MAGKS

Joint Discussion Paper Series in Economics

by the Universities of

Aachen · Gießen · Göttingen
Kassel · Marburg · Siegen

ISSN 1867-3678

No. 25-2018

Stefan Büchele

Bridging the Gap - how Effective are Remedial Math Courses in Germany?

Coordination: Bernd Hayo • Philipps-University Marburg
School of Business and Economics • Universitätsstraße 24, D-35032 Marburg
Tel: +49-6421-2823091, Fax: +49-6421-2823088, e-mail: hayo@wiwi.uni-marburg.de

Bridging the Gap - how Effective are Remedial Math Courses in Germany?

Stefan Büchele*

July 30, 2018

Abstract

Almost all German universities complain that the math skills of students entering the higher education system do not equal the math skills that are demanded by the universities. To bridge the gap, almost all universities offer some remedial math courses. However, there is only weak empirical evidence for the effectiveness of these courses in Germany. This paper aims to fill this gap by evaluating some math bridging courses given at the Economics Department of the University of Kassel. A key finding: taking a math bridging course on a regular basis will enhance students' math skills and increase the probability of passing the final math exam by 35 percent.

*Department of Economics, University of Kassel, Nora-Platiel-Str. 4, D-34127 Kassel, Germany; E-mail: buechele@uni-kassel.de.

1 Introduction

Lecturers of different subjects all over Germany are reporting poor math skills of freshman students at German universities and universities of applied sciences. This is not only a problem in math-related studies such as Engineering or natural Sciences. Students also struggle with the math required in courses of Economics and Business Administration (EBA). Studies show that many (German) freshman students have a particular lack of secondary school math skills and do not fulfill the demands of post-secondary math education (Abel and Weber, 2014; Laging and Voßkamp, 2017). Therefore, in the case of Economics, the problem is not the classical transition from secondary to tertiary mathematics (Luk, 2005; Gueudet, 2008) as is expected in other courses of study but rather the lack of math skills usually required for a secondary school degree. The consequences are poor grades, overextension and motivational issues in math and math-related subjects as well as an overall high dropout rate for those critical students (Georg, 2009; Heublein, 2014).

For that reason, many higher education institutions are trying to bring unprepared students to a level of skill that allows them to participate more successfully in first-year university (math) courses. In the US, so called remedial courses are widely spread but controversial (Bahr, 2008). In Europe, remedial courses are less common and systematic but, especially in German higher education, math remediation has become a much more frequent subject of discussion and implementation over the last decades. For example, the German government provided two billion Euros from 2011 to 2020 for university teachers and consultants to ensure, among other things, the funding of remedial courses.

Considering the costs and effort of math remediation as well as the amount of theoretical and background research (Bausch et al., 2014; Hoppenbrock et al., 2016) in math remediation, there is very little evidence for the impact of such courses in Germany. In contrast to US studies, the few German studies examining math remediation do not allow causal estimations of treatment effects. There are different reasons for why German research does not focus on that. Firstly, attendance in remedial courses is mostly voluntary in Germany. Unlike at universities in other countries, there are hardly any entry tests or cut-off rates that automatically result in students being placed in remediation. Secondly, students mostly do not have to pass the remedial offers in order to continue their studies in developmental courses. Finally, due to German data privacy regulations and the fact that remediation is (mostly) an optional component of students' education, gathering data is complicated. That is why controlled experiments as well as many other mainly used methods like discontinuity approaches are no options for an impact evaluation. Besides, unlike in the United States, German remedial courses are typically designed for one special course of study. For example, at the University of Kassel, there are different preparatory and bridging courses in math for studies in EBA, Mathematics or Engineering instead of one offer that covers all degree programs.

To estimate causal effects of a remedial math course I will run a difference-in-difference regression that is controlled through baseline propensity score matching and study specific variables. To this end, a sample of 155 students enrolled in courses of Economics and Business Administration is taken into account. A second sample of 98 students is used to measure a medium-term effect within the final exam pass rate of the developmental math course. I find that attending the remedial course on a regular basis has a significant positive effect on the improvement of secondary school math skills. Furthermore, it results in a higher probability to pass the final exam in the developmental math course.

The paper is organized as follows. An overview on the common literature and a description of math remediation in Germany is given in Section 2. Section 3 provides information about the design of the study and the data used in the empirical approach. Section 4 explains the methodical procedure. The results are presented in Section 5 and section 6 discusses and concludes the paper.

2 Background

2.1 Literature

As mentioned, there is hardly any evidence of an impact of remedial courses for German universities. Studies are either poorly designed and do not measure a quantitative effect (e.g. Greefrath and Neugebauer, 2017; Greefrath and Hoever, 2016) or are not searching for a treatment effect in the first place. For example, in Laging and Voßkamp (2017), positive effects of a math preparatory course are just noticed as side effects of other examinations and not suitable for causal interpretations. German research in math remediation is done more on a theoretical and educational basis. Most studies are examining how to structure math remediation (e.g. blended-learning formats), not whether the programs have any impact on math skills or related factors (see e.g. Bausch et al., 2014 ; Hoppenbrock et al., 2016).

International research, mainly US studies, give an insight into a possible impact of math remediation. Ahead of an extensive number of small-scaled and poorly designed studies[1] are current large-scaled and methodologically and statistically strong studies with mixed results regarding the effect of math remediation or remediation in general. Bettinger and Long (2009), for example, look at 28,000 US students controlled via an Instrumental Variable approach and find that remediation has a positive effect on students' persistence. For instance, students enrolled in math remediation are less likely to drop out. In addition to a lower dropout rate, students who successfully pass math remediation in community colleges have the same probability of transferring to a four-year-college as students that achieve the required math skills without remediation (Bahr, 2008). Evidence from Europe is given by De Paola and Scoppa (2014) who estimate effects of remediation at an Italian university with a regression discontinuity approach. They conclude that students with remedial background have a lower dropout rate as well as a higher number of credits after two years. Boatman and Long (2018) run a regression discontinuity approach and, although they do not find an impact

[1] For an overview see Bahr, 2008

of math remediation on math skills for students near the cut-off rate, they are able to measure positive effects for students with lower math skills. While a study from Scott-Clayton and Rodriguez (2015) does not find that difference between students of various skill-levels, they do report an overall positive effect of math remediation on math skills.

