

Reken op Zuidoost

Een evaluatie van **het tweede loopjaar** 2020/2021 Bridge High Dosage Tutoring op vijf basisscholen in Amsterdam Zuidoost

31 december 2021

Jurgen Tijms (Universiteit van Amsterdam)

Joppe de Ree (Erasmus Universiteit Rotterdam, JDR Analytics)

Thabiso Epema (The Bridge Learning Interventions¹)



¹ Thabiso Epema is formeel in dienst van The Bridge. Door regelgeving met betrekking tot de Algemene Verordening Gegevensbescherming (AVG) is het onderzoeksteam deels afhankelijk voor dataverwerking (inclusief pseudonimisering) van The Bridge. Naast dataverwerking heeft Thabiso Epema ook een rol gespeeld bij de ontwikkeling van nieuwe manieren van dataverwerking en bij verkennende data-analyses. Deze inputs zijn belangrijk geweest voor het onderzoeksproject en maken het coauteurschap van Thabiso Epema wenselijk. De uiteindelijke verantwoordelijkheid voor de inhoud van dit rapport ligt bij de contracthouder (AISSR, de Universiteit van Amsterdam) en de PI Jurgen Tijms.

Inhoudsopgave

Samenvatting	3
Dankwoord	6
Leeswijzer	7
1. Context en de belangrijkste resultaten in vogelvlucht	8
2. De interventie Bridge HDT	14
3. Het onderzoek naar de effectiviteit van Bridge HDT	16
3.1 De onderzoeksopzet (design)	17
4. Resultaten van het gerandomiseerde onderzoek	20
4.1 Operationele bevindingen	20
4.2 Effecten van Bridge HDT	21
4.3 Vervolgonderzoek	27
5. Conclusies en aanbevelingen	29
Literatuurlijst	30
Appendix A: Ontwikkelingen in middelbare schooladviezen	31
Appendix B: Operationalisering van statistische analyses en validiteitsanalyses	35
Appendix C: Beschrijvende analyses op basis van CBS Microdata	37
Appendix D: Resultaten op basis van de TOA niet-talige rekentoets	40

Samenvatting

Dit verslag bespreekt de resultaten van een gerandomiseerd onderzoek (een *randomized controlled trial*, of RCT) naar de effecten van het High Dosage Tutoring (HDT) programma van Stichting The Bridge Learning Interventions (The Bridge) op vijf basisscholen in Amsterdam Zuidoost. Na de tussenevaluatie van het eerste loopjaar in 2019/2020² richt dit verslag zich op het tweede loopjaar 2020/2021 van het HDT programma. In december 2022 verschijnt er nog een laatste onderzoeksverslag dat bericht over het derde en laatste loopjaar (schooljaar 2021/2022).

Het HDT programma van The Bridge (Bridge HDT) heeft plaatsgevonden in nauwe samenwerking met de vijf deelnemende basisscholen³, de schoolbesturen, de Gemeente Amsterdam en de uitvoerende organisatie The Bridge. Bij loopjaar 1 en 2 van Bridge HDT in Amsterdam Zuidoost krijgen leerlingen in groep 7 een schooljaar lang, vijf dagen in de week, 45 minuten per dag, extra rekenondersteuning op maat. Het programma draait onder schooltijd in vaste teams van één tutor en twee leerlingen. Daarnaast onderhouden de tutoren wekelijks telefonisch contact met de ouders/verzorgers van de leerlingen en is er nauw contact met het schoolpersoneel. De tutoren worden dagelijks begeleid en ondersteund door de *Site Director* (projectleider). Het *central office* zorgt voor monitoring, projectontwikkeling en kwaliteitswaarborging. Bridge HDT richt zich op leerlingen die beneden gemiddeld scoren op rekentoetsen en met name op groepen met een verhoogde kans op onderpresteren. The Bridge is om die reden actief in buurten en scholen met relatief veel huishoudens met lagere inkomens. In Amsterdam Zuidoost zijn de deelnemende scholen samen met de Gemeente geselecteerd op basis van percentages ‘doelgroepleerlingen’.⁴ The Bridge is ook actief op basis- en middelbare scholen in Haarlem Schalkwijk/Oost, Amsterdam Nieuw-West en Zaandam Poelenburg. In januari 2022 start er tevens een nieuw project op een middelbare school in Amsterdam Zuidoost.

Het onderzoek naar het tweede loopjaar laat zeer substantiële effecten zien op de rekenprestaties van de deelnemers. We vinden effecten in de orde van 0.6

² Paulle en De Ree (2020).

³ In 2019/2020 en 2020/2021 doen dezelfde vijf basisscholen mee met Bridge HDT.

⁴ OIS en de Gemeente spreken hierbij over ‘doelgroepleerlingen’, leerlingen uit gezinnen die leven van een inkomen op bijstandsniveau, etc. (Achtergronddocument berekening doelgroepleerlingen 2017/'18, Onderzoek, Informatie en Statistiek — Gemeente Amsterdam)

populatiestandaarddeviaties op basis van een schooljaar lang Bridge HDT. Deze effectgroottes zijn spectaculair en worden maar zelden gerapporteerd (zie bijvoorbeeld Kraft (2020)). Met dergelijke effecten kunnen belangrijke beleidsdoelen worden gerealiseerd. In Amsterdam Zuidoost zijn leerlingen met een laag rekenniveau substantieel oververtegenwoordigd. Met effecten van 0.6 populatiestandaarddeviaties kan deze oververtegenwoordiging van laagpresteerders binnen een schooljaar worden weggewerkt (zie hoofdstuk 1 en Appendix C van dit rapport).

De resultaten leggen twee belangrijke aspecten bloot: 1) Basisschoolleerlingen in Amsterdam Zuidoost kunnen veel beter presteren op basis van gerichte ondersteuning. 2) In samenwerking met het reguliere programma op school, biedt Bridge HDT dergelijke gerichte ondersteuning.

Achterstanden in Amsterdam Zuidoost hebben een structureel karakter. Trends in de schooladviezen laten zien dat leerlingen in Zuidoost al ruim een decennia achterblijven, zonder noemenswaardige verbeteringen (zie hoofdstuk 1 en Appendix A). Het is dan ook van belang om Bridge HDT als onderdeel van een oplossing te zien in dit bredere kader van achterstanden. Bridge HDT staat niet op zichzelf. Bridge HDT is complementair aan het reguliere programma op school. Het kan de leerkracht ontlasten om zo systematisch en met meer energie, grotere slagen te maken. Ook in de Verenigde Staten, waar HDT-projecten het eerst zijn ingezet en systematisch geëvalueerd, gaan nu stemmen op om tutoring een vast(er) onderdeel te laten zijn van de schooldag (zie bijvoorbeeld: Kraft en Falken, 2021). Niet in plaats van het reguliere programma, maar als onderdeel ervan.

Zoals ook bij de evaluatie van het eerste loopjaar is opgemerkt, laat recent onderzoek overtuigend zien dat de lockdowns tijdens het schooljaar 2019/2020 de prestaties op school negatief hebben beïnvloed. De orde van grootte van deze *Covid learning-loss* is echter veel kleiner (+/- 0.1 populatiestandaarddeviaties) dan de effecten van een jaar lang Bridge HDT (Engzell, Freyd & Verhagen, 2020). Bridge HDT kan dus worden ingezet om dergelijke achterstanden weg te werken, maar moet in eerste instantie worden gezien als oplossing voor grotere en meer structurele vormen van achterstanden.

De in dit rapport beschreven effecten van Bridge HDT zijn niet eenmalig en niet specifiek voor het primair onderwijs. De eerste Bridge HDT interventie vond plaats in 2017/2018 op het Mundus College, een school voor vmbo en praktijkonderwijs in Amsterdam Nieuw-West. Dit programma leverde effecten op de rekenprestaties vergelijkbaar aan de in dit rapport

beschreven effecten (De Ree et al., 2021). Mede op basis van de eerder behaalde resultaten is The Bridge in gesprek met de gemeente Amsterdam over verdere opschaling in Zuidoost en Nieuw-West, zowel in het primair- als in het voortgezet onderwijs.

Bridge HDT is niet per se het enige middel tegen onderwijsachterstanden. Bridge HDT heeft echter wel een voorsprong op de meeste andere interventies omdat het zich rigoureuus op effectiviteit laat toetsen, met behulp van RCT's. Met RCT's kan structureel en doelgericht worden gezocht naar effectieve, schaalbare en betaalbare oplossingen voor de onderwijsachterstandenproblematiek.

Tot slot bieden we in dit rapport nog een inkijkje in andere aspecten van het onderpresteren. Naast het overaanbod van laagpresteerders in Amsterdam Zuidoost is er ook een relatief tekort aan hoogpresteerders in Amsterdam Zuidoost. Er zijn met name relatief weinig leerlingen die op basis van de toetsscores op de basisschool in aanmerking komen voor een vwo-advies. In de huidige toepassing richt Bridge HDT zich specifiek op de laagpresteerders. Maar onderpresteren is niet alleen een probleem van laagpresteerders. Ook de hoogpresteerders kunnen, waarschijnlijk, met passende ondersteuning beduidend beter presteren dan ze nu al doen.

Dankwoord

Dit onderzoek is mogelijk gemaakt door financiële steun van de Gemeente Amsterdam. Verder profiteert het onderzoek van de inzet van leerlingen, leerkrachten, ander personeel en het bestuur van de deelnemende vijf basisscholen in Amsterdam Zuidoost. Ook wil het onderzoeksteam het personeel van The Bridge bedanken voor hun inzet bij het mogelijk maken van dit onderzoek.

Op basis van de samenwerking tussen wetenschap, overheid, uitvoeringsinstanties en scholen kan de samenleving lessen trekken over de werkzaamheid van Bridge HDT in deze specifieke context. De lessen leveren opties voor het bestrijden van onderwijsachterstanden. Niet alleen in Amsterdam, maar ook in de rest van Nederland. We danken hiervoor The Bridge, de Gemeente Amsterdam en de deelnemende scholen (waaronder de directeuren, intern begeleiders, betrokken leerkrachten, leerlingen en hun ouders/verzorgers) die hiermee bijdragen aan het zichtbaar maken van oplossingen voor de problematiek van onderwijsachterstanden onder een breder publiek.

Leeswijzer

Voor een gedetailleerde beschrijving van de inhoud van het Bridge HDT-programma, verwijzen we naar het onderzoeksverslag van het eerste loopjaar van Bridge HDT in Amsterdam Zuidoost (Paulle & De Ree, 2020) en de inhoudelijke procesverslagen van The Bridge zelf ('Inhoudelijk verslag over Bridge High Dosage Tutoring in Amsterdam Zuidoost - loopjaar 2 (schooljaar 2020-2021)', ingediend op 31 oktober 2021 via het subsidieportaal). Het Bridge HDT project is bekend onder de naam 'Bridge HDT - Amsterdam PO'.

In hoofdstuk 1 gaan we in op de bestaande onderwijsachterstanden in Amsterdam Zuidoost. Hiervoor gebruiken we onder andere de CBS Microdata die onder voorwaarden voor onderzoek beschikbaar zijn. We bespreken achtereenvolgens de bestaande onderwijsachterstanden, de hardnekkigheid van deze achterstanden, en wat we kunnen verwachten als Bridge HDT (of een andere gelijkwaardige interventie) Zuidoost-breed zou worden uitgerold. In deze laatste stap nemen we een voorschot op de resultaten van dit onderzoek. Hoofdstuk 1 is ook voor beleidsmakers het meest toegankelijke en belangrijke hoofdstuk. In de hoofdstukken 3-5 beschrijven we verdere details van het onderzoek.

In hoofdstuk 2 gaan we in op de inhoud van het programma, waarbij we voornamelijk verwijzen naar het onderzoeksverslag van het eerste loopjaar Bridge HDT in Amsterdam Zuidoost (Paulle & De Ree, 2020). In Hoofdstuk 3 beschrijven we de onderzoeksopzet (*research design*). In hoofdstuk 4 beschrijven we de resultaten van het onderzoek en in hoofdstuk 5 bespreken we een aantal conclusies en aanbevelingen.

1. Context en de belangrijkste resultaten in vogelvlucht

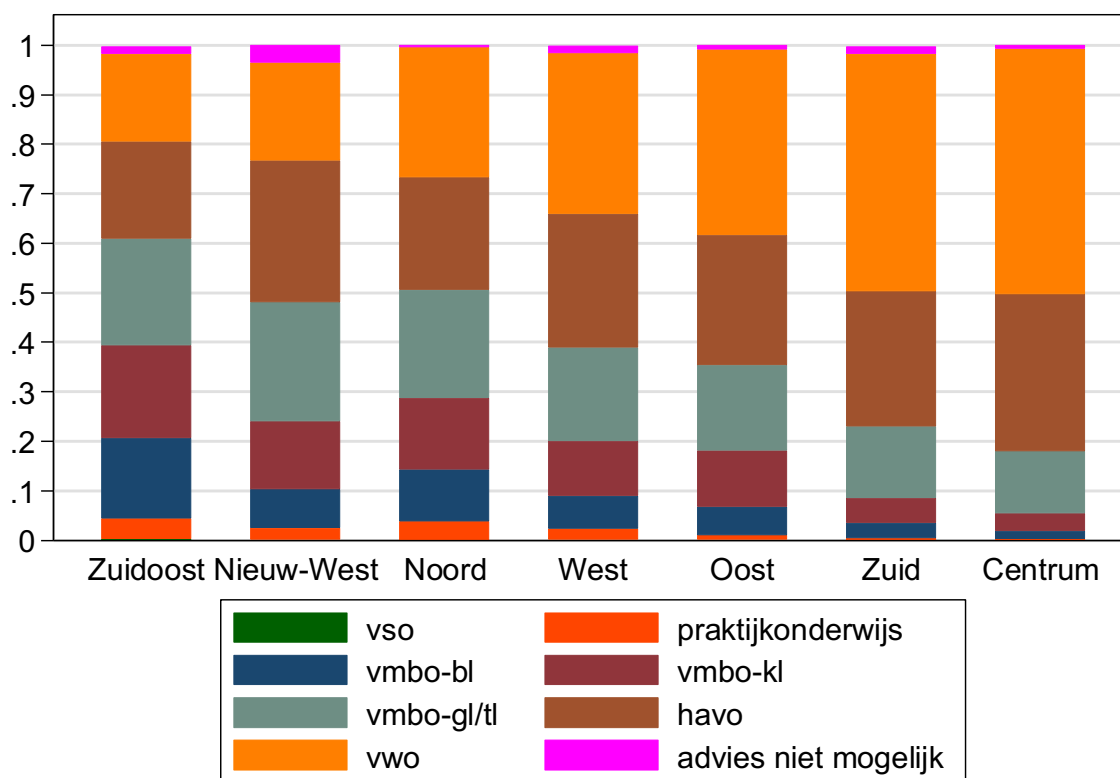
Amsterdam Zuidoost is een stadsdeel in Amsterdam met een bovengemiddeld percentage huishoudens met lage tot zeer lage inkomens. Zoals bekend blijven de schoolprestaties van leerlingen uit gezinnen met lage inkomens vaak achter. Uitdagingen thuis of in de bredere sociale context kunnen hierbij een rol spelen.

