pedocs
Open Access Erziehungswissenschaften

DIPF

Guill, Karin; Bos, Wilfried

Effectiveness of private tutoring in mathematics with regard to subjective and objective indicators of academic achievement. Evidence from a German secondary school sample

Journal for educational research online 6 (2014) 1, S. 34-67

Quellenangabe/ Reference:
Guill, Karin; Bos, Wilfried: Effectiveness of private tutoring in mathematics with regard to subjective and objective indicators of academic achievement. Evidence from a German secondary school sample - In: Journal for educational research online 6 (2014) 1, S. 34-67 - URN: urn:nbn:de:0111-opus-88400 - DOI: 10.25656/01:8840

https://nbn-resolving.org/urn:nbn:de:0111-opus-88400
https://doi.org/10.25656/01:8840

in Kooperation mit / in cooperation with:

WAXMANN
www.waxmann.com

http://www.waxmann.com

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.
This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

pedocs
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der Leibniz-Gemeinschaft

Journal for Educational Research Online
Journal für Bildungsforschung Online
Volume 6 (2014), No. 1, 34–67
© 2014 Waxmann

Karin Guill & Wilfried Bos

Effectiveness of private tutoring in mathematics with regard to subjective and objective indicators of academic achievement

Evidence from a German secondary school sample

Abstract

In addition to the mainstream education system many students draw on private tutoring to improve their academic achievement. However, evidence about the effectiveness of private tutoring is rare and contradictory. The identification of positive or rather neutral effects of private tutoring seems to depend at least partly, on the underlying concept of academic success.

Therefore, we analyzed the effects of private tutoring using various indicators of academic success within the same sample. The data originated from the German longitudinal study KESS ("Competencies and Attitudes of Students"), which includes a complete cohort of students at the beginning of Grade 7 and at the end of Grade 8. The effect of private tutoring was evaluated using multiple criteria: The majority of the parents (N = 447) as well as of the students (N = 618) stated that private tutoring had improved the students' mathematics achievement. In contrast, there was neither a significant improvement of math marks of tutored students compared to the entire sample of non-tutored students nor an improvement in math achievement test results due to private tutoring when controlling for prior knowledge, motivational variables and school level effects (HLM, N = 4,701). Further analyses with matched samples of tutored and non-tutored students (PSM, N = 1,236) confirmed these results. We compared the effects of private tutoring on the different criteria and discussed possible reasons for the contradictory results.

Dr. Karin Guill (corresponding author), Institute for School Development Research, TU Dortmund University & Leibniz Institute for Science and Mathematics Education at the University of Kiel, Olshausenstraße 62, 24118 Kiel, Germany
e-mail: guill@ipn.uni-kiel.de

Prof. Dr. Wilfried Bos, Institute for School Development Research, TU Dortmund University, Vogelpothsweg 78, 44227 Dortmund, Germany
e-mail: officebos@ifs.tu-dortmund.de

Keywords
Private tutoring; Shadow education; Mathematics achievement; School achievement; Secondary schooling; HLM

Effektivität von Nachhilfeunterricht in Mathematik im Hinblick auf subjektive und objektive Indikatoren der Schulleistung

Befunde aus einer deutschen Sekundarschulstichprobe

Zusammenfassung
Neben dem öffentlichen Schulsystem nutzen viele Schülerinnen und Schüler privaten Nachhilfeunterricht, um ihre Schulleistung zu verbessern. Es gibt jedoch nur wenige und widersprüchliche Befunde bezüglich der Effektivität von Nachhilfeunterricht. Der Nachweis positiver oder eher neutraler Effekte von Nachhilfeunterricht scheint zumindest teilweise vom zugrunde liegenden Konzept von Schulerfolg abzuhängen.

Daher wurde in der vorliegenden Studie der Effekt von Nachhilfeunterricht innerhalb derselben Stichprobe anhand verschiedener Indikatoren des Schulerfolgs untersucht. Die Daten sind Teil der deutschen Längsschnittstudie KESS (Kompetenzen und Einstellungen von Schülerinnen und Schülern), die eine vollständige Schülerkohorte am Beginn von Jahrgangsstufe 7 und am Ende von Jahrgangsstufe 8 umfasst. Der Effekt von Nachhilfeunterricht wurde anhand mehrerer Kriterien untersucht: Die Mehrheit der Eltern (N = 447) wie der Schülerinnen und Schüler (N = 618) gab an, dass der Nachhilfeunterricht die Mathematikleistung der Jugendlichen verbessert habe. Dagegen ließ sich weder eine signifikante Verbesserung der Mathematiknoten der Nachhilfeschülerinnen und -schüler im Vergleich zur vollständigen Stichprobe der Schülerinnen und Schüler ohne Nachhilfeunterricht nachweisen noch eine Verbesserung der Mathematiktestleistung aufgrund von Nachhilfeunterricht, wenn das Vorwissen, motivationale Variablen und Effekte der Schulebene kontrolliert werden (HLM, N = 4701). Weitere Analysen mit parallelisierten Stichproben (PSM, N = 1236) von Schülerinnen und Schülern mit und ohne Nachhilfeunterricht bestätigten diese Ergebnisse. Die Effekte von Nachhilfeunterricht auf die verschiedenen Kriterien wurden verglichen und mögliche Ursachen der widersprüchlichen Ergebnisse diskutiert.

Schlagworte
Nachhilfeunterricht; Mathematikleistung; Schulleistung; Sekundarstufe; HLM

1. Introduction

Private tutoring is a wide-spread phenomenon all over the world. In addition to mainstream education many students attend private lessons in academic subjects which are already covered at school. For example in Germany, where the presented empirical analyses were executed, according to a recent representative study nearly 24% of 12 to 21 year old secondary school students currently attend private tutoring (Leven, Quenzel, & Hurrelmann, 2010). This number resembles the proportion of tutored students in neighboring countries like England and France. In many Eastern European, African and East Asian countries up to 50% to 80% of the students, especially in higher grades, are privately tutored (Bray, 2009).

This tutoring may be provided on a one-to-one basis as well as in small or even large groups, though this last form is not common in Germany. Students may attend private tutoring for different reasons: as a kind of additional day-care after school, to meet up with their peers, or to avoid tensions at home caused by academic problems. But the main reason for private tutoring is to improve academic achievement at school or in high stakes exams at the end of secondary school (Bray, 2009). Therefore, on the school system level in the middle of secondary schooling private tutoring has mostly remedial purposes: With the exception of few countries like, e.g., South Korea, on average privately tutored students perform below the non-tutored students (Baker, Akiba, LeTendre, & Wiseman, 2001). A specific feature of the school system in Germany, is tracking according to the students' academic achievement. Consequently, private tutoring is not simply most prevalent between low performing students of an age or grade cohort but between low performing students within each academic track and relative to its specific achievement level (Guill, 2012).

Despite the large number of students who take part in private tutoring, there is little research whether private tutoring reaches its main goal of improving academic achievement. Therefore, the advantages privately tutored students have compared to their counterparts can hardly be quantified. Even if it is well-known that private tutoring is more easily accessible for students from wealthy families (Bray, 2009), it remains unclear whether or to what extent private tutoring or specific types of private tutoring increase social disparities in academic achievement.

In this paper we would like to contribute to closing this research gap by analyzing the effects of private tutoring in a large sample of German students in secondary school. The following outline is given: Section 2 sketches a theoretical model of private tutoring on academic achievement at school. Further, it gives a literature review about current research on the effectiveness of private tutoring. Section 3 outlines the leading research questions for the empirical analyses. Section 4 gives the description of the methodological approach. Section 5 describes the results of the empirical analyses. Finally, Section 6 discusses these results.

2. Theoretical background and literature review

Up to now, there is no well-established comprehensive theoretical model how private tutoring effects the tutees' academic achievement. Consequently, studies focusing on the global question of whether private tutoring is worthwhile or not usually propose only an eclectic array of reasons why private tutoring might support the students' learning process (e.g., Ryu & Kang, 2010; Smyth, 2009; Ünal, Özkan, Milton, Price, & Curva, 2010).

Looking for a theoretical framework of the effects of private tutoring for our own analyses and given the close relationship of private tutoring and classroom teaching, we discuss the adoption of a model of classroom teaching. Here, we opted for Helmke's (2009) offer-usage model of instructional effects for the simplicity of its basic idea: On the one side, school offers learning opportunities to the students depending on the instructional quality of the lessons. On the other side, the students use these learning opportunities during their active learning time in and out of school which results in academic and non-academic outcomes. The students' learning activities depend on their learning potential. The students' families offer an additional learning environment.

Private tutoring interacts with this system. We elaborate on this interaction for the typical German tutoring setting where between one and five students are tutored by more advanced secondary school or university students or teachers who in most cases are not their teachers in school. The private tutor has to diagnose his or her tutee's competences and what the classroom teacher expects from above mentioned tutee. Tutees as well as their parents have academic and non-academic expectations from the tutoring lessons which the tutor has to take into account when setting or negotiating the goals for the tutoring lessons. Once these goals are set, the private tutoring lessons can be conceptualized according to the offer-usage model: During the tutoring lessons the tutor offers additional learning time to the students. He or she can close gaps in the tutees' prior knowledge and teach them learning strategies. Due to the individual instruction, it is easier for the tutor to recognize the tutees' progress than for a classroom teacher teaching twenty or more students at the same time. By giving individual feedback according to an individual reference norm orientation, the tutor can enhance the tutees' learning motivation and ability self-concept as well as decrease their achievement anxiety (Mischo & Kessel, 2005). This way, the tutor may increase the tutees' learning potential and as a result, the tutees' benefit from classroom instruction may increase. However, private tutors may differ in their academic as well as pedagogic education, and therefore in the instructional quality of their tutoring lessons. Tutees use the tutoring lessons depending on their learning potential: their general cognitive and motivational abilities as well as their private tutoring motivation. If students are forced by their parents to attend private tutoring, their motivation to profit from the tutoring lessons might be low. Also, when using private tutoring extensively, students are not obliged to develop and practice self-regulated

learning styles, and therefore might not manage them after stopping with tutoring. Additional problems arise especially concerning the relationship between private tutoring and mainstream schooling: Students who are additionally tutored may become bored and inattentive in the classroom as they rely mostly on their tutors' support and make less use of the learning opportunities in their normal classroom. Also, different pedagogical approaches of the classroom teacher and the private tutors might be a challenge for the tutees. Hence, learning gains from private tutoring lessons – or classroom instruction – might be hindered (Bray, 2009; Kenny & Faunce, 2004).

