begrijpend lezen	20
		Woordenschat	20
Rekenen	60	Getallen en bewerkingen	25
		Verhoudingen, breuken en procenten	20
		Meten, meetkunde, tijd en geld	15
Studievaardigheden	40	Studieteksten	10
		Informatiebronnen	10
		Kaartlezen	10
		Tabellen en grafieken	10
Wereldoriëntatie	90	Aardrijkskunde	30
		Geschiedenis	30
		Natuuronderwijs	30

Tabel 1: Overzicht Cito-toets schooljaar 2007/2008, 2010/2011 en 2013/2014

Score	Schooladvies
501-520	basisberoepsgerichte leerweg
519-525	basis- en kaderberoepsgerichte leerweg
523-528	kaderberoepsgerichte leerweg
529-533	gemengde theoretische leerweg
533-536	kaderberoepsgerichte leerweg en havo
537-540	havo
540-544	havo/vwo
545-550	vwo

Tabel 2: Schooladvies Cito-toets uitgebreid

Score	Schooladvies
501-523	basisberoepsgerichte leerweg
524-528	kaderberoepsgerichte leerweg
529-536	gemengde theoretische leerweg
537-544	havo
545-550	vwo

Tabel 3: schooladvies Cito-toets



Variabele	categorie	frequentie
schoolscore	100-109	6846
	110-119	2020
	120-139	1717
	140-159	1284
	>159	1315
weegfactor	0,0	9403
	0,3	1171
	1,2	1110
	0,25	2009
	0,4	44
	0,7	39
schooljaar	2007/2008	8021
	2010/2012	5161
	2013/2014	5474
Opleiding vader	lager onderwijs	1832
	maximaal lager beroepsonderwijs	4735
	maximaal mbo	6055
	hbo/wo	4189
Opleiding moeder	lager onderwijs	2355
	maximaal lager beroepsonderwijs	4033
	maximaal mbo	7804
	hbo/wo	3512
geslacht	Jongen	9002
	Meisje	9162
provincie	Zuid-Holland	2018
	Groningen	272
	Drenthe	456
	Friesland	532
	Noord-Holland	1934
	Flevoland	268
	Utrecht	322
	Gelderland	1496
	Overijssel	1053
	Zeeland	554
	Noord-Brabant	3043
	Limburg	1234
stedelijkheid	zeer stedelijk	2140
	sterk stedelijk	3113
	matig stedelijk	2940



	weinig stedelijk	3394
	niet stedelijk	1595
schooltype	openbaar	4302
	protestants-christelijk	2536
	rooms-katholiek	4999
	overig bijzonder	1345
Geheel algemeen bijzondere school	nee	17445
	Montessori	356
	Dalton	669
	Jenaplan	186
Gedeeltelijke Montessori	nee	18599
	ja	57
Gedeeltelijke Dalton	nee	17595
	ja	1061
Gedeeltelijke Jenaplan	nee	18450
	ja	206

Tabel 7: Frequenties per variabele en categorie

Variabele	Categorie	Coëfficiënt	Gestandaardiseerd effect
Constante		531,549*** (0,475)	
Groepsgrootte		-0,026*** (0,010)	-0,255
Schoolscore	110-119	-1,210*** (0,261)	-11,879
	120-139	-1,980*** (0,294)	-19,438
	140-159	-2,504*** (0,355)	-24,583
	>159	-2,759*** (0,387)	-27,086
Weefactor	0,3	-0,994*** (0,333)	-9,758
	1,2	-1,809*** (0,358)	-17,760
	0,25	-1,304*** (0,310)	-12,802
	0,4	-3,487** (1,429)	-34,233
	0,7	-5,972*** (1,529)	-58,629
schooljaar	2010/2011	0,927*** (0,192)	9,101
	2013/2014	-0,426 (0,389)	
Opleiding vader	maximaal lager beroepsonderwijs	-0,125 (0,235)	
	maximaal mbo	1,770*** (0,238)	17,377
	hbo/wo	4,426*** (0,266)	43,452
Opleiding moeder	maximaal lager beroepsonderwijs	-1,190*** (0,264)	-11,683
	maximaal mbo	1,782*** (0,253)	17,495
	hbo/wo	4,972*** (0,297)	48,812
geslacht	Meisje	-0,161 (0,137)	
provincie	Groningen	-1,321** (0,632)	-12,969
	Drenthe	-0,314 (0,530)	
	Friesland	0,304 (0,501)	
	Noord-Holland	0,248	

