

Effectiviteit van facultatief aansluitonderwijs wiskunde in de transitie van voortgezet naar hoger onderwijs

D. T. Tempelaar, B. Rienties, W. Kaper, B. Giesbers, L. van Gastel, E. van de Vrie,
H. van der Kooij, en H. Cuypers

Samenvatting

Wat is het effect van vrijwillig aansluitonderwijs, zoals het remediërend wiskundeonderwijs dat op veel instellingen voor hoger onderwijs wordt gegeven ter verbetering van de doorstroom? Bepaling van de doelmatigheid van dit onderwijs wordt bemoeilijkt door selectie-effecten die kunnen optreden: deelnemers aan aansluitonderwijs zullen in het algemeen andere achtergrondkenmerken hebben dan niet-deelnemers. In deze empirische studie wordt gebruik gemaakt van een breed spectrum van achtergrondkenmerken om een gecorrigeerd behandelingseffect te berekenen op basis van de *propensityscore*-methode. Analyse van vijf cohorten met 4.500 studenten geeft aan dat selectie-effecten inderdaad optreden, maar dat na correctie daarvoor een substantieel behandelingseffect resteert, in de orde van grootte van 50% van het effect veroorzaakt door het volgen van wiskunde op B-niveau in de vooropleiding. Een resultaat dat aangeeft dat aansluitproblemen veroorzaakt door internationalisering van hoger onderwijs op adequate wijze kunnen worden aangepakt.

1 Inleiding

Deze bijdrage richt zich op een vorm van onderwijs waarvoor verschillende aanduidingen worden gebruikt: aansluitonderwijs, bijspijkeronderwijs, opfrisonderwijs, of algemener, remediërend onderwijs. Het betreft onderwijs gericht op het verbeteren van aansluiting tussen voortgezet en hoger onderwijs en het bevorderen van succeschansen in het eerste jaar van hoger onderwijs, overeenkomstig adviezen van de Onderwijsraad (2006, 2007, 2008). Datering van deze adviezen geeft aan dat Nederlandse aandacht voor aansluitonderwijs buiten het 'open onderwijs' recent is. Het eerste landelijke project op dit gebied,

het door Stichting SURF ondersteunde WebSpijkeren I, dateert van 2004. Samen met vervolgpriject WebSpijkeren II had dit project ten doel te verkennen welke didactische scenario's geschikt zijn voor remediërend onderwijs ter voorbereiding van een universitaire studie (Rienties, Dijkstra, Rehm, Tempelaar, & Blok, 2005; Tempelaar, 2007; Wieland et al., 2006). Een opschaling van deze verkenningen van aansluitonderwijs naar genoeg alle Nederlandse universiteiten en een aantal hbo-instellingen vond plaats in het Nationaal Actieplan e-Learning van Stichting SURF. Deze Nationale Kennisbank Basisvaardigheden Wiskunde-projecten, NKBW I, II (2007-2010), richtten zich specifiek op wiskunde, maar gingen verder dan het onderzoeken van bruikbare scenario's van aansluitonderwijs. Daarenboven beoogden ze een stelsel van landelijke door voortgezet en hoger onderwijs als maatgevend beoordeelde diagnostische instaptoetsen te ontwikkelen en een wiskunde monitor samen te stellen op basis van uitkomsten van die landelijke toetsen.

Een belangrijke rol voor de instelling van aansluitonderwijs wiskunde is weggelegd door de IOWO WO-Instroommonitor, die sinds 1998 jaarlijks verschijnt en gegevens bevat over hoe studenten de aansluiting tussen voortgezet en hoger onderwijs ervaren. Onderdeel van die monitor is het percentage van studenten die ontevreden zijn over de aansluiting van de afzonderlijke vo-vakken met het eerste jaar op de universiteit, per sector van studie. Voor de sector economie liggen die aansluitproblemen volgens studentpercepties vooral bij een tweetal vakken: wiskunde en Engels. Het percentage over de aansluiting van wiskunde ontevreden studenten fluctueert vrij stabiel rond 40% en overtreft ruimschoots dat van alle andere vakken. Analyses over waarom wiskunde zo'n groot aandeel heeft in aansluitproblemen die studenten ervaren zijn terug te vinden in Tweede Fase Adviespunt (2005). Daar wordt op basis van

onderzoek onder ho-opleiders geconcludeerd dat:

Wat de rekenvaardigheden betreft constateren vooral de natuurwetenschappelijke, technische en economische opleidingen toenemende problemen. Die problemen zijn niet zo zeer afleesbaar aan de examenprogramma's voor de verschillende wiskundeprogramma's, het gaat er meer om dat de studenten basisvaardigheden niet goed beheersen en niet zelf goed kunnen bepalen waar ze welke kennis en vaardigheden voor kunnen inzetten. Men heeft het gevoel dat de vaardigheden niet echt doorleefd zijn. Het is wel aan de orde geweest, maar men heeft er niet genoeg mee geoefend, het geleerde nooit hoeven toepassen in andere situaties etc. Al met al verwoordt men dat binnenkomende studenten meer en meer geleerd hebben 'trucjes' toe te passen, zonder dat ze werkelijk beseft hebben van wat men werkelijk aan het doen is. De basisvaardigheden die ontbreken lijken overigens vaak tot het programma van de onderbouw te behoren (Tweede Fase Adviespunt, 2005, p. 89).

In (continentaal) Europa neemt Nederland op het gebied van aansluitonderwijs met computergebaseerde leeromgevingen een leidende positie in. Europese initiatieven, zoals de EU-projecten M.A.S.T.E.R., MathBridge en S.T.E.P., zijn van recente datum en bouwen deels voort op eerder Nederlands onderzoek (Rienties et al., in druk). Die voortrekkersrol is op zich niet verwonderlijk: in de twee aspecten waarin voor aansluitonderwijs een functie is weggelegd, loopt het Nederlandse onderwijs voorop in Europese ontwikkelingen. Het eerste aspect is de typisch Europese invalshoek: hoger onderwijs internationaliseert in hoog tempo, met grote consequenties (sommige opleidingen in 'grenslanduniversiteiten' zoals deze Maastrichtse casus, kennen een internationale instroom die is opgelopen tot 75%). Europese systemen van voortgezet onderwijs, zelfs in buurlanden als Nederland, Duitsland en België, verschillen aanzienlijk, met als gevolg een grote diversiteit aan instroomkennis. Daarnaast is er een tweede, nationaal aspect, dat in de hierboven geciteerde verkenningen en adviezen van de Onderwijsraad vooral wordt uitgewerkt, en

dat te maken heeft met de onvolkomen aansluiting van ketens in het onderwijs die wel verondersteld worden elkaar naadloos op te volgen. Door onder andere de gevolgen van onderwijsvernieuwingen treden kennisdeficiënties op in de zin van het niet bereiken van alle geprogrammeerde onderwijsdoelen. In de terminologie van Onderwijsraad (2006) en de daaraan ten grondslag liggende studie van Kools en Van der Neut (2006): kennisdeficiënties die voortkomen uit het niet aansluiten van de programma's van vooropleiding en vervolgopleiding, en welke voortkomen uit het feit dat programma's op zich wel aansluiten, maar waarbij het beoogde niveau van de vooropleiding niet wordt gehaald. Boven genoemde EU-projecten richten zich bij uitstek op het internationale aspect van niet aansluitende programma's. Ze benutten daarbij een vraagstelling die sterk overeenkomt met die van de eerste Nederlandse projecten: een inventarisatie van didactische scenario's die worden gebruikt in de verschillende landen, met daaraan gekoppeld een eerste aanzet tot het formuleren van succesindicatoren.

Het conceptuele kader dat uit bovenstaande studies kan worden afgeleid is nogal specifiek; Kennisdeficiënties hebben een doorslaggevend effect op academisch succes en doorstroming. Een volledige conceptueel kader is natuurlijk rijker; Naast kennis tekorten zijn er vele andere determinanten van studiesucces (Boekaerts & Simons, 1995). Alternatieve kaders, ontleend aan modellen voor motivatieleer, doeloriëntaties en leerstijlen, zullen in deze studie een belangrijke rol spelen. Ze fungeren namelijk als referentie in de statistische effectanalyse en zullen in de methodesectie worden geïntroduceerd.