In summary, private tutoring as an additional offer of instruction time has the potential to increase the students' learning potential and consequently their academic achievement but can easily fail if one component of the complex interaction system does not work. Based on the offer-usage model of private tutoring the leading research question of our own analyses is whether private tutoring improves the students' academic achievement.

2.1 Effects of private tutoring on different indicators of academic achievement

Upon closer investigation there are numerous ways to conceptualize and operationalize the improvement of academic achievement. This section will elaborate upon five criteria of an improved academic achievement in relation to the offer-usage model of private tutoring and give an overview of empirical results concerning the effectiveness of private tutoring on the respective criterion.

As a first and second criterion one can look at the parents' and students' assessment whether private tutoring has improved the students' academic achievement. To make such an assessment, parents and students on the one hand have to evaluate whether the student's academic achievement increased at all. Here they can rely on marks and individual feedback from the classroom teacher, feedback from the tutor and the students' individual ability self-concept. On the other hand, they have to decide whether to attribute the improvement of the students' academic achievement to the private tutoring lessons or other factors like an increased effort of the students and therefore a better usage of the classroom instruction or even an improved, e.g., more adapted classroom instruction. This might be especially difficult for parents as they usually do not have direct access to the tutoring and classroom lessons, and must rely on their children's and the tutors' reports. However, the statements of parents as well as students generally give a positive impression regarding the effectiveness of private tutoring. Parents and students from Germany (Rudolph, 2002; Synovate Kids+Teens & Bundesverband Nachhilfe- und Nachmittagsschulen e.V., 2007), the UK (Ireson & Rushforth, 2005) and eight former socialist countries (Silova & Bray, 2005) for the most part indicated satisfaction, and agreed that private tutoring improved their child's/their own achieve-

ment at school or in the university entrance examinations. However, it remains unclear to what extent these assessments are biased by errors in judgment.

In third place the improvement of the students' school marks is a self-evident criterion for the effectiveness of private tutoring, given that the marks are important for the students' educational careers. Private tutoring can improve the students' marks by implementing instructional offers on different levels. These offers require different degrees of qualifications from the tutor and different amounts of effort from the students. Simply focusing on correct homework completion could already improve the students' marks. Taking it a step further, private tutoring might also improve the students' marks when focusing on short-term training for the next test, without a long-term improvement of underlying competencies (Guill & Bonsen, 2010). With a high level of instructional quality and effort on the part of the student, private tutoring might improve the students' marks by sufficiently improving the underlying competencies so that they are noticed by the teacher. Empirically, the effect of private tutoring on students' marks in tests at school or in report cards is still positive overall but the findings are less consistent. Mischo and Haag (2002) found large and positive effects of a rather intensive German tutoring program on students' marks in teacher-designed tests in a pre-post-control-group design (N = 244, e.g., d = 0.72 in math). Studies from different countries like Spain (Elvira et al., 2006), Greece (Polydorides, 1986), Kenya (Buchmann, 2002), and Vietnam (Dang, 2007) confirm positive statistical effects of private tutoring or the spending on private tutoring on marks or, as depending from marks, the likelihood to repeat a grade at school. In contrast, studies from other countries (Cheo & Quah, 2005; Kenny & Faunce, 2004) revealed partly negative effects of private tutoring on students' marks and their end-of-year examination results. However, these studies investigated very different private tutoring settings.

A fourth indicator of the effectiveness of private tutoring is the results of national school leaving exams or university entrance exams because some students mainly attend private tutoring to improve their results in these exams. Here the students' assessment is less dependent on their teachers. Still, tutors can focus their instructional offer more on test-taking skills or more on an improvement of the competencies required for passing the exam. Studies using national exams as indicators generally give a less consistent insight into the effectiveness of private tutoring. They suggest positive effects (Hamid, Susses, & Khan, 2009; Tansel & Bircan, 2005), as well as mixed (Ireson & Rushforth, 2005; Kenny & Faunce, 2004; Stevenson & Baker, 1992), or missing (Kang, 2007; Smyth, 2008) effects of private tutoring on the corresponding exams. In Greece, there was even a weak negative relationship between private tutoring and the scores of the national university entrance examination (Polydorides, 1986).

Finally, as a fifth criterion, one can focus on the question whether private tutoring improves the students' actual competencies or only their test taking skills, their own competency assessment or the competency assessment of the teacher. Therefore, the effects of private tutoring have also been analyzed in studies which employ their own achievement or competence tests. These tests vary in the degree

to which they follow the school curriculum, and focus more strongly on competencies necessary for daily-life problems and social inclusion. Again, there are quite mixed findings concerning the effectiveness of private tutoring to improve academic achievement. These vary from rather neutral (Guill & Bonsen, 2010; Ryu & Kang, 2010) to positive statistical effects of private tutoring on test achievement (Hamid, Sussex, & Khan, 2009; Nath, 2008; Paviot, Heinsohn, & Korkman, 2008; Ünal et al., 2010). Interestingly, in Japan (Sawada & Kobayashi, cited from Bray, 1999) attending private tutoring resulted in higher scores in problems requiring arithmetic calculation and algebra but not higher scores in arithmetic application and geometry.

In summary, parents' and students' statements generally reveal positive effects of private tutoring on academic achievement. However, studies regarding the effects of private tutoring on marks, examination performance, and competence test results lead to more contradictory results. The cultural and educational settings of the studies presented here differ significantly. Therefore, it is an open question how this pattern of results is formed when the effect of private tutoring on different criteria is investigated within the same sample in a constant cultural and educational setting. Consequently, our differentiated research questions focus on the effects of private tutoring on several distinct achievement related criteria and the relationship of the results on these criteria.

2.2 Methodological approaches to evaluate the effectiveness of private tutoring

Before further investigating this question, one has to take a close look at the methodological approaches of studies on the effectiveness of private tutoring. Therefore, the aim of this section is to describe and discuss these approaches in reference to the offer-usage model of private tutoring.

Studies on the effectiveness of private tutoring differ a lot in their methodological approach. This is due to their sample sizes, the students' age cohorts, the statistical approach, and the control variables employed. The main approaches to analyze the effects of private tutoring are cross-sectional and longitudinal correlation studies. Secondary analyses of existing data sets from large-scale studies are widely spread (e.g., Guill & Bonsen, 2010; Ryu & Kang, 2010; Smyth, 2008; Ünal et al., 2010). The elaborated quasi-experimental design employed by Mischo and Haag (2002, v. s.) is an exception.

The combination of cross-sectional studies and (linear) regression analyses mainly revealed positive effects of private tutoring. (Buchmann, 2002; Dang, 2007; Elvira et al., 2006; Hamid et al., 2009; Nath, 2008; Smyth, 2009, data from 2004; Ünal et al., 2010). Only analyses from Singapore (Cheo & Quah, 2005) with the same approach showed partly negative statistical effects of private tutoring. Most of the studies controlled for family characteristics like income and educational resources. In the offer-usage model this is the additional learning environment of-

fered by the students' families. But most studies failed to take into account systematic differences between privately tutored and non-tutored students concerning cognitive and motivational variables like their prior achievement or motivation which according to the proposed model influence the degree to which the students use the instructional offer of the classroom and the tutoring lessons. Therefore, it is unclear whether private tutoring really improves the students' academic achievement or whether simply the more capable and motivated students more often tend to seek additional support in private tutoring classes, for example to increase their chances in high stakes exams (Bray, 2009).

Therefore, longitudinal studies which control for these systematic differences are methodically more adequate. Interestingly, a considerable number of these longitudinal studies reveal no effect of private tutoring at all (Guill & Bonsen, 2010; Kang, 2007; Ryu & Kang, 2010; Smyth, 2008). In contrast, longitudinal analyses from further countries (Ireson & Rushforth, 2005; Polydorides, 1986; Stevenson & Baker, 1992; Tansel & Bircan, 2005) revealed at least partly positive relations between private tutoring and the students' academic achievement.

In summary, one has to conclude that more elaborate study designs and statistical methods imply more skeptical – or at least more mixed – conclusions about the power of private tutoring to improve academic achievement. Further research should therefore not fall behind these methodological standards.

3. Research questions

All in all, the theoretical model previously lined out presents a number of reasons why and how private tutoring can improve the tutees' academic achievement. Given the multiple requirements of private tutoring lessons, they can also easily fail to achieve this goal.

Looking at the small number of methodologically adequate studies, and the different countries, cultural contexts, and school systems it is still an open question whether private tutoring pays off and improves the students' academic achievement. Therefore, the main research question of our own analyses is whether private tutoring improves the tutees' academic achievement.

However, the theoretical model and further explanations should have clarified that there are different criteria or indicators of an improved academic achievement which do not necessarily lead to the same result. There is hardly any research which has evaluated the effectiveness of private tutoring on more than one criterion at the same time (an exception is, e.g., Ireson & Rushforth, 2005). Therefore, the aim of our study is to analyze the effects of private tutoring more precisely by looking at multiple achievement-related criteria within the same sample. This way we keep other possible factors influencing achievement like age, educational system or cultural effects constant.

In detail, following the differentiation in Section 2.1 – and considering the limitations of our data sets – we focused on the following research questions: (1) Does private tutoring improve students' academic achievement from their parents' point of view? (2) Does private tutoring improve students' academic achievement from their own point of view? (3) Does private tutoring improve students' marks at school? (4) Does private tutoring improve students' academic achievement in a school achievement test?

In accordance with previous research, we suppose that parents and students will be satisfied overall with their private tutoring lessons. As the previous research on effects on marks and achievement test results is less consistent we did not formulate any hypothesis concerning the third and fourth research question.