This provides a general problem in the evaluation of math remediation. Discontinuity approaches are widely spread, because they allow an easy measurement of treatment effects without controlling for too many confounding variables if there is a hard cut-off. Since only students near the cut-off rate are taken into account, one can only estimate effects for students at the margins of needing, but not for those who are generally weaker and way below the average skill-level.

Dadgar (2012), for example, examines a group of students in remediation with the lowest math skills. She finds that students required to take only two instead of three remedial classes have a significant higher likelihood of gaining an associate degree. In addition, other studies question the effect of math remediation. Unlike in Germany, remediation is highly controversial in the United States. Critics argue that students who need remediation should not be allowed to attend higher education institutions in the first place; they fear a decrease in the quality of academic outcomes. Students could also suffer from negative peer-effects and the stigma associated with remediation could harm educational outcomes as well as student efforts. Furthermore, it seems clear that remediation causes high costs for students and taxpayers in the form of tuition fees and paying twice for the same schooling[2]. In contrast to the studies finding a positive effect of remediation, the results of other papers not only question the above-mentioned aspects, but also the positive academic outcome in general. Especially Lagerlöf and Seltzer (2009), investigating the effectiveness of math remediation on subjects of economics, could not find a significant effect by estimating a difference-in-difference regression. Di Pietro (2014), Martorell and McFarlin (2011) and Calcagno and Long (2008) do not find any positive evidence for short- or long-term effects of remedial course takers as well. However, with the exception of

[2]For more literature of this controversial discussions see Bahr (2008)

Lagerlöf and Seltzer (2009)), they are all estimating the effects with regression discontinuity approaches which do not consider average treatment effects. In contrast, I will operate with a difference-in-difference estimation and will measure the treatment effects for all students in remediation.

2.2 Math remediation in Germany

Although there is a large number of studies evaluating the effectiveness of remediation in other countries, it cannot be taken for granted that the results can be transferred to the German higher education system as different structures and conditions apply to German remediation. Scott-Clayton and Rodriguez (2015) categorize remediation in three, generally discussed topics, namely skill development, the discouraging effect of remediation and lower heterogeneity in college classrooms. For German research, only the first and third aspect is of relevance. Generally, it is assumed that remediation courses have positive effects on students' math skills whereas the degree of heterogeneity is decreased. However, empirical evidence is rare. Due to voluntary participation, one should not assume discouragement or negative peer effects through remediation particularly because remedial courses and support offers in general are, after all, often used by students of all skill level (Voßkamp and Laging, 2014). Therefore, the paper will help investigate the effect of math remediation on math skills while the topics of heterogeneity and discouragement are no longer seen as a central matter of the examination.

There are mainly two kinds of remedial courses in Germany. A preparatory course is a block event of usually 2-5 weeks set prior the beginning of the first semester. Topics are mostly revisions of secondary school maths or preparations for the upcoming course of study. A so-called bridging course takes place during the semester (up to 14 times and once a week) and also runs parallel to the main math lecture. Topics are typically coordinated with the main lecture and can widely spread from secondary to tertiary math. Both types of remediation aim to improve students understanding of math in order to succeed in their studies; but while a preparatory course refreshes math skills for the whole degree program, a bridging course is designed to help students better understand the contents of the main math lecture by repeating basics at a point of time when they are especially relevant to the lectures' topics.

Typically, remediation at larger universities is not centralized. Consequently, every department or course of study related to math offers its own preparation or bridging course. Participation and attendance in remedial courses is usually not mandatory or controlled by the lecturers. This aspect complicates the measurement of causal effects many times over. First, due to the students' choice of taking remediation, evaluations and estimations of treatment effects are biased by self-selection. Secondly, there is no structural benefit in math entry tests if students are not forced to attend or pass remediation courses. Common methods like discontinuity approaches cannot be applied for the estimation of causal effects in such cases. And finally, due to the local organisation in each department and also through the voluntary character of remediation, there is no control mechanism for which students attend or pass these courses. Data has to be gathered by the researcher trough questionnaires and skill tests in every course separately. Additional complications are posed by the German data regulation which requires that the data is collected on a completely anonymous basis. Therefore, one cannot easily fall back on university databases. Especially the estimation of long-term outcomes or labour market effects of remediation is impossible.

2.3 Math remediation at the Department of Economics at the University of Kassel

The Department of Economics at the University of Kassel offers math remediation since 2009. Freshman students can take part in a two-week preparatory course before, and a bridging course during the semester. The courses are mainly designed for students in studies of EBA and Educational Economics (EE). Up to 500 students enroll in these degree programs every year and all of them have to pass the final math exam in some point in their studies, but not necessarily in their first semester. Although there is a prescribed plan of study that recommends to take the math course during the first year of study, students at the University of Kassel, as well as in many other universities in Germany are free with the regards to the point in time when they take certain classes. Therefore, students can decide on their own whether they want to take the math course in their first or their sixth semester, for example.

As usual, enrollment and attendance in math remediation is not compulsory. It is the students' choice whether they take the offer of remediation or not. An official registration is not required and, even if enrolled, course attendance is not mandatory. The students' acceptance of math remedial courses is usually high, but on average, enrolled students just attend about two thirds of the lessons.

In addition to remedial courses the department of Economics offers several other support programs (e.g. tutorials or online tests) and an optional entry test that is taken during the first session of the main math lecture. An additional questionnaire collects information about the students' educational, social and motivational variables. The acceptance of the test and questionnaire is very high. Every year, about 400 students take part and get feedback on their math skills. The results show that the need for math remediation at the University of Kassel is high. On average, freshman students only reach about 20% of the overall points, while the heterogeneity of the student population is extremely high. Most of all, students' educational biographies differ. At most universities in the federal state of Hessen and hence at the University of Kassel, students can enroll after a shortened secondary school track of 12 instead of the usual 13 years which they can complete at vocational schools. Usually, students with the shortened 12-year-degree only have access to universities of applied sciences but not to classical universities which makes the Universities in Hessen a special case in Germany. The issue for students from the short track (German: "Fachhochschulreife") is that they suffer from a reduced and simplified secondary school curriculum which does not only apply to math but other subjects as well.