De echte voortrekkers van aansluitonderwijs zijn te vinden in het Angelsaksische onderwijs, zeer in het bijzonder de VS. 'Developmental education for underprepared students', zoals het gangbaar wordt aangeduid, heeft daar een enorme omvang verkregen: niet minder dan 52% van eerstejaars van publieke colleges in de VS volgt enige vorm van aansluitonderwijs. Algemeen wordt geoordeeld dat een te groot deel van publiek onderwijsgeld aan aansluitonderwijs wordt besteed, en dat de belangrijkste discussie is hoe dat kan worden teruggebracht (zie bij-

voorbeeld het themanummer van *New Directions for Community Colleges*, Schuetz & Barr, 2008). Uit de VS komen dan ook veel van de empirische studies naar waargenomen effecten van aansluitonderwijs, zoals Bahr (2008), Bettinger en Long (2008), Calcagno en Long (2008), Jamelske (2009); voor overzichten, zie Bettinger en Long (2008) en Brants en Struyven (2009). De specifieke Amerikaanse context van deze studies is bepalend voor de uitwerking van de onderzoeksvraag naar effectiviteit van aansluitonderwijs. Amerikaans hoger onderwijs is gebaseerd op plaatsing en selectie, en onderdeel daarvan is dat aspirantstudenten een plaatsingstoets (*placement test*) maken en verplicht deelnemen aan aansluitonderwijs indien de score beneden een bepaalde grenswaarde valt. Amerikaanse effectstudies vergelijken het studiesucces van studenten die net onder die grenswaarde scores (en dus verplicht deelnemen aan aansluitonderwijs) met dat van studenten die net boven die grenswaarde scores (en niet mogen deelnemen aan aansluitonderwijs) met gebruik van zogenaamde regressie-discontinuïteitmodellen. In de Europese context is een dergelijke aanpak onbruikbaar, want in ons onderwijs is plaatsing en plaatsingstoetsing ongewoon. Als gevolg is er geen stelsel dat sommige studenten verplicht tot, en andere studenten juist uitsluit van, deelname aan aansluitonderwijs (Brants & Struyven, 2009). Wel heeft de probleemstelling van effectonderzoek in beide contexten belangrijke overeenkomsten. Van een toevalsexperiment is geen sprake omdat deelname aan aansluitonderwijs niet op basis van gerandomiseerde toewijzing plaatsvindt, maar op basis van ofwel prestatie in een toets (VS) of zelfselectie (Europa). Als gevolg kan effectonderzoek niet gebeuren door een rechtstreekse vergelijking van academisch succes van studenten die wel en niet aansluitonderwijs hebben gevolgd: De kenmerken van beide groepen zullen in het algemeen verschillen. In een quasi-experimentele onderzoekopzet moeten gevonden verschillen tussen experimentele en controlegroep worden verklaard met behulp van verschillen in achtergrondkenmerken van de twee groepen, de covariaten (Hart, Boeije, & Hox, 2005). Anders dan in Amerikaanse empirische stu-

dies is er niet één enkel duidelijk aanwijsbaar kenmerk dat experimentele en controlegroep onderscheidt, waardoor methoden die gebruik maken van zo'n discontinuïteit niet toepasbaar zijn. In plaats daarvan zal gebruik moeten worden gemaakt van methoden die zijn ontwikkeld voor de quasi-experimentele onderzoekopzet zonder voormeting en met niet-equivalente groepen: de *propensityscore*-methode (Fraas, 2007; Shadish, Cook, & Campbell, 2002; Yanovitzky, Zanutto, & Hornik, 2005).

Het effectonderzoek van deze bijdrage richt zich op ervaringsgegevens opgedaan in de aansluitcursus wiskunde voor studenten bedrijfskunde en economie van de Universiteit Maastricht (UM). Deze cursus is vormgegeven als een facultatieve zomercursus voorafgaand aan het reguliere programma. Het is één van de langst bestaande cycli van aansluitonderwijs. Vanaf 2003 wordt de cursus in nagenoeg constant formaat iedere zomer gegeven, met in totaal 750 deelnemers in zeven zomers. De casus heeft een belangrijke positie ingenomen in de hierboven beschreven nationale en Europese projecten en heeft omgekeerd in belangrijke mate geprofiteerd van die projecten. Tegelijkertijd is de casus niet representatief voor de vele initiatieven die gebundeld zijn in de vier nationale projecten, en die een meer nationale focus hebben. Deelnemers zijn daar overwegend Nederlandse (aspirant) studenten die een vwo-vooropleiding hebben gevolgd en een deficiëntie hebben opgelopen. In de Maastrichtse casus staat het andere aspect van kennisdeficiëntie centraal, aangezien 90% van alle zomercursus deelnemers van internationale herkomst is. Is zo'n facultatieve zomercursus effectief in het bevorderen van doorstroom, is de centrale vraag van deze bijdrage, tegen de achtergrond van cumulerende Amerikaanse inzichten dat aansluitonderwijs wel erg duur, maar qua effect van twijfelachtige betekenis is.

2 Methode

2.1 De UM-zomercursus wiskunde

Aan de kenmerken van de UM-studenteninstroom, inclusief de uitkomsten van de diag-

nostische toetsing die verder in sectie 3.1 worden besproken, is een aantal ontwerppunten af te leiden, die bepalend zijn geweest voor de keuze voor een zomercursusformaat. De belangrijkste daarvan zijn:

- Als gevolg van grote diversiteit in studenteninstroom, voortkomend uit het grote aandeel van internationale studenten, moet ruimte gecreëerd worden voor een cursus wiskunde met een forse omvang (tot 100 studie-uren afhankelijk van vooropleiding). Zo'n cursus wijkt wat het aspect van studieomvang betreft af van voorbeelden van zomerscholen die in Onderwijsraad (2008) worden besproken, die veel beperkter in omvang zijn.
- Bij een cursus van die omvang is de mogelijkheid tot individualisering cruciaal: voorkom enerzijds dat studenten modules moeten doorlopen die onder hun kunnen liggen, anderzijds dat ze met modules moeten beginnen die daarboven liggen.
- Om deze individualisering mogelijk te maken, moet (herhaalde) diagnostische en adaptieve toetsing een grote rol spelen en moet het instructieaanbod zich kunnen aanpassen aan de uitkomsten van die toetsing.
- Door de omvang van de aansluitcursus en de ongelijke belasting ervan voor studenten kan die cursus niet in de poort van de reguliere studie worden aangeboden, zoals bij veel Nederlands aansluitonderwijs het geval is. In plaats daarvan moet hij voor de poort worden aangeboden: in de vorm van een zomercursus die aan de reguliere studie voorafgaat.
- Omdat onder de deelnemers vooral ook internationale studenten zijn, die in de zomer voorafgaand aan de studie nog niet in Maastricht leven, moet het model van afstandsonderwijs gebruikt worden.
- Aangezien deelname moet kunnen samenvallen met andere zomerse activiteiten, moet de invulling van de zomercursus flexibel zijn. De student moet zelf kunnen kiezen op welke momenten er gewerkt wordt.
- Uit de nog te bespreken afnamen van de instaptoetsen onder alle eerstejaars bleek dat belangrijke deficiënties zich bevinden op terreinen die eerder tot het leerstof-

aanbod van de onderbouw gerekend moeten worden, dan bovenbouw, zoals algebraïsche rekenvaardigheden. Dit stelt vergaande eisen aan de breedte van het instructieaanbod.

- De taal van de instructie van het aansluitonderwijs moet gelijk zijn aan die van het reguliere programma: Engels.

Op basis van deze overwegingen is besloten het UM-aansluitonderwijs aan te bieden in de vorm van een online zomercursus met als leeromgeving de adaptieve *e-tutorial* ALEKS, module College Algebra (<http://www.aleks.com/highered/math/>). ALEKS staat voor Assessment and Learning in Knowledge Spaces (zie ook Doignon & Falmange, 1999; Falmange, Cosyn, Doignon, & Thiéry, 2004). Kennisruimtetheorie is de tak van kunstmatige intelligentie die zich bezig houdt met de vraag hoe de totale kennis binnen een wetenschapsgebied, de 'kennisruimte', kan worden gerepresenteerd en hoe vervolgens de kennis van een individuele student (zijn 'kennistoestand') kan worden beschreven en door toetsing kan worden bepaald. ALEKS is een op die principes gebaseerde elektronische leeromgeving waarin adaptieve toetsing een cruciale rol speelt (in jargon: een ATS of *Adaptive Tutoring System*). Door die toetsing bepaalt ALEKS het kennisniveau van de student, en geeft het de student aan welke keuzen die kan maken uit een beperkt aantal op dat gemeten kennisniveau aansluitende vervolgonderwerpen (zie ook Tempelaar et al., 2011).

Qua inhoud dekt ALEKS het typische Amerikaanse high-schoolprogramma af, hetgeen in grote lijnen overeenkomt met de onderbouwleerstof plus een aanvang van de bovenbouwleerstof van het vwo-WiA1,2-programma. Om deficiënties voor het hele vwo-programma te remediëren is de inhoud dus niet voldoende. Echter, zoals aangegeven in analyses als verwoord in Tweede Fase Adviespunt (2005) en de uitkomsten van eigen instaptoetsing, liggen voor de doorstroom cruciale deficiënties juist bij meer basale onderwerpen. Door de breedte van de leerstofinhoud is het onvermijdelijk dat de uitleg die ALEKS in de *learning mode* geeft, bondig is. In de literatuur wordt ALEKS dan ook eerder als 'opfris- en trainingsomgeving' dan als

leeromgeving voor initieel leren aangeduid. Voor het opsporen en verhelpen van kennisachterstanden en het opfrissen van niet-para-te kennis is het echter bij uitstek geschikt.

2.2 Deelnemers en niet-deelnemers

Deze studie is gebaseerd op analyses van vijf cohorten eerstejaarsstudenten bedrijfskunde en economie, de vijf meest recente instroomlichtingen: 2005/2006 t/m 2009/2010. Dit zijn de cohorten waarvoor zowel alle zomercursusgegevens, als student achtergrondkenmerken beschikbaar zijn. Samen omvatten deze vijf cohorten zo'n 4.500 eerstejaars, waarvan 68% internationale studenten. Hier-van schrijven 13% van die eerstejaars zich in voor de facultatieve zomercursus (578 partici-panten in vijf uitvoeringen).