As described above, one would expect that parents and students take objective measures of achievement like the students' marks into account when evaluating the private tutoring lessons. However, the relationship between the different criteria of effectiveness of private tutoring has not yet been analyzed, and is therefore a final research question of our analyses: (5) (How) are the different criteria of effectiveness of private tutoring related?

As private tutoring in Germany is mostly subject-specific, this evaluation of its effects will focus on private tutoring in one subject and therefore also on students' corresponding marks and achievement test results in this subject. We chose mathematics as the subject focused on, as this is the most widespread tutoring subject in Germany as well as internationally (Guill, 2012).

4. Method

4.1 Sample

The data are part of the German panel study KESS ("Competencies and Attitudes of Students"). A complete student cohort in the city of Hamburg, a metropolis in Northern Germany, took part in this study. There were three measurement points, each approximately two years apart: at the end of Grade 4 in 2003, at the beginning of Grade 7 in 2005 and at the end of Grade 8 in 2007. Effects of private tutoring in the lower age group have already been analyzed (Guill & Bonsen, 2010). For the analyses at hand, data from the second (KESS 7, N = 14,200) and third (KESS 8, N = 13,871) point of measurement will be used (N = 11,735 students in the panel). At that point students were usually 12 to 15 years old. As part of the study, parents and students answered questionnaires about the students' learning conditions, and students were asked to take part in achievement tests in the domains of mathematics (participation rate: 95.6%), reading, English, and sciences (Bos, Gröhlich, Guill, Scharenberg, & Wendt, 2010).

4.2 Missing values

There are many missing values in the data set due to drop-outs in the panel and item non-response. (Multiple) imputation is often described as the useful approach to the missing value problem (Schafer, 1997). However because of a large number of students and especially parents, who did not answer the KESS 8 questionnaires (response rates 74.4% and 47.3% respectively), no adequate statistical background model could be created to replace all missing data in the KESS 8 cohort. Therefore, the sample is reduced to those KESS 8 students who answered the students' questionnaire, and whose parents answered the KESS 8 parents questionnaire (N = 4,701 students in 164 schools), i.e., those students and parents who mainly answered the tutoring questions. As the private tutoring information from the parent and the student questionnaire is not interchangeable (see Section 4.3), we opted for this restricted criterion of inclusion.

Still missing data in this sample were replaced by imputation with the software Norm 2.03 (Schafer, 1999) using additional variables and cases as a statistical background model. The sample was divided into different subsamples to impute the students' marks within their frame of reference, which are the different tracks. Due to this complex division of the sample only a single imputation was performed. This way, we were able to handle the missing data aspect in advance of our main analyses and apply an adequate statistical background model. Meanwhile, we accepted a possible underestimation of the standard errors because a single imputation does not account for the uncertainty caused by the imputation process (Schafer, 1997). Even with a similar KESS 8 subsample praxis shows that the between-imputation variance is small compared to the within-imputation variance, and therefore the underestimation of the standard errors is rather low (Guill, 2012). Hence, only predictors close to the significance level are affected.

As all continuous variables are z-standardized in the complete KESS 8 sample (see Section 4.3), the sample statistics in Table 1 illustrate that the analyzed subsample is slightly biased towards students with a higher achievement in mathematics and a higher socio-economic background.

4.3 Measures

Private tutoring. In the KESS 8 students' questionnaire, students stated whether and since when they were taking part in private tutoring in mathematics at the time of the study. This was dummy-coded to no (0) or ongoing private tutoring (1). Parents stated in the KESS 8 parents' questionnaire whether and how long their child had taken part in private tutoring in mathematics during the past two years in Grade 7 or 8. This was dummy-coded to no (0) or any private tutoring (1) in the past for those students who did not mention any ongoing private tutoring. In the parents' questionnaire, private tutoring (*Nachhilfeunterricht*) was measured in contrast to in-school tutoring (*Förderunterricht*) and defined as out-of-school

support to improve academic achievement by persons in some way qualified (e.g., more advanced students or teachers) who are not family members. Three groups can be differentiated: students with no private tutoring in Grade 7 or 8, students with ongoing private tutoring at the end of Grade 8 and students with private tutoring in the course of Grade 7 and 8, which ended before the data collection in Grade 8. On a 4-point rating scale parents and students who attended private tutoring stated whether it improved their child's, respectively their own performance in mathematics.

The KESS 7 and KESS 8 *mathematics achievement* tests are composed of tasks from different mathematics domains like algebra, geometry and probability and focus on the capability to solve mathematical problems related to daily life. Curricular validity was intended but not systematically investigated. The data collection followed a multi-matrix design with 72 tasks for the KESS 7 test and 51 tasks for the KESS 8 test. The math tests were scaled with the unidimensional dichotomous Rasch model (Guill, Gröhlich, Scharenberg, Wendt, & Bos, 2010).

The students' *marks* in mathematics were collected at the end of the first and second term in Grade 8. According to the German grading system, there are six different marks from 1 (excellent) to 6 (insufficient/fail). We recoded them from 1 (insufficient/fail) to 6 (excellent) to allow for easier interpretation. Teachers were not informed about the students' test results when marking performance. Marks in the second term and test achievement in Grade 8 are correlated between $r = .39$ and $r = .54$ depending on the track. The highest correlation is found in the academic track while additional streaming in the other tracks lowers the correlation.

The students' *intelligence* was only measured with the subscale "figural analogies" from the KFT 4-12+R (cognitive abilities test; Heller & Perleth, 2000) during the KESS 8 data collection. Therefore, we used this measure for our analyses.

The students' *socioeconomic background* is operationalized by the families' highest socio-economic index (HISEI) which is based on the parents' statements about their profession (Ganzeboom & Treiman, 2003).

The students' motivation at the beginning of Grade 7 is measured by several constructs: their *achievement anxiety* (4 items, e.g., "When my teacher says we are going to write a test I am afraid of it", $\alpha = .80$, Valtin & Darge, 2001), their *interest in mathematics* (6 items, e.g., "I like solving word problems.", $\alpha = .83$, Bos et al., 2005) and their *ability self-concept in mathematics* (3 items, e.g., "I have always been good at math.", $\alpha = .92$, Marsh, 1990a). All items were part of the KESS 7 student questionnaire and were to be answered on a 4-point rating scale. Scales were formed by factor scores.

Gender is dummy-coded to differ between girls (0) and boys (1).

The students' *migration background* is differentiated according to whether one parent, both, or neither (reference group) were born abroad.

On the school level, schools are described by their students' mean prior knowledge, i.e., the students' average math achievement test result in Grade 7, and students' average socioeconomic status. Additionally, due to the German school system the schools belong to different tracks, namely the higher secondary or academ-

ic track (*Gymnasium*), the lower secondary track as reference group (*Haupt- und Realschule*), and comprehensive schools (*Integrierte Gesamtschule*).

All continuous variables have been *z*-standardized in the complete KESS 8 sample.

4.4 Analyses

The analyses are organized in two steps: First, the effects of private tutoring on four different indicators of academic achievement were analyzed separately. Second, the effects on the different indicators were compared.

As only parents and students who attended private tutoring stated their opinion regarding the effectiveness of their tutoring lessons on single items, results are reported using descriptive statistics. To improve comparability, parents' statements are only reported for those students being tutored at the time of the study.

As presented in Helmke's offer-usage model of instructional effects marks and achievement, test results are influenced by individual predictors as well as the context that is the class or school level (Helmke, 2009; q.v. Bonsen, Gröhlich, & Bos, 2009; Ingenkamp, 1993). Therefore, we analyzed the effect of private tutoring on marks and the achievement test results by hierarchical regression analyses. On the individual level, well established predictors of school achievement such as prior knowledge (mark in Term 1 resp. math test achievement in Grade 7), intelligence, gender, migration background, socioeconomic status and motivational constructs were introduced to the model directly after the students' tutoring status in mathematics. On the school level, we controlled for tracking, the students' average prior knowledge and their average socioeconomic status. Due to the short time period between the marks of Term 1 and the data collection we did not analyze effects of past private tutoring (Grade 7/8) on the marks of Term 2 because this tutoring mostly ended before the marking in Term 1. Therefore, the HL models to predict marks only include ongoing private tutoring as a predictor while the HL models to predict test achievement include ongoing and past private tutoring as predictors. All HLM analyses were run with HLM 6.08 (Raudenbush, Bryk, & Congdon, 2004) using restricted maximum likelihood estimation.

During recent years, propensity score matching (PSM) became established in educational sciences as an additional promising way to handle selection bias and to estimate treatment effects (Thoemmes & Kim, 2011). We therefore used analyses based on PSM as an additional validation of our HLM analyses (cf. Smyth, 2008). The methodically interested reader will find a short explanation of this approach to estimate the effect of private tutoring in Appendix A.1.

5. Results

5.1 Descriptive statistics

At the end of Grade 8, 13.2% of students attended private tutoring. Additionally, 11.4% of students took part in private tutoring in Grade 7 or Grade 8 at an earlier point in time.

Earlier descriptive analyses for those students with complete data confirm that the private tutoring settings are typical for Germany: The tutoring lasts mainly up to one academic year, and is given by more advanced secondary and university students, teachers, or private tutoring institutes. Private tutoring institutes usually employ staff with similar qualification (Guill, 2012). Main activities include the preparation for tests and going over previous tests, doing and checking of homework, training for current lessons and repetition of earlier lessons. Concentration training and the training of study techniques are less important (Guill, 2010). Table 1 gives an overview of similarities and differences of the sample students tutored and non-tutored at the time of the study. As to be seen here, the tutored students showed significantly lower achievement scores in terms of marks, test achievement, and intelligence than the non-tutored students. Their interest in math as well as their ability self-concept was located significantly below those of the non-tutored students, while they show higher scores in achievement anxiety. Differences concerning the proportion of boys and girls, students with migration background and academic tracks as well as the students' socio-economic background failed statistical significance. To sum up, tutored and non-tutored students differed systematically in their academic achievement as well as in their motivational characteristics. These differences in our sample coincide with the mostly remedial nature of private tutoring in middle school described in the introduction section. Additionally and also in line with current research, students with higher socio-economic status are more likely to attend private tutoring when academic achievement is controlled for (Guill, 2012).