Although there are two remedial courses in Economics at the University of Kassel, this study takes a closer look at the bridging course which takes place every week during the semester. The preparatory course is controlled as a confounding variable. With a total number of four weekly lecture hours, the bridging course has the same class time as the main math lecture. Students can attend up to 12 lectures and are supplied with exercises for each session. The course's topics are coordinated with the main course's content, with a focus on secondary school math. Each student's attendance and participation is voluntary.

3 Design and Data

The purpose of this study is to give causal evidence for the effects of a remedial math course at a German university. One questions of interest is whether students attending remedial courses have an advantage over their fellow students, or, can at least compensate their skill differences to match the level of the students who had no need to participate in remedial courses. That focus on the students' heterogeneity and its possible decrease due to remediation can, because of the small scale of the study, not be followed. It is, however, a topic for future large-scale examinations taking place at the University of Kassel's Department of Economics at the moment. Another, more fundamental question and one this study does seek to examine is whether the bridging course affects students' short- and medium-term math outcomes.

Therefore, I will compare the math outcomes of students who attended the bridging course on a regular basis (where attendance rate equals or exceeds 2/3 of the lectures) with those who attended fewer sessions or did not take part at all. The effects are measured with a controlled difference-in-difference approach under propensity score-matched premises.

3.1 Design

The study is built on a quasi-experimental design with a treatment group (bridging course participants) and a control group (non-participants). The collection of data took place in the winter semester of 2016 (October 2016 to February 2017) at the the Departments of Economics at the University of Kassel.

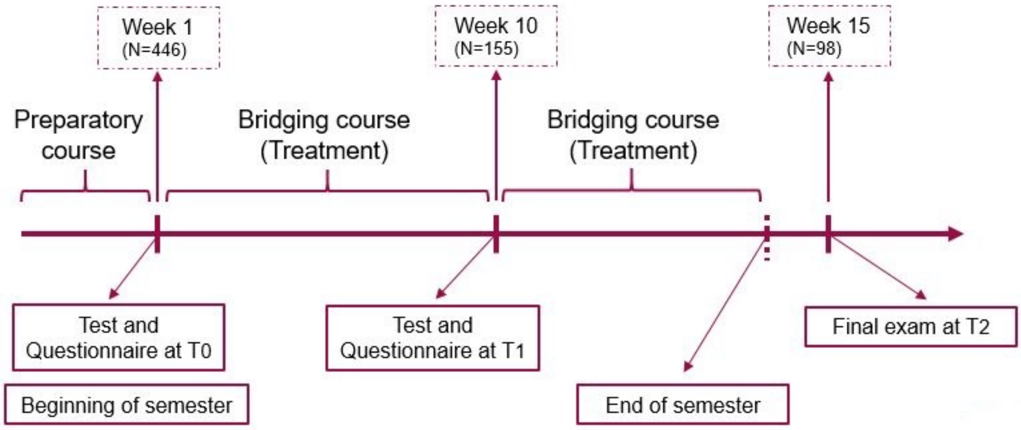

Figure 1: Overview of the study design

Data was first gathered anonymously in the first lecture of the main math course (T_0) at the beginning of the semester, then again after 10 weeks (T_1), also in the main course, and once more one week after the end of the semester in the form of the final exam results (T_2). Participation was completely optional and 446 students answered the questionnaire and took the math skill test at T_0, while 155 of them also participated in a similar test and questionnaire at T_1. Proof of passing the final exam (T_2) could only be matched to 98 of the remaining 155 students. Because of the three points of time, the effect of the bridging course can be estimated for short-term (at T_1) and mid-term (at T_2) math outcomes. Therefore, I will work with two samples. The first sample (N=155) will give information about the short-term effect, while the second sample (N=98) provides insight into whether the bridging course raises the probability of passing the final exam in the development math course.

The entry test at T_0 and midterm test at T_1 are math skill tests, consisting of 30 tasks of secondary math schooling (e.g. terms, equations, functions and calculus). Both tests are about equally difficult, being composed of different but comparable tasks (also see Laging and Voßkamp, 2017).

The students were also asked to answer questionnaires at both points in time. The collected data includes different confounding variables needed for the causal estimation of the bridging course. In addition to the outcome of the skill tests and final exam, this provides information on study-specific variables and educational-biographical variables as well as variables measuring the use of math support during the semester.

3.2 Data

In this section, I will give an overview of the variables and their measures. Controlling for these variables reduce the selection-bias. Due to the present study design, the variables are sorted in three blocks. The first of which comprises the outcome variables of the skill tests and final exam. Secondly, there are time-independent baseline variables that do not change during the semester. The third block consists of the semester variables which measure the students' ongoing activities during T_0 and T_1. Information will include the coding, value, number of items, means and standard derivation (SD) of all variables.

3.2.1 Outcome variables

In Table 1, one can find the outcome measures of the students' math skills. Y_0 and Y_1 describe how many points (out of a maximum of 30) a student gathered in the skill tests at T_0 and T_1. The dummy variable Y_2 depicts whether a student did pass the final exam (1) or not (0). The pass-rate of the final exam is 65%. The variable ΔY gives the difference in points between Y_1 and Y_0 and is mainly used for further analysis.

Interesting but not surprising is the difference in means of the two samples. In both tests, the second sample (N=98) performed better than the first one (N=155), which is to be expected. Since the second sample will still be taking the final exam, while other students dropped out of the math course during the semester (by their own choice), the students have ex-ante different characteristics (see Tables 2 and 3).