Participatie is gebaseerd op zelfselectie, die wordt voorafgegaan door de volgende voorlichtingsactiviteiten:

- In maart t/m mei worden aspirantstuden-ten geïnformeerd over mogelijkheid een gratis zomercursus te volgen. Onderdeel van de informatie is een korte, elektroni-sche kennismakingstoets, die de studenten een globaal beeld verschaft van hun wis-kunde beheersing.
- Jaarlijks doen zo'n 300-500 personen aan de kennismakingstoets mee. Serieuze deelnames worden daaruit geselecteerd en aangeschreven met de vraag of ze belang-stelling hebben voor deelname aan de zo-mercursus, plus de intentie hebben aan de UM te gaan studeren én zich bereid ver-klaren tenminste 80 uur aan de zomercur-sus te besteden. Jaarlijks ontvangen tussen 150 en 250 studenten uiteindelijk een uit-nodiging.
- Ongeveer de helft daarvan, jaarlijks ruim 100 studenten, aanvaardt die uitnodiging tot deelname aan de zomercursus en ont-vangt een licentie tot het gebruik van ALEKS College Algebra.

Overigens lopen de toezegging serieus aan de zomercursus deel te nemen, en de realisatie ervan, voor nogal wat studenten uiteen: slechts 52% van de deelnemers slaagt voor de zomercursus, in de zin dat tenminste 55% van de leerstof wordt bestudeerd. Die typisch Nederlandse norm van 5,5 is soepel, omdat ook onderwerpen die zoals blijkt uit de eerste

diagnostische toets al beheerst worden en dus in de leermodus worden overgeslagen, daar-toe worden meegerekend (gemiddeld instap-niveau ligt op 15% beheersing). Studie-inzet loopt tussen de beide groepen sterk uiteen: zo is de gemiddelde *connect*tijd van de cursisten die slagen 52,1 uur, die van de cursisten die zakken 15,1 uur (*connect*tijd is een onder-grens van de studie-inzet, omdat het aangeeft hoeveel de student binnen ALEKS heeft ge-leerd, zonder dat gemeten wordt hoeveel er buiten ALEKS om wordt geleerd).

Na afloop van de zomercursus, eind au-gustus, beginnen de reguliere programma's van de bachelorstudies bedrijfskunde en eco-nomie in september. Beide programma's star-ten in de eerste week van september met de-zelfde twee 'blokken' van ieder acht weken en een 50% studiebelasting (20 uur per week). Het ene blok is een inleiding in de vakken marketing en organisatieleer. Het an-dere blok, in het kader van deze bijdrage van bijzonder belang, heet QM1 of Kwantitatieve Methoden 1: een eerste kennismaking met wiskunde en statistiek. Dit blok speelt in onze studie de rol van afnemend vak voor de zomercursus. Eerste activiteit in het blok is het doen van de instaptoets, voor zowel mo-nitordoeleinden als de vormgeving van de zomercursus en de onderwerpsbepaling in QM1 zelf. De inhoud van QM1 valt kort aan te duiden als primair een herhaling van de leerstof uit de laatste twee jaren van het vwo-A1,2-programma, met beperkte aandacht voor een aantal nieuwe onderwerpen. QM1 vertoont inhoudelijk geen overlap met de zo-mercursus, die immers meer basale kennis afdekt. Uit eerdere analyses van prestaties in het QM1 onderwijs is bekend dat het niveau van voortgezet wiskunde onderwijs een be-langrijke voorspeller is. In deze studie wor-den twee niveaus onderscheiden: basis en uit-gebred. Voor Nederlandse vwo-aspiranten komt dat overeen met vwo-wiskunde A1,2 (minimum eis voor studies bedrijfskunde en economie) versus wiskunde B1 of B1,2. In Duitstalige landen is er een enigszins ver-gelijkbaar onderscheid: *Grundkurs* versus *Leistungskurs*. Hetzelfde geldt voor studen-ten met een Internationaal Baccalaureaat: *Mathematics SL* versus *HL*. Deze laatste, naar Nederlandse verhoudingen vrij grote,

groep is ingedeeld bij de internationale studenten, terwijl de groep natuurlijk ook Nederlandse aspiranten omvat: internationaal verwijst dus primair naar type vooropleiding, niet naar nationaliteit. Tabel 1 verschaft de indeling naar cohorten en per cohort naar vooropleiding van die participanten aan de zomercursus, waarvan vooropleidinggegevens bekend zijn.

Overeenkomstig de doelstelling van de zomercursus, is participatie het grootst bij internationale studenten, en studenten met een basis wiskundevooropleiding. Qua omvang de tweede groep zijn internationale studenten met een uitgebreide wiskunde vooropleiding die deelnemen aan het aansluitonderwijs. Dat valt te verklaren uit de traditie in Duitstalige landen om, gedwongen (dienstplicht) dan wel vrijwillig, de studie te onderbreken na voortgezet onderwijs; voor deze studenten is de zomercursus primair een 'opfriscursus'.

2.3 Statistische analyses

Een belangrijk aspect van deze bijdrage is de methode van effectonderzoek. Omdat participatie in het aansluitprogramma op vrijwillige basis plaatsvindt, is een quasi-experimenteel ontwerp voor effectonderzoek van toepassing. Tevens geldt dat voormetingen ontbreken, zodat de relevante onderzoeksopzet die is van een quasi-experimenteel ontwerp met niet-equivalente groepen en enkele nameting (Hart et al., 2005; Shadish et al., 2002). Dit ontwerp brengt risico's met zich mee, zoals de mogelijke aanwezigheid van zelfselectie. In lijn met recente adviezen inzake het opsporen van causale effecten in observatio-

nele studies (Schneider, Carnoy, Kilpatrick, Schmidt, & Shavelson, 2007), zijn quasi-experimentele elementen in het onderzoeksontwerp benut, waarvan de belangrijkste het opnemen omhelst van een breed spectrum van studentkenmerken welke gerelateerd zijn aan potentiële selectie-effecten. Deze achtergrondgegevens van studenten zijn verzameld zowel bij studenten die aan de zomercursus hebben deelgenomen, als bij studenten die hebben geopteerd niet aan de zomercursus deel te nemen. De gegevens zijn ontleend aan langdurig onderzoek dat binnen de UM plaatsvindt naar studiesucces in het eerste studiejaar (zie Tempelaar, Rienties, & Gijsselaers, 2006, 2007) en dat is gebaseerd op een brede schakering van theoretische kaders voor verklaring van studiesucces. Het betreft: type voortgezet onderwijs (vwo of internationaal), niveau voortgezet wiskundeonderwijs (basis of uitgebreid), leerstijlen, doeloriëntaties, metacognitie, academische motivaties en onderwerpspecifieke prestatiemotivaties, gemeten met behulp van zelfrapportage-instrumenten.

Traditionele benaderingen voor effectonderzoek in observationele studies bepalen het behandelingseffect aan de hand van een multipel (logistisch) regressiemodel of ANCOVA met als verklarende variabelen, naast de behandeling, ook covariaten die het effect corrigeren voor variatie in de effectvariabele die niet het gevolg is van de behandeling (maar bijvoorbeeld van selectie-effecten). In specifieke toepassingen, vooral wanneer experimentele en controlegroep nogal verschillen in die achtergrondkenmerken en

Tabel 1

Samenstelling participatie zomercursus in vijf cohorten eerstejaars, naar vooropleiding

Zomercursus participatie	Vwo-vooropleiding		Internationale vooropleiding		Totaal
	wiskunde basis	wiskunde uitgebreid	wiskunde basis	wiskunde uitgebreid	
Cohort 2005/2006	2 (0,9%)	0 (0,0%)	41 (12,9%)	9 (7,5%)	52 (6,7%)
Cohort 2006/2007	1 (0,6%)	1 (1,2%)	71 (19,7%)	21 (17,9%)	94 (12,9%)
Cohort 2007/2008	14 (7,4%)	0 (0,0%)	88 (21,6%)	15 (10,2%)	117 (14,4%)
Cohort 2008/2009	8 (4,9%)	1 (1,2%)	79 (17,6%)	25 (12,9%)	113 (12,7%)
Cohort 2009/2010	19 (9,8%)	8 (8,3%)	115 (24,5%)	34 (15,7%)	176 (18,1%)
Alle vijf cohorten	44 (4,3%)	10 (2,7%)	394 (18,2%)	104 (12,8%)	552 (13,2%)

Noot. Percentages hebben betrekking op aandeel zomercursus participatie in totaal van vooropleiding. Van 36 van de 578 participanten zijn geen vooropleidinggegevens bekend.

sommige cellen in de variantieanalyse tabel slechts spaarzaam gevuld zijn, hebben deze benaderingen echter tekortkomingen (Fraas, 2007; Yanovitzky et al., 2005). Deze problemen worden voorkomen door gebruik te maken van de methode van *propensityscores*, waarbij het behandelingseffect wordt gecorrigeerd (Fraas, 2007; Shadish et al., 2002; Yanovitzky et al., 2005). Die correctie vindt plaats op basis van de *propensityscores*: de conditionele kans dat een individu tot de experimentele groep behoort, c.q. de controlegroep, gegeven de set van covariaten. *Propensityscores* worden meestal geschat met logistische regressieanalyse. De correctie van het behandelingseffect kan op verschillende manieren gebeuren: het gebruik van de *propensityscore* als matchingvariabele, stratificatievariabele, of covariaat. In deze studie wordt de stratificatiebenadering gevolgd omdat experimentele- en controlegroep nogal ongelijk van omvang zijn, zodat matching met omvangrijk verlies aan data gepaard gaat. Als intermediaire stap wordt het model met de *propensityscore* als covariaat afgeleid. Effectbepaling met een model waarin alle achtergrondvariabelen als covariaat worden opgenomen, is in deze toepassing minder geschikt, gezien het grote aantal gemeten kenmerken.