		(0,308)	
	Flevoland	-0,050 (0,637)	
	Utrecht	0,350 (0,583)	
	Gelderland	0,042 (0,362)	
	Overijssel	-0,274 (0,396)	
	Zeeland	0,185 (0,478)	
	Noord-Brabant	0,612** (0,311)	6,008
	Limburg	1,500*** (0,371)	14,726
stedelijkheid	zeer stedelijk	1,103** (0,447)	10,829
	sterk stedelijk	0,553 (0,342)	
	matig stedelijk	0,569* (0,334)	5,586
	weinig stedelijk	0,600** (0,300)	5,890
$R^2$	0.156		

Tabel 10: Model 3 met Cito-score als afhankelijke variabele

\*\*\* significant op 1% niveau

\*\* significant op 5% niveau

\* significant op 10% niveau

Variabele	Categorie	Coëfficiënt	Gestandaardiseerd effect
Constante		531,675*** (0,498)	
Groepsgrootte		-0,027*** (0,010)	-0,265
Schoolscore	110-119	-1,104*** (0,265)	-10,838
	120-139	-1,955*** (0,299)	-19,193
	140-159	-2,461*** (0,361)	-24,161
	>159	-2,790*** (0,393)	-27,391
Weefactor	0,3	-0,968*** (0,333)	-9,503
	1,2	-1,802*** (0,358)	-17,691
	0,25	-1,338*** (0,310)	-13,136
	0,4	-2,989** (1,467)	-29,344
	0,7	-5,772*** (1,537)	-56,666
schooljaar	2010/2011	0,898*** (0,194)	8,816
	2013/2014	-0,446 (0,417)	
Opleiding vader	maximaal lager beroepsonderwijs	-0,138 (0,235)	
	maximaal mbo	1,758*** (0,238)	17,259
	hbo/wo	4,423*** (0,266)	43,422
Opleiding moeder	maximaal lager beroepsonderwijs	-1,205*** (0,264)	-11,830
	maximaal mbo	1,777*** (0,253)	17,446
	hbo/wo	4,947*** (0,297)	48,567
geslacht	Meisje	-0,148 (0,137)	
provincie	Groningen	-1,278** (0,637)	-12,547
	Drenthe	-0,102 (0,541)	
	Friesland	0,363 (0,506)	
	Noord-Holland	0,222 (0,311)	

	Flevoland	0,135 (0,645)	
	Utrecht	0,531 (0,590)	
	Gelderland	0,209 (0,366)	
	Overijssel	-0,186 (0,399)	
	Zeeland	0,242 (0,479)	
	Noord-Brabant	0,918*** (0,327)	9,012
	Limburg	1,765*** (0,389)	17,328
stedelijkheid	zeer stedelijk	1,016** (0,456)	9,974
	sterk stedelijk	0,283 (0,351)	
	matig stedelijk	0,424 (0,338)	
	weinig stedelijk	0,447 (0,302)	
schooltype	protestants- christelijk	0,513** (0,256)	5,036
	rooms-katholiek	-0,344 (0,233)	
	overig bijzonder	0,304 (0,317)	
Geheel algemeen bijzondere school	Montessori	0,589 (0,531)	
	Dalton	-0,660* (0,388)	-6,479
	Jenaplan	-0,974 (0,701)	
Gedeeltelijke Montessori	ja	-1,834 (1,314)	
Gedeeltelijke Dalton	ja	-1,071*** (0,310)	-10,514
Gedeeltelijke Jenaplan	ja	-0,910 (0,673)	
R <sup>2</sup>	0.158		