3.2.2 Baseline variables

The block of baseline variables is pooled by social, educational and biographic variables that are time-independent and used to be typical determi-

Table 1: Outcome variables compared by different samples

Code	Description	Type	Value	Items	Sample 1 Mean (SD)	Sample 2 Mean (SD)
Y_0	math skill at T_0	metric	0 to 30	30	8.50 (5.06)	9.66 (5.38)
Y_1	math skill at T_1	metric	0 to 30	30	11.16 (5.56)	12.75 (5.54)
Y_2	pass final exam	binary	0 or 1			0.65
ΔY	$Y_1 - Y_0$	metric			2.66 (3.78)	3.09 (3.97)
N					155	98

nants of academic performance and math performance of students (Laging and Voßkamp, 2017; Mallik and Shankar, 2016; Mallik and Lodewijks, 2010; Byrne and Flood, 2008; Krohn and O'Connor, 2005).

Some variables require further explanations. B_3 is a dummy variable that checks if a student has already participated in the main math course during a previous semester and did not take or pass the final exam. The course of study is controlled by B_4. Students are either enrolled in the degree program of EBA or EE. As mentioned above, students can enroll in both courses of study with a short-track secondary school degree (B_5) that is quite different compared to the regular degree. The structure of grades and GPAs in Germany is, compared to other countries, different. Variables B_6 and B_7 measure the prior high-school GPA and the math grade over the last two to three years in secondary education. The lower the value of the grade, the better the students are in their academic outcomes. The education gap (B_8) measures the time between the high school degree and the beginning of higher education studies in years. For B_9 students were asked on a scale from one to five how they judge their own maths skill in general. Whether a student participated in the math preparatory course (two weeks block lecture set prior T_0) is included with B_{10}.

Table 2: Baseline variables compared by different samples

Code	Description	Type	Value	Sample 1 Mean (SD)	Sample 2 Mean (SD)
B_1	Gender (female = 1; male=0)	binary	0 or 1	0.53	0.54
B_2	Year of study	metric	1 to 3	1.16 (0.45)	1.20 (0.50)
B_3	Math course already taken yes = 1)	binary	0 or 1	0.08	0.11
B_4	Course of study (EBA = 1; EE = 0)	binary	0 or 1	0.78	0.77
B_5	Graduation type (13-year-degree = 1)	binary	0 or 1	0.78	0.86
B_6	Prior GPA (lower is better)	metric	1 to 4	2.58 (0.56)	2.50 (0.57)
B_7	Math grade in sec. school (lower is better)	metric	1 to 5	2.62 (0.88)	2.55 (0.85)
B_8	Education gap	metric	0 to 9	1.57 (1.83)	1.64 (1.93)
B_9	Math self-efficacy	metric	1 to 5	3.13 (0.81)	3.11 (0.82)
B_{10}	Participation preparation course (yes = 1)	binary	0 or 1	0.71	0.71
N				155	98

3.2.3 Semester variables

With the time-dependent semester variables, I will control the students' learning and engagement habits during T_0 and T_1. By means of the questionnaire in T_1, students were asked for their use of the given learning opportunities and their weekly learning hours. Besides the math remedial offers, students can use several support services which are all optional. Of course, they can attend the math main lecture (S_1). In addition, they can participate in weekly math tutorials which are held by senior students (S_2). Further training possibilities are posed by the exercise sheets (S_3) and online math exercises (S_4). The data for (S_1) to (S_4) was raised on a scale from one to

Table 3: Semester variables compared by different samples

Code	Description	Type	Value	Sample 1 Mean (SD)	Sample 2 Mean (SD)
S_1	Main math lecture attendance	metric	1 to 6	4.59 (0.64)	5.63 (0.60)
S_2	Math tutorial attendance	metric	1 to 6	3.65 (1.68)	4.95 (1.45)
S_3	Completion of exercise sheets	metric	1 to 6	4.02 (1.31)	5.23 (1.12)
S_4	Completion of online tests	metric	1 to 6	1.35 (1.47)	2.55 (1.55)
S_5	Study hours (not counting lecture attendance)	metric	0 to 20	4.85 (3.66)	5.05 (3.08)
S_6	Number of other courses during semester (excl. math)	metric	0 to 6	4.30 (7.75)	3.76 (1.24)
S_7	Hours working	metric	0 to 35	4.10 (7.66)	4.48 (7.89)
N				155	98

six. Therefore, students were asked how often they make use of the support services (1 = no time; 6 = at all times). The variable S_5 provides information about the weekly hours students study for math, not counting the attendance in lectures or tutorials, while S_7 measures the weekly working hours spent to earn one's livelihood. Because one cannot assume that these variables are comparable in their means within the treatment and control group, but most certainly have some impact on the test results at T_1, they will be controlled for in the final model.

4 Method

Due to the design of the study, a difference-in-difference approach is considered as a proper method to estimate the causal effects of the math bridging course. Although, having a fixed effects estimation with the skill tests at T_0 and T_1, one should not assume causal effects without any control of the given variables. Because the students are free in their treatment choice, one has

Table 4: Variables compared by treatment and control groups

Code	short description	Sample 1		Sample 2	
		Treatment	Control	Treatment	Control
Y_0	Test outcome at T_0	7.30	9.08	8.09	10.53
Y_1	Test outcome at T_1	11.50	11.00	12.24	13.03
Y_2	Final exam pass-rate			0.57	0.70
ΔY		4.20	1.91	4.16	2.50
B_1	Gender	0.58	0.50	0.54	0.54
B_2	Year of study	1.18	1.15	1.26	1.17
B_3	Course already taken	0.08	0.09	0.14	0.10
B_4	Study program	0.78	0.78	0.74	0.78
B_5	Graduation type	0.74	0.80	0.83	0.87
B_6	Prior GPA	2.74	2.50	2.74	2.37
B_7	Math grade	2.86	2.51	2.92	2.35
B_8	Education gap	2.08	1.32	2.43	1.21
B_9	Self-efficacy	3.51	2.94	3.54	2.87
B_{10}	Prep. course participation	0.74	0.70	0.74	0.70
S_1	Lecture attendance	4.70	4.54	5.66	5.62
S_2	Tutorial attendance	3.96	3.50	5.00	4.92
S_3	Completion of exercise sheets	4.32	3.88	5.49	5.10
S_4	Completion of online tests	1.46	1.30	2.86	2.38
S_5	Study hours	6.00	4.31	6.59	4.19
S_6	Number of other courses	3.70	3.69	3.60	3.84
S_7	Working hours	3.37	4.45	5.46	3.93
N		50	105	35	63

to deal with self-selection and therefore, biased estimations. Table 4 gives an insight into the means of all variables compared by treatment groups and control groups of both samples.