Eén achtergrondvariabele zal overigens niet worden gebruikt ter bepaling van de *propensityscore*, maar afzonderlijk in het model opgenomen worden, naast de *propensityscore*: het niveau van de wiskunde vooropleiding. Dit omdat het verschil in prestatie tussen studenten met en zonder wiskunde vooropleiding op uitgebreid niveau, de enige expliciete referentiemaatstaf is voor een beoordeling van het behandelingseffect: "Hoe verhoudt het effect van succesvolle participatie in de zomercursus zich tot het verschil tussen wiskunde A en B in de vooropleiding?".

2.4 Effectmaten en covariaten: de student achtergrondkenmerken

Het effectonderzoek richt zich op de academische prestaties van deelnemers en niet-deelnemers van de zomercursus in het vak QM1. Als effectmaten worden gebruikt de totaalscore voor het vak (op een schaal van 0 ... 40) en de dichotome variabele die sla-

gen of zakken voor het vak aangeeft. Keuze voor andere effectvariabelen, zoals academisch succes in het gehele eerste studiejaar, geven overigens vergelijkbare uitkomsten, door specifieke eigenschappen van het programma (het bestaan van een bindend studieadvies waarvoor het slagen voor QM1 de facto een noodzakelijke voorwaarde is; zo is de correlatie van vakscore en GPA gelijk aan 0,88).

De gebruikte covariaten zijn ontleend aan lopend onderzoek naar studiesucces in het eerste studiejaar. Voor dat onderzoek zijn een aantal instrumenten geselecteerd die in toegepast onderwijskundig onderzoek naar voren komen als belangrijke determinanten van studiegedrag en studieprestaties. Deze covariaten refereren naar alle aspecten van de rol van individuele verschillen in leerprocessen. Onder verwijzing naar de onderverdeling in Boekaerts en Simons (1995): de covariaten omvatten zowel cognitieve strategieën en leerstrategieën, metacognitieve kennis en vaardigheden, motivationele kenmerken, affectieve kenmerken en domeinspecifieke variabelen. Elders, zie Tempelaar, Rienties, & Gijsselaers (2006, 2007), is de keus voor de gebruikte instrumenten en hun rol in het leerproces meer uitvoerig toegelicht. Van alle instrumenten is de Engelstalige versie gebruikt. Betrouwbaarheden van alle gebruikte schalen variëren van redelijk tot goed (liggen alle in de range van 0,6 tot 0,9, in grote meerderheid tussen 0,7 en 0,8) en komen overeen met betrouwbaarheden gerapporteerd in de studies waar hieronder naar wordt verwezen.

Leerstijlen

Alle studenten zijn bevraagd naar leerbepalingen met behulp van het leerstijlmodel van Vermunt (Vermunt, 1992; zie ook Picarelli, Slaats, Bouhuijs, & Vermunt, 2006; Van Hout-Wolters, 2009) en het daarop gebaseerde instrument ILS-HO (Engelstalige versie). Vermunt onderscheidt in zijn model vier domeinen of componenten van leren: cognitieve verwerkingsstrategieën, metacognitieve regulatiestrategieën, leerconcepties en leeroriëntaties. Het domein van de cognitieve verwerkingsstrategieën betreft de leerstrategieën die studenten gewoonlijk gebruiken om leerstof te verwerken. Twee van die leerstrategieën, lateren & structureren en kritisch ver-

werken, vormen samen een hoofdschaal die Vermunt met diepteverwerking aanduidt. De subschalen memoriseren & herhalen en analyseren vormen samen de hoofdschaal stapsgewijze verwerking. Een derde leerstrategie is die van concrete verwerking. Het tweede domein betreft de metacognitieve regulatiestrategieën. De eerste twee schalen zijn zelfsturing van leerproces & leerresultaat en zelfsturing van leerinhoud; samen vormen deze twee subschalen de hoofdschaal zelfsturing. De veronderstelde tegenpool van *zelfsturing* is *externe sturing*, uitgedrukt in de subschalen externe sturing van leerproces en externe sturing van leerresultaten. Gebrek aan enige sturing wordt vertegenwoordigd door de schaal stuurloos leergedrag. De laatste twee domeinen van het ILS-HO-instrument betreffen de leeroriëntaties en de leerconcepties. De vijf onderscheiden leeroriëntaties zijn: persoonlijk geïnteresseerd, certificaatgericht, testgericht, beroepsgericht en ambivalent. De vijf leerconcepties tenslotte zijn: opbouw van kennis of constructivisme, opname van kennis, gebruik van kennis, stimulerend onderwijs en samen of coöperatief leren.

Doeloriëntaties

De relatie tussen impliciete theorieën over intelligentie, opvattingen over inzet in het leren, en doeloriëntaties staan centraal in het onderzoek van Dweck (1999). Al deze studentkenmerken zijn onderzocht, maar enkel de doeloriëntaties bleken een rol te spelen in deze studie. Doeloriëntaties zijn gemeten met het instrument van Grant en Dweck (2003) dat op gebruikelijke wijze leerstreven en prestatiestreven veronderstelt, en vervolgens binnen het prestatiestreven normatieve doelen tegenover intrapersoonlijke of niet-normatieve doelen plaatst, en competentiedoelen (*ability goals*) tegenover uitkomstdoelen (*outcome goals*). Op deze wijze worden zes typen doelen onderscheiden: intrapersoonlijke uitkomstdoelen, intrapersoonlijke competentiedoelen, normatieve uitkomstdoelen, en twee typen leerstreven die verschillen in de mate waarin de student op zoek is naar uitdaging, leerstreefdoel (in enge zin), en een op uitdaging gerichte doeloriëntatie (*challenge-mastery*).

Metacognitie

Onderwijskundigen van de Universiteit van Amsterdam (Elshout-Mohr, Van Daalen-Kaptein, & Meijer, 2001) hebben een instrument ontwikkeld om vast te stellen hoe goed studenten uit de voeten kunnen binnen een leeromgeving die zich baseert op constructivistische leerbenaderingen, zoals het studiehuis. In dat instrument, de Rapportage Autonoom Studeren of RAS, worden studenten gevraagd een oordeel te geven over de eigen metacognitieve vaardigheden. In ons onderzoek hebben we gebruik gemaakt van de Engelstalige versie van dit instrument, dat met de naam Awareness of Independent Learning Inventory (AILI) wordt aangeduid (zie ook Tempelaar, 2006). De AILI onderscheidt drie verschillende typen van metacognitieve kwaliteiten: metacognitieve kennis (kennis met betrekking tot leren en studeren), metacognitieve regulatie (de vaardigheid om het studeren systematisch te sturen) en metacognitieve ontwikkeling of responsiviteit (informatiegevoelige en onderzoekende houding).

Academische motivaties

Motivatietheorieën in de traditie van de zelfdeterminatietheorie maken gebruik van het onderscheid tussen intrinsieke motivatie (leren om meer te weten) en extrinsieke motivatie (leren voor de beloning; zie bijvoorbeeld Martens & Bastiaens, 2005; Ryan & Deci, 2000). Intrinsieke motivatie kent een drietal componenten: om te weten, om te volvoeren en om een stimulerende omgeving te ervaren. Er zijn drie typen extrinsieke motivatie, die zich bovendien laten ordenen van puur extrinsiek, via mengvormen van extrinsiek tot intrinsiek. Externe regulatie is de pure vorm van extrinsieke motivatie: het gedrag wordt geheel gestuurd door externe prikkels. Dicht daar tegenaan ligt introjectie: de externe prikkels zijn zo zeer geïnternaliseerd, dat ze voor het uitlokken van het gedrag niet meer nodig zijn. De vorm die het meest verwant is met intrinsieke motivatie is identificatie, ook wel als integratie aangeduid: de waarden en opvattingen die aan de externe prikkels ten grondslag liggen, worden overgenomen, zodat het gedrag intern gereguleerd gaat worden. Ten slotte is er de

toestand van amotivatie of gebrek aan enige motivatie. Deze zeven typen motivatie worden gemeten met de Academic Motivation Scale (AMS; zie Ratelle, Guay, Vallerand, Larose, & Senécal, 2007; Vallerand et al., 1993).

Onderwerpspecifieke prestatiemotivaties

De attitudes die studenten hebben tegenover de vakken die ze in hun studie tegenkomen is een belangrijke determinant van het leren. Die vakattitudes voor het vak QM1 hebben we gemeten met een in internationaal onderwijsonderzoek veelgebruikte vragenlijst, Survey of Attitudes Toward Statistics (SATS; zie Tempelaar, Gijselaers, Schim van der Loeff & Nijhuis (2007) voor een meer uitvoerige bespreking ervan. De vragenlijst is gebaseerd op het verwachting-waardemodel van Eccles en Wigfield (Eccles & Wigfield, 2002) en onderscheidt voor een vakgebied de volgende zes attitudes: affectie voor het vak, zelfperceptie van cognitieve competentie, de aan het vak toegekende waarde, gepercipieerde moeilijkheid, interesse en tenslotte voorgenomen inzet tot leren.