Table 1: Characteristics of all students in the analyzed sample (N = 4,701) and in subsamples of students with (N = 618) and without (N = 4,083) ongoing private tutoring in mathematics

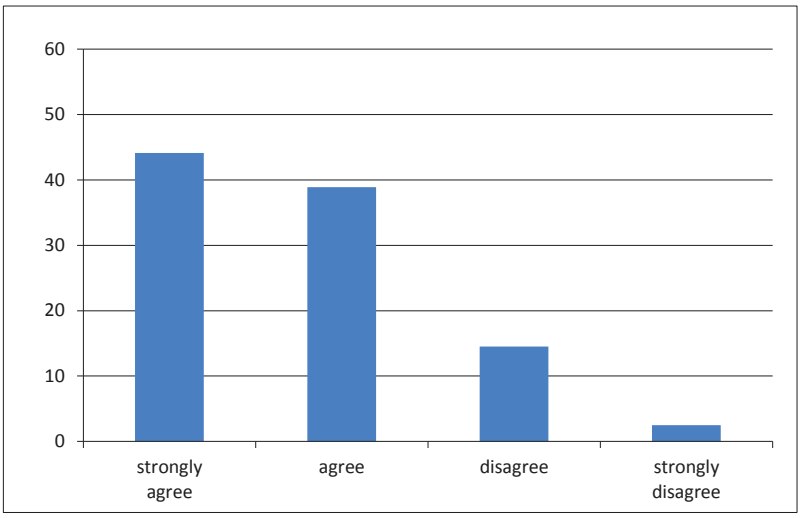
	Complete Sample		Ongoing private tutoring		No ongoing private tutoring		Group differences (c. 2 to c. 3)		Effect size
	M	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*t*	(*p*)	*d*
Marks in Grade 8									
- Term 1	3.89	(1.00)	**3.35**	**(0.89)**	**3.97**	**(0.99)**	-14.65	(< .01)	-0.63
- Term 2	3.86	(0.98)	**3.36**	**(0.82)**	**3.94**	**(0.97)**	-13.91	(< .01)	-0.61
Math test achievement									
- Grade 7	0.21	(1.00)	**-0.17**	**(0.81)**	**0.27**	**(1.01)**	-10.26	(< .01)	-0.45
- Grade 8	0.25	(0.98)	**-0.09**	**(0.85)**	**0.31**	**(0.99)**	-9.46	(< .01)	-0.41
Intelligence	0.20	(0.93)	**-0.12**	**(0.93)**	**0.25**	**(0.93)**	-9.15	(< .01)	-0.40
ISEI	0.14	(0.98)	0.08	(0.96)	0.15	(0.98)	-1.63	(.10)	-0.07
Interest in math	-0.01	(1.00)	**-0.25**	**(1.01)**	**0.03**	**(0.99)**	-6.65	(< .01)	-0.28
Self-concept math	0.02	(1.00)	**-0.44**	**(0.95)**	**0.09**	**(0.99)**	-12.45	(< .01)	-0.54
Achievement anxiety	-0.08	(0.98)	**0.05**	**(1.02)**	**-0.11**	**(0.97)**	3.78	(< .01)	0.16
Gender (% boys)	45.3%		40.9%		45.9%		-2.33	(.20)	
Migration background									
- One parent	10.7%		12.9%		10.4%		1.90	(.06)	
- Both parents	20.2%		18.9%		20.4%		-0.86	(.39)	
Tracking									
- Lower secondary school	22.7%		21.7%		22.8%		-0.62	(.54)	
- Comprehensive school	27.0%		29.0%		26.7%		1.21	(.23)	
- Academic track	50.4%		49.4%		50.6%		-0.56	(.58)	

Notes. Range of marks from 1 (insufficient/fail) to 6 (excellent). Other continuous variables are *z*-standardized within the KESS 8 cohort. Significant group differences are printed in bold numbers.

5.2 Parents' and students' statements on the effectiveness of private tutoring

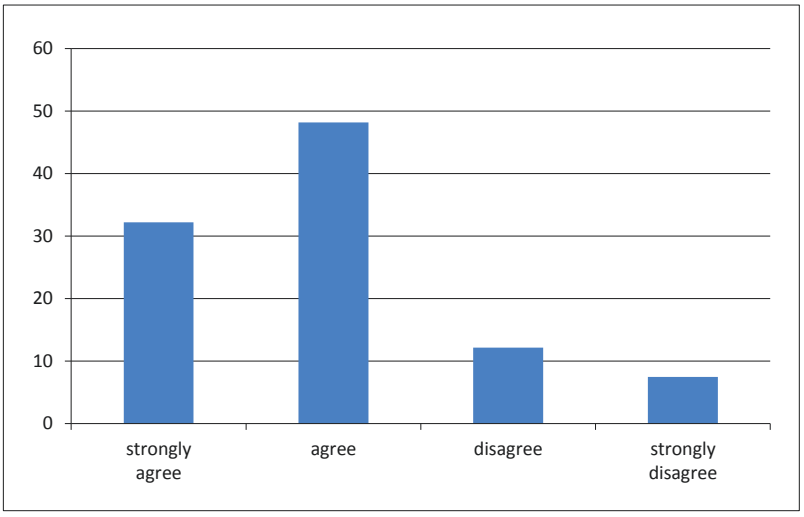
The vast majority, i.e., 76.7% of the parents whose children had ongoing private tutoring lessons at the time of the data collection (N = 447), agreed or strongly agreed that private tutoring improved their child's performance in mathematics, 23.3% disagreed or strongly disagreed.

Figure 1: Parents' approval of "Private tutoring improved my child's performance in mathematics." (Percentages, N = 447)



An even greater majority, i.e., 80.4% of the students who were privately tutored in mathematics at the time of the study, agreed or strongly agreed that private tutoring improved their performance in mathematics, 19.6% disagreed or disagreed strongly.

Figure 2: Students' approval to "Private tutoring improved my performance in mathematics." (Percentages, N = 618)



5.3 Estimation of group differences by hierarchical linear regression analyses

5.3.1 Marks in mathematics at the end of Grade 8

The effect of private tutoring on the marks in mathematics is analyzed using hierarchical linear regression models. We first outline the sequencing of our four models and then describe them in more detail.

Model 1 aims to show the difference in marks in mathematics between tutored and non-tutored students. Model 2 focuses on the effects of all control variables and the variance they explain. The re-introduction of private tutoring in Model 3 makes it possible to estimate its effect on the marks and the incremental variance it explains compared to Model 2. Interaction effects between private tutoring and the control variables were analyzed in Model 4 as private tutoring might have differential effects in specific subgroups.

Table 2: Effects of private tutoring on mathematics marks at the end of Grade 8 – Hierarchical linear regression analysis (N = 4,701)

	Model 1	Model 2	Model 3	Model 4
Intercept	**3.89**	**3.92**	**3.93**	**3.93**
Individual level				
Mark in Term 1[a]		**.65**	**.65**	**.66**
Intelligence[b]		**.10**	**.09**	**.09**
ISEI[b]		.01	.01	.01
Interest in math[b]		.01	.01	.01
Self-concept math[b]		**.10**	**.10**	**.10**
Achievement anxiety[b]		**-.03**	**-.03**	**-.03**
Boys		.00	.00	.00
Migration background				
- One parent		-.05	-.05	-.05
- Both parents		**-.12**	**-.12**	**-.12**
Private tutoring (ongoing)	**-.56**		**-.06**	**-.09**
Interaction				
PT x mark in Term 1[a]				**-.06**
PT x ISEI[a]				-.05
R^2	.039	.621	.621	.622
School level				
Mean prior knowledge[b]		**.10**	**.11**	**.11**
Mean ISEI[b]		.01	.01	.01
Tracking				
- Comprehensive school[c]		**-.10**	**-.10**	**-.09**
- Academic track[c]		**-.10**	**-.10**	**-.11**
R^2	.027	.982	.984	.983

Notes. 7.6% of variance on school level (unconditional model). Dependent variable (math mark at the end of Grade 8) not standardized. Range of marks from 1 (insufficient/fail) to 6 (excellent).
[a]centered at the grand mean of the analyzed sample. [b]z-standardized within the KESS 8 cohort. [c]reference: lower secondary schools. PT = private tutoring. Significant predictors are printed in bold numbers.

In Model 1, students being tutored at the time of the study performed more than half a mark (b = -0.56) below those students who were not tutored at the time. In this model, the intercept is the mean (reversed) mark of the non-tutored students at the end of Grade 8 (b = 3.89).

In Model 2, on the individual level the most important predictor were the marks in Term 1 (b = 0.65). There were also small positive effects of the students' intelligence and their ability self-concept (both b = 0.10) and small negative effects of the students' migration background (b = -0.12, if both parents born abroad) and their achievement anxiety (b = -0.03). On the school level, a higher mean prior knowledge in a school was associated with a slightly higher mark at the end of Grade 8 (b = 0.10), while the attendance of an academic track as well as of a comprehensive school was associated with a slightly worse mark (b = -0.10). The predictors in this model explained R^2 = .62 of the variance on the individual and R^2 = .98 on the school level.

In Model 3 with the multiple control variables, the negative effect of private tutoring was much smaller. But still, students who were being tutored at the time of the study received a significantly poorer mark than those students who were not being tutored at the time (b = -0.06).

When checking the interaction terms individually, two significant effects were revealed. In the common Model 4 there remained only a small but significant negative interaction term between private tutoring and the students' mark in Term 1 (b = -0.06).

Both complete Models 3 and 4 hardly explain any additional variance on the individual level (R^2 = .62). Hence, despite the significance of the effect of private tutoring, this effect is very small.

5.3.2 Mathematic achievement test results

The effect of private tutoring on the mathematics achievement test results was also analyzed using hierarchical linear regression models. The model sequencing is organized as in the section before.