The means of essential confounding variables, which can be assumed to have an influence on the treatment choice as well as on the math performance in the skill tests, differ in the treatment and control groups. As treated students show fewer math skills in the math entry test (Y_0), the treatment group has overall worse preconditions in both samples.

4.1 The need to control variables

Table 5 shows that the baseline variables affecting math performance (as example in the skill test at T_0) as well as students' treatment choice. Therefore,

Table 5: OLS and logistic regression results for confounding variables at T_0

Baseline variables	short description	Standardized coefficients (OLS)	Coefficients (LR)
B_1	Gender	-0.29***	0.47
B_2	Year of study	-0.08	-0.12
B_3	Course already taken	0.22**	-0.30
B_4	Study program	-0.05	0.21
B_5	Graduation type	0.37***	-0.54
B_6	Prior GPA	-0.31***	0.75
B_7	Math grade	0.09	-0.23
B_8	Education gap	0.15*	0.21*
B_9	Self-efficacy	-0.25**	0.94**
B_{10}	Prep. course part.	0.14*	0.21
Dependent variable		Points at T_0 (Y_0)	Treatment choice
N		155	155
Adj. R^2		0.428	
Pseudo R^2			0.221 (Nagelkerkes)

*** p<0.001; ** p<0.01; *p<0.05

the baseline variables correlations with the test outcome Y_0 and treatment choice (dummy-variable) were estimated within an ordinary-least-square regression (OLS) and a logistic regression (LR).

It is crucial to consider the baseline variables, particulary the differences in B_1, B_3, B_5, B_6, B_8 and B_9, as they show mostly a negative correlation between the math performance and the students' treatment choice and can thus be defined as confounding variables. Although these variables are time-independent and fixed (influencing Y_0 as well as Y_1), one can assume that the control group has a higher learning speed and a higher understanding which can, in turn, have an influence on the parallel trend of the difference-in-difference approach. Ignoring these differences could result in a biased estimation.

Besides the baseline variables, the semester variables can affect the parallel trend as well. Comparing the means from both samples (Table 4), it stands out that the students in the treatment group show more attendance in lectures and tutorials as well as more engagement in learning math and every other support offer, while working less to make their living. Although the second sample shows less heterogeneity, there are still differences that can result in a biased estimation. All in all, one can easily assume that the treatment effect will be overestimated if semester variables were not controlled. Altogether, there are two issues that are violating central assumptions of the difference-in-difference approach.

1. The outcome variable (Y_0) at T_0 does not affect the treatment choice.

This assumption is clearly violated, not only because a logistic regression shows that the results of the skill test at T_0 have a significant influence on the treatment choice, but also because the students are urged to attend the bridging course if they perform poorly in the entry test.

2. The treatment and control group exhibit a parallel trend over time.

As seen in Table 4, the not-randomized groups differ in their means of confounding baseline and semester variables. But because these variables are, in addition to the bridging course, responsible for the growing math skills, this difference in means will affect the parallel trend between treatment and control groups. While the semester variables can be easily controlled for in the difference-in-difference regression, the baseline variables have to be checked in an additional step.

4.2 Control for baseline variables

To control for the bias caused by the baseline variables I run a logistic regression and calculate propensity scores (Rosenbaum and Rubin, 1983; Stuart, 2010). The treatment and control group will be matched with a 1:1 nearest neighbour algorithm. Dependent variable of the logistic regression is the treatment participation, while all baseline variables are taken into account as independent variables. After various samplings, taking all baseline variables and not only the clearly confounding variables into the propensity score calculation leads to the best matching results. The distribution of propensity scores can be seen in figure 2.

Table 6 shows that the propensity score matching reduces major differences in means in both samples. Compared to Table 4, especially the means

Sample 1 Sample 2

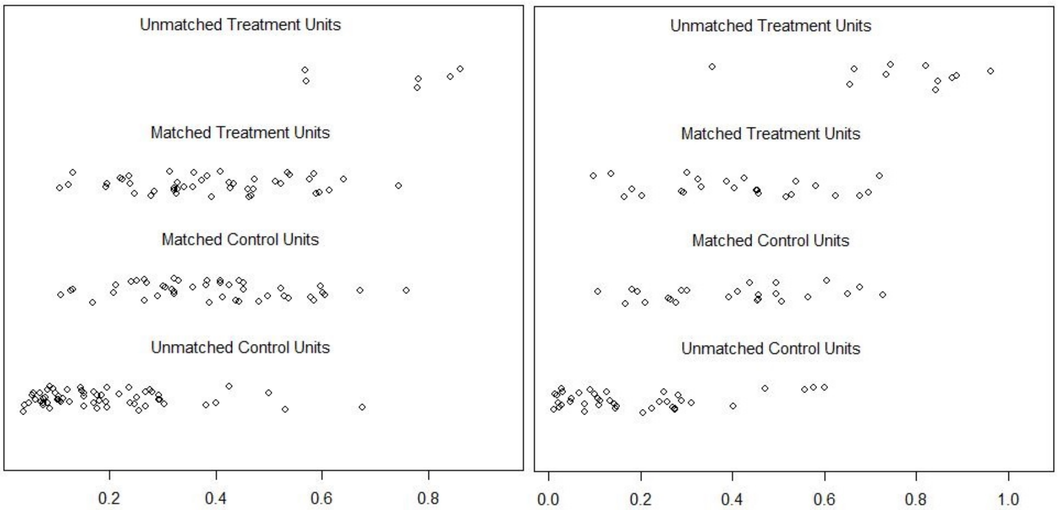

Figure 2: Distribution of propensity scores for both samples

of key variables such as B_3, B_5, B_6, B_8 and B_9 are now similar. Unfortunately, the gender (B_1) cannot be matched perfectly; however, due to the control for further variables and the fixed effects design, this is not assumed to have a strong influence on the parallel trend.