3 Resultaten

3.1 De landelijke instaptoets

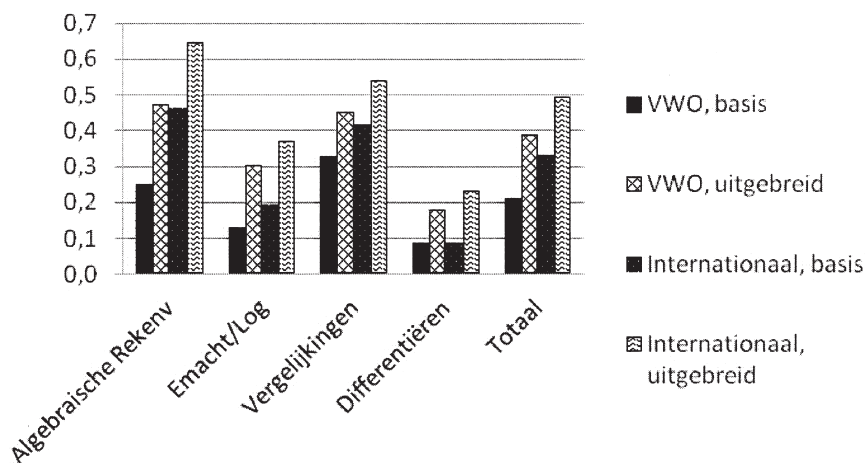
In het kader van de verschillende landelijke projecten voor aansluiting en doorstroming is in het laatste onderzochte jaar voor het eerst de Nationale KennisBank Wiskunde (NKBW) instaptoets afgenomen, aan de hand waarvan de landelijke wiskunde monitor wordt samengesteld. In deze studie betreft dat echter maar één van de vijf onderzochte cohorten. Als voorloper van de NKBW-instaptoets kan de 3TU-instaptoets worden beschouwd: een diagnostische toets ontworpen door de drie technische universiteiten. Een specifieke editie van die 3TU-instaptoets is in drie van de vijf hier onderzochte cohorten afgenomen: in totaal ruim 2.600 eerstejaars studenten (zie voor een beschrijving van de toets en toetskenmerken Tempelaar & Caspers, 2008). De breedte van de toets is ingeperkt tot de vereiste wiskunde voorkennis voor de bedrijfskunde en economie opleiding: het wiskunde vwo-A1,2-programma.

Toetsvragen zijn globaal in te delen in twee basale onderwerpen, Algebraïsche rekenvaardigheden en E-macht & logaritme, en twee geavanceerde onderwerpen, Vergelijkingen en Differentiëren. In Figuur 1 zijn scores voor beheersing van die vier onderwerpen en een totaalscore weergegeven voor vier groepen studenten: studenten met een vwo-vooropleiding dan wel internationale vooropleiding, gekruist met wiskunde in de vooropleiding op basis- dan wel uitgebreid niveau. Scores betreffen proportie goed-waarden na correctie voor gokken (instaptoets is een vierkeuzetoets).

Belangrijke uitkomst van de instaptoetsing voor het ontwerp van het aansluitonderwijs is dat naast deficiënties op het gebied van bovenbouwonderwerpen als Vergelijkingen en Differentiëren, onderwerpen die vaak al deel uitmaken van regulier eerstejaars-onderwijs, ook aanzienlijke deficiënties bestaan in onderwerpen als Algebraïsche rekenvaardigheden en E-macht & logaritme, die overwegend in de onderbouw worden onderwezen en die geen onderdeel vormen van regulier universitair onderwijs. Deze observatie ondersteunt de analyse in Tweede Fase Adviespunt (2005) over de cruciale rol van basale vaardigheden. Andere observaties zijn dat alle scores teleurstellend laag liggen en dat bij een vergelijking binnen vooropleidingsniveaus de vwo-aspiranten het duidelijk zwakker doen dan de internationale aspiranten. Selectie-effecten kunnen hier echter niet uitgesloten worden, zodat een niveauvergelijking enige terughoudendheid vereist. Wel is onbetwist dat, relatief ten opzichte van de verschillende onderwerpen, vwo-aspiranten het zwakst zijn in de basale onderwerpen, zoals Algebraïsche rekenvaardigheden: op dat terrein scoren vwo-B-aspiranten niet hoger dan internationale studenten met een basis wiskundevooropleiding. Het focussen van Nederlandse aansluittrajecten op deze primaire deelonderwerpen lijkt daarom een correct uitgangspunt.

3.2 Beschrijvende analyses

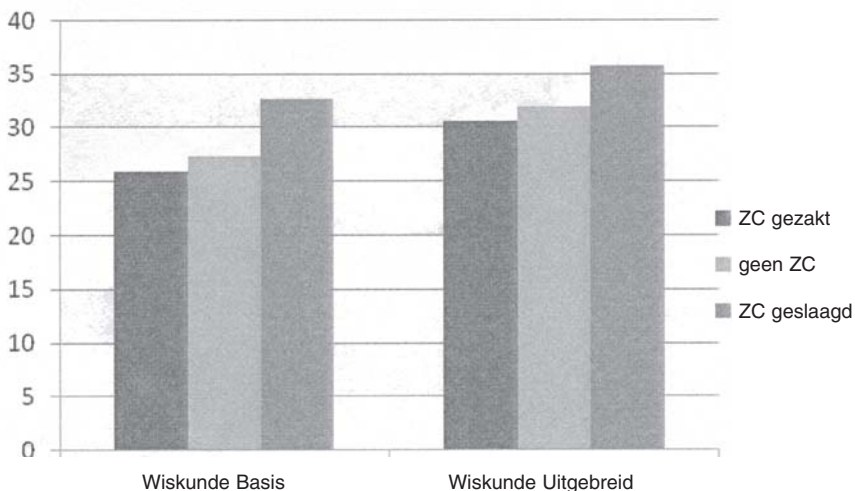
Figuren 2 en 3 geven de ongecorrigeerde behandelingseffecten van deelname aan de zomercursus weer, respectievelijk op de QM1-totaalscore (Figuur 2) en de QM1-slaagkans



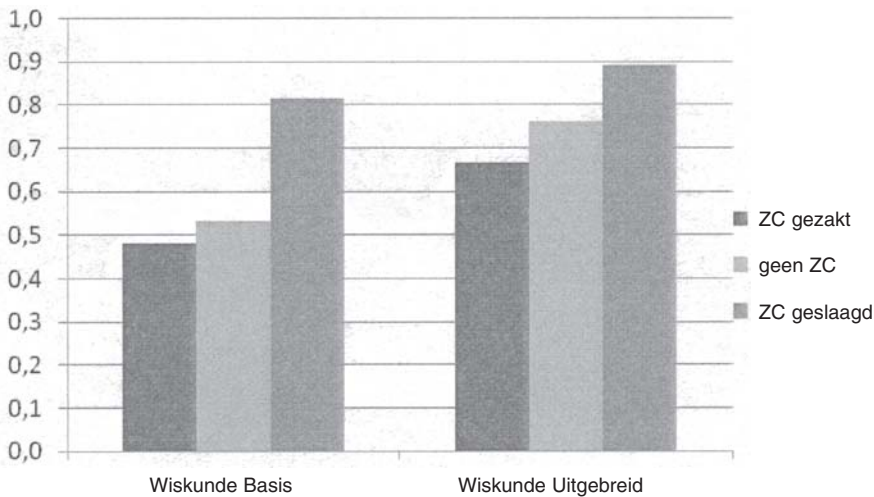
Figuur 1. Proportie goed scores op totale instaptoets en deelonderwerpen, uitgesplitst naar instroom, in drie cohorten.

(Figuur 3) in de eerste kans van de toets. Om het substantiële interactie-effect met vooropleiding zichtbaar te maken, worden studenten met basis wiskundevoorbereiding en uitgebreide wiskundevoorbereiding onderscheiden. In totaal nemen 4.188 van de onderzochte studenten aan die eerste toetskans deel, waarvan 280 succesvolle zomercursus deelnemers, 231 niet-succesvolle deelnemers en 3.677 niet-deelnemers. Het effect van een uitgebreide wiskundevoorbereiding in de referentiegroep van niet aan de zomercursus deelnemende studenten op de totaalscore is 4,6 punt (of, uitgedrukt als effectgrootte, 0,64

standaarddeviatie), en op de slaagkans 23% (0,53 standaarddeviatie). Het effect van succesvolle deelname aan de zomercursus, ten opzichte van geen deelname, is 5,3 punt (0,76 standaarddeviatie) respectievelijk 28% (0,56 standaarddeviatie) voor studenten met een basis wiskundeopleiding, en 3,8 punt (0,54 standaarddeviatie) respectievelijk 13% (0,31 standaarddeviatie) voor studenten met een uitgebreide wiskundeopleiding. Zoals te verwachten is het behandelingseffect veel groter voor studenten met een basis wiskundeopleiding, dan voor studenten met een uitgebreide wiskundeopleiding. De forse effectgroottes



Figuur 2. Ongecorrigeerd behandelingseffect QM1 totaalscore van zomercursusdeelname, per vooropleiding.



Figuur 3. Ongecorrigeerd behandelingseffect QM1 slaagkans van zomercursusdeelname, per vooropleiding.

geven al aan dat de ongecorrigeerde behandelingseffecten statistisch significant zijn, met significantieniveaus onder 0,001 voor alle studenten gezamenlijk en voor studenten met basis wiskundeopleiding, en onder 0,01 voor studenten met een uitgebreide wiskundeopleiding. Voor die primaire doelgroep van studenten met een basisopleiding is het effect van succesvolle deelname aan de zomercursus zelfs zo groot, dat die het effect van een uitgebreide wiskundeopleiding ruimschoots overtreft, zowel wat betreft eindscore als wat betreft slaagkans.