Tabel 11: Model 4 met Cito-score als afhankelijke variabele

\*\*\* significant op 1% niveau

\*\* significant op 5% niveau

\* significant op 10% niveau

Variabele	Categorie	Coëfficiënt	Gestandaardiseerd effect
Constance		535,088*** (0,634)	
Groepsgrootte		-0,076*** (0,014)	-0,746
Weegfactor	0,3	0,125 (0,613)	
	1,2	-1,994 (1,241)	
	0,25	-3,882*** (0,807)	-38,111
	0,4	-0,676 (2,935)	
	0,7	-3,512 (6,178)	
schooljaar	2010/2011	0,566** (0,231)	5,557
Opleiding vader	maximaal lager beroepsonderwijs	-0,632 (0,476)	
	maximaal mbo	1,724*** (0,470)	16,925
	hbo/wo	4,152*** (0,492)	40,762
Opleiding moeder	maximaal lager beroepsonderwijs	-2,914*** (0,518)	-28,608
	maximaal mbo	0,250 (0,491)	
	hbo/wo	3,270*** (0,532)	32,103
geslacht	Meisje	-0,298 (0,211)	
provincie	Groningen	-1,848*** (0,683)	-18,143
	Drenthe	-1,216* (0,664)	-11,938
	Friesland	-0,766 (0,564)	
	Noord-Holland	-0,646 (0,444)	
	Flevoland	-4,878*** (1,375)	-47,889
	Utrecht	-0,776 (0,683)	
	Gelderland	-0,289 (0,431)	
	Overijssel	-0,117 (0,538)	

	Zeeland	-0,130 (0,574)	
	Noord-Brabant	-0,175 (0,419)	
	Limburg	1,147** (0,491)	11,261
stedelijkheid	zeer stedelijk	1,120* (0,587)	10,995
	sterk stedelijk	1,048*** (0,396)	10,289
	matig stedelijk	0,304 (0,378)	
	weinig stedelijk	0,691** (0,301)	6,784
$R^2$	0.144		

Tabel 12: Model 5 met Cito-score als afhankelijke variabele en schoolscore categorie 1

\*\*\* significant op 1% niveau

\*\* significant op 5% niveau

\* significant op 10% niveau

Variabele	Categorie	Coëfficiënt	Gestandaardiseerd effect
Constante		531,247*** (0,537)	
Groepsgrootte		-0,020* (0,011)	-0,196
Schoolscore	110-119	-1,088*** (0,274)	-10,681
	120-139	-1,962*** (0,306)	-19,262
	140-159	-2,462*** (0,371)	-24,170
	>159	-2,864*** (0,402)	-28,117
Weefactor	0,3	-0,974*** (0,344)	-9,562
	1,2	-1,795*** (0,368)	-17,622
	0,25	-1,201*** (0,318)	-11,791
	0,4	-3,119** (1,517)	-30,620
	0,7	-5,573*** (1,542)	-54,712
schooljaar	2010/2011	0,928*** (0,201)	9,111
	2013/2014	-0,190 (0,437)	
Opleiding vader	maximaal lager beroepsonderwijs	-0,147 (0,241)	
	maximaal mbo	1,829*** (0,244)	17,956
	hbo/wo	4,506*** (0,274)	44,237
Opleiding moeder	maximaal lager beroepsonderwijs	-1,205*** (0,272)	-11,830
	maximaal mbo	1,777*** (0,260)	17,446
	hbo/wo	4,894*** (0,305)	48,046
geslacht	Meisje	-0,152 (0,141)	
provincie	Groningen	-1,455** (0,687)	-14,284
	Drenthe	-0,066 (0,569)	
	Friesland	0,332 (0,534)	
	Noord-Holland	0,284 (0,318)	



	Flevoland	0,344 (0,673)	
	Utrecht	0,640 (0,608)	
	Gelderland	0,213 (0,377)	
	Overijssel	-0,093 (0,408)	
	Zeeland	0,282 (0,506)	
	Noord-Brabant	0,938*** (0,333)	9,209
	Limburg	1,815*** (0,398)	17,819
stedelijkheid	zeer stedelijk	1,111** (0,474)	10,907
	sterk stedelijk	0,387 (0,368)	
	matig stedelijk	0,526 (0,355)	
	weinig stedelijk	0,551* (0,321)	5,409
schooltype	protestants- christelijk	0,567** (0,266)	5,566
	rooms-katholiek	-0,313 (0,242)	
	overig bijzonder	0,414 (0,328)	
Geheel algemeen bijzondere school	Montessori	0,975 (0,608)	
	Dalton	-0,502 (0,406)	
	Jenaplan	-1,540** (0,771)	-15,119
Gedeeltelijke Montessori	ja	-1,866 (1,467)	
Gedeeltelijke Dalton	ja	-1,001*** (0,326)	-9,827
Gedeeltelijke Jenaplan	ja	-0,616 (0,714)	
R <sup>2</sup>	0.159		

Tabel 13: Model 6 met Cito-score als afhankelijke variabele met groepsgrootte 11 tot en met 40

\*\*\* significant op 1% niveau

\*\* significant op 5% niveau

\* significant op 10% niveau