Apart from the baseline variables, the outcome of the entry test Y_0 is nearly the same for the first sample. These results solve the violation that the test outcome does affect the students' treatment choice as well as the baseline variables affecting the parallel trend of math skill growth during the semester. The outcome difference for the second sample is only reduced from about 2.5 points (see table 4) to 1.6 points but with the effect that after the matching is completed, the treatment group performs better than the control group. Because only students who had performed poorly were urged to participate the bridging course, the treatment choice of the students is not affected by the result at T_0 anymore.

Although the issue of self-selection seems solved, propensity score matching results in a major loss of degrees of freedom. After the matching is done, the population of sample 1 is down to N=88 and sample 2 only has N=48 students left. Having had only limited number of cases to begin with, the

Table 6: Comparison of matched treatment and control groups

Baseline variables	Sample 1		Sample 2	
	Treatment	Control	Treatment	Control
Y_0	7.85	7.50	8.88	7.25
Y_1	11.92	8.93	12.38	9.06
ΔY	4.07	1.43	3.50	1.81
B_1	0.52	0.45	0.58	0.63
B_2	1.20	1.18	1.29	1.25
B_3	0.09	0.09	0.13	0.13
B_4	0.80	0.80	0.83	0.88
B_5	0.77	0.77	0.83	0.83
B_6	2.72	2.86	2.68	2.71
B_7	2.79	2.96	2.69	2.91
B_8	1.86	1.80	1.67	1.29
B_9	3.41	3.36	3.29	3.38
B_{10}	0.70	0.70	0.75	0.75
N	44	44	24	24

loss of nearly half of the sample's size is not acceptable for further examinations. A more promising approach to check for the baseline variables under propensity score conditions but without running a matching algorithm and consequently losing cases is given by Hirano, Imbens, and Ridder (2003).

4.3 The Model

To ensure that the baseline variables are taken into account, I will estimate a fixed effects weighted least square (WLS) model. Hirano, Imbens, and Ridder (2003) show that taking $\frac{PS(X)}{1-PS(X)}$, with $PS(X)$ as the propensity score of student X, as the regression weight for non-treated individuals and 1 for treated individuals, results in an efficient difference-in-difference estimator (see also Khandker, Koolwal, and Samad, 2010). But first, for comparison, the treatment effect is measured without any controls for the given variables.

Model 1:

$$\Delta Y = \alpha + \beta T + \varepsilon \quad (OLS)$$

The first model is estimated within a standard OLS regression with ΔY

as the outcome difference between T_1 and T_0, α as the constant and ε as the error term. T is a treatment dummy and β the difference-in-difference estimator.

The second model is estimated with a WLS regression including the above mentioned propensity score weighting. The regression function is the same as in the first model but the estimation now controls for the variables B_1 to B_{10}, as all of them were taken into account for the calculation of $PS(X)$.

Model 2:

$$\Delta Y = \alpha + \beta T + \varepsilon \quad (WLS)$$

The third model is complemented with the block of semester variables to control the students learning behavior between the points of time T_0 and T_1.

Model 3:

$$\Delta Y = \alpha + \beta T + \sum_{i=1}^{7} \gamma_i S_i + \varepsilon \quad (WLS)$$

5 Results

5.1 Short-term effect

Table 7 shows the short-term treatment effects of the math bridging course for different controls of the variables.

Table 7: Short-term treatment effects for the first sample

	Model 1 - OLS	Model 2 - WLS	Model 3 - WLS
Constant (α)	1.914*** (0.355)	1.643*** (0.386)	0.005 (1.968)
treatment effect (β)	2.286*** (0.625)	2.557*** (0.547)	2.051*** (0.575)
Controlled for baseline variables	No	Yes	Yes
Controlled for semester variables	No	No	Yes
N	155	155	155
Adj. R^2	0.074	0.119	0.166

*** p<0.001; ** p<0.01; *p<0.05

In all models, the treatment effect is positive and highly significant at a 0.1% level. Students attending the bridging course on a regular basis do exhibit, on average, a higher increase in their math skills. The difference-in-difference effect varies, depending on which model is used, from 2.05 up to 2.56 points. The OLS estimation in the first model calculates an average treatment effect of about 2.3 points, without any control for the confounding variables. Taking the baseline variables into account as propensity score weighting, the effect increases. The WLS regression in model 2 estimates an effect of 2.56 points. This means that the not controlled effect of the first model is, at first, slightly underestimated. That makes sense, keeping in mind that the treatment group had worse preconditions with regards to their math skills as indicated by their test outcomes such as their secondary school math grade. Model 2 controls for these variables and, therefore, gives a more accurate estimation.

The third model controls for the semester variables as well, and reduces the effect from 2.3 points (model 1) or 2.56 points (model 2) to 2.05 points. This picture matches the information in Table 4 because the treatment group shows, on average, more engagement in attending lectures, tutorials or other math support programs. This influences the difference-in-difference estimator in a positive manner, although it cannot be attributed to the treatment. Altogether, the sample shows the importance of controlling for confounding variables since the effects can easily be over- or underestimated, even if the range of the bias lies just within 0.15 standard derivations.

5.2 Mid-term effect

For the estimation of the mid-term effect I am following the same method as for the short-term effect at first but using the second sample. Table 8 shows the average short-term treatment effects of the math bridging course for this sample.