De selectie van covariaten voldoet aan de doelstelling: van de 44 covariaten, hangen er 38 significant samen met de effectmaten. Uitzonderingen zijn de leerstrategie *memoriseren*, de oriëntaties *persoonlijke interesse* en *certificaat-gerichtheid*, *extrinsieke motivatie*, *introjectie* en *intrinsieke motivatie door stimulering*.

3.3 Propensityscores

Conditionele kansen om deel te nemen aan de zomercursus voor de 3.240 studenten waarvan alle achtergrondvariabelen bekend zijn, zijn geschat met behulp van de logistische regressiemethode. Van de 44 achtergrondvariabelen blijken zes een statistisch significant effect te hebben in het simultane regressiemodel. Verreweg de sterkste predictor is, overeenkomstig de gekozen focus van de zomercursus, de 'dummy' die internatio-

nale studenten onderscheidt van studenten met een vwo-opleiding. Daarna, in volgorde van afnemend belang: de *beroepsgerichte leeroriëntatie* uit het instrument ILS-HO, *zelfperceptie van cognitieve competentie* (negatief) als onderwerpspecifieke prestatie-motivatie, *metacognitieve kennis* uit het AILI-instrument, de *constructivistische leerconceptie* uit ILS-HO en de variabele *amotivatie* (negatief) als academische motivatie. De uitkomsten van de logistische regressie, zowel in termen van significantie van achtergrondvariabelen en het teken in de regressie, zijn intuïtief. Een internationale vooropleiding, een sterke focus op de latere beroepsuitoefening, het denkbeeld dat je vooral kennis verwerft door zelf actief met de leersof om te gaan en goede metacognitieve vaardigheden versterken de kans op deelname aan de zomercursus. Gebrek aan motivatie en de perceptie dat je al zeer competent bent op het vlak van de kwantitatieve methoden, verkleinen de kans op deelname.

Overeenkomstig aanbevelingen in de literatuur (Fraas, 2007; Shadish et al., 2002; Yanovitzky et al., 2005) zijn de *propensityscores* berekend op basis van het volledige model, dus met inbegrip van zowel significante als niet-significante covariaten.

3.4 Propensityscore als covariaat

Na berekening van de *propensityscores* is de effectanalyse herhaald, met de *propensity-*

score als additionele predictor. Afhankelijk van de omstandigheid of QM1-totaalscore dan wel QM1-slaagkans als resultaatvariabele wordt genomen, is de aangewezen methode die van multipele regressie respectievelijk logistische regressie. Tabel 2 bevat de uitkomsten van multipele regressie van QM1-totaalscore op de predictorvariabele *propensityscores* en drie indicatorvariabelen (dummies): wiskunde op uitgebreid niveau, succesvolle deelname aan de zomercursus, en niet-succesvolle deelname. De keuze van de indicatorvariabelen impliceert dat wiskunde vooropleiding op basisniveau, en niet deelnemen aan de zomercursus, als referentiegroepen fungeren. *Propensityscores* en de drie indicatorvariabelen tezamen verklaren 11,2% van de variatie in eindscore.

Tabel 2 bevestigt het beeld van de vorige resultaatsectie: deelnemers aan de zomercursus onderscheiden zich van niet-deelnemers op grond van achtergrondkenmerken die een positieve invloed hebben op leergedrag. Het gevolg is dat in de gecorrigeerde berekening van het effect van de zomercursus, de *propensityscores* een deel van de verklaring van succesvolle deelname aan de zomercursus overnemen ten opzichte van het niet-gecorrigeerde model. Daardoor wordt de bijdrage aan verklaarde variantie van de zomercursusdeelname kleiner en is het niet meer de belangrijkste determinant. De indicatorvariabele die aangeeft of wiskunde op het hoogste niveau is gevolgd in voortgezet onderwijs neemt die plaats nu over. Tegelijkertijd resteert er nog

Tabel 2
Uitkomsten effectonderzoek zomercursusdeelname op QM1 totaalscore met *propensityscore* als covariaat

	β	<i>t</i>	<i>p</i>
<i>Propensityscore</i>	0,072	4,116	<0,001
Wiskunde uitgebreid	0,271	15,978	<0,001
Succesvolle deelname zomercursus	0,154	8,899	<0,001
Niet-succesvolle deelname zomercursus	-0,086	-5,002	<0,001

Tabel 3

Uitkomsten effectonderzoek zomercursusdeelname op QM1 slagingskans met *propensityscore* als covariaat

	<i>B</i> (S.E.)	<i>p</i>	Exp(<i>B</i>)
<i>Propensityscore</i>	1,298 (0,456)	0,004	3,663
Wiskunde uitgebreid	0,967 (0,091)	<0,001	2,629
Succesvolle deelname zomercursus	1,097 (0,181)	<0,001	2,996
Niet-succesvolle deelname zomercursus	-0,494 (0,149)	0,001	0,610

steeds een substantieel effect van succesvolle deelname aan de zomercursus: de β (gestandaardiseerde regressiecoëfficiënt) is groter dan 50% van de β van wiskunde op uitgebreid niveau.

Wanneer de focus wordt verlegd naar het slagen of zakken voor het vak QM1 als resultaatvariabele, ontstaat een vergelijkbaar beeld. De aangewezen methode is nu die van logistische regressie; Tabel 3 bevat daarvan de uitkomsten. De verklaarde variatie, uitgedrukt als de Nagelkerke R^2 , bedraagt 8,1%.

Voor de interpretatie van de uitkomst van de logistische regressie is vooral de laatste kolom van Tabel 3 van belang: die geeft de verandering van de *odds ratios* van het slagen voor het vak QM1, als gevolg van een eenheidswijziging in de predictorvariabele. Studentachtergrondkenmerken die van invloed zijn op de deelname aan de zomercursus zijn de belangrijkste determinant voor de verklaring van de kans (*odds*) om voor QM1 te slagen. Onmiddellijk gevolgd door de twee indicatorvariabelen wiskunde uitgebreid en succesvolle deelname zomercursus, waarbij opvalt dat de *odds-ratio* van slagen voor de zomercursus zelfs de waarde van de *odds-ratio* voor wiskunde op het hoogste niveau overtreft.

3.5 *Propensityscore* als stratificatie variabele

De meest zorgvuldige wijze om te corrigeren voor mogelijke selectie-effecten in een quasi-

experimentele onderzoeksofzet met niet-equivalente groepen is die waarbij op basis van de *propensity* scores stratificatie wordt toegepast (Fraas, 2007; Shadish et al., 2002; Yanovitzky et al., 2005). Aanbevolen wordt om een vijftal strata te creëren, gebaseerd op de quintielen van de verdeling van de *propensity* scores. Ieder van die vijf strata bevat dan subjecten met een vergelijkbare *propensity* score, zodat effectonderzoek binnen elk stratum minimaal wordt beïnvloed door de waarde van de *propensity* scores en op deze wijze voor selectie-effecten wordt gecorrigeerd. Deze aanpak is ook hier gevolgd, en na de creatie van vijf strata is de multi-pele regressie beschreven in de vorige sectie voor ieder stratum herhaald. De uit-

komsten van die analyse zijn verzameld in Tabel 4.

Stratificatie blijkt exact te bewerkstelligen wat ermee wordt beoogd: de invloed van de studentachtergrondkenmerken, uitgedrukt als *propensity* score, is in ieder van de strata insignificant. Uitkomsten verkregen in de vorige sectie, het regressiemodel geschat op de gehele populatie, herhalen zich nu, met enkel een afwijkend beeld voor het eerste stratum. In dat eerste stratum, het quintiel van studenten met de laagste score op de achtergrondkenmerken die bijdragen aan deelname aan de zomercursus, overheerst het negatieve effect van het staken van de zomercursus het positieve effect van het succesvol deelnemen aan de zomercursus. Dit enigszins afwij-

Tabel 4

Uitkomsten effectonderzoek zomercursusdeelname op QM1 totaalscore met propensity score als stratificatievariabele

	β	t	p
Stratum 1: <i>propensity</i> score < 0,055 (ZCp=6, ZCf=17)			
<i>Propensity</i> score	0,062	1,715	0,087
Wiskunde uitgebreid dummy	0,365	10,065	<0,001
Succesvolle deelname zomercursus dummy	0,061	1,674	0,095
Niet-succesvolle deelname zomercursus dummy	-0,118	-3,251	0,001
Stratum 2: 0,055 < <i>propensity</i> score < 0,117 (ZCp=31, ZCf=27)			
<i>Propensity</i> score	0,022	0,563	0,573
Wiskunde uitgebreid dummy	0,310	8,124	<0,001
Succesvolle deelname zomercursus dummy	0,173	4,511	<0,001
Niet-succesvolle deelname zomercursus dummy	-0,050	-1,308	0,191
Stratum 3: 0,117 < <i>propensity</i> score < 0,166 (ZCp=58, ZCf=4)			
<i>Propensity</i> score	0,008	0,201	0,841
Wiskunde uitgebreid dummy	0,222	5,697	<0,001
Succesvolle deelname zomercursus dummy	0,127	3,257	0,001
Niet-succesvolle deelname zomercursus dummy	-0,090	-2,315	0,021
Stratum 4: 0,166 < <i>propensity</i> score < 0,217 (ZCp=64, ZCf=55)			
<i>Propensity</i> score	0,041	1,063	0,288
Wiskunde uitgebreid dummy	0,195	4,990	<0,001
Succesvolle deelname zomercursus dummy	0,146	3,735	<0,001
Niet-succesvolle deelname zomercursus dummy	-0,117	-2,996	0,003
Stratum 5: 0,217 < <i>propensity</i> score (ZCp=93, ZCf=82)			
<i>Propensity</i> score	0,007	0,191	0,849
Wiskunde uitgebreid dummy	0,268	7,071	<0,001
Succesvolle deelname zomercursus dummy	0,210	5,485	<0,001
Niet-succesvolle deelname zomercursus dummy	-0,068	-1,769	0,077

