Are We Wasting Our Time At School? Causal Evidence of the Impact of Shortening Secondary School Duration

Bettina Büttner*

University of Magdeburg

Stephan L. Thomsen†

University of Magdeburg & ZEW Mannheim

This version: November 18, 2009

Abstract

This paper analyzes the impact of shortening the length of higher secondary schooling on human capital accumulation. A policy reform in the German state of Saxony-Anhalt implemented in 2003 provides a natural experiment by abolishing the 13th year for students in the 9th grade and leaving students in the 10th grade unaffected. The curriculum remained almost unchanged. Using data from the double cohort of graduates in 2007, we find significant negative effects of shortening secondary school duration on student performance in math for both gender and in English for females, whereas the effects on literature are not statistically significant.

Keywords: student performance, school duration, learning intensity, natural experiment

JEL Classification: I21, J18, C21

*We would like to thank Narine Karakanyan, Andrea Mühlenweg, Louis-Philippe Morin, Patrick Puhani, Guido Schwerdt as well as seminar participants in Hannover and Magdeburg for helpful comments. Christian Rusche and Christoph Wiese provided important research assistance. Bettina Büttner is a Research Associate at Otto-von-Guericke-University Magdeburg, Department of Economics and Management, PO Box 4120, D-39016 Magdeburg, e-mail: bettina.buettner@ovgu.de, phone: +49 391 6718175.

†Stephan L. Thomsen is an Assistant Professor of Labor Economics at Otto-von-Guericke-University Magdeburg and Research Associate at ZEW, Mannheim. Address: Otto-von-Guericke-University, Department of Economics and Management, PO Box 4120, D-39016 Magdeburg, e-mail: stephan.thomsen@ovgu.de, phone: +49 391 6718431, fax: +49 391 6711700. Stephan L. Thomsen thanks the Stifterverband für die Deutsche Wissenschaft (Claussen-Simon-Stiftung) for financial support.

1 Introduction

Education policy which raises student performance is on the top of the public agenda because schooling is one of the most important factors in human capital accumulation and given the accelerating technological changes and fastering global economic competition the importance of human capital for economic growth and for individual income has been increased.

In the past education policy concentrated more on the quantity of schooling. Compulsory school laws, minimum school leaving laws and extension of schooling time for reaching the university entrance qualification were implemented to enhance educational outcome.[1] But the opportunity costs of additional schooling are high: less time for further education, less time for work experience, less time to earn income, less time for starting a family.

Education policy should therefore stronger focus on the substitution of quantity for quality.[2] Therefore an important question is whether an increase of the learning intensity, i.e. the ratio of curriculum per instruction time, by shortening schooling leaving the curriculum unchanged could be a chance to return to shorter school duration without affecting the level of education. Since little is known about the relationship between learning intensity as one factor of educational quality and student performance as a measure for human capital skills, it has to be regarded, how variation of the learning intensity affects human capital accumulation.

In this paper, we estimate