The results are almost comparable to the short-term effects of the first sample. Small differences can be seen in the first model and in the transition from the second to the third model. The first model only shows a treatment effect of 1.66 points and underestimates the actual effect by 0.5 points. Model

Table 8: Short-term treatment effects for the second sample

	Model 1 - OLS	Model 2 - WLS	Model 3 - WLS
Constant	2.500*** (0.493)	2.006*** (0.552)	-0.573 (4.101)
treatment effect	1.657* (0.824)	2.151** (0.776)	2.083* (0.892)
Controlled for baseline variables	No	Yes	Yes
Controlled for semester variables	No	No	Yes
N	98	98	98
Adj. R^2	0.03	0.064	0.127

*** p<0.001; ** p<0.01; *p<0.05

Table 9: Skill test performance and final exam passing probability

	Beta	Exp(B)
Constant	-1.336* (0.049)	0.263
Y_1	0.164** (0.601)	1.178
N	98	
Pseudo R^2	0.183 (Nagelkerkes)	

*** p<0.001; ** p<0.01; *p<0.05

2 shows similar effects as above but the gap between the second and third model is almost nonexistent. That is because the treatment and control group of the second sample, while showing more heterogeneity in their preconditions (B_i), have similar values with regards to their semester variables. All in all, the short-term effect is still significant and seems very steady, stabilizing at around 2 points in the third model.

To estimate the mid-term effect of the bridging course, I will take a closer look at the final exam result of the students in the second sample. Table 9 shows the logistic regression results for the correlation between the test performance (Y_1) and the probability to pass the final exam (Y_2).

The results show that the performance Y_1 has a significant and positive impact on the students' chance of passing the final exam. More precisely, having one additional point in the skill test leads to a 17.8% higher probability of passing the final exam. Combined with the casual treatment effect of about two additional points in the skill test at T_1, one can assume an about 35% higher probability of passing the final exam in the case of treated students.

6 Discussion and concluding words

This study shows how causal effects of higher educational treatments can be measured even if one cannot rely on popular methods like discontinuity approaches. A difference-in-difference design combined with propensity scores for weighting leads to an efficient estimator, solves self-selection issues and can also control for time-dependent variables. Even more, one can estimate average treatment effects for the whole sample and not only for students at the margin of needing the remedial offers. The results show that regular participation in the bridging course affects secondary math skills and raises the probability of passing the final exam. Furthermore, one can assume that the bridging course as well as all other math offers (S_i) of the Department of Economics reduce the heterogeneity of the students, since the test results in both the treatment and control group are nearly the same at T_1 (see Table 4). Although the results are clear so far, they need further discussion.

The differences of treatment effects between the first (not controlled) and third (fully controlled) model are just at about 0.2 or 0.4 points, with an average treatment effect of about 2 points for both samples which means that, at least in the first sample, there is hardly any selection-bias. This seems astonishing at first, since the preconditions of the students (B_i) in the treatment group are much worse (see also Table 4), but is, after further consideration, not surprising. At first, the fixed-effects model covers for most of the preconditional bias. Even if one does not control the baseline variables for the parallel trend of the treatment and control group, this should result in a minor bias. Second, as can be seen by means of the model 2, the worse prerequisites are compensated by the students' learning behavior during the semester. This compensation leads to a reduction of the selection-bias in these cases. But it cannot be taken for granted that students always behave like this and it should be pointed out that this small-scale study

Table 10: Comparison of difference-in-difference effects

Model	Sample 1	Sample 2
WLS Model 2	2.56	2.15
Nearest Neighbour Matching	2.64	1.69
Difference	-0.08	0.46

offers no universal proof that the selection-bias of such treatments always is compensated by students' learning habits.

Furthermore, with the approach of Hirano, Imbens, and Ridder (2003), an unusual method was applied. Due to the limited scope of the study there are no further options to control for time-independent variables, since the matching algorithm leads to a major loss in degrees of freedom. Having a look at the results of the second model and compare that with the estimated difference-in-difference effect of the matched samples, it is pointed out how efficient that method can be (see Table 10). As the WLS-model estimates a short-term difference-in-difference effect of 2.56 for sample 1, the effect estimated with the matched sample is at 2.64 points, showing a minor difference of 0.08 points. The difference of the second sample is higher but, due to the small matched sample that remains, this cannot be seen as clear evidence against propensity score weighting.

Bridging courses seem to be a appropriate remedial measure to raise students' skills to a level matching that of their non-treated fellow students. But even if remedial courses have positive effects and can reduce heterogeneity, students are often still not adequately prepared for higher education, especially in maths. Overall, having in mind that both skill tests require secondary school math, with an average of about 12.2 points for treated and 13 points for non-treated students (out of 30 maximum points), the students' second skill test results are still poor.

All in all, there are limitations for this study. Although the methodical approach is robust and allows causal estimation of a treatment effect, the samples sizes are small. With samples of only 155 and 98 students it cannot be taken for granted that these effects could be shown in the same way in other semesters. Furthermore, even if the results are good for this specific bridging course, the success of the treatment, of course, highly depends on the lecturer and the topics that are taught. That means that this study shows how one can evaluate this kind of remedial course but the results should not be easily generalized. Therefore, larger samples and evaluations of not only one but various semesters, teachers and institutions are required.

Meanwhile, math remediation is an important part of higher education systems, in particular, having the growing numbers of students in mind. Therefore, remediation is used to compensate the insufficient math skills of students after secondary school. The discussion of the use of math remediation is not as heated in Germany as it is in the US at the moment. However, in Germany a higher graduation rate is wanted by higher education institutes themselves and federal policy. This results in a higher number of (unprepared) students. For this reason, remedial offers are highly required in Germany and are not questioned in general. All in all, the math remediation in Germany needs further examinations. Especially pre-university preparatory courses are extremely popular, but their beneficial effect is unclear, in particular with regards to medium- or long-term effects, since there are no evaluations as to whether the skill improvements of these courses are lasting.

References

Abel, Heinrich and Bruno Weber (2014). "28 Jahre Esslinger Modell – Studienanfänger und Mathematik". In: *Mathematische Vor- und Brückenkurse: Konzepte, Probleme und Perspektiven.* Ed. by Bausch, I.; Biehler, R.; Bruder, R.; Fischer, P.; Hochmuth, R.; Koepf, W.; Schreiber, S.; Wassong, T., pp. 9–19.

Bahr, Peter Riley (2008). "Does mathematics remediation work?: A comparative analysis of academic attainment among community college students". In: *Research in Higher Education* 49.5, pp. 420–450.

Bausch I.; Biehler, R.; Bruder R.; Fischer P.; Hochmuth R.; Koepf W.; Schreiber S.; Wassong T., ed. (2014). *Mathematische Vor- und Brückenkurse. Konzepte, Probleme und Persepektiven.*

Bettinger, Eric P. and Bridget Terry Long (2009). "Addressing the Needs of Underprepared Students in Higher Education: Does College Remediation Work?" In: *The Journal of Human Resources* 44.3, pp. 736–771.

Boatman, Angela and Bridget Terry Long (2018). "Does Remediation Work for All Students? How the Effects of Postsecondary Remedial and Developmental Courses Vary by Level of Academic Preparation". In: *Educational Evaluation and Policy Analysis* 40, pp. 29–58.

Byrne, Marann and Barbara Flood (2008). "Examining the relationships among background variables and academic performance of first year accounting students at an Irish University". In: *Journal of Accounting Education* 26, pp. 202–212.

Calcagno, Juan Carlos and Bridget Terry Long (2008). "The Impact of Postsecondary Remediation Using a Regression Discontinuity Approach: Addressing Endogenous Sorting and Noncompliance". In: *National Bureau of Economic Research Working Paper Series.*

Dadgar, Mina (2012). "Essays on the Economics of Community College Students' Academic and Labor Market Success". PhD thesis.

De Paola, Maria and Vincenzo Scoppa (2014). "The effectiveness of remedial courses in Italy: A fuzzy regression discontinuity design". In: *Journal of Population Economics* 27, pp. 365–386.

Di Pietro, Giorgio (2014). “The short-term effectiveness of a remedial mathematics course: Evidence from a UK university”. In: *Manchester School* 82, pp. 363–384.

Georg, Werner (2009). “Individual and institutional factors in the tendency to drop out of higher education: A multilevel analysis using data from the Konstanz Student Survey”. In: *Studies in Higher Education* 34, pp. 647–661.

Greefrath, Gilbert and Georg Hoever (2016). “Was bewirken Mathematik-Vorkurse? Eine Untersuchung zum Studienerfolg nach Vorkursteilnahme an der FH Aachen”. In: *Lehren und Lernen von Mathematik in der Studieneingangsphase.* Ed. by A. Hoppenbrock; R. Biehler; R. Hochmuth; H.G. Rück, pp. 517–530.

Greefrath Gilbert; Koepf, Wolfram and Christoph Neugebauer (2017). “Is there a link between Preparatory Course Attendance and Academic Success? A Case Study of Degree Prgrammes in Electrical Engineering and Computer Science”. In: *International Journal of Research in Undergraduate Mathematics Education* 3, pp. 143–167.

Gueudet, Ghislaine (2008). “Investigating the secondary-tertiary transition”. In: *Educational Studies in Mathematics* 67, pp. 237–254.

Heublein, Ulrich (2014). “Student Drop-out from German Higher Education Institutions”. In: *European Journal of Education* 49, pp. 497–513.

Hirano, Keisuke, Guido W. Imbens, and Geert Ridder (2003). “Efficient estimation of average treatment effects using the estimated propensity score”. In: *Econometrica* 71, pp. 1161–1189.

Hoppenbrock A.; Biehler, R.; Hochmuth R.; Rück H.G., ed. (2016). *Lehren und Lernen von Mathematik in der Studieneingangsphase. Herausforderungen und Lösungsansätze.*

Khandker, Shahidur R., G. B Koolwal, and Hussian a. Samad (2010). *Handbook on Impact Evaluation: Quantitative Methods and Practices.*

Krohn, Gregory and Catherine O’Connor (2005). “Student effort and performance over the semester”. In: *Journal of Economic Education* 36, pp. 3–28.

Lagerlöf, Johan and Andrew Seltzer (2009). "The effects of remedial mathematics on the learning of economics: Evidence from a natural experiment". In: *Journal of Economic Education* 40, pp. 115–136.

Laging, Angela and Rainer Voßkamp (2017). "Determinants of Maths Performance of First-Year Business Administration and Economics Students". In: *International Journal of Research in Undergraduate Mathematics Education* 3, pp. 108–142.

Luk, Hing Sun (2005). "The gap between secondary school and university mathematics". In: *International Journal of Mathematical Education in Science and Technology* 36, pp. 161–174.

Mallik, G. and S. Shankar (2016). "Does prior knowledge of economics and higher level mathematics improve student learning in principles of economics?" In: *Economic Analysis and Policy* 49, pp. 66–73.

Mallik, Girijasankar and John Lodewijks (2010). "Student Performance in a Large First Year Economics Subject: Which Variables are Significant?" In: *Economic Papers* 29, pp. 80–86.

Martorell, Paco and Isaac McFarlin (2011). "Help or Hindrance? The Effects of College Remediation on Academic and Labor Market Outcomes". In: *Review of Economics and Statistics* 93, pp. 436–454.

Rosenbaum, Paul R. and Donald B. Rubin (1983). "The central role of the propensity score in observational studies for causal effects". In: *Biometrika* 70, pp. 41–55.

Scott-Clayton, Judith and Olga Rodriguez (2015). "Development, Discouragement, or Diversion? New Evidence on the Effects of College Remediation Policy." In: *Education Finance and Policy* 10, pp. 4–45.

Stuart, Elizabeth A. (2010). "Matching Methods for Causal Inference: A Review and a Look Forward". In: *Statistical Science* 25, pp. 1–21.

Voßkamp, Rainer and Angela Laging (2014). "Teilnahmeentscheidungen und Erfolg". In: *Mathematische Vor- und Brückenkurse: Konzepte, Probleme und Perspektiven.* Ed. by Bausch, I.; Biehler, R.; Bruder, R.; Fischer, P.; Hochmuth, R.; Koepf, W.; Schreiber, S.; Wassong, T., pp. 67–83.