Noot. ZCp=aantal succesvolle zomercursus deelnemers; ZCf=aantal niet-succesvolle zomercursus deelnemers.

kende beeld voor het eerste stratum is een artefact van de wijze waarop de strata zijn geformeerd. Door de lage score op de achtergrondkenmerken bevat het eerste stratum verreweg het minste aantal deelnemers aan de zomercursus en daarbinnen nog een meerderheid aan zomercursusstakers. Van de 679 studenten in dit stratum, zijn slechts 23 deelnemers aan de zomercursus, met daaronder 17 die de cursus niet afmaken (in Tabel 4 zijn ook voor de andere strata deze aantallen studenten opgenomen). De vier andere strata, met veel meer zomercursusdeelnemers en vooral ook succesvolle deelnemers, laten steeds eenzelfde beeld zien. Het grootste effect is dat van de dummyvariabele wiskundevooropleiding op uitgebreid niveau, terwijl het behandelingseffect van succesvolle deelname aan de zomercursus daarop volgt en telkens tenminste 50% bedraagt van het effect van de vooropleiding.

Herhaling van de analyse ter bepaling van het behandelingseffect op slaagkans voor QM1 op basis van stratificatie in vijf quintielen van de *propensityscores* levert ook een bevestiging op van de eerdere analyses. Binnen de strata hebben de *propensityscores* geen significant effect meer op de slaagkans. En behoudens het eerste quintiel, waar succesvol zijn in de zomercursus insignificant is, zijn er in de vier andere strata significante effecten van zowel vooropleiding als succesvolle zomercursus deelname, waarbij de *odds-ratio* van de laatste steeds die van de eerste overtreft.

Om een overall maatstaf voor de effectomvang te verkrijgen, zijn als laatste stap de stratumspecifieke effecten geaggregeerd tot een gemiddeld behandelingseffect (ATE; Guo & Fraser, 2010). Deze ATE bedraagt $t = 4,75$, hetgeen significant is op 5% significantieniveau.

In bovenstaande analyses onderscheidt de geschatte *propensityscore* functie twee groepen studenten: deelnemers en niet-deelnemers aan de zomercursus. Voor de berekening van het behandelingseffect zijn echter drie groepen studenten van belang: succesvolle deelnemers, niet-succesvolle deelnemers en niet-deelnemers. Deze indeling in drie groepen had ook de basis kunnen zijn voor een toepassing van de *propensityscore*-

methode, waarbij de *propensityscore*functie dan met behulp van multinomiale logistische regressie geschat zou worden. Vervolgens kan dan stratificatie plaatsvinden op basis van zowel de geschatte conditionele kans om succesvol zomercursus deelnemer te zijn als de kans om niet-succesvol deelnemer te zijn. De uitkomst van die stratificatie is een tweedimensionale clustering. Als onderdeel van deze studie is deze analyse verricht, met stratificatie op basis van tertielen van beide geschatte *propensityscores*. Dit leidt tot negen strata, waarbij in acht van de negen een behandelingseffect is waar te nemen. Slechts in één stratum, dat van het tertiel van de studenten met laagste kans op succesvolle deelname, en hoogste kans op niet-succesvolle deelname, is een afwijkend patroon van een niet-significant, negatief behandelingseffect waar te nemen. Het gemiddelde behandelingseffect, ATE, is nu omvangrijk: $t = 25,8$. Om pragmatische redenen is gekozen om in de rapportage de nadruk te leggen op de indeling in twee groepen. Enerzijds zijn de uitkomsten inzichtelijker door de eendimensionale structuur van Tabel 4, anderzijds voegt het onderscheiden van succesvolle van niet-succesvolle deelnemers qua uitkomsten weinig toe. Dit omdat de covariaten meer invloed hebben op de beslissing al dan niet deel te nemen aan de zomercursus, dan dat ze voorspellend zijn voor het succes van deelname. Dat het succesvol afronden van de zomercursus zo moeilijk te voorspellen is correspondeert met de uitkomsten van de evaluaties. Daarin verwijzen studenten die afhaken overwegend naar contextuele oorzaken, niet naar kenmerken van de cursus of eigen vaardigheden. Met als belangrijkste factor: het nog niet weten hoe de zomer ingevuld gaat worden, wanneer in het begin van de zomer de keuze gemaakt moet worden om deel te nemen.

4 Conclusie en discussie

In veel eerstejaars universitaire programma's zijn elementen van remediërend onderwijs ingebouwd: slechts na het opruimen van een aantal onderwerpen die ook in de laatste leerfase van het voortgezet onderwijs zijn onderwezen, wordt vervolgd met de behandeling

van nieuwe onderwerpen. Uitkomsten van onze instaptoetsing geven aan dat deze benadering niet voldoende is. Belangrijke deficiënties liggen op het vlak van meer basale, in de onderbouw behandelde onderwerpen. Dit correspondeert met de brede invulling die voor de UM-zomercursus is gekozen, waarvan het ingangsniveau wordt bepaald door de individuele resultaten behaald op een adaptieve, diagnostische toets.

Het ongecorrigeerde effect van succesvolle participatie in de UM-zomercursus is zeer substantieel: het overstijgt het effect van het volgen van wiskunde op uitgebreid niveau in het voortgezet onderwijs. De onderzoekopzet van deze studie is echter, in de hiërarchie van mogelijke ontwerpen, een zwakke: geen randomisatie, geen voormeting. Derhalve is correctie van het behandelingseffect voor mogelijke selectie-effecten vereist. Correctie op basis van de *propensity-score*-methode wijst uit dat inderdaad een deel van het behandelingseffect kan worden toegeschreven aan de omstandigheid dat deelnemers aan de zomercursus over gunstiger achtergrondkenmerken beschikken dan niet-deelnemers. Tegelijkertijd resteert er een substantieel behandelingseffect, in orde van grootte van de helft van het effect van het volgen van wiskunde op uitgebreid niveau tijdens de vooropleiding. De uitgevoerde statistische analyse maakt enerzijds duidelijk dat correctie voor selectie-effecten een cruciale analysestap is; Was die achterwege gebleven, dan had dat tot een forse overschatting van het effect geleid. Anderzijds suggereert het dat dit type van extracurriculair, remediërend onderwijs zeer effectief is. Daarbij moet bedacht worden dat de Maastrichtse casus, door het grote aantal internationale studenten, niet als representatief voor het hele hoger onderwijs kan worden gezien, maar dankzij de sterke internationalisering waarschijnlijk wel maatgevend is voor de situatie die vele instellingen binnenkort kunnen verwachten.

De uitkomsten van het effectonderzoek suggereren dat het gekozen formaat van aansluitonderwijs, dat van een online zomercursus met een brede dekking van onderwerpen en sturing van leeraanbod door adaptieve toetsing, een efficiënte methode is om ken-

nisdeficiënties te overbruggen. De gemiddelde studielast voor succesvolle zomercursusdeelnemers, een ruime 50 uur, is evident veel geringer dan het verschil in studielast tussen uitgebreide en basis wiskunde in het voortgezet onderwijs. Desondanks wordt een leer-effect behaald in de orde van grootte van 50% van dat laatste effect. Een belangwekkende onderzoeksvraag lijkt te zijn of andere vormen van aansluitonderwijs, zoals de in Nederland gebruikelijke vorm van remediërend onderwijs in de poort, tot vergelijkbare effecten leiden.

Het substantiële effect dat in deze studie naar voren komt steekt gunstig af bij de meer bescheiden uitkomsten die in Angelsaksisch onderzoek zijn gevonden. Dat laatste onderzoek behelst vooral remediërend onderwijs dat, ook in die gevallen waarin het de student geen studiepunten verschaft, doorgaans ingebed is in regulier onderwijs. De hier beschreven zomercursus wordt extracurriculair aangeboden, en kan zo afwijken in een sterke adaptiviteit van het leerstofaanbod: Geen twee studenten hebben hetzelfde programma doorlopen. Dit kan een belangrijke sleutel zijn in de verklaring van het gevonden effect, en vraagt om verder onderzoek dat zich specifiek richt op de rol van individuele flexibiliteit in remediërend onderwijs in de verklaring van effecten.

Noot

Het beschreven onderzoek is mogelijk gemaakt door ondersteuning van Stichting SURF, onder andere vanuit het Nationaal ActiePlan e-learning, en de EU programma's Lifelong Learning en Minerva.

Literatuur

- Bahr, P. R. (2008). Does mathematics remediation work?: A comparative analysis of academic attainment among community college students. *Research in Higher Education*, 49, 420-450.
- Bettinger, E. P., & Long, B. T. (2009). Addressing the needs of underprepared students in higher education: Does college remediation work?