Table 3: Effects of private tutoring on mathematics achievement test results in Grade 8 – Hierarchical linear regression analysis (N = 4,701)

	Model 1	Model 2	Model 3	Model 4
Intercept	**0.22**	0.05	**0.07**	**0.07**
Individual level				
Prior knowledge[a] (Math test – Grade 7)		**.46**	**.45**	**.45**
Intelligence[a]		**.21**	**.21**	**.21**
ISEI[a]		**.03**	**.03**	**.04**
Interest in math[a]		-.01	-.01	-.01
Self-concept math[a]		**.09**	**.09**	**.09**
Achievement anxiety[a]		**-.03**	**-.03**	**-.03**
Boys		.03	.03	.03
Migration background				
- One parent		-.03	-.03	-.02
- Both parents		-.03	-.03	-.03
Private tutoring				
- In the past (Grade 7/8)	**-.46**		**-.09**	**-.08**
- Ongoing	**-.47**		**-.08**	**-.08**
Interaction				
PT in the past x ISEI[a]				**-.03**
R^2	.069	.473	.475	.475
School level				
Mean prior knowledge[a]		**.46**	**.47**	**.47**
Mean ISEI[a]		**-.15**	**-.14**	**-.14**
Tracking				
- Comprehensive school[b]		.02	.02	.02
- Academic track[b]		.02	.01	.01
R^2	.024	.983	.982	.982

Notes. 43.5% of variance on school level (unconditional model). Dependent variable (math test – Grade 8) z-standardized within the KESS 8 cohort.
[a] z-standardized within the KESS 8 cohort. [b] reference: lower secondary schools. PT = private tutoring. Significant predictors are printed in bold numbers.

In Model 1, students tutored at the time of the study or who had been tutored in the course of Grade 7 or 8, performed nearly half a standard deviation of the KESS 8 cohort (b = -0.47 and b = -0.46) below those students who had never been tutored during this time period. In this model, the intercept was the mean achievement of the non-tutored students at the end of Grade 8 which is slightly above the average of the KESS-8-cohort (b = 0.22).

In Model 2 on the individual level the most important predictor was the students' test achievement in Grade 7 (b = 0.46). There were also small positive effects of the students' intelligence (b = 0.21), their socio-economic status (b = 0.03), and their ability self-concept (b = 0.09) as well as a small negative effect of their achievement anxiety (b = -0.03). On the school level, there was a considerable effect of the mean prior knowledge in school (b = 0.46). Surprisingly, a higher mean socio-economic status in school was associated with a lower individual test achieve-

ment in Grade 8 (b = -0.21). The predictors in this model explain R^2 = .47 of the variance on individual and R^2 = .98 on the school level.

In Model 3 with the multiple control variables, the negative effect of private tutoring was much smaller. But still, students who were tutored at the time of the study or had been tutored during Grade 7 or 8, performed significantly worse than those students who had not been tutored during this time period (b = -0.09 and b = -0.08).

As presented in Model 4, there was only a small but significant interaction term between private tutoring in the past and the students' socio-economic status (b = -0.03).

Both, complete models 3 and 4 scarcely explain any additional variance on the individual level (R^2 = .48). Hence, despite the significance of the effect of private tutoring, this effect is very small.[1]

The analyses based on propensity score matching revealed no significant effects of private tutoring on the students' marks and test achievement. A detailed presentation of these results is given in Appendix A.2.

5.4 Comparing the criteria of effectiveness of private tutoring

Evaluating the effectiveness of private tutoring on different criteria obviously yields different results. Therefore, it is interesting to directly compare these different criteria. While in-depth analyses would go beyond the scope of this article, it was done on a descriptive level for those indicators which were easiest to contrast, i.e., the parents' and the students' assessments of the effectiveness of private tutoring, and the change in students' marks from Term 1 to Term 2. Parents' and students' assessments were pooled into two groups: those who (strongly) agreed with the effectiveness of the tutoring lessons, and those who (strongly) disagreed. Cross tabulations were used to compare the estimations because they more clearly reveal, than correlation analyses, whether parents and students evaluated only an improved or also a constant mark as evidence for the effectiveness of private tutoring.

1 Following the suggestion of an anonymous reviewer we checked private tutoring effects on marks and test achievement in each track separately, too. There were no significant differences to the private tutoring effects in the complete sample.

Table 4: Comparison of parents' and students' assessment of the effectiveness of private tutoring (N = 447)

Parents: performance improved by private tutoring	Students: performance improved by private tutoring					
	(strongly) agree		(strongly) disagree		Sum	
(strongly) agree	328	(73.4%)	43	(9.6%)	371	(83.0%)
(strongly) disagree	39	(8.7%)	76	(8.3%)	76	(17.0%)
Sum	367	(82.1%)	80	(17.9%)	447	(100.0%)

In their statement parents and students mostly agreed that private tutoring improved the students' performance in school (81.7% of corresponding statements). Only in a relatively small proportion of cases students saw no improvement, while their parents did (9.6%), or vice versa (8.7%).

Table 5: Parents' (N = 447) and students' (N = 618) assessment of the effectiveness of private tutoring compared to development of the students' marks from Term 1 to Term 2

Marks from Term 1 to Term 2	Parents: performance improved by PT						Students: performance improved by PT					
	(strongly) agree		(strongly) disagree		Sum		(strongly) agree		(strongly) disagree		Sum	
upgraded	71	(15.9%)	8	(1.8%)	79	(17.7%)	95	(15.4%)	20	(3.2%)	115	(18.6%)
no change	234	(52.3%)	54	(12.1%)	288	(64.4%)	305	(49.4%)	83	(13.4%)	388	(62.8%)
degraded	66	(14.8%)	14	(3.1%)	80	(17.9%)	97	(15.7%)	18	(2.9%)	115	(18.6%)
sum	371	(83.0%)	76	(17.0%)	447	(100.0%)	497	(80.4%)	121	(19.6%)	618	(100.0%)

Note. PT = private tutoring.

Most parents assessed their child's private tutoring as effective if the student's mark improved (15.9%), or remained the same (52.3%). But there were also a number of parents who saw no improvement in their child's performance, when his or her mark remained the same (12.1%), or – less surprisingly – when it went down (3.1%). Interestingly, there were some parents (14.8%) who assessed their child's private tutoring as effective despite the negative development of the student's marks. Parents, who did not report positive effects resulting from their child's private tutoring while the students' mark improved, were negligible (1.8%).

Most students assessed their private tutoring as effective if the mark improved (15.4%) or remained the same (49.4%). But there were also a number of students who saw no improvement in their performance when their marks remained the same (13.4%), or – less surprisingly – when they went down (2.9%). Interestingly, there were some students (15.7%) who assessed their private tutoring as effective despite the negative development of their marks. The percentage of students who did not state positive effects of their private tutoring while their mark improved, was negligible (3.2%).

6. Discussion

6.1 Effectiveness of private tutoring

Concerning the first and second research question our results from descriptive analyses confirmed prior research as well as our hypotheses: Parents and students were generally satisfied with the effect of the student's tutoring lessons. Concerning the third and fourth research question, neither mathematics marks nor mathematics achievement test results significantly improved through private tutoring. These later results are quite independent from the statistical approach. Neither the hierarchical regression analyses nor the propensity score matching approach revealed any positive effect of private tutoring on the students' achievement when a number of cognitive, motivational, family, and school context covariates were controlled for. That means, in line with previous research, effects of private tutoring often do not last if elaborated methodological approaches are employed. In comparison to previous research, the missing effect of private tutoring on students' marks is an exception and especially contradicts the positive effects of tutoring found by Mischo and Haag (2002), and Haag and Jäger (2009) in two German samples. The effects on achievement test results were already very mixed. Especially for the Grade 5/6 sample of the KESS study, effects of private tutoring on test achievement were also slightly negative (Guill & Bonsen, 2010). Limitations of our analyses will be discussed in the next section.

Concerning the fifth research question, the four criteria of the effectiveness of private tutoring employed in this study paint a different picture of the same phenomenon. As cross tabulations showed, parents and students mostly agreed in their statements (over 80% corresponding statements) about the effectiveness of private tutoring. Interestingly, students' and parents' satisfaction with the effects of private tutoring are opposite the absent effects of private tutoring regarding the more objective criteria marks and test results. As shown by cross tabulations, on a descriptive level parents and students were often already satisfied with the tutoring lessons if the students' marks showed no negative development from Term 1 to Term 2. In contrast, in England the high level of satisfaction of the parents went along with students' at least partly better GCSE results (Ireson & Rushforth, 2005). Reconsidering the offer-usage model of private tutoring, there are three possible explanations for our diverging results:

Firstly, parents' and students' expectations influence the focus of the private tutoring lessons. Improvements of marks might go beyond their level of aspiration. Maybe they decided on private tutoring when fearing lower marks, and therefore considered maintaining the marks as an improvement. Hence, a more detailed analysis of students' and parents' motives for choosing private tutoring would be interesting. Secondly, students and parents invest a lot of effort in private tutoring lessons in terms of money and spare time. Therefore, when evaluating the tutoring lessons they might not want to admit that these investments were all in vain and

overrate small positive effects. Overstating the effects of private tutoring might be an effect of dissonance reduction (Festinger, 1962). Thirdly, checking the students' homework and going over previous tests are part of the main activities during the tutoring lessons (Guill, 2010; Rudolph, 2002). They might give the students a false impression of competence. The students might be able to follow the tutor's explanations but fail to transfer their understanding to the mainstream classroom where they cannot rely on the tutor's assistance.

Still, a major result of this study is that private tutoring has not shown to improve academic achievement per se but its effects depend a lot on the operationalization of academic success. Future research on the effects of private tutoring should take this into account.

6.2 Limitations

On a methodological level, our study design has some limitations. General achievement tests as well as end of term mark might be too vague a measure to capture small improvements in academic achievement by private tutoring. Therefore, a more detailed diagnosis of the students' achievement development when attending private tutoring is desirable as it is employed, e.g., by Mischo and Haag (2002). They used marks in teacher-designed tests instead of end of term marks. One might also discuss class retention as a more adequate indicator of the effectiveness of private tutoring. But this is not an option for our analyses for two reasons: Firstly private tutoring is not only chosen to avoid class retention but also to improve already sufficient marks (Guill, 2012), and secondly we focus only on private tutoring in mathematics while class retention depends on a combination of insufficient marks in several subjects. As analyzed in Section 2.1, it might also be difficult for parents and students to assess the net effect of private tutoring. An alternative is asking them before and after the tutoring lessons whether they are satisfied with their own or their children's academic achievement.

A further limitation of our study is the large amount of missing data. Efforts to avoid systematic sample drop-outs are necessary in future studies. We opted for single imputation to replace missing values and therefore risked underestimating the standard errors of our estimates. This is mainly relevant for effects close to the significance level. However, for the given analyses only some small negative private tutoring effects are close to the significance level and these were not interpreted.

A specific form of systematic sample drop-out not mentioned yet is class repetition: Students – and possibly mainly non-tutored students – have left the sample to repeat a grade and are therefore missing as a relevant comparison group to the privately tutored students. Hence, at least the tutoring effect for the students whose private tutoring ended before the data collection and might have taken place in Grade 7 might be underestimated. But the analyses of the tutoring effect for the students still tutored during the data collection at the end of Grade 8 should be less affected by this problem. Tutoring mainly lasted up to one year (Guill, 2012)

and grade repetition is most common at the end of Term 2. Therefore, the relevant comparison group, that is students with the same achievement level, should still be in the sample when the data collection took place before the summer vacation.

The HLM as well as the PSM analyses validate each other and demonstrate that there are no positive effects of private tutoring on the students' achievement (see Appendix A.3 for a short discussion of PSM specific aspects).

Both forms of analysis assume that there are no omitted variables concerning systematic differences between privately tutored and non-tutored students (Ho, Imai, King, & Stuart, 2007). Therefore, it is necessary to control for further systematic differences between the groups as long as randomly controlled designs are hard to establish. The offer-usage model of private tutoring implies that the students' motivation and willingness to exert themselves and their ability to self-regulate learning might be important differences between tutored and non-tutored students and an explanation of the nominal small but negative effects of private tutoring revealed by the HL analyses.

Even if the HL models to predict marks and test achievement include the most important determinants of academic achievement, they are still parsimonious. The cognitive and motivational variables focus mainly on math related variables while more general indicators like the general academic self-concept are not included. Still we would defend that these domain specific predictors capture most of the variance of the dependent variables math marks and test achievement as they are more closely related to them. Adding additional variables like the general academic self-concept would mainly increase collinearity between predictors as it also contains aspects of the ability self-concept in mathematics (Marsh, 1990b).

Furthermore, all motivational indicators measured at the beginning of Grade 7 as well as the students' marks are influenced by reference group effects. These reference group effects are mainly caused by the tracking in the German school system which the students have already experienced for two years before the KESS 7 data collection. We see this as an advantage of our control variables as they are temporarily closer to the dependent variables than any variable measured before tracking started, and as they are measured in the context in which we are interested in their effects. Still, for reasons of parsimony our HL models imply that the predictors work similarly in the different reference groups on the school level which is supported by small and mainly non-significant amounts of variance in the slopes of these predictors in the HLM analyses.

6.3 Future outlook

Beyond these limitations and keeping in mind the existing variety of tutoring settings, our results should be validated with samples from different age cohorts, for additional school subjects, and taking into account the international context, in different cultural contexts. It would be interesting to know if it is a generalizable pattern that parents' and students' satisfaction with the tutoring lessons surpass the

objectively measurable results, or if other aspects are responsible for the internationally diverging results on the effectiveness of private tutoring. One might assume, e.g., that private tutoring is especially effective in underfunded educational systems where every additional support for the students fills a large gap.

Above all, research on private tutoring should not only check lists of potential factors working during the tutoring lessons but develop and test comprehensive models of its mode of operation and its effectiveness. The offer-usage model of private tutoring sketched in Section 2 seems to be a fruitful base for the discussion of our findings as well as for the development of further hypotheses on the effectiveness of private tutoring.

For example, at a first glance it is contra-intuitive that private tutoring should not help given the additional time on task it offers to the students. However, the offer-usage model suggests that not all tutees will be motivated to use the additional instructional offer. Additionally, even if the tutees are motivated they might not be capable to transfer learning gains from the tutoring lessons to the mainstream classroom, e.g., due to diverging didactic approaches in the two settings.

Essentially, the offer-usage model suggests three directions for future research on the effects of private tutoring: Firstly, the parents', tutees' and tutors aims concerning the tutoring lessons and their effects on the instructional process should be analyzed in more detail. For example, the hypothesis that private tutoring focuses on short-term training effects for teacher-designed tests instead of lasting competencies still has to be tested (Guill & Bonsen, 2010). Therefore, a closer look at the effects of tutoring on teacher dependent marks compared to the more distal achievement tests would be interesting.

Secondly, the offer-usage process during the tutoring lessons should be focused. Analyzing teacher-student dialogues (Hohoff, 2002) could be a way to learn more about the teaching and learning processes during the tutoring lessons. It would be interesting to know which teaching methods are employed, and to research their effectiveness in connection with the learning potential of the tutees. Our own operationalization of private tutoring is rather distal and we focus on the summary effect of all forms of private tutoring in the whole sample. For future analyses, evaluating the duration and intensity of private tutoring would allow discovering non-linear effects of private tutoring, as, e.g., findings from Singapore (Cheo & Quah, 2005) suggest. Additionally, while there is growing research on the cognitive effects of private tutoring, hardly anything is known – apart from anecdotal evidence – about its motivational and emotional effects, e.g., on the students' ability self-concept or their achievement anxiety (an exception are Mischo & Haag, 2002).

Thirdly, a further important point of investigation is the interface between learning in public school and private tutoring: how does the tutor diagnose his tutee's learning deficiencies at school? How does he know what his or her tutee is supposed to perform at school? How do students transfer competencies learned in the protected setting of the tutoring lessons to their classroom?

Taken together, focusing on these three areas of research would considerably improve our theoretical and empirical knowledge on the effects of private tutoring.

6.4 Practical implications

Due to the diverging results regarding the effectiveness of private tutoring, it is difficult to formulate practical implications at this point. At least parents and students should be advised to look very critically at their tutoring lessons in terms of meeting their expectations. Looking at the neutral effects of private tutoring on the students' marks and test achievement, this study gives no indication that private tutoring is an important mediator to convert economic capital into educational advantages and increase social disparities in academic achievement. Political initiatives aiming to subsidize private tutoring to support underachievers in school (Bray, 2009), should therefore be critically evaluated.

Acknowledgments

The authors would like to thank Katrin Lintorf for valuable comments and feedback to earlier drafts of this article. We also wish to thank Cornelia A. Gerigk for proofreading and correcting the English text.

References

Baker, D. P., Akiba, M., LeTendre, G. K., & Wiseman, A. W. (2001). Worldwide shadow education: Outside-school learning, institutional quality of schooling, and cross-national mathematics achievement. *Educational Evaluation and Policy Analysis, 23*(1), 1–17.

Bonsen, M., Gröhlich, C., & Bos, W. (2009). Differentielle Lern- und Entwicklungsmilieus in der Hamburger Beobachtungsstufe? [Differential environments for learning and development in the Hamburg observation stage?] In W. Bos, M. Bonsen, & C. Gröhlich (Eds.), *KESS 7. Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen zu Beginn der Jahrgangsstufe 7* (pp. 113–122). Münster, Germany: Waxmann.

Bos, W., Gröhlich, C., Guill, K., Scharenberg, K., & Wendt, H. (2010). Ziele und Anlage der Studie KESS 8 [Aims and design of the KESS 8 Study]. In W. Bos & C. Gröhlich (Eds.), *KESS 8. Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8* (pp. 9–20). Münster, Germany: Waxmann.

Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Valtin, R., Voss, A., & Walther, G. (2005). *IGLU. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente* [IGLU. Handbook of scales to document the instruments]. Münster, Germany: Waxmann.

Bray, M. (1999). *The shadow education system: Private tutoring and its implications for planners*. Paris, France: UNESCO International Institute for Educational Planning (IIEP).

Bray, M. (2009). *Confronting the shadow education system: What government policies for what private tutoring?* Paris, France: UNESCO International Institute for Educational Planning (IIEP).

Buchmann, C. (2002). Getting ahead in Kenya: Social capital, shadow education and achievement. In A. Nutt, E. Hannum, & B. Fuller (Eds.), *Schooling and social capital in diverse cultures* (pp. 133–159). Amsterdam, Netherlands: JAI/Elsevier.

Cheo, R., & Quah, E. (2005). Mothers, maids and tutors: An empirical evaluation of their effect on children's academic grades in Singapore. *Education Economics, 13*(3), 269–285.

Dang, H.-A. (2007). The determinants and impact of private tutoring classes in Vietnam. *Economics of Education Review, 26*, 683–698.

Elvira, J. A. M., Cívico, F. A., Cabrera, R. A., Osuna, J. P., Cabrera, J. H., & Olivares, R. R. (2006). Actividades extraescolares y rendimiento académico en alumnos de educación secundaria [Extra-curricular activities and private tutoring between secondary school students]. *Revista Electrónica de Investigación Psicoeducativa, 4*(1), 35–46.

Festinger, L. (1962). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Ganzeboom, H. B. G., & Treiman, D. J. (2003). Three internationally standardised measures for comparative research on occupational status. In J. H. P. Hoffmeyer-Zlotnik & C. Wolf (Eds.), *Advances in cross-national comparison* (pp. 159–193). New York, NY: Kluwer Academic/Plenum Publishers.

Guill, K. (2010, July). *Effectiveness of private tutoring*. Poster presented at the conference of Junior Researches of EARLI (JURE), Frankfurt, Germany.

Guill, K. (2012). *Nachhilfeunterricht. Individuelle, familiäre und schulische Prädiktoren* [Private tutoring. Individual, family and school level predictors]. Münster, Germany: Waxmann.

Guill, K., & Bonsen, M. (2010). Die Effektivität von Nachhilfeunterricht in Mathematik am Beginn der Sekundarstufe I – Ergebnisse der Hamburger Schulleistungsstudie KESS 7 [Achievement gains through private tutoring in mathematics at the beginning of secondary education? – Results from the German (Hamburg) School Achievement Study KESS]. *Unterrichtswissenschaft, 38*(2), 117–133.

Guill, K., Gröhlich, C., Scharenberg, K., Wendt, H., & Bos, W. (2010). Die mathematischen Kompetenzen der Schülerinnen und Schüler [Mathematical competencies of students]. In W. Bos & C. Gröhlich (Eds.), *KESS 8. Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8* (pp. 37–48). Münster, Germany: Waxmann.

Haag, L., & Jäger, R. S. (2009). Potentiale von Nachhilfeunterricht [Capabilities of private tutoring]. *Empirische Pädagogik, 23*(3), 277–289.

Hamid, M. O., Sussex, R., & Khan, A. (2009). Private tutoring in English for secondary school students in Bangladesh. *Tesol Quarterly, 43*(2), 281–308.

Heller, K. A., & Perleth, C. (2000). *KFT 4-12+R – Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision* [Cognitive abilities test from Grade 4 to 12, revision]. Göttingen, Germany: Beltz Test.

Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität* [Instructional quality and teacher professionality]. Seelze-Velber, Germany: Klett.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as a nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis, 15*, 199–236.

Hohoff, S. (2002). Mathe auf Rezept. Eine Interaktionsanalyse von Nachhilfeunterricht in Mathematik [Prescription for Math. An interaction analysis of private tutoring in mathematics]. *Journal für Mathematik-Didaktik, 23*, 3–27.

Ingenkamp, K. (1993). Der Prognosewert von Zensuren, Lehrergutachten, Aufnahmeprüfungen und Tests während der Grundschulzeit für den Sekundarschulerfolg [The projected value of marks, teachers' recommendations, entrance exams and tests during primary school regarding success in secondary school]. In R.

Olechowski & E. Persy (Eds.), *Frühe schulische Auslese* (pp. 68–85). Frankfurt, Germany: Peter Lang.

Ireson, J., & Rushforth, K. (2005). *Mapping and evaluating shadow education* (ESRC Research Project RES-000-23-017 – End of Award Report). London, UK: Institute of Education, University of London.

Kang, C. (2007). *Does money matter? The effect of private educational expenditures on academic performance*. Singapore: Department of Economics, National University of Singapore.

Kenny, D. T., & Faunce, G. (2004). Effects of academic coaching on elementary and secondary school students. *Journal of Educational Research, 98*, 115–126.

Leven, I., Quenzel, G., & Hurrelmann, K. (2010). Familie, Schule, Freizeit: Kontinuitäten im Wandel [Family, school, leisure time: Transitioning continuities]. In Shell Deutschland Holding (Ed.), *Jugend 2010. Eine pragmatische Generation behauptet sich* (pp. 53–128). Frankfurt, Germany: Fischer.

Marsh, H. W. (1990a). *Self Description Questionnaire (SDQ) II: A theoretical and empirical basis for the measurement of multiple dimensions of adolescents self-concept: An interim test manual and a research monograph*. San Antonio, TX: The Psychological Corporation.

Marsh, H. W. (1990b). The structure of academic self-concept: The Marsh/Shevelson model. *Journal of Educational Psychology, 82*(4), 623–636.

Mischo, C., & Haag, L. (2002). Expansion and effectiveness of private tutoring. *European Journal of Psychology of Education, 17*(3), 263–273.

Mischo, C., & Kessel, M. von (2005). Wie wirkt Nachhilfe? Mögliche Wirkfaktoren im Wissen der Nachhilfelehrer [How does private tutoring work? Possible effective factors in the knowledge base of tutoring teachers]. *Empirische Pädagogik, 19*(1), 28–46.

Nath, S. R. (2008). Private supplementary tutoring among primary students in Bangladesh. *Educational Studies, 34*(1), 55–72.

Painter (2004). *SPSS macro for propensity score matching*. Retrieved from http://ssw.unc.edu/VRC/Lectures/index.htm

Paviot, L., Heinsohn, N., & Korkman, J. (2008). Extra tuition in Southern and Eastern Africa: Coverage, growth, and linkages with pupil achievement. *International Journal of Educational Development, 28*, 149–160.

Polydorides, G. (1986, April). *The determinants of educational achievement at the end of secondary schooling: The case of Greece*. Paper presented at The Annual Meeting of the American Educational Research Association (AERA). San Francisco, CA.

Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2004). *HLM 6 for Windows* [Computer software]. Lincolnwood, IL: Scientific Software International, Inc.

Rudolph, M. (2002). *Nachhilfe – Gekaufte Bildung? Empirische Untersuchung zur Kritik der außerschulischen Lernbegleitung* [Private tutoring – Purchased education?]. Bad Heilbrunn, Germany: Klinkhardt.

Ryu, D., & Kang, C. (2010). *Do private tutoring expenditures raise academic performance? Evidence from middle school students in South Korea*. Retrieved from: http://yeri.yonsei.ac.kr/new/bbs/data/workshop/101202_01.pdf

Schafer, J. L. (1997). *Analysis of incomplete multivariate data* (Monographs on statistics and applied probability, Vol. 72). London, UK: Chapman & Hall.

Schafer, J. L. (1999). *NORM: Multiple imputation of incomplete multivariate data under a normal model, version 2. Software for Windows 95/98/NT*. Retrieved from http://www.stat.psu.edu/~jls/misoftwa.html

Silova, I., & Bray, M. (2005). The hidden marketplace: Private tutoring in former socialist countries. In I. Silova, V. Budiene, & M. Bray (Eds.), *Education in a hidden marketplace, monitoring of private tutoring. Overview and country reports* (pp. 71–98). New York, NY: Open Society Institute.

Smyth, E. (2008). The more, the better? Intensity of involvement in private tuition and examination performance. *Educational Research and Evaluation, 14*(5), 465–476.

Smyth, E. (2009). Buying your way into college? Private tuition and the transition to higher education in Ireland. *Oxford Review of Education, 35*(1), 1–22.

Stevenson, D. L., & Baker, D. P. (1992). Shadow education and allocation in formal schooling: Transition to university in Japan. *American Journal of Sociology, 97*(6), 1639–1657.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science, 25*(1), 1–21.

Synovate Kids+Teens, & Bundesverband Nachhilfe- und Nachmittagsschulen e. V. (2007). *Mit Nachhilfe kommt man weiter – Fakten zur Nachhilfesituation in Deutschland* [Advancing through tutoring – Facts about the private tutoring situation in Germany]. Retrieved from http://www.nachhilfe.de/presse_nachhilfe_studien.html

Tansel, A., & Bircan, F. (2005). *Effect of private tutoring on university entrance examination performance in Turkey*. Bonn, Germany: Forschungsinstitut zur Zukunft der Arbeit.

Thoemmes, F. J., & Kim, E. S. (2011). A systematic review of propensity score methods in the social sciences. *Multivariate Behavioral Research, 46*, 90–118.

Ünal, H., Özkan, E. M., Milton, S., Price, K., & Curva, F. (2010). The effect of private tutoring on performance in mathematics in Turkey: A comparison across occupational types. *Procedia Social and Behavioral Sciences, 2*, 5512–5517.

Valtin, R., & Darge, K. (2001). *Gute Noten für die sechsjährige Grundschule* [Good grades for the six-year elementary school]. Retrieved from http://www.gew-berlin.de/blz/314.htm

Appendix

A.1 Analyses: Propensity score matching

The basic idea of propensity score matching is to match each tutored student with a non-tutored student having identical characteristics. The aim is to eliminate or to reduce systematic differences between the two groups not due to private tutoring. More precisely, PSM aims to balance the distributions of the covariates in the groups of tutored and non-tutored students. The students are matched based on the propensity score, i.e., their probability to attend private tutoring. The propensity score summarizes the covariates into one scalar (Stuart, 2010; Ho et al., 2007). It was estimated with logistic regression analysis.

We chose those variables as predictors of the propensity score which reflected the decision for or against private tutoring (Guill, 2012) and further variables which might cover systematic differences between tutored and non-tutored students (Guill, 2012). These are the individual predictors in the regression model, tracks, and the students' reading interest, their English achievement test results, and their ability self-concept. These later variables are meant to cover differences in verbal skills and motivation. As we replaced missing values only using single imputation we added missing data indicators for all predictors as additional predictors of the propensity score thereby following Stuart's (2010) advice. Naturally, PSM cannot reduce the bias by unmeasured confounders unless they are correlated with the measured covariates (Ho, Imai, King, & Stuart, 2007).

Following Stuart's (2010) guidance for practice, we employed *k*:1, precisely 1:1 nearest neighbor matching without replacement as matching algorithm because there are about five times more non-tutored or control students than tutored students. The matching procedure aimed to match a comparable non-tutored student to each tutored student while excluding the non-matched non-tutored students from further analyses. Matching was performed with a SPSS macro adapted for PSM by Painter (2004). It was performed for the complete sample and separately within each track to control for differences on school level between tutored and non-tutored students.

Math marks at the end of Term 2 and math achievement test results of the tutored and non-tutored students in the parallelized samples were compared by simple *t* tests for independent samples. Thus, we estimate the average treatment effect on the treated effect (ATT) that is the effect of private tutoring on those students eligible for private tutoring.

A.2 Results: Estimation of group differences by propensity score matching

As explained in the method section, propensity score matching was chosen as a second methodological approach to validate the results from the hierarchical linear regression analyses. A first matching was performed for the whole sample of tutored students in all secondary tracks. In a second step, the matching was performed for each academic track separately. These separate data sets are important because a substantial ratio of variance of the dependent variables is on school level and mainly caused by tracking effects.

When performing propensity score matching one first has to check whether the matching procedure eliminated systematic bias between the samples of tutored and non-tutored students. There are several established measures of balance. Firstly, using *t* tests we checked that there are few significant differences between tutored and non-tutored students on all matching variables while there were a number of significant differences in the complete sample (see Table 1). Secondly, as *t* tests are insufficient as a measure of balance, the standardized differences in means between the tutored and the non-tutored groups were calculated for the propensity score and all covariates. As required, the absolute values are smaller than 0.25 and the variance ratios of the two groups lay between 0.5 and 2.0 (Stuart, 2010). The standardized differences of means of the propensity score and most of the covariates decreased in the matched samples compared to the complete samples. They increased for a few covariates whose differences were already very small in the complete sample (see Table A1).

Table A1: Balance diagnostics of the matched samples of tutored and non-tutored students I: Standardized differences in means and variance ratios

		All tracks		Lower secondary school		Comprehensive school		Academic track	
		M Diff	$\sigma^2_{PT}/\sigma^2_{nPT}$	*M* Diff	$\sigma^2_{PT}/\sigma^2_{nPT}$	*M* Diff	$\sigma^2_{PT}/\sigma^2_{nPT}$	*M* Diff	$\sigma^2_{PT}/\sigma^2_{nPT}$
Propensity score	before matching	0.845	(1.761)	0.646	(0.013)	0.766	(0.015)	0.952	(2.635)
	after matching	0.025	(1.084)	0.030	(1.120)	-0.013	(0.965)	0.140	(1.600)
Math: prior knowledge (Grade 7)	before matching	-0.540	(0.642)	-0.287	(0.843)	-0.530	(0.679)	-0.848	(0.632)
	after matching	0.052	(0.845)	0.054	(1.063)	-0.032	(0.843)	-0.030	(0.913)
Math mark - Term 1	before matching	-0.689	(0.821)	-0.387	(1.125)	-0.406	(0.859)	-1.000	(0.758)
	after matching	-0.003	(1.037)	-0.022	(1.404)	-0.047	(1.143)	-0.042	(1.213)
Intelligence	before matching	-0.394	(1.006)	-0.274	(0.911)	-0.416	(0.983)	-0.500	(1.217)
	after matching	0.003	(0.885)	-0.050	(0.977)	-0.098	(0.856)	-0.040	(0.914)
HISEI	before matching	-0.072	(0.961)	-0.049	(1.066)	0.020	(0.901)	-0.130	(1.034)
	after matching	0.033	(0.954)	-0.014	(1.054)	-0.029	(0.990)	0.024	(1.095)
Interest in math	before matching	-0.282	(1.042)	-0.084	(1.241)	-0.334	(1.032)	-0.360	(0.939)
	after matching	-0.005	(1.050)	-0.027	(1.285)	0.062	(0.970)	-0.030	(1.087)
Self-concept in math	before matching	-0.556	(0.926)	-0.156	(1.073)	-0.494	(0.832)	-0.829	(0.831)
	after matching	-0.012	(1.000)	-0.056	(1.150)	0.078	(0.868)	-0.040	(1.014)
Achievement anxiety	before matching	0.155	(1.127)	0.168	(1.091)	0.058	(1.043)	0.209	(1.236)
	after matching	0.013	(0.979)	-0.020	(1.015)	-0.090	(0.900)	0.042	(1.081)
Gender	before matching	-1.102	(0.975)	0.005	(1.006)	0.047	(1.009)	-0.251	(0.879)
	after matching	-0.099	(0.974)	-0.045	(1.006)	0.056	(1.005)	-0.057	(0.959)

Notes. *M* Diff = standardized difference in means between samples of tutored and non-tutored relative to *SD* of the tutored student sample.
PT = private tuoring. nPT = no private tutoring.

continued

Table A1: *(continued)*

		All tracks		Lower secondary school		Comprehensive school		Academic track	
		M Diff	$\sigma^2_{PT}/\sigma^2_{nPT}$	*M* Diff	$\sigma^2_{PT}/\sigma^2_{nPT}$	*M* Diff	$\sigma^2_{PT}/\sigma^2_{nPT}$	*M* Diff	$\sigma^2_{PT}/\sigma^2_{nPT}$
Migration (one parent)	before matching	0.075	(1.210)	-0.323	(0.406)	0.079	(1.213)	0.166	(1.498)
	after matching	0.048	(1.122)	0.039	(1.240)	0.000	(1.000)	0.044	(1.090)
Migration (both parents)	before matching	-0.038	(0.946)	-0.095	(0.916)	-0.052	(0.930)	0.002	(1.007)
	after matching	-0.045	(0.935)	0.051	(1.065)	0.014	(1.023)	0.045	(1.092)
Comprehensive school	before matching	0.051	(1.054)						
	after matching	0.011	(1.010)						
Academic track	before matching	-0.024	(1.001)						
	after matching	0.039	(1.003)						
English: achievement (Grade 8)	before matching	-0.143	(1.060)	-0.150	(1.249)	-0.219	(1.183)	-0.168	(0.827)
	after matching	0.096	(1.063)	0.035	(1.073)	0.043	(1.018)	0.083	(0.861)
English: prior knowledge (Grade 7)	before matching	-0.132	(0.911)	-0.059	(1.136)	-0.188	(0.851)	-0.187	(0.829)
	after matching	0.057	(0.900)	0.051	(1.277)	-0.025	(0.806)	0.060	(0.934)
Self-concept in English	before matching	-0.002	(0.967)	-0.040	(0.907)	-0.110	(0.953)	0.088	(0.944)
	after matching	0.036	(0.906)	0.134	(0.801)	0.069	(0.910)	0.063	(1.104)
Reading interest	before matching	0.845	(1.761)	0.124	(1.278)	-0.020	(0.951)	-0.073	(1.103)
	after matching	0.011	(1.037)	0.150	(1.051)	0.067	(1.026)	0.033	(1.043)

Notes. *M* Diff = standardized difference in means between samples of tutored and non-tutored relative to *SD* of the tutored student sample. PT = private tuoring. nPT = no private tutoring.

Thirdly, we compared the ranges of the propensity scores for the tutored and the matched non-tutored students in each sample and checked their overlap, that is, the region of common support. These measures are given in Table A2. As there are very few tutored students (0% to 2.3%) outside of the region of common support they are not eliminated from the samples to keep the sample sizes of tutored students equal between descriptive, HLM, and PSM analyses.

Table A2: Balance diagnostics of the matched samples of tutored and non-tutored students II: Ranges of propensity scores for matched students and regions of common support

	All tracks	Lower secondary school	Comprehensive school	Academic track
N	2 x 618 = 1236	2 x 134 = 268	2 x 179 = 358	2 x 305 = 610
Propensiy score range private tutored students	0.008–0.843	0.020–0.802	0.020–0.656	0.009–0.922
Propensiy score range non-tutored students	0.008–0.878	0.020–0.779	0.021–0.673	0.009–0.811
Region of common support	0.008–0.843	0.020–0.779	0.020–0.673	0.009–0.811
% of tutored students outside of the region of common support	0.0%	1.5%	0.6%	2.3%

Given the sufficient balance between the samples of tutored and non-tutored students it is possible to estimate the effect of private tutoring. Therefore we compared the mathematics marks in Term 2 of Grade 8, and the Grade 8 mathematics test achievement of the tutored and non-tutored students using *t* tests. Table A3 shows that privately tutored students and their matched counterparts without private tutoring, neither differed in their marks in Term 2 of Grade 8, nor in their mathematics achievement test results. The small but nominal differences never reached statistical significance. This holds true for the sample of students from all tracks as well as the subsamples in the different tracks.

Table A3: Marks and achievement test results of tutored and non-tutored students in Grade 8 in matched samples

	Ongoing private tutoring		No ongoing private tutoring		*t*	(*p*)	*N*
	M	(*SD*)	*M*	(*SD*)			
All tracks							
Marks in Term 2 (Grade 8)	3.36	(0.82)	3.44	(0.85)	1.67	(0.10)	618
Math test achievement 8	-0.09	(0.85)	-0.09	(0.91)	-0.00	(1.00)	618
Lower secondary school							
Marks in Term 2 (Grade 8)	3.25	(0.93)	3.40	(0.81)	1.48	(0.14)	134
Math test achievement 8	-0.57	(0.72)	-0.58	(0.76)	-0.14	(0.89)	134
Comprehensive school							
Marks in Term 2 (Grade 8)	3.40	(0.75)	3.39	(0.87)	0.29	(0.78)	179
Math test achievement 8	-0.57	(0.73)	-0.55	(0.73)	0.20	(0.85)	179
Academic track							
Marks in Term 2 (Grade 8)	3.39	(0.81)	3.41	(0.77)	0.26	(0.80)	305
Math test achievement 8	0.40	(0.67)	0.44	(0.72)	0.63	(0.53)	305

Notes. Range of marks from 1 (insufficient/fail) to 6 (excellent). Test achievement *z*-standardized within the KESS 8 cohort. Groups matched by marks in math (Term 1), math and English test achievement (Grade 7), intelligence, ISEI, achievement anxiety, interest in mathematics, reading interest, ability self-concept in mathematics and English, gender, migration background and tracking.

A.3 Discussion of limitations of the PSM analyses

In summary, the numerical balance diagnostics of the matching result were satisfying. However, during the matching process a greater number of cases were eliminated. Alternative matching algorithms like *k*:1-matching or full matching would also allow the use of information of these cases, and improve further analyses (Stuart, 2010). Additional variables to improve the modeling of the selection procedure into private tutoring and thereby the estimation of the propensity score are discussed in the main section (see Sections 6.2–6.3). Although the track-specific PS models should have captured most of the variance of the selection procedure on a school level, hierarchical logistic regression analysis to estimate the propensity score is an alternative approach to reduce selection bias due to school level variables.

Karin Guill

Nachhilfeunterricht

Individuelle, familiäre und schulische Prädiktoren

2012, 220 Seiten, br., 29,90 €
ISBN 978-3-8309-2693-1
E-Book-Preis: 26,99 €

In diesem Buch werden auf Basis der Rational-Choice-Theorie Modelle zur Auswirkung von Nachhilfeunterricht auf den erwarteten Nutzen vorgestellt. Mittels dieser Modelle werden Hypothesen generiert, wie spezifische individuelle, familiäre und schulische Faktoren die Entscheidung für Nachhilfeunterricht beeinflussen. Diese Hypothesen werden mit (hierarchischen) logistischen Regressionsanalysen auf Basis von Daten der Hamburger Längsschnittstudie „Kompetenzen und Einstellungen von Schülerinnen und Schülern“ (KESS) empirisch überprüft.

WAXMANN