Figuur 1 laat zien dat ‘lagere’ schooladviezen relatief oververtegenwoordigd zijn in Amsterdam Zuidoost. In vergelijking met de andere stadsdelen heeft Amsterdam Zuidoost bijvoorbeeld het hoogste percentage leerlingen met een vmbo-bl of vmbo-kl advies. In Amsterdam Zuidoost krijgt 40% van de groep 8-leerlingen een dergelijk advies. In Amsterdam Oost of West is dit 20% en in de stadsdelen Zuid en Centrum maar ongeveer 5-10% van alle groep 8 leerlingen. We zien ook dat de resultaten in Zuidoost achterblijven bij scholen en groepen met een vergelijkbaar sociaal-economisch profiel. De afdeling Onderzoek Informatie en Statistiek (OIS) van de Gemeente Amsterdam laat zien dat scholen met vergelijkbare percentages ‘doelgroepleerlingen’, in Zuidoost minder presteren dan bijvoorbeeld in Amsterdam Nieuw-West of Noord.^{5 6}

⁵ OIS en de Gemeente spreken hierbij over ‘doelgroepleerlingen’, leerlingen uit gezinnen die leven van een inkomen op bijstandsniveau, etc. (Achtergronddocument berekening doelgroepleerlingen 2017/’18, Onderzoek, Informatie en Statistiek — Gemeente Amsterdam)

⁶ OIS rapporteert over de relatieve achterstand in Zuidoost, ten opzichte van leerlingen met een vergelijkbaar profiel in hun inputs voor een klankbordgroep wetenschap en beleid, getiteld *Gelijke kansen in het onderwijs in Zuidoost*.

Figuur 1. Middelbare schooladviezen in Amsterdam per stadsdeel



Amsterdam 2019/20 per stadsdeel

Noot: Eigen berekeningen op basis van publiek beschikbare gegevens van de Dienst Uitvoering Onderwijs (DUO, https://duo.nl/open_onderwijsdata/). Een deel van de leerlingen krijgt een gemengd advies, bijvoorbeeld havo-vwo. Voor de leesbaarheid van de Figuur hebben we de gemengde adviezen gelijk verdeeld over de relevante enkelvoudige adviezen. Leerlingen met een havo-vwo advies worden dus voor de helft toegeschreven aan havo en de andere helft aan vwo.

Achterstanden van leerlingen in Amsterdam Zuidoost zijn in de laatste tien jaar ook toegenomen ten opzichte van achterstanden van leerlingen in de rest van de stad. Dit patroon is het resultaat van een afname van het percentage lagere adviezen in de rest van Amsterdam. De gemiddelde leerprestaties in Amsterdam Zuidoost, afgemeten aan de gerealiseerde middelbare schooladviezen, zijn niet noemenswaardig verbeterd of verslechterd (zie Appendix A).

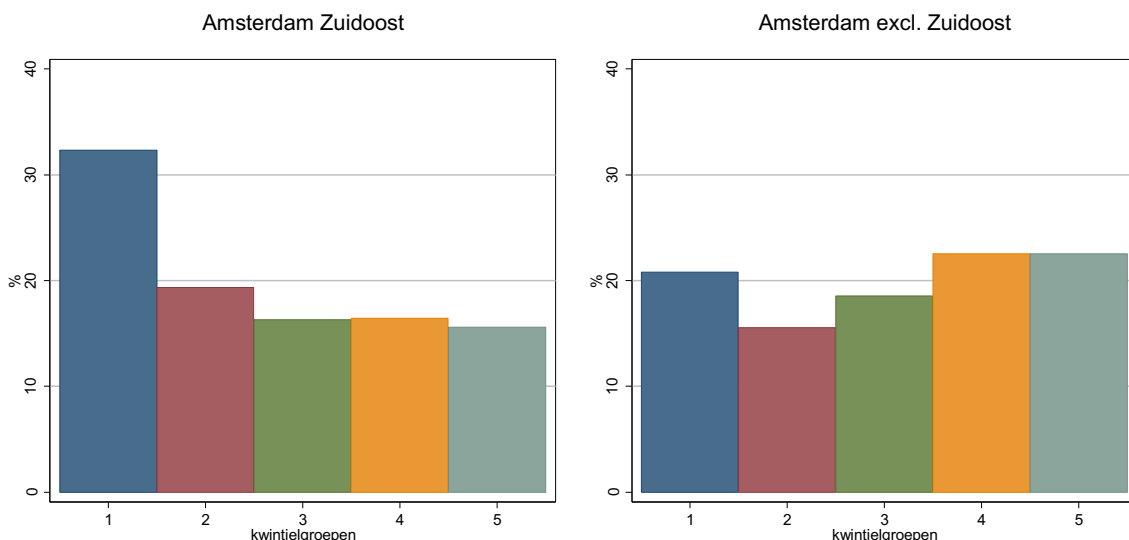
Prestaties in taal en rekenen zijn de belangrijkste basis voor het middelbare schooladvies. Op basis van Cito-onderdelen taal en rekenen zien we de achterstanden in Zuidoost nogmaals terug. Voor de figuren op basis van de Cito-scores richten we ons op het schooljaar 2013/2014. De resultaten uit 2013/2014 zijn belegen, maar het is, gegeven de stabiliteit in de advisering, aannemelijk dat ze voldoende representatief zijn voor de situatie van vandaag. We

kiezen voor het schooljaar 2013/2014 als basis, omdat vanaf 2014/2015 scholen over kunnen stappen naar andere toetsaanbieders en omdat de eindtoetsen een andere rol hebben gekregen in het adviseringsproces. Het is bijvoorbeeld na 2014/2015 niet meer voor elke leerling van belang om goed te scoren op de eindtoets, omdat het voorlopige advies al ruim voor de eindtoets wordt vastgesteld.

In Figuur 2 proberen we de vergelijking in rekenprestaties tussen Amsterdam Zuidoost en de rest van Amsterdam intuïtief te maken. Hiervoor delen we alle groep 8-leerlingen in Nederland op in kwintielen, op basis van hun scores op het rekenonderdeel van de Cito-toets 2013/2014. De blauwe staaf in de staafdiagrammen, bijvoorbeeld, refereert aan de 20% laagst scorende leerlingen in Nederland. De grijze staaf in de staafdiagrammen refereert aan de 20% hoogst scorende leerlingen in Nederland.

Figuur 2 laat zien dat laagst scorende groep 8-leerlingen in Zuidoost sterk zijn oververtegenwoordigd, met ongeveer 32% van alle leerlingen in Zuidoost. In de rest van Amsterdam valt ongeveer 21% van de leerlingen in het eerste kwintiel. Deze groep leerlingen rekt op een niveau dat past bij een vmbo-bl of vmbo-kl middelbare schooladvies.

Figuur 2. Rekenprestaties op basis van het rekenonderdeel van de Cito-eindtoets (schooljaar 2013/2014) in Amsterdam Zuidoost [links] en Amsterdam (exclusief Amsterdam Zuidoost) [rechts] en, in vergelijking met de rekenprestaties landelijk.



In het huidige onderzoek hebben we met een RCT de effecten van Bridge HDT geëvalueerd, we vinden effecten van 0.6 populatiestandaarddeviaties op basis van een schooljaar lang Bridge HDT. Het is van belang om deze effecten inzichtelijk te maken. We doen dit door de

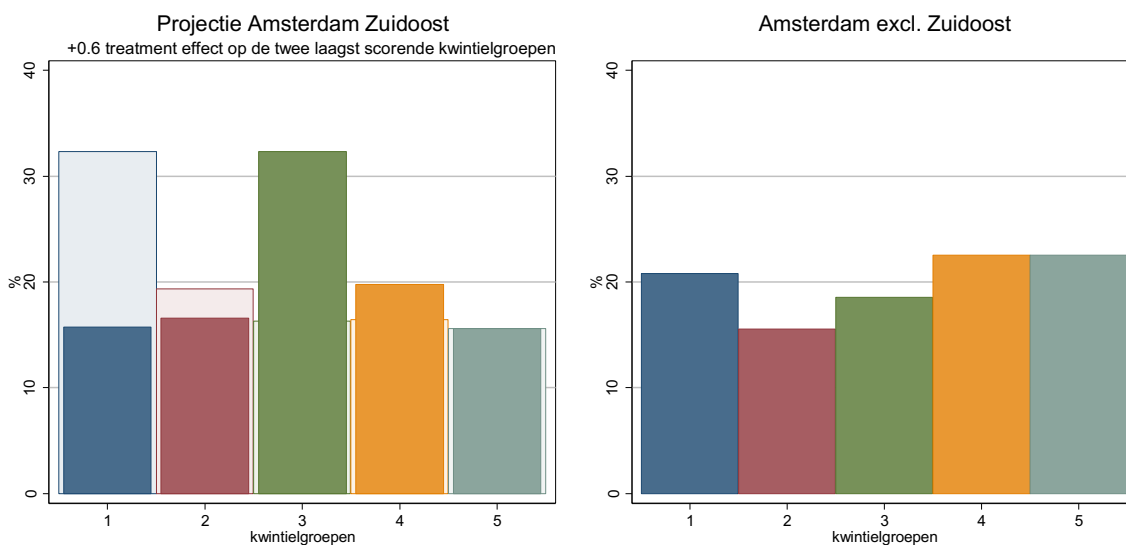
gevonden effecten op te tellen bij de scores van leerlingen uit de eerste twee kwintielen, dat wil zeggen leerlingen met rekencores in de laagst scorende 40% landelijk (de blauwe en bruine staven in Figuur 2).⁷ Vervolgens rekenen we nogmaals uit in welke van de vijf kwintielgroepen de leerlingen zouden vallen. Figuur 3 laat de invloed van deze projecties zien. Met andere woorden, in Figuur 3 zien we wat we kunnen verwachten als Bridge HDT Zuidoost-breed zou worden uitgerold onder leerlingen uit de twee laagst scorende kwintielen. Figuur 3 laat zien dat de achterstanden in termen van het aandeel laag scorende leerlingen in zijn geheel kan worden ingelopen. De resultaten zijn zelfs gekanteld in het voordeel van Amsterdam Zuidoost. In Figuur 3 zien we dat effectgroottes van 0.6 zeer substantieel zijn, zoals ook opgemerkt in de relevante literatuur (Kraft, 2020).

Met de projecties laten we zien op welke manier en in welke mate de onderwijsachterstanden kunnen worden weggewerkt. Bridge HDT richt zich in de huidige toepassing op laagpresteerders. Wanneer Bridge HDT in deze vorm verder zou worden uitgerold, verwachten we met name dat leerlingen aan de onderkant van de vaardigheidsverdeling veel vorderingen maken. Logischerwijs biedt Bridge HDT in de huidige toepassing dus geen oplossing voor het aanpakken van het relatief lage percentages leerlingen met vwo advies bijvoorbeeld. Het is belangrijk om op te merken dat onderpresteren niet alleen een zaak is van laag scorende leerlingen.

The Bridge werkt samen met de Gemeente Amsterdam aan de voorbereiding van een gedeeltelijke opschaling. Het gaat hierbij om het opschalen van de financiële en personele middelen en deze vervolgens in te passen in de bestaande projectstructuur. The Bridge heeft als een van de prioriteiten bij het opschalen het op peil houden van de kwaliteit van de implementatie. Met effectgroottes van 0.6 populatiestandaarddeviaties is er echter ook enige marge. Ook met effecten in de orde van 0.3-0.4 populatiestandaarddeviaties kan een groot deel van de achterstanden in de rekenprestaties worden weggewerkt.

Figuur 3. [links] Geprojecteerde rekenprestaties op basis van het rekenonderdeel van de Cito-eindtoets in het schooljaar 2013/2014. We projecteren *treatment* effecten van 0.6 populatiestandaarddeviaties op de twee laagst scorende kwintielen in Amsterdam Zuidoost. De projecties zijn donker gekleurd. De werkelijke situatie in 2013/2014 is licht gekleurd in de achtergrond. [rechts] Rekenprestaties op basis van het rekenonderdeel van de Cito-eindtoets in Amsterdam exclusief stadsdeel Amsterdam Zuidoost.

⁷ Appendix C bespreken we de details van deze projecties en de beschikbaarheid van data over Cito-scores.



Deze verbeteringen in de leerprestaties zijn zeer substantieel en geven aan dat 1) achterstanden niet *'fixed'* zijn en dat substantiële veranderingen mogelijk zijn, en dat 2) Bridge HDT onderdeel van een oplossing kan zijn voor het bestrijden van onderwijsachterstanden. Dit laatste betekent niet dat Bridge HDT de enige interventie is die helpt. Ook andere interventies kunnen kansrijk zijn. Het is echter belangrijk om ook andere interventies te onderwerpen aan gerandomiseerd onderzoek (RCT's), om zo doelgericht, effectieve, betaalbare, en schaalbare oplossingen voor onderwijsachterstanden te ontwikkelen.

Voor dit rapport onderzoeken we ook de effecten van Bridge HDT op het begrijpend lezen. Hiervoor zijn twee redenen. Omdat Bridge HDT onder schooltijd plaatsvindt, missen deelnemers aan Bridge HDT elke dag 45 minuten van het reguliere (misschien taliger) onderwijsaanbod. Dit kan negatieve effecten hebben op de prestaties in het begrijpend lezen, bijvoorbeeld. Ook kan Bridge HDT positieve effecten heeft op andere domeinen van het leren, omdat leerlingen tijdens Bridge HDT meer vertrouwen in eigen kunnen opbouwen bijvoorbeeld. In dit onderzoek vinden we geen effecten van Bridge HDT op de prestaties in het begrijpend lezen, niet positief en niet negatief. Hiermee sluiten we aan bij eerder onderzoek in Nederland en in de Verenigde Staten (De Ree en co-auteurs, 2021a-b, en Guryan en co-auteurs, 2021)

Het onderzoek laat de kracht zien van RCT's. Het gebruik van RCT's heeft onder aanvoering van nobelprijswinnaars Abhijit Banerjee, Esther Duflo en Michael Kremer een ontwikkeling doorgemaakt. De trend heeft zich vervolgens vanuit de ontwikkelingseconomie doorgezet

naar de Verenigde Staten, en nu ook naar Europa. Het gebruik van RCT's voor het evalueren van onderwijsinterventies in Nederland is echter nog beperkt. Het CPB doet bijvoorbeeld in de publicatie *Kansrijk Onderwijsbeleid* niet of nauwelijks een beroep op Nederlands onderzoek. Het commitment van de Gemeente Amsterdam, The Bridge en de deelnemende scholen aan het ondersteunen en opbouwen van wetenschappelijke kennis op basis van RCT's heeft dus een voortrekkersrol en een voorbeeldfunctie in Nederland. Zo speelt deze evaluatie een rol bij het ontwikkelen van effectieve en schaalbare interventies voor minder kansrijke groepen leerlingen, bijvoorbeeld in Amsterdam Zuidoost.

2. De interventie Bridge HDT

High Dosage Tutoring (HDT) is een interventie ontworpen en ontwikkeld in de Verenigde Staten (VS). De belangrijkste aanbieder van HDT in de VS is Saga Education.⁸ Saga Education, en ‘Saga HDT’ is ook herhaaldelijk met behulp van RCT’s onderzocht, en ook op grotere schaal succesvol gebleken (Guryan en co-auteurs. 2021). Sinds 2017 voert Stichting The Bridge Learning Interventions (The Bridge) HDT interventies uit in Nederland. Het HDT programma van The Bridge (Bridge HDT) is voor een belangrijk deel gebaseerd op het HDT programma van Saga Education. Het is echter aangepast aan de Nederlandse context en bevat meer sociaal-emotionele componenten. In de huidige toepassing richt Bridge HDT zich op leerlingen op of onder het landelijk gemiddelde, in termen van de rekenprestaties. Saga Education is sinds het eerste begin betrokken geweest als consultant van The Bridge.

De inhoud en methodes van Bridge HDT hebben we uitvoerig beschreven in de effectrapportage van het eerste loopjaar Bridge HDT in Amsterdam Zuidoost (Paulle & De Ree, 2020). Bridge HDT houdt kort gezegd in dat leerlingen in het basisonderwijs (groepen 7-8) of voortgezet onderwijs (klas 1-2) een schooljaar lang onder schooltijd (40-50 minuten) twee tot vijf keer per week extra rekenbegeleiding ontvangen in een vaste 2-op-1 setting (twee leerlingen per tutor). Tutoren zijn betaalde krachten van The Bridge en worden professioneel getraind en op dagelijkse basis gecoacht en begeleid door de *Site Director*.

De lessen verlopen volgens een vaste structuur. De tutoren maken de lessen op maat aan de hand van het eigen rekencurriculum van The Bridge (op dit moment gebaseerd op het 1F-niveau). Ook begeleidt de tutor de leerling in andere vaardigheden, zoals plannen, samenwerken, zelfstandig werken en doorzetten op de momenten dat het nodig is. Daarnaast hebben de tutoren (twee-)wekelijks telefonisch contact met de ouders/verzorgers van de leerlingen. Het ontwikkelen van een band tussen de leerling, tutor en ouders staat centraal in de benadering van The Bridge. Daarnaast is Bridge HDT een verlengstuk van de school en is er nauw contact met school- en zorgpersoneel. Het *central office* coördineert alle projecten en zorgt voor afstemming met het onderzoeksteam, externe partners, HRM, financiële zaken, projectontwikkeling en kwaliteitswaarborging. Meer informatie is te vinden op de website

⁸ Zie ook: <https://www.sagaeducation.org/>

van The Bridge (www.tbli.nl) en in de rapportage van het eerste loopjaar Bridge HDT in het primair onderwijs in Amsterdam Zuidoost (Paulle & De Ree, 2020).

In het eerste en tweede loopjaar van Bridge HDT in Amsterdam Zuidoost doen dezelfde vijf basisscholen mee. De interventie in het eerste en tweede loopjaar is gebaseerd op het ‘originele model’, waarbij deelnemers elke schooldag van de week meedoen aan de tutorsessies. Het derde loopjaar van dit project in het schooljaar 2021/2022 is anders. Er doen dan 8 scholen mee, en de tutorsessies vinden drie keer per week plaats. Op basis van dit alternatieve model, kunnen met vrijwel dezelfde middelen, meer leerlingen van de ondersteuning profiteren.

3. Het onderzoek naar de effectiviteit van Bridge HDT

Bridge HDT in Amsterdam Zuidoost wordt geëvalueerd op basis van een gerandomiseerd onderzoek met controlegroep, te weten een *randomized controlled trial* (of RCT).⁹ Bij gerandomiseerd onderzoek worden deelnemers willekeurig (aselect, of *random*) getrokken uit de doelpopulatie. Vervolgens wordt het geselecteerde en het niet-geselecteerde deel van de doelpopulatie met elkaar vergeleken. De consequentie van het willekeurig selecteren van deelnemers is dat het geselecteerde deel niet systematisch verschilt van het niet-geselecteerde deel. Systematische verschillen tussen de geselecteerde groep (de *treatmentgroep*) en de niet-geselecteerde groep (de controlegroep) die we na afloop van de interventie terugzien in de data, kunnen dus worden beoordeeld als een effect van de interventie. RCT's worden gezien als de gouden standaard in het onderzoek naar de (causale) effectiviteit van o.a. beleidsinterventies.

In dit hoofdstuk kiezen we ervoor om de resultaten van het eerste loopjaar (in het schooljaar 2019/2020) en de resultaten van het tweede loopjaar (in het schooljaar 2020/2021) gezamenlijk te bestuderen. Een duidelijk verschil tussen beide loopjaren is de invloed van lockdowns als gevolg van de Covid-19 pandemie. De lockdowns hebben vooral invloed gehad op de uitvoering van het eerste loopjaar van Bridge HDT, in de tweede helft van het schooljaar 2019/2020. De effecten van de lockdowns zagen we in de eerste rapportage ook terug in de resultaten. Het tweede loopjaar heeft minder last gehad van lockdowns, zoals later in het verslag in meer detail wordt beschreven. In dit rapport laten we zien dat de substantieel grotere effecten die we vinden voor dit tweede loopjaar, kunnen worden toegeschreven aan de grotere intensiteit en stabiliteit van de interventie. De belangrijkste oorzaak van de verschillen in resultaten tussen het eerste en het tweede loopjaar is het verschil in invloed van de Covid-19 pandemie.

⁹ Het onderzoek is geregistreerd in het *social science registry* voor *randomized controlled trials* van de American Economic Association (<https://www.socialscienceregistry.org/trials/4643>).

3.1 De onderzoeksofzet (design)

Vijf basisscholen in Amsterdam Zuidoost zijn - in samenwerking met de Gemeente Amsterdam - voorgeselecteerd op basis van voldoende aantallen 'doelgroepleerlingen' voor Bridge HDT. De doelgroep van Bridge HDT in Amsterdam Zuidoost bestaat uit leerlingen die halverwege groep 6 rekenen met vaardigheidsniveaus III, IV of V. Dit zijn leerlingen die op of onder het landelijk gemiddelde scoren op de Cito LVS-toetsen Rekenen-Wiskunde. Het gaat dus niet alleen om leerlingen met grote leerachterstanden.

De procedure voor de randomisatie is als volgt opgezet en uitgevoerd. The Bridge heeft van de vijf scholen een overzicht ontvangen met namen en toetsscores van leerlingen die aan de bovengenoemde voorwaarden voldoen. Op basis van deze lijsten zijn de ouders/verzorgers benaderd middels een toestemmingsformulier. Uiteindelijk heeft dit een doelgroep van 102 leerlingen in loopjaar 2 opgeleverd, verdeeld over de vijf deelnemende basisscholen.¹⁰ De data-manager van The Bridge heeft van deze leerlingen de persoonsgegevens uit de lijsten verwijderd en vervangen door pseudoniemen. Vervolgens heeft de data-manager deze lijst gedeeld met onderzoekers van de Universiteit van Amsterdam via een beveiligde online verbinding. Op basis van deze lijst is vervolgens de randomisatie uitgevoerd door het onderzoeksteam van de UvA.

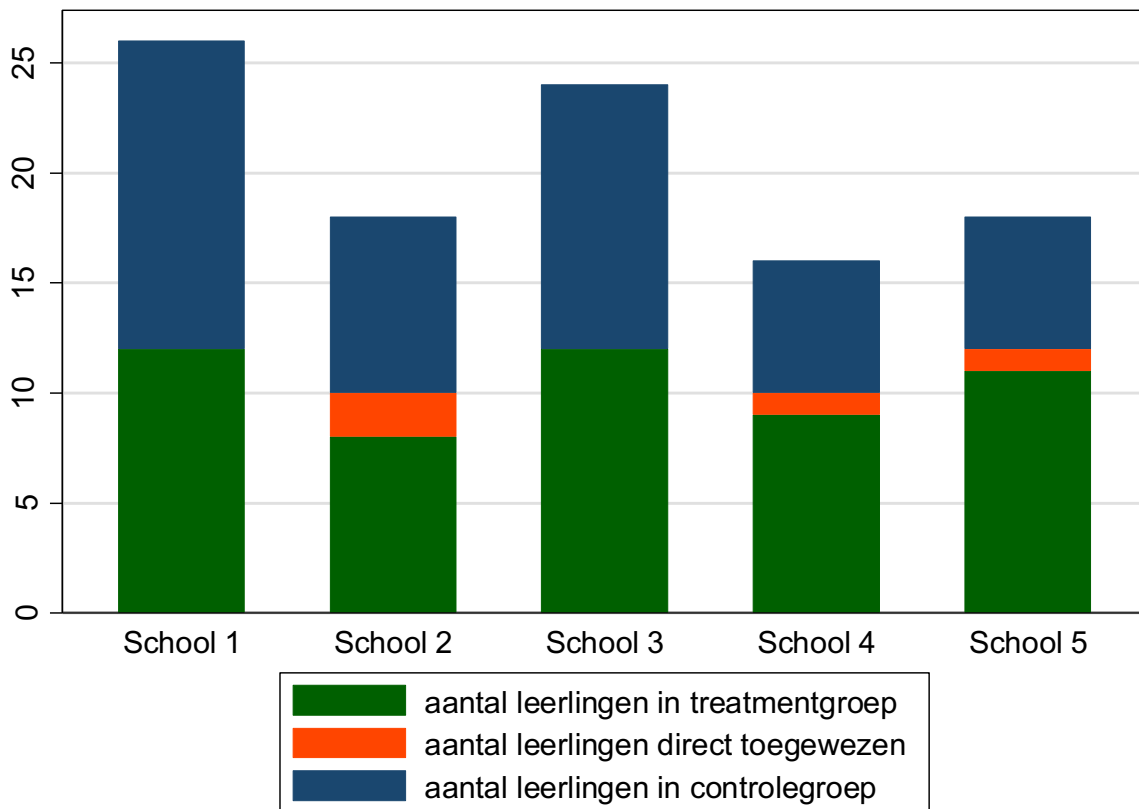
Randomisatie is uitgevoerd op basis van een *block*-randomisatieprocedure waarbij *sampling blocks* zijn aangemaakt op basis van de rekenscores in groep 6, het leerjaar voorafgaand aan de HDT-interventie in groep 7. *Block*-randomisatie zoals hier uitgevoerd heeft het voordeel dat voorafgaand aan de randomisatie meer balans (gelijkheid) wordt geforceerd tussen de *treatment*- en controlegroep in termen van rekenscores, scholen en klassen binnen de scholen. *Block*-randomisatieprocedures leveren over het algemeen preciezere schattingen van effecten van interventies.

Van de 102 leerlingen zijn er 4 direct (met $p = 1$) toegewezen aan het HDT-programma.¹¹ De overgebleven 98 leerlingen zijn met een kans kleiner dan 1 ($p < 1$) geselecteerd en dus onderdeel van het gerandomiseerde onderzoek. Van de 98 leerlingen zijn er 52 toegewezen aan de *treatmentgroep* en de rest, $(98-52=)$ 46, aan de controlegroep.

¹⁰ Het gaat hier om een duidelijke meerderheid van het totaal aan doelgroepleerlingen.

¹¹ Af en toe weken baseline toetsscores sterk af van andere leerlingen in de klas. Voor het goed functioneren van het onderzoeksdesign hebben we ervoor gekozen om in een aantal gevallen leerlingen direct toe te wijzen aan het HDT programma.

Figuur 4. Aantallen leerlingen in *treatment*- en controlegroep, per school



Figuur 4 laat zien hoe de aantallen leerlingen in de *treatment*- en controlegroepen zijn verdeeld over de deelnemende scholen. Met zes beschikbare tutores, bestaan de tutorgroepen uit maximaal twaalf leerlingen per school.¹² Tijdens de randomisatie leggen we op dat er per tutorgroep maximaal twaalf leerlingen worden geselecteerd, mits er voldoende leerlingen beschikbaar zijn. Voor twee van de vijf deelnemende scholen hebben we voor een andere verdeling gekozen door de lagere aantallen doelgroepeleringen. Het feit dat het aantal leerlingen in de *treatmentgroep* groter is dan het aantal leerlingen in de controlegroep heeft geen invloed op de validiteit van de onderzoeksopzet.

Scores op de Cito leerlingvolgsysteem-toetsen (LVS-toetsen) Rekenen-Wiskunde zijn meestal sterk gecorreleerd door de tijd heen. De scores halverwege groep 6 zijn bijvoorbeeld zeer voorspellend voor de scores aan het einde van groep 7. Kennis over deze correlaties zijn meegenomen bij de powerberekeningen. Details over powerberekeningen en, daaraan

¹² De zes tutores en de *Site Director* reizen dus elke dag van school naar school.

gerelateerde, ‘*minimum detectable effects*’ of MDE’s, zijn beschreven in het *social science registry* van de American Economic Association.¹³

We meten de effecten van Bridge HDT aan de hand van scores op de Cito LVS-toetsen Rekenen-Wiskunde en Begrijpend lezen. Naast het gebruik van Cito-toetsen hebben we ook een niet-talige TOA rekentoets afgenomen. Bij de toetsafnames van de TOA rekentoets zijn door het onderzoeksteam georganiseerde externe observanten aanwezig geweest. We onderzoeken de effecten na een half schooljaar Bridge HDT en na een volledig schooljaar. Bij de eindrapportage na afloop van het derde loopjaar Bridge HDT, zullen we ook de toetsresultaten van halverwege groep 8 analyseren.

De Cito LVS-vaardigheidsscores worden eerst gestandaardiseerd ten opzichte van het landelijk gemiddelde en de landelijke spreiding van de betreffende leeftijdsgroep. Vervolgens gebruiken we de gestandaardiseerde scores als afhankelijke variabelen in de regressieanalyses. Als onafhankelijke variabelen gebruiken we een indicator variabele¹⁴ voor de *treatmentgroep* en een volledige set aan indicator variabelen voor de verschillende *sampling blocks*. In Appendix B lichten we een aantal details van de analyses verder toe. Bij de bespreking in Appendix B hebben we ervoor gekozen om het gebruik van formules te vermijden. Extra (meer formele) details van deze analyses zijn op aanvraag beschikbaar.

¹³ <https://www.socialscienceregistry.org/trials/4643>

¹⁴ Indicator variabelen worden ook wel dummy variabelen of 0/1 variabelen genoemd. Deze variabelen hebben twee mogelijke uitkomsten, een 1 (als aan de voorwaarde wordt voldaan) en een 0 (als niet aan de voorwaarde wordt voldaan).

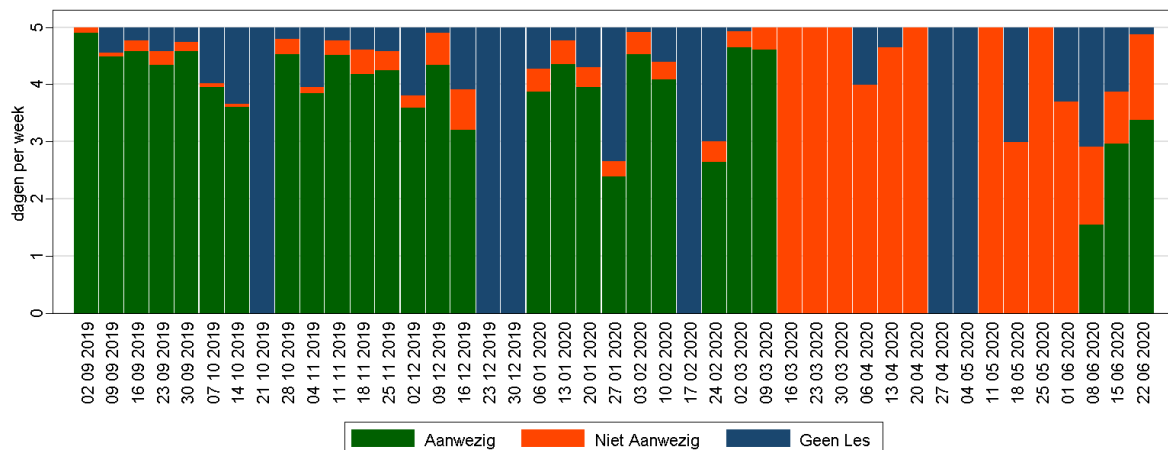
4. Resultaten van het gerandomiseerde onderzoek

In dit hoofdstuk beschrijven we de resultaten van het gerandomiseerde onderzoek naar de effecten van het tweede loopjaar Bridge HDT in stadsdeel Amsterdam Zuidoost in 2020/2021. Ook rapporteren we nogmaals de resultaten van het eerste loopjaar Bridge HDT, om de vergelijking tussen de twee loopjaren inzichtelijker te maken.

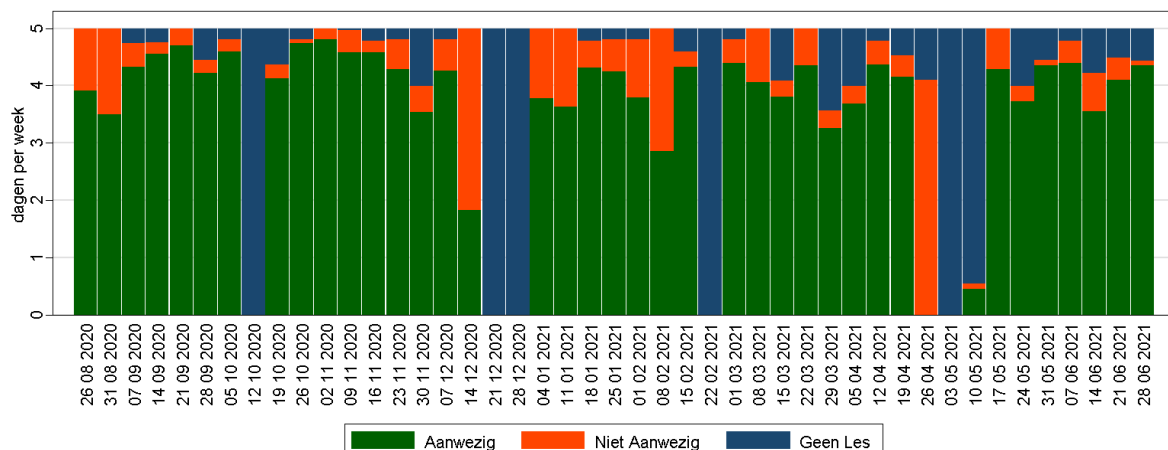
4.1 Operationele bevindingen

Het tweede loopjaar heeft in tegenstelling tot het eerste loopjaar veel minder last gehad van lockdowns. In Figuur 5 (eerste loopjaar) en Figuur 6 (tweede loopjaar) beschrijven we de aanwezigheid bij de tutor sessies op basis van de aanwezigheidsregistraties.

Figuur 5. Aanwezigheid bij tutor sessies, in dagen per week (eerste loopjaar)



Figuur 6. Aanwezigheid bij tutorsessies, in dagen per week (tweede loopjaar)



In Figuur 5 is duidelijk de invloed van de lockdowns te zien. De oranje delen geven afwezigheid aan. Vanaf halverwege maart 2020 (tijdens het eerste loopjaar) gingen de scholen dicht. The Bridge heeft in die fase een plan opgesteld om door middel van weektaken Bridge HDT voort te zetten. Dit sloot aan bij de manier waarop veel scholen toen hun eigen lessen vormgaven. Vanaf 13 april zijn tutores gaan videobellen met de leerlingen, als aanvulling op de weektaken. Deze videolessen vonden echter nog steeds een stuk minder frequent plaats dan oorspronkelijk gepland was. In juni zijn op vier van de vijf scholen de tutorsessies opnieuw fysiek opgestart.

In Figuur 6 zien we dat de aanwezigheid bij de tutorsessies in het tweede loopjaar vrijwel altijd op peil is gebleven. Hoewel er tijdens het tweede loopjaar (tussen 14 december en 8 februari) ook een schoolsluiting was, hebben leerlingen in deze periode wel online de tutorsessies kunnen volgen volgens het reguliere rooster. The Bridge en de scholen waren in deze tweede lockdown ook veel beter voorbereid op onderwijs op afstand. Deelnemers hebben dus, ziekte en andere omstandigheden daargelaten, altijd vier tot vijf dagen per week kunnen meedoen.

Over de invulling van Bridge HDT gedurende deze twee loopjaren, heeft The Bridge verslag gedaan bij de Gemeente Amsterdam. In deze verslagen staat ook beschreven hoe The Bridge is omgegaan met de schoolsluitingen. Zie hiervoor het verslag over loopjaar 1 (verslag genaamd “Inhoudelijk Verslag over Bridge High Dosage Tutoring in Amsterdam Zuidoost”, ingediend op 31 oktober 2020 via het subsidieportaal) en het verslag over loopjaar 2 (verslag genaamd “Inhoudelijk verslag over Bridge High Dosage Tutoring in Amsterdam Zuidoost - loopjaar 2 (schooljaar 2020-2021)”, ingediend op 31 oktober 2021 via het subsidieportaal)

4.2 Effecten van Bridge HDT

De invloed van de lockdowns zien we ook duidelijk terug in de effecten van het HDT-programma. Tabel 1 rapporteert de effecten van Bridge HDT gemeten aan de hand van de Cito LVS-scores Rekenen-Wiskunde, voor het eerste loopjaar (Tabel 1A) en het tweede loopjaar (Tabel 1B) respectievelijk.¹⁵

Tabel 1. Effecten van Bridge HDT op de rekenprestaties voor loopjaar 1 (1A) en loopjaar 2 (1B) op basis van Cito LVS-toetsen Cito Rekenen/Wiskunde (in termen van populatiestandaarddeviaties.)

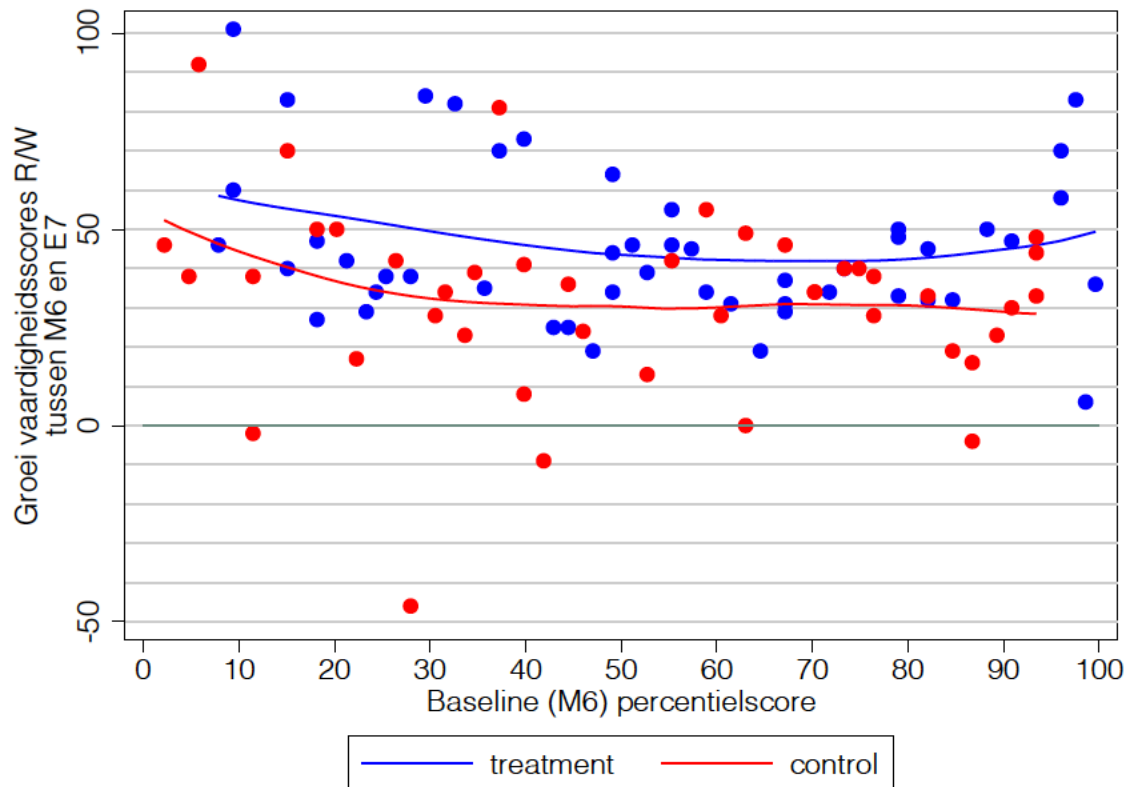
	(1)	(2)
A: Loopjaar 1		
	Midline M7	Endline E7
Treatment effect	0.28**	0.28
Standaardfout	(0.12)	(0.17)
Aantal observaties	[98]	[94]
B: Loopjaar 2		
	Midline M7	Endline E7
Treatment effect	0.40***	0.65***
Standaardfout	(0.15)	(0.18)
Aantal observaties	[90]	[92]
<p>Noten: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. De effecten van de interventie zijn geschat op basis van een regressiemodel. Als afhankelijke variabele gebruiken we de gestandaardiseerde toetsscores. Als onafhankelijke variabelen gebruiken we 1) dummy/indicator variabelen voor alle <i>sampling blocks</i>/strata, 2) en een dummy/indicator variabele voor de treatment groep, 3) toetsscores bij baseline, halverwege groep 6. De geschatte parameter op de indicator voor de treatmentgroep is gerapporteerd in deze tabel als schatting voor het treatment effect. Verder wegen we de data zodat de treatmentgroep en de controlegroep beide representatief zijn voor de gehele experimentele dataset. Observaties in de treatmentgroep worden gewogen met $1/p$, en observaties in de controlegroep worden gewogen met $1/(1-p)$. De factor p is de kans op selectie in de treatmentgroep.</p>		

¹⁵ De gerapporteerde resultaten voor het eerste loopjaar wijken licht af (minder dan 0.01) van de resultaten die we eerder voor het eerste loopjaar hebben gerapporteerd (Paulle & De Ree (2020)). De reden hiervoor is een kleine update in de analysestrategie. De resultaten veranderen door deze update niet fundamenteel.

In lijn met de verschillen in de aanwezigheidscijfers tussen het eerste en tweede loopjaar, zien we in het eerste deel van de interventie (van *baseline* naar *midline* [halverwege het schooljaar]) een beperkt verschil in de effecten van de twee loopjaren. Het tweede loopjaar komt hoger uit, maar wanneer we specifiek kijken naar de trend vanaf de eerste *baseline* meting zijn de resultaten vergelijkbaar. In het tweede deel van het schooljaar zien we, zoals verwacht, een duidelijk verschil ontstaan. In het tweede loopjaar worden de effecten in het tweede deel van het schooljaar verder uitgebouwd. Uiteindelijk levert de analyse geschatte *treatment* effecten op in de orde van 0.6 populatiestandaarddeviaties, op basis van een volledig schooljaar Bridge HDT.

We zien geen duidelijke verschillen in het treatmenteffect afhankelijk van het startniveau. Het startniveau heeft dus geen invloed in de mate waarin deelnemers profiteren van de deelname aan de tutorsessies. In Figuur 7 presenteren we de groei in termen van Cito LVS Vaardigheidsscores Rekenen/Wiskunde, tussen baseline (halverwege groep 6) en endline (aan het eind van groep 7). Hierbij kijken we apart naar de treatmentgroep (blauw) en de controlegroep (rood), en maken we onderscheid naar het startniveau bij baseline (halverwege groep 6). Over de hele linie, van relatief hoog scorende leerlingen tot de relatief laag scorende leerlingen, zien we dat de deelnemers aan Bridge HDT (de *treatmentgroep*) meer vooruitgang boeken. Deze bevindingen komen overeen met eerdere bevindingen (De Ree en co-auteurs, 2021b).

Figuur 7. Voortgang in rekenprestaties tussen baseline (halverwege groep 6) en endline (eind van groep 7) voor treatmentgroep en controlegroep, afhankelijk van startniveau bij baseline.

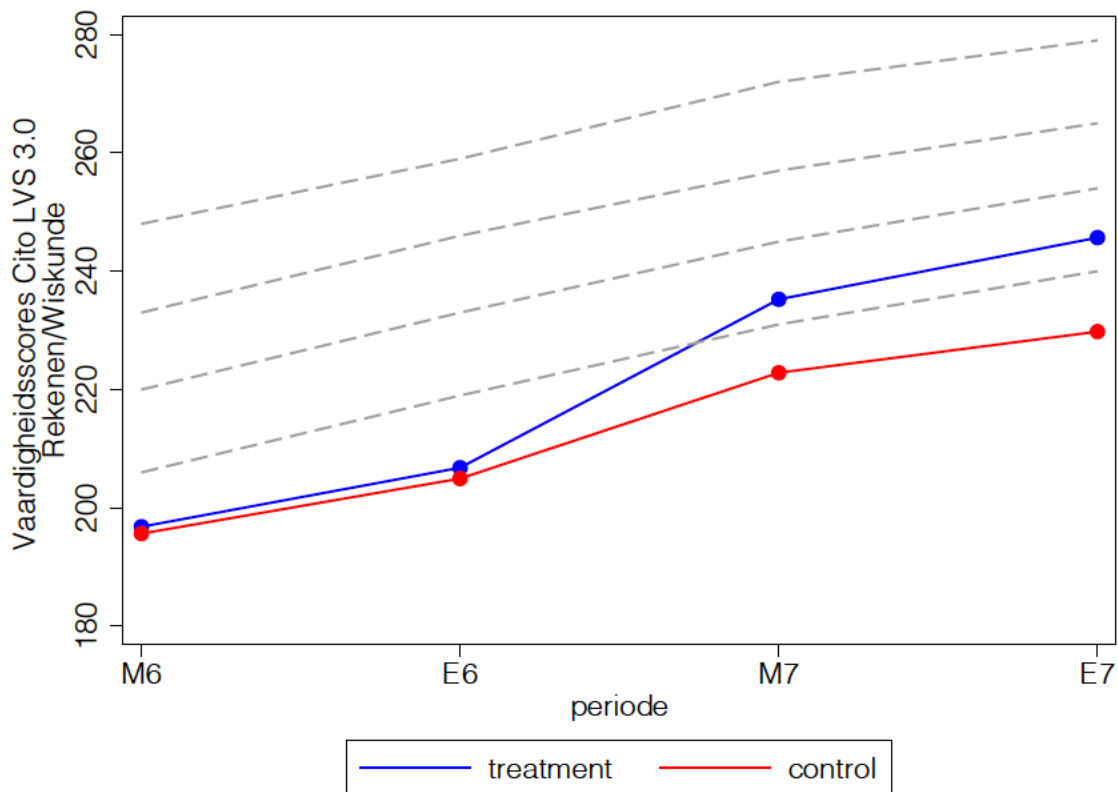


Afgemeten aan de beschikbare literatuur over effectieve interventies in het onderwijs, zijn de gemeten effecten zeer indrukwekkend (Kraft, 2020). Dergelijke effecten zijn voldoende om de verschillen in achterstanden tussen Zuidoost en de rest van Amsterdam, in termen van percentages laag scorende leerlingen, binnen een jaar in te lopen. Het onderzoek laat zien dat de huidige achterstanden er niet hoeven te zijn. Het bestaan van onderwijsachterstanden is een kwestie van politieke wil en voldoende financiële en personele middelen. Niet van onduidelijkheid over hoe dit probleem moet (of kan) worden aangepakt.

Om de resultaten van het tweede loopjaar ook op een andere manier inzichtelijk te maken volgen we de ontwikkelingen in de rekenprestaties van de treatmentgroep en de controlegroep, in de periode vanaf baseline (halverwege groep 6, of M6) tot aan de endline (aan het einde van groep 7, of E7). In Figuur 8 zetten we de ontwikkelingen van de twee groepen af tegen grenswaardes voor de functioneringsniveaus V, IV, III, II en I. Deelnemers aan Bridge HDT en aan het onderzoek zijn relatief laag scorende leerlingen, met rekenscores op of onder het landelijk gemiddelde (oftewel de functioneringsniveaus V, IV en III). De

deelnemers (treatmentgroep en controlegroep) scoren gemiddeld een ‘niveau V score’ bij baseline. Dit betekent een gemiddelde score ruim beneden het 20^e percentiel. Vanaf eind groep 6, na de start van Bridge HDT, begint de treatmentgroep gemiddeld beter te scoren. Uiteindelijk halen ze gemiddeld een ruime IV-score, zo rond het 30^e percentiel landelijk.

Figuur 8. Ontwikkelingen in rekenvaardigheid in termen van Cito LVS-R/W vaardigheidsscores (toetsgeneratie 3.0), van de treatmentgroep en controlegroep.



We vinden effecten van vergelijkbare grootte op een andere, specifiek voor dit onderzoek afgenomen niet-talige TOA-rekentoets. Bij de afnames van de TOA rekentoets zijn ook externe, door het onderzoeksteam aangestelde, observanten aanwezig geweest. De resultaten worden kort besproken in Appendix D.

In Tabel 2 presenteren we de effecten van Bridge HDT op de scores op de Cito LVS-toetsen Begrijpend lezen. We vinden geen meetbare (=significante) effecten. Hierbij moet worden opgemerkt dat de effecten voor de 2 loopjaren afzonderlijk niet heel precies zijn geschat, en dat we een *range* aan relevante effectgroottes niet kunnen uitsluiten. Echter, we zien ook de resultaten voor het eerste en tweede loopjaar verschillen in ‘teken’. In loopjaar 1 doet de controlegroep het iets beter, en in loopjaar 2 doet de treatmentgroep het iets beter. Deze verschillen zijn afzonderlijk niet van toeval te onderscheiden. Het samenvoegen van de data

van beide loopjaren levert een schatting van het treatmenteffect dicht in de buurt van nul (Tabel 1C). Tabel 1C laat zien dat leerlingen die deelnemen aan Bridge HDT niet minder zijn gaan lezen door het missen van een deel van het reguliere programma op school. Ze zijn echter ook niet beter gaan lezen.

Tabel 2. Effecten van Bridge HDT op de prestaties in het *Begrijpend Lezen* voor loopjaar 1 (1A) en loopjaar 2 (1B) en loopjaar 1 en 2 gecombineerd (1C) op basis van Cito LVS-toetsen Cito Begrijpend Lezen (in termen van populatiestandaarddeviaties)

	(1)	(2)
A: Loopjaar 1		
	Midline M7	Endline E7
Treatment effect	-0.13	-0.19
Standaardfout	(0.20)	(0.26)
Aantal observaties	[93]	[72]
B: Loopjaar 2		
	Midline M7	Endline E7
Treatment effect	0.18	0.16
Standaardfout	(0.19)	(0.17)
Aantal observaties	[89]	[92]
C: Loopjaar 1 en 2 samen		
Treatment effect	0.01	0.00
Standaardfout	(0.14)	(0.15)
Aantal observaties	[182]	[164]

Noten: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. De effecten van de interventie zijn geschat op basis van een regressiemodel. Als afhankelijke variabele gebruiken we de gestandaardiseerde toetscores. Als onafhankelijke variabelen gebruiken we 1) dummy/indicator variabelen voor alle *sampling blocks/strata*, 2) en een dummy/indicator variabele voor de treatment groep, 3) toetscores bij baseline, halverwege groep 6. De geschatte parameter op de indicator voor de treatmentgroep is gerapporteerd in deze tabel als schatting voor het treatment effect. Verder wegen we de data zodat de treatmentgroep en de controlegroep beide representatief zijn voor de gehele experimentele dataset. Observaties in de treatmentgroep worden gewogen met $1/p$, en observaties in de controlegroep worden gewogen met $1/(1-p)$. De factor p is de kans op selectie in de treatmentgroep.

Niet voor alle deelnemers aan het onderzoek hebben we toetsscores kunnen verzamelen. In Appendix B rapporteren we validiteitsanalyses, waarbij we onderzoeken of het missen van een toets wordt beïnvloed door deelname aan het HDT programma. We vinden bij deze analyses geen significante verschillen tussen treatmentgroep en de controlegroep. De validiteit van het onderzoek wordt door deze analyses dus niet in twijfel getrokken.

Tot slot merken we op dat de resultaten moeten worden geduid als kortetermijneffecten. Literatuur laat zien dat resultaten van succesvolle interventies in het onderwijs, na twee of drie jaar vaak weer (deels) zijn weggezakt. Cascio & Staiger (2012) bijvoorbeeld concluderen op basis van literatuuronderzoek dat *'test score impacts of early educational interventions almost universally "fade out" over time'*. Op basis van de literatuur is het dan ook aannemelijk dat ook deelnemers aan Bridge HDT na een aantal jaren tenminste een deel van hun voorsprong weer moeten inleveren. Om de verbeterde prestaties vast te houden kan worden gedacht aan *'light touch'* onderhoudsprogramma's, waarbij doelgroepleerlingen voor langere tijd worden ondersteund. Een onderzoek waarbij de leerlingen over langere termijn gevolgd worden, zou waardevolle inzichten kunnen geven in de mate van consolidatie van de effecten en nut van eventuele light touch onderhoudsprogramma.

4.3 Vervolgonderzoek

Dit is de tweede tussenrapportage van het onderzoek naar de effecten van Bridge HDT in het primair onderwijs in Amsterdam Zuidoost. Het onderzoek bestaat uit drie loopjaren. De eerste twee loopjaren gaan uit van dezelfde interventie, waarbij leerlingen vijf dagen in de week deelnemen aan de tutorsessies. Vanuit de gemeente is er altijd een wens geweest om ook goedkopere HDT modellen te testen. Op deze manier zou kunnen worden onderzocht of meer leerlingen bereikt kunnen worden met dezelfde financiële investering. In het derde loopjaar 2021/2022 wordt daarom een minder intensieve variant van Bridge HDT uitgevoerd, waarbij leerlingen drie, in plaats van vijf, dagen per week deelnemen aan de tutorsessies. Hierdoor kunnen meer scholen en meer leerlingen deelnemen op basis van een vergelijkbare investering. Met dit nieuwe model doen acht scholen mee, in plaats van vijf. Op elk van de

acht scholen kunnen 12 leerlingen deelnemen zodat het nieuwe model maximaal 96 leerlingen bereikt.

In de eindrapportage zullen we de drie loopjaren gezamenlijk bestuderen. Een belangrijke onderzoeksvraag voor de eindrapportage is of een proportionele afname in de intensiteit (en in de kosten) ook leidt tot een proportionele afname in de effectiviteit.

5. Conclusies en aanbevelingen

Het onderzoek laat zien dat met Bridge HDT zeer substantiële leerwinsten kunnen worden geboekt. Op basis van de in dit rapport geëvalueerde toepassing, verwachten we dat het breed opschalen van Bridge HDT in het basisonderwijs in Amsterdam Zuidoost kan leiden tot het volledig wegwerken van bepaalde relevante onderwijsachterstanden. Het gaat dan specifiek over het wegwerken van achterstanden in termen van percentages laag-scorende leerlingen tussen Amsterdam Zuidoost en de rest van de stad.

In dit rapport evalueren we specifiek de effectiviteit van Bridge HDT in het basisonderwijs in Amsterdam Zuidoost. Bridge HDT heeft echter ook, op basis van eerder onderzoek, laten zien dat vergelijkbare substantiële leerwinsten kunnen worden geboekt in het vmbo bijvoorbeeld (De Ree en co-auteurs, 2021a). Het totaal aan onderzoeksresultaten laat dus zien dat Bridge HDT niet sterk afhankelijk is van de context (basisschool of vmbo), en/of van het startniveau van deelnemers. Dit nieuwe rapport bevestigt daarom de resultaten van eerder onderzoek, en verstevigt het fundament voor verder opschalen van het Bridge HDT programma.

Gezien de hier gepresenteerde onderzoeksresultaten en de inzet van de stad om ongelijkheid in onderwijsuitkomsten te verminderen, adviseren we om het (eventuele) proces van opschalen ook specifiek te monitoren. Dit onderzoek geeft aan welke projectdoelen (in termen van verbeterde prestaties op school) redelijkerwijs kunnen worden gesteld.

Literatuurlijst

Cascio, E.U. & Staiger, D.O. (2012). *Knowledge, Tests and Fadeout in Educational Interventions*. NBER Working Paper 18038.

Centraal Planbureau. (2016). *Kansrijk onderwijsbeleid*. Den Haag.

De Ree, Joppe, Mario A. Maggioni, Bowen Paulle, Domenico Rossignoli and Dawid Walentek (2021a). *High dosage tutoring in pre-vocational secondary education: Experimental evidence from Amsterdam*, SocArxiv, <https://osf.io/preprints/socarxiv/r56um/>

De Ree, Joppe, Mario A. Maggioni, Bowen Paulle, Domenico Rossignoli, Nienke Ruijs and Dawid Walentek (2021b). *Closing the income-achievement gap? Experimental evidence from high-dosage tutoring in Dutch primary education*. SocArxiv, <https://osf.io/preprints/socarxiv/qepc2/>

Engzell, P., Freyd, A. & Verhagen, M. (2021). *Learning loss due to school closures during the COVID-19 pandemic*. PNAS 118 (17)

Guryan, Jonathan, Jens Ludwig, Monica P. Bhatt, Philip J. Cook, Jonathan M.V. Davis, Kenneth Dodge, George Farkas, Roland G. Fryer Jr, Susan Mayer, Harold Pollack & Laurence Steinberg (2021). *Not Too Late: Improving Academic Outcomes Among Adolescents*. National Bureau of Economic Research (No. 28531).

Kraft, Matthew A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher* 49 (4)

Kraft, Matthew A. & Grace T. Falken (2021). A Blueprint for Scaling Tutoring and Mentoring Across Public Schools. *AERA Open* 7(1)

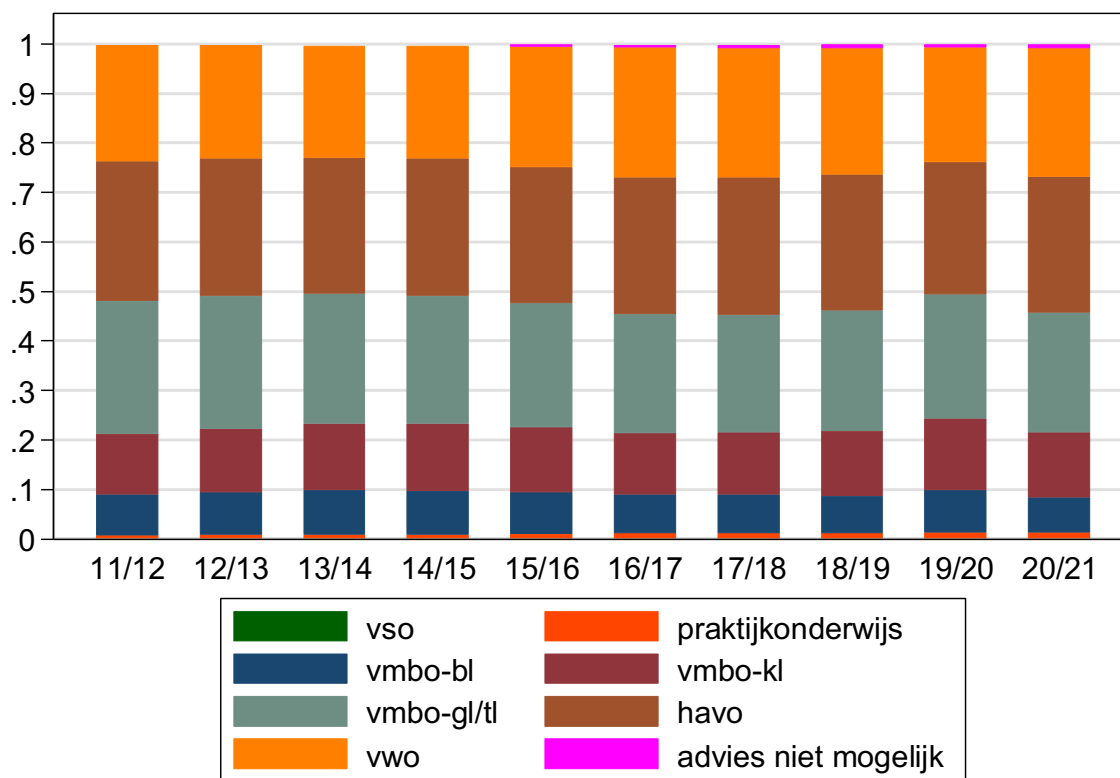
Onderzoek, Informatie en Statistiek (2020). *De Staat van het Amsterdamse Primair Onderwijs*. Gemeente Amsterdam.

Appendices

Appendix A: Ontwikkelingen in middelbare schooladviezen

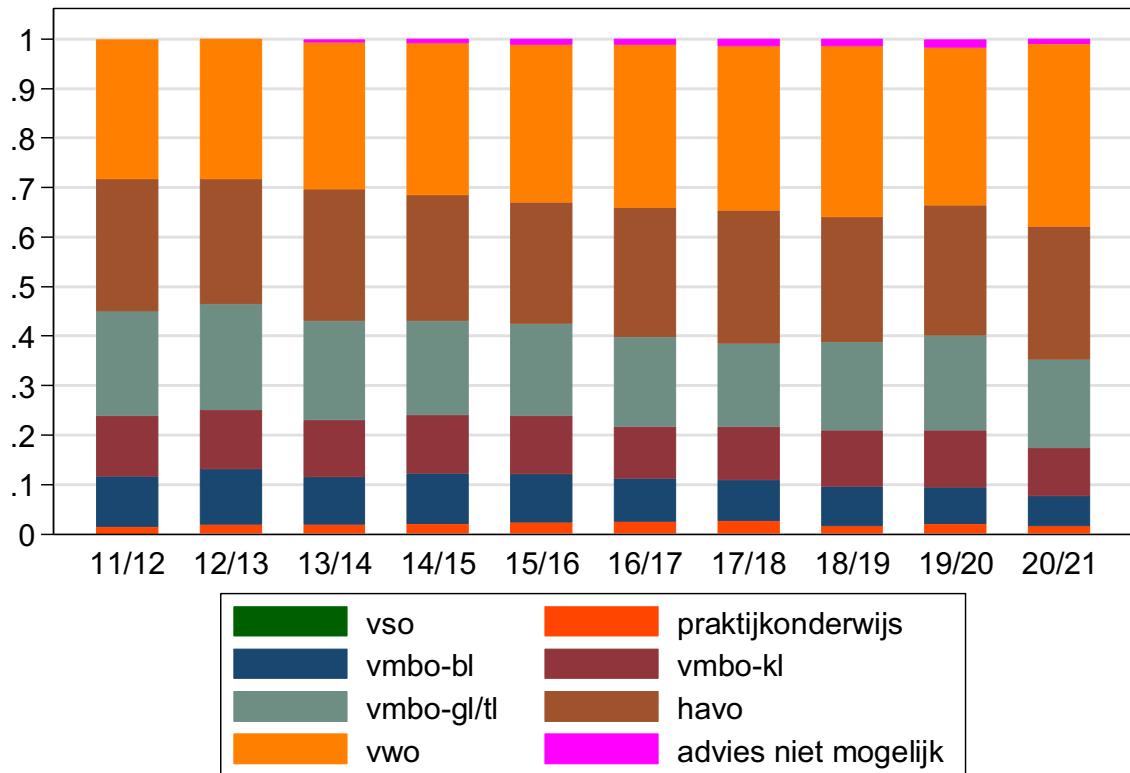
Figuur A.1 presenteert de ontwikkelingen van de middelbare schooladviezen in Nederland. De Figuur laat zien dat de adviezen door de tijd heen niet of nauwelijks veranderen.

Figuur A.1: Ontwikkelingen middelbare schooladviezen in groep 8, door de tijd heen. De Figuur is aangemaakt op basis van publiek beschikbare data op schoolniveau van de Dienst Uitvoering Onderwijs (DUO). De Figuur heeft betrekking op Nederland als geheel.



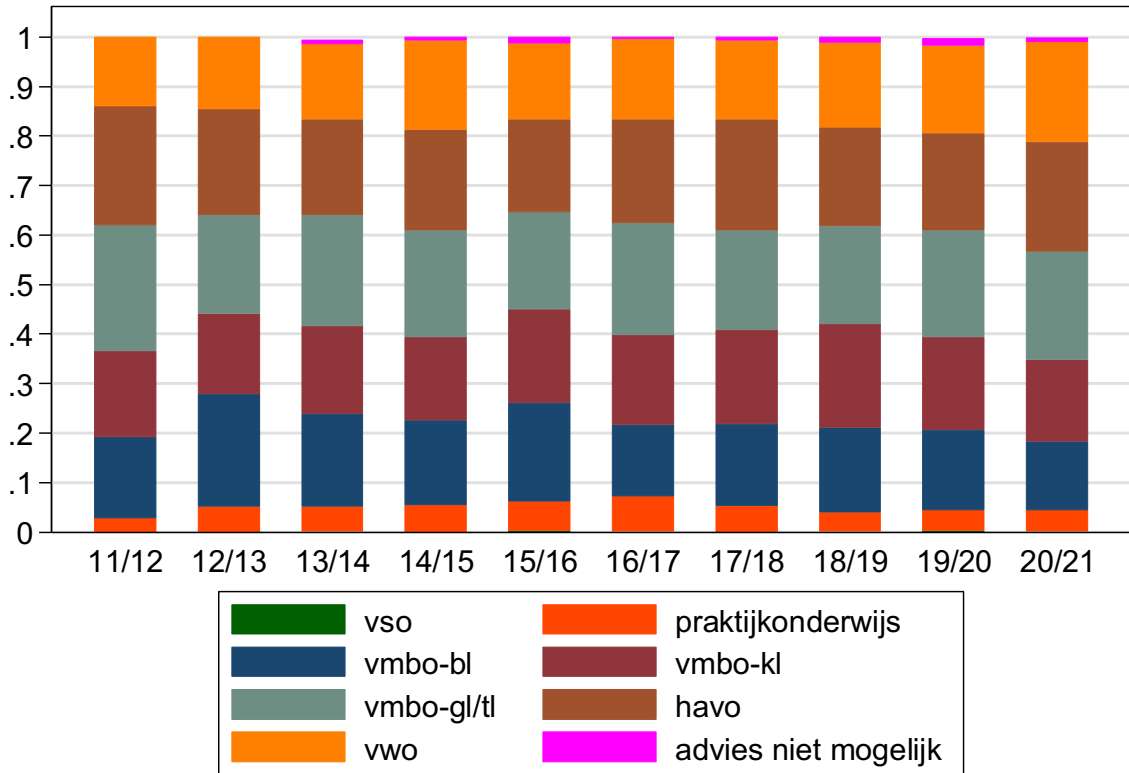
Nederland

Figuur A.2: Ontwikkelingen middelbare schooladviezen in groep 8, door de tijd heen. De Figuur is aangemaakt op basis van publiek beschikbare gegevens op schoolniveau van Dienst Uitvoering Onderwijs (DUO). De Figuur heeft betrekking op Amsterdam als geheel.



Amsterdam

Figuur A.3: Ontwikkelingen middelbare schooladviezen in groep 8, door de tijd heen. De Figuur is aangemaakt op basis van publiek beschikbare gegevens op schoolniveau van Dienst Uitvoering Onderwijs (DUO). De Figuur heeft betrekking op Amsterdam stadsdeel Zuidoost.



Amsterdam Zuidoost

Appendix B: Operationalisering van statistische analyses en validiteitsanalyses

De uitkomstmaten op basis waarvan we de effecten van Bridge HDT meten zijn vaardigheidsscores, gebaseerd op de Cito LVS-toetsen Rekenen/Wiskunde en Begrijpend Lezen. De vaardigheidsscores zijn vervolgens geschaald in afwijking van een schatting van het landelijk gemiddelde van de betreffende leeftijdsgroep en gedeeld door een schatting van de populatiestandaarddeviaties van de betreffende leeftijdsgroep. Het landelijk gemiddelde schatten we met het middelpunt van het 20e en 80e percentiel. De populatiestandaarddeviatie schatten we door het verschil van het 20e en het 80e percentiel te delen door 1.683. Het 20e en 80e percentiel zijn respectievelijk de grenzen tussen de functioneringsniveaus V en IV, en, II en I. De factor 1.683 is de afstand in standaarddeviaties tussen het 20e en 80e percentiel in de normale verdeling.

Vervolgens maken we indicator variabelen (of dummy variabelen). De *treatment* indicator is 1 voor leerlingen in de *treatmentgroep* en 0 voor leerlingen in de controlegroep. De *block* indicatoren zijn 1 voor leerlingen in een bepaald *sampling block* of stratum, en anders 0.

De effecten van Bridge HDT worden geschat met een regressiemodel, met de gestandaardiseerde toetsscores als afhankelijke variabelen. De onafhankelijke variabelen zijn de *treatment* indicator, de volledige set aan *sampling block* of stratum indicatoren, en baseline toetsscores (die halverwege groep 6 zijn behaald). De geschatte parameter bij de *treatment* indicator is de schatting van het effect van Bridge HDT. De gerapporteerde standaardfouten bij deze schattingen zijn robuust tegen heteroscedasticiteit in de foutterm.

De randomisatie zorgt voor vergelijkbaarheid van de *treatment*- en controlegroep. Echter, omdat we niet altijd van alle 98 leerlingen alle toetsscores hebben kunnen verzamelen, is er een mogelijkheid dat de *treatment*- en controle observaties van de geobserveerde dataset niet meer vergelijkbaar zijn, in de gebruikelijke zin. We kunnen toetsen voor deze mogelijke verminderde vergelijkbaarheid door een indicator voor een missende toetsscore aan te maken, en deze te regresseren op de *treatment* indicator, de *sampling block* indicatoren. We vinden hierbij geen evidentie voor verminderde vergelijkbaarheid. Het wel of niet kunnen meten van een toetsscore wordt dus niet significant beïnvloed door deelname aan Bridge HDT. Er is op basis van deze analyses geen reden om te twijfelen aan de validiteit van het onderzoek.

Tabel B.1. Effecten van Bridge HDT op niet deelnemen aan de Cito LVS Rekenen/Wiskunde

	(1)	(2)
	Midline M7	Endline E7
Treatment effect	-0.03	-0.02
Standaardfout	(0.04)	(0.04)
Gemiddelde in controleconditie	<0.10>	<0.08>
Aantal observaties	[98]	[98]
<p>Noten: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. De effecten zijn geschat op basis van een regressiemodel. Als afhankelijke variabele gebruiken we een dummy variabele. De dummy variabele is 1 wanneer de rekentoetsscore niet wordt geobserveerd, en 0 wanneer de rekentoetsscore wel is geobserveerd. Als onafhankelijke variabelen gebruiken we 1) dummy/indicator variabelen voor alle <i>sampling blocks/strata</i>, 2) en een dummy/indicator variabele voor de treatment groep. Verder wegen we de data zodat de treatmentgroep en de controlegroep beide representatief zijn voor de gehele experimentele dataset. Observaties in de treatmentgroep worden gewogen met $1/p$, en observaties in de controlegroep worden gewogen met $1/(1-p)$. De factor p is de kans op selectie in de treatmentgroep.</p>		

Appendix C: Beschrijvende analyses op basis van CBS Microdata

We starten de beschrijvende analyse met alle leerlingen in groep 8 in Nederland, van 2013/2014 tot 2018/2019. Het aantal scholen met groep 8-leerlingen in de data daalt licht in deze periode, van 6761 naar 6296. Het aantal scholen dat gebruikmaakt van de Cito-toets daalt harder, van 5700 naar 3229. We hebben toegang tot twee bronnen voor de resultaten op de Cito-toets, de CITOTAB en de inschrijvingen op de basisschool WPOINSCHRTAB. Beide bronnen overlappen voor het grootste deel. We maken nu gebruik van de CITOTAB, omdat deze bron afzonderlijk de scores op de reken- en taalonderdelen rapporteert.

In Amsterdam zijn ongeveer 200 basisscholen (met groep 8-leerlingen). In 2013/2014 maken bijna al deze scholen gebruik van de Cito-eindtoets. Ook in Amsterdam daalt het gebruik van de Cito-eindtoets, maar minder snel dan landelijk. In 2018/2019 maakt nog 146 van de 200 scholen gebruik van de Cito-eindtoets.

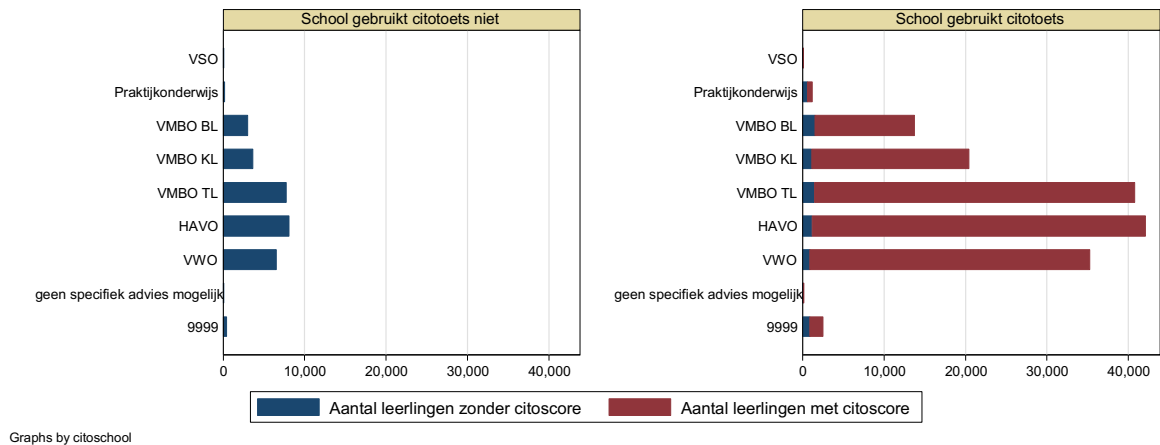
We nemen de resultaten van het schooljaar 2013/2014 als basis voor de beschrijvende analyses. In 2013/2014 had de Cito-toets nog een duidelijkere rol in de aanloop naar het de vorming van het middelbare schooladvies. Vanaf 2014/2015 is de rol van de Cito-toets veranderd en is het niet voor alle leerlingen meer even belangrijk om op de toets erg 'hun best' te doen. (In Appendix A laten we zien dat de adviseringen door de tijd heen stabiel zijn. Het is om deze reden aannemelijk dat de prestaties op taal en rekenen ook niet fundamenteel zijn verschoven.)

Niet alle leerlingen maken de Cito-toets, of een andere eindtoets. Dit gaat deels om leerlingen van scholen die überhaupt geen gebruikmaken van de toets, maar ook (soms) om leerlingen op scholen die wel van de Cito-toets gebruik maken. Met name leerlingen met een advies voor praktijkonderwijs doen regelmatig niet mee met de eindtoets. Voor een deel van deze groep wordt deze toets te moeilijk geacht.

Figuur C.1 laat zien dat er landelijk vrij veel groep 8-leerlingen op scholen zitten waar de Cito-toets (in 2013/2014) niet werd gebruikt. Als de Cito-toets op school wordt gebruikt, maakt maar een kleine minderheid de toets niet. Voor leerlingen met een advies voor praktijkonderwijs, maakt een groter percentage de toets niet (ongeveer de helft). Echter, relatief weinig leerlingen krijgen een advies voor praktijkonderwijs. Verder zijn de leerlingen

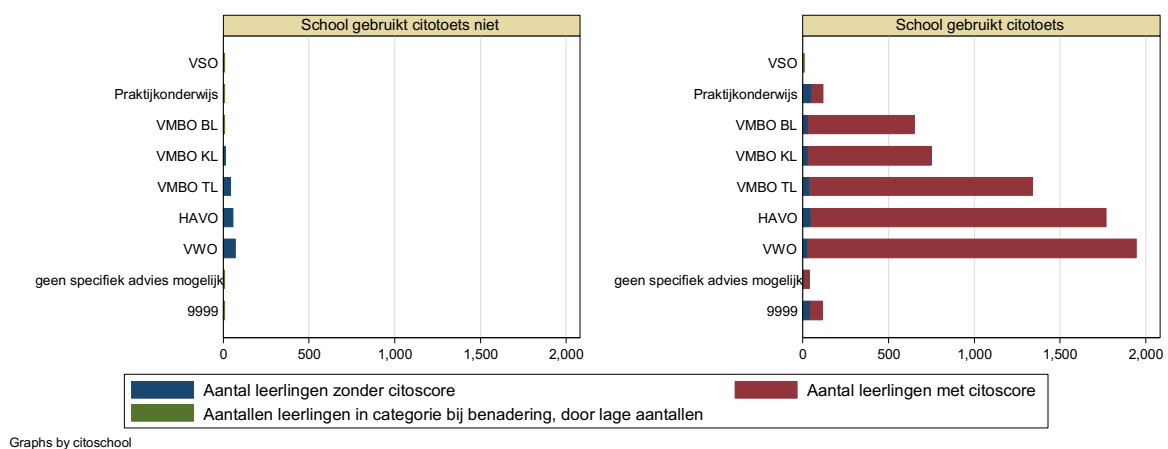
zonder toetsresultaten vrij gelijk verdeeld over de verschillende adviezen. Naar ons inzicht zijn de geobserveerde Cito-scores dus voldoende representatief voor de leerlingenpopulatie als geheel, met uitzondering van scholen die de Cito-toets niet gebruiken en leerlingen met een advies voor praktijkonderwijs.

Figuur C.1 Aantallen leerlingen met (rood) en zonder (blauw) Cito-score, bij scholen die de Cito-toets niet gebruiken (links) en bij scholen die de Cito-toets wel gebruiken (rechts). De Figuur heeft betrekking op Nederland als geheel.



Figuur C.2 presenteert de resultaten voor Amsterdam. Zoals eerder opgemerkt, maken in 2013/2014 bijna alle scholen gebruik van de Cito-toets. Ook hier zien we dat de non-respons klein is, met uitzondering van leerlingen met een advies voor praktijkonderwijs.

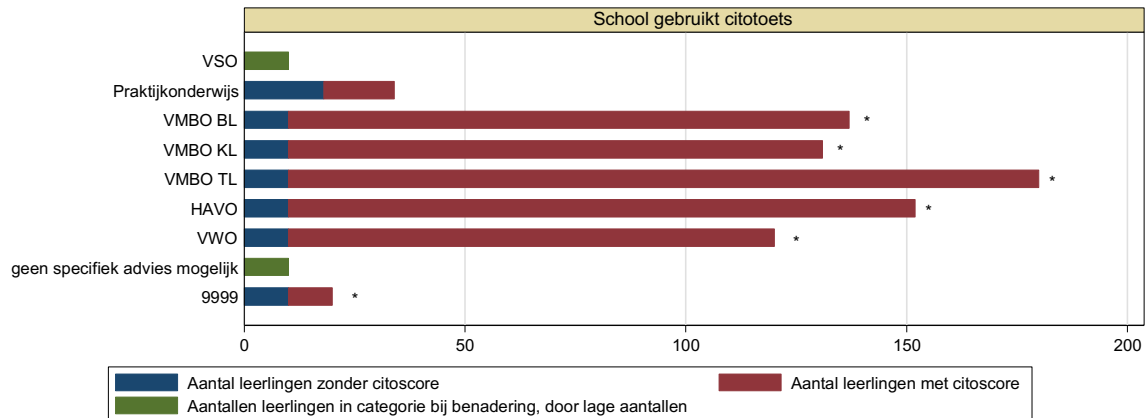
Figuur C.2 Aantallen leerlingen met (rood) en zonder (blauw) Cito-score, bij scholen die de Cito-toets niet gebruiken (links) en bij scholen die de Cito-toets wel gebruiken (rechts). Figuur heeft betrekking op Amsterdam als geheel.



Figuur C.3 presenteert de resultaten voor het stadsdeel Amsterdam Zuidoost. In 2013/2014 maken minder dan 3 (0, 1 of 2) scholen geen gebruik van de Cito-toets. Het CBS legt beperkingen op in gevallen van kleine aantallen. We kunnen dus op basis van de CBS Microdata niet precies rapporteren hoeveel scholen in Amsterdam Zuidoost in 2013/2014 geen gebruik maken van de Cito-toets. Verder kunnen we ook geen resultaten rapporteren over leerlingen op scholen in Amsterdam Zuidoost die de Cito-toets niet gebruiken. Zoals gezegd is het aantal scholen dat in 2013/2014 de Cito-toets niet gebruikt een minderheid: minder dan 3 op een totaal van 27. Ook voor de meerderheid van de scholen die wel van de Cito-toets gebruik maken, kunnen we niet alle cijfers volledig rapporteren. De staven in Figuur C.3 gemarkeerd met een * zijn benaderingen.

De leerlingen met advies voor praktijkonderwijs maken ook in Amsterdam Zuidoost relatief vaak de Cito-toets niet. In Zuidoost is het aandeel leerlingen met een advies voor praktijkonderwijs ook veel groter dan de aandelen in Amsterdam als geheel bijvoorbeeld.

Figuur C.3 Aantallen leerlingen met (rood) en zonder (blauw) Cito-score, bij scholen die de Cito-toets gebruiken. Figuur heeft betrekking op stadsdeel Amsterdam Zuidoost als geheel.



Graphs by citoschool

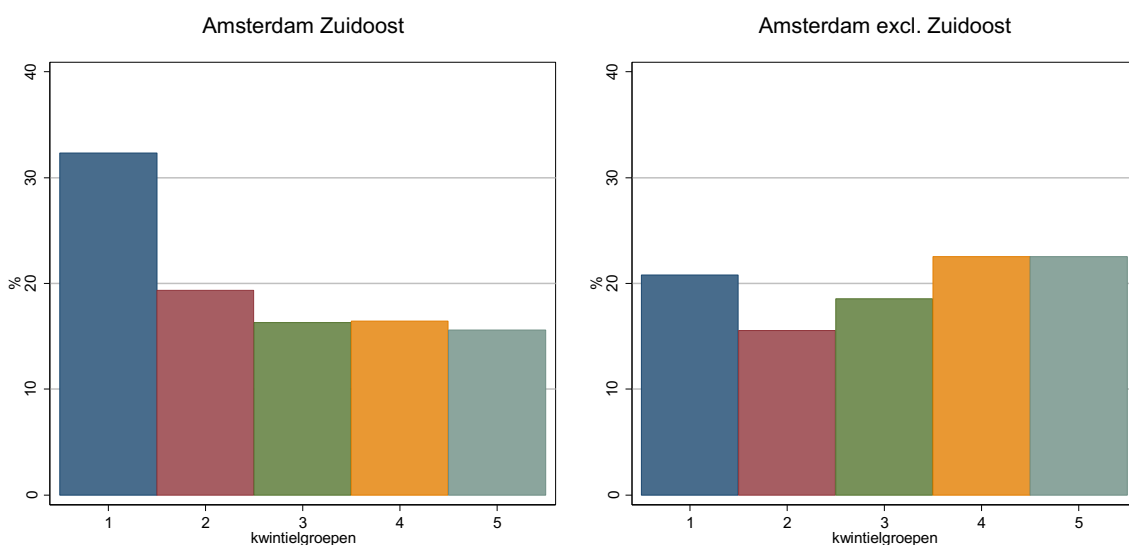
Op basis van deze beschrijvingen concluderen we dat de data van de Cito-scores vrij compleet zijn en dat een vergelijking tussen bijvoorbeeld de prestaties in Zuidoost en de rest van Amsterdam goed mogelijk is.

De volgende figuur laat zien hoe de prestaties van leerlingen in Amsterdam Zuidoost [linker paneel] zich verhouden tot prestaties in de rest van Amsterdam Amsterdam exclusief

stadsdeel Zuidoost) [rechter paneel]. En hoe deze prestaties zich verhouden tot de rekenprestaties landelijk. Voor deze figuur rangschikken we eerst alle groep 8-leerlingen in Nederland (in 2013/2014) op basis van de score op het rekenonderdeel van de Cito-toets. Op basis van deze *ranking* maken we vijf groepen van gelijke grootte, de kwintielen, van de laagst scorende 20% tot de hoogst scorende 20%. Vervolgens bekijken we in deze figuren in welke verhouding leerlingen in Amsterdam als geheel, of in Amsterdam Zuidoost zich in deze vijf kwintielen bevinden. Voor Amsterdam exclusief stadsdeel Zuidoost zien we dat de rekenprestaties van leerlingen vrij sterk overeenkomen met de leerprestaties van alle leerlingen in Nederland. We zien echter een lichte ‘U-vorm’. Er zijn in Amsterdam relatief iets meer laag scorende leerlingen en iets meer hoog scorende leerlingen dan in de rest van Nederland, en relatief iets minder leerlingen in het midden van de verdeling.

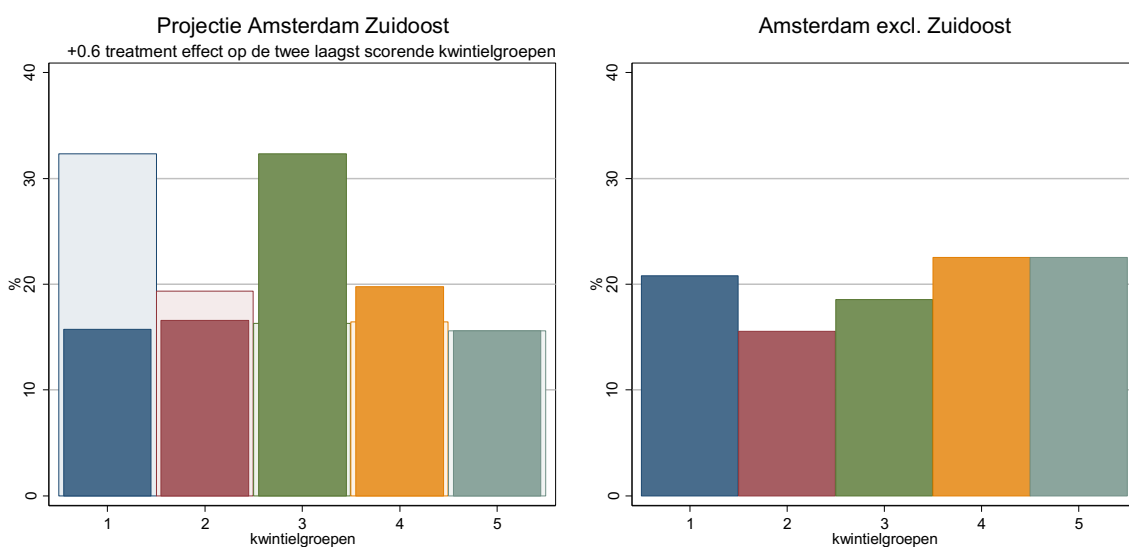
In stadsdeel Amsterdam Zuidoost zien we een ander beeld. In Zuidoost zien we een sterke oververtegenwoordiging van leerlingen in de ‘laagste’ categorie, het eerste kwintiel, oftewel de laagst scorende 20% landelijk. Dit is de categorie leerlingen waarvan een meerderheid met een vmbo-bl of vmbo-kl advies uitstroomt. In Zuidoost zijn ook leerlingen met hoge toetsscores ondervertegenwoordigd. In Zuidoost zit ongeveer 16% van de leerlingen in het 5^e kwintiel, dat wil zeggen, leerlingen die rekenen op een niveau dat past bij een vwo-advies. Landelijk is dit per definitie 20% (per definitie) en in de rest van Amsterdam (Amsterdam, exclusief groep 8 leerlingen uit het stadsdeel Amsterdam Zuidoost) is dit ongeveer 23%.

Figuur C.4 Rekenprestaties op basis van rekenonderdeel van de Cito-eindtoets (schooljaar 2013/2014) in Amsterdam Zuidoost [links] de rest van Amsterdam (exclusief Zuidoost) [rechts], in vergelijking met de rekenprestaties landelijk.



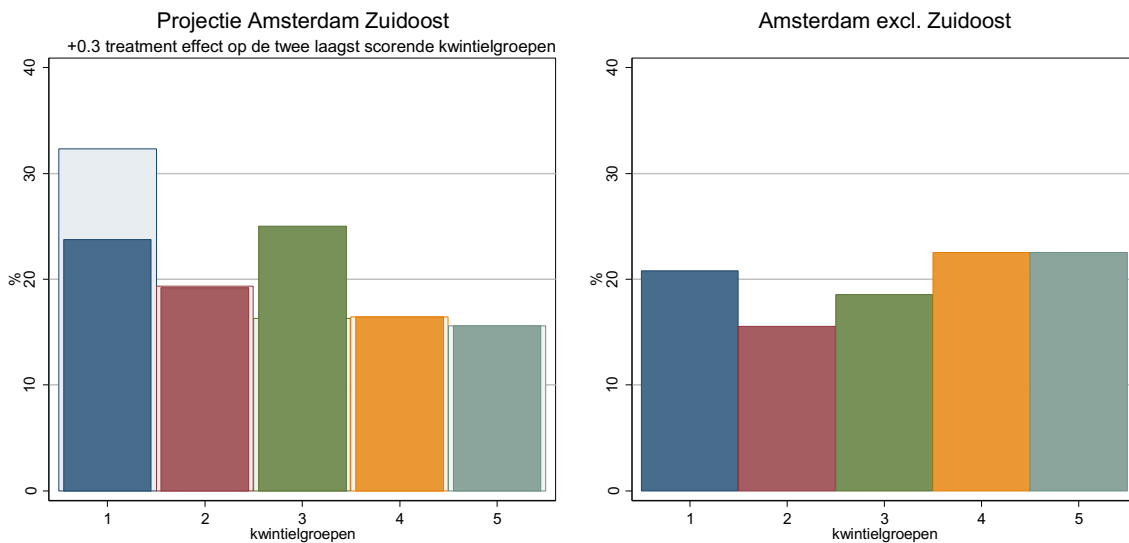
Dit onderzoek laat zien dat leerlingen gemiddeld veel baat hebben bij het HDT-programma van The Bridge. In Figuur C.5 maken we een projectie van de effecten van een Zuidoost-brede uitrol van het Bridge HDT programma in de toepassing waarbij de eerste twee kwintielen, de laagst scorende 40% landelijk, deelnemen. We nemen hiervoor als startpunt de rekenscores van leerlingen, en tellen hier voor de doelgroep het gevonden *treatment effect* bij op. Leerlingen in Amsterdam Zuidoost in de laagst scorende twee kwintielen krijgen er dus 0.6 populatiestandaarddeviaties bij. Een *treatment effect* van 0.6 populatiestandaarddeviaties voor alle leerlingen in de laagst scorende twee categorieën in Amsterdam Zuidoost, zorgt voor een substantiële verschuiving van het beeld. De achterstand in termen van het aandeel laag scorende leerlingen is in deze projectie geheel ingelopen en zelfs gekanteld in het voordeel van Amsterdam Zuidoost. We merken hierbij op dat Bridge HDT zich in de huidige toepassing richt op de laagst scorende 40% of soms 60% (zoals bij het in dit rapport geëvalueerde programma) van de leerlingen. Logischerwijs zien we in de huidige toepassing dus geen verschuivingen in het relatief lage aandeel hoog scorende leerlingen in Amsterdam Zuidoost, de leerlingen die rekenen op vwo-niveau.

Figuur C.5 [links] Geprojecteerde rekenprestaties op basis van het rekenonderdeel van de Cito-eindtoets in het schooljaar 2013/2014. We projecteren *treatment* effecten van 0.6 populatiestandaarddeviaties op de twee laagst scorende kwintielen in Amsterdam Zuidoost. De resultaten van de projecties zijn donker gekleurd. De werkelijke situatie in 2013/2014 is licht gekleurd in de achtergrond. [rechts] Rekenprestaties op basis van het rekenonderdeel van de Cito-eindtoets in Amsterdam exclusief stadsdeel Amsterdam Zuidoost.



In Figuur C.6 laten we zien dat ook met kleinere effecten een groot deel van de achterstand in het aandeel laag scorende leerlingen kan worden weggewerkt.

Figuur C.6 [links] Geprojecteerde rekenprestaties op basis van het rekenonderdeel van de Cito-eindtoets in het schooljaar 2013/2014. We projecteren *treatment* effecten van 0.3 populatiestandaarddeviaties op de twee laagst scorende kwintielen in Amsterdam Zuidoost. De resultaten van de projecties zijn donker gekleurd. De werkelijke situatie in 2013/2014 is licht gekleurd in de achtergrond. [rechts] Rekenprestaties op basis van het rekenonderdeel van de Cito-eindtoets in Amsterdam exclusief stadsdeel Amsterdam Zuidoost.



Appendix D: Resultaten op basis van de TOA niet-talige rekentoets

In Tabel D.1 rapporteren we effectmetingen op basis van de TOA niet-talige rekentoets. De geschatte effecten zijn positief en zeer significant. De effecten lijken echter iets kleiner te zijn als de effecten die we in kolom 2 van Tabel 1 hebben gerapporteerd (0.65 in Tabel 1 en 0.49 in Tabel F.1). Echter, bij de analyses op basis van de TOA rekentoets passen we een andere schaling toe, waardoor de resultaten van Tabel 1 en Tabel D.1 niet 1-op-1 vergelijkbaar zijn. Om de resultaten toch te kunnen vergelijken passen we in Tabel D.1 en D.2 dezelfde schaling toe, waardoor we de resultaten op basis TOA rekentoets en de Cito LVS Rekenen/Wiskundetoets toch naast elkaar kunnen zetten. Op basis van Tabel D.1 en D.2

concluderen we dat de gebruikte rekentoetsen weinig verschil maken, bij het vergelijken van de effectgroottes.

Tabel D.1. Effecten van Bridge HDT op de rekenprestaties voor loopjaar 2 (1B) op basis van TOA niet-talige rekentoets (in termen van controlegroep standaarddeviaties, *Glass' Delta*)

	(1)	(2)
	Midline M7	Endline E7
Treatment effect	-	0.43***
Standaardfout	-	(0.15)
Aantal observaties	-	[88]
<p>Noten: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. De effecten van de interventie zijn geschat op basis van een regressiemodel. Als afhankelijke variabele gebruiken we de gestandaardiseerde TOA rekenscores, waarbij de standaardiseren door de ruwe scores te verminderen met de gemiddelde score in de controlegroep en te delen voor de standaarddeviatie in de controlegroep. Als onafhankelijke variabelen gebruiken we 1) dummy/indicator variabelen voor alle <i>sampling blocks</i>/strata, 2) en een dummy/indicator variabele voor de treatment groep, 3) toetsscores bij baseline, halverwege groep 6. De geschatte parameter op de indicator voor de treatmentgroep is gerapporteerd in deze tabel als schatting voor het treatment effect. Verder wegen we de data zodat de treatmentgroep en de controlegroep beide representatief zijn voor de gehele experimentele dataset. Observaties in de treatmentgroep worden gewogen met $1/p$, en observaties in de controlegroep worden gewogen met $1/(1-p)$. De factor p is de kans op selectie in de treatmentgroep.</p>		

Tabel D.2. Effecten van Bridge HDT op de rekenprestaties voor loopjaar 2 (1B) op basis van Cito LVS Rekenen/Wiskunde (in termen van controlegroep standaarddeviaties, *Glass' Delta*)

	(1)	(2)
	Midline M7	Endline E7
Treatment effect	0.32***	0.49***
Standaardfout	(0.12)	(0.14)
Aantal observaties	[90]	[92]
<p>Noten: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. De effecten van de interventie zijn geschat op basis van een regressiemodel. Als afhankelijke variabele gebruiken we gestandaardiseerde Cito LVS scores op het onderdeel Rekenen/Wiskunde, waarbij de standaardiseren door de ruwe scores te verminderen met de gemiddelde score in de controlegroep en te delen voor de standaarddeviatie in de controlegroep. Als onafhankelijke variabelen gebruiken we 1) dummy/indicator variabelen voor alle <i>sampling blocks</i>/strata, 2) en een dummy/indicator variabele voor de treatment groep, 3) toetsscores bij baseline, halverwege groep 6. De geschatte parameter op de indicator voor de treatmentgroep is gerapporteerd in deze tabel als schatting voor het treatment effect. Verder wegen we de data zodat de treatmentgroep en de controlegroep beide</p>		

representatief zijn voor de gehele experimentele dataset. Observaties in de treatmentgroep worden gewogen met $1/p$, en observaties in de controlegroep worden gewogen met $1/(1-p)$. De factor p is de kans op selectie in de treatmentgroep.