- The Journal of Human Resources*, 44, 736-771.
- Boekaerts, M., & Simons, P. R.-J. (1995). *Leren en instructie: Psychologie van de leerling en het leerproces*. Assen, Nederland: Van Gorcum.
- Brants, L., & Struyven, K. (2009). Literature review on online remedial education: A European perspective. *Industry and Higher Education*, 23, 269-276.
- Calcagno, J. C., & Long, B. T. (2008). The Impact of postsecondary remediation using a regression discontinuity approach: Addressing endogenous sorting and noncompliance. *NBER working paper series, no. 14194*.
- Doignon, J., & Falmagne J. C. (1999). *Knowledge spaces*. Berlin: Springer.
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Reviews Psychology*, 53, 109-132.
- Elshout-Mohr, M., Daalen-Kapteijns, M. M. van, & Meijer, J. (2001). *Constructie van het instrument "Rapportage Autonom Studeren"*. Amsterdam: SCO-Kohnstamm Instituut en Instituut voor de Leraren Opleiding (ILO).
- Falmagne, J., Cosyn, E., Doignon, J., & Thiéry, N. (2004). *The assessment of knowledge, in theory and in practice*. Opgehaald op 26 mei 2011 van http://www.aleks.com/about_aleks/Science_Behind_ALEKS.pdf.
- Fraas, J. W. (2007, april). *A comparison of propensity score analysis to analysis of covariance: a case illustration*. Paper gepresenteerd op de jaarlijkse bijeenkomst van de American Educational Research Association, Chicago.
- Grant, H., & Dweck, C. S. (2003). Clarifying achievement goals and their impact. *Journal of Personality and Social Psychology*, 85, 541-553.
- Guo, S., & Fraser, M. W. (2010). *Propensity score analysis: statistical methods and applications*. Los Angeles: Sage.
- Hart, H. 't, Boeije, H., & Hox, J. (red.). (2005). *Onderzoeksmethoden* (zesde druk). Amsterdam: Boom onderwijs.
- Hout-Wolters, B. H. A. M. van. (2009). Leerstrategieën meten. Soorten meetmethoden en hun bruikbaarheid in onderwijs en onderzoek. *Pedagogische Studiën*, 86, 110-127.
- Jamelske, E. (2009). Measuring the impact of a university first-year experience program on student GPA and retention. *Higher Education*, 57, 373-391.
- Kools, Q. H., & Neut, A. C. van der. (2006). *Oplossingen voor kennisdeficiënties*. Tilburg, Nederland: IVA Beleidsonderzoek en Advies. Opgehaald op 26 mei 2011 van http://www.onderwijsraad.nl/upload/publicaties/314/documenten/studie_iva_oplossingen_voor_kennisdeficiënties.pdf.
- Martens, R. & Bastiaens, T. (2005). De relatie tussen autonomie en motivatie. *Onderzoek van Onderwijs*, 14, 37-40.
- Onderwijsraad. (2006). *Versteviging van kennis in het onderwijs; verkenning*. Den Haag, Nederland: Onderw sraad. Opgehaald op 26 mei 2011 van http://www.onderwijsraad.nl/upload/publicaties/313/documenten/versteviging_van_kennis_in_het_onderwijs.pdf.
- Onderwijsraad. (2007). *Versteviging van kennis in het onderwijs II; advies*. Den Haag, Nederland: Onderw sraad. Opgehaald op 26 mei 2011 van http://www.onderwijsraad.nl/upload/publicaties/299/documenten/versteviging_van_kennis_in_het_onderwijs_ii.pdf.
- Onderwijsraad. (2008). *Een succesvolle start in het hoger onderwijs; advies*. Den Haag, Nederland: Onderw sraad. Opgehaald op 26 mei 2011 van http://www.onderwijsraad.nl/upload/publicaties/282/documenten/een_succesvolle_start_in_het_hoger_onderwijs.pdf.
- Picarelli, A., Slaats, M., Bouhuijs, P. A. J., & Vermunt, J. D. (2006). Leerstijl en leeromgeving in het voortgezet onderwijs: Nederland en Vlaanderen vergeleken. *Pedagogische Studiën*, 83, 139-155.
- Ratelle, C. F., Guay, F., Vallerand, R. J., Larose, S., & Senécal, C. (2007). Autonomous, controlled, and amotivated types of academic motivation: A person-oriented analysis. *Journal of Educational Psychology*, 99, 734-746.
- Rienties, B., Dijkstra, J., Rehm, M., Tempelaar, D. T., & Blok, G. (2005). Online bijspijkeronderwijs in de praktijk. *Tijdschrift voor Hoger Onderwijs*, 23, 239-253.
- Rienties, B., Kaper, W., Struyven, K., Tempelaar, D. T., Gastel, L. van, Vrancken, S., et al. (in druk). A review of the role of ICT and course design in transitional education practices. *Interactive Learning Environments*.
- Ryan, R. M., & Deci, E. L. (2000). Self-determina-

- tion theory and the facilitation of intrinsic motivation, social development, and well being. *American Psychologist*, 55, 68-78.
- Schneider, M., Carnoy, B., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). Estimating causal effects: *Using experimental and observational designs*. Washington, DC: AERA.
- Schuetz, P., & Barr, J. (Eds.). (2008). Are community colleges underprepared for underprepared students? Special issue. *New Directions for Community Colleges*, 144.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental Designs for generalized causal inference*. Boston, NY: Houghton Mifflin Company.
- Tempelaar, D. T. (2006). The role of metacognition in business education. *Industry & Higher Education*, 20, 291-298.
- Tempelaar, D. T. (2007). Onderwijzen of bijspijkeren? *Nieuw Archief voor Wiskunde*, 8 (1), 55-59.
- Tempelaar, D. T., Caspers, W. (2008). De rol van de instaptoets. *Nieuw Archief voor Wiskunde*, 9 (1), 66-71.
- Tempelaar, D. T., Gijsselaers, W.H., Schim van der Loeff, S., & Nijhuis, J. F. H. (2007). A structural equation model analyzing the relationship of student personality factors and achievement motivations, in a range of academic subjects. *Contemporary Educational Psychology*, 32, 105-131.
- Tempelaar, D. T., Rienties, B., & Gijsselaers, W. (2006). Internationalisering: En de Nederlandse student? *Onderzoek van Onderwijs*, 35, 40-45.
- Tempelaar, D. T., Rienties, B., & Gijsselaers, W. (2007). Internationalisering: Leerbenaderingen van Nederlandse en Duitse studenten. *Onderzoek van Onderwijs*, 36, 4-9.
- Tempelaar, D. T., Rienties, B., Kaper, W., Giesbers, B., Schim van der Loeff, S., Gastel, L. van, et al. (2011). Mathematics bridging education using an online, adaptive e-tutorial: preparing international students for higher education. In A.A. Juan, M.A. Huertas, S. Trenholm, & C. Steegmann (Eds.), *Teaching mathematics online: Emergent technologies and methodologies*. Hershey, PA: IGI Global.
- Tweede Fase Adviespunt. (2005). *Zeven jaar Tweede Fase, een balans. Evaluatie Tweede Fase*. Den Haag, Nederland: Tweede Fase Adviespunt.
- Vallerand, R. J., Pelletier, L. G., Blais, M. R., Brière, N. M., Sénécal, C., & Vallières, E. F. (1993). On the assessment of intrinsic, extrinsic, and amotivation in education: evidence on the concurrent and construct validity of the Academic Motivation Scale. *Educational and Psychological Measurement*, 53, 159-172.
- Vermunt, J. D. (1992). *Leerstijlen en sturen van leerprocessen in het Hoger Onderwijs*. Amsterdam: Swets & Zeitlinger.
- Yanovitzky, I., Zanutto, E., & Hornik, R. (2005). Estimating causal effects of public health education campaigns using propensity score methodology. *Evaluation and Program Planning*, 28, 209-220.
- Wieland, A., Brouwer, N., Kaper, W., Tempelaar, D. T., Leijen, M. van, Rienties, B., et al. (2007). Didactische scenario's voor remediërend wiskundeonderwijs. *Tijdschrift voor Hoger Onderwijs*, 25, 2-15.

Manuscript aanvaard op: 11 april 2011

Auteurs

Dirk Tempelaar, **Bart Rienties** en **Bas Giesbers** zijn verbonden aan de Maastricht University School of Business & Economics, **Wolter Kaper** en **Leendert van Gastel** aan de Faculteit van Natuurwetenschappen van de Universiteit van Amsterdam, **Evert van de Vrie** aan de Faculteit Informatica, Open Universiteit Nederland, **Henk van der Kooij** aan het Freudenthal Instituut, Universiteit Utrecht, en **Hans Cuypers** aan de Faculteit Wiskunde & Informatica, Technische Universiteit Eindhoven.

Correspondentieadres: Dirk Tempelaar, Maastricht University School of Business & Economics, Departement Kwantitatieve Economie, Postbus 616, 6200 MD Maastricht. E-mail: D.Tempelaar@MaastrichtUniversity.nl.

Abstract

Effectiveness of a voluntary postsecondary remediation program in mathematics

This contribution evaluates a postsecondary remediation program in mathematics, aiming to ease the transition from high school to university and to improve the success rates in the first year of bachelor studies. The remediation program consists of the administration of an entry test and the organisation of voluntary bridging education in the format of an online summer course. Participants are prospective students of the university programs business and economics of Maastricht University, and are mostly students with an international background. Effect analysis suggests a strong treatment effect of successful participation in the summer course, and with lower effect sizes, of being educated in advanced math versus basic math in high school. However, given the quasi-experimental setup of this study, with non-equivalent groups, selection effects may be responsible for a part of that effect. For that reason, the treatment effect is corrected by applying the propensity score method, indicating that indeed a selection effect is present, but that after correction, a substantial treatment effect remains, of about 50% the size of the prior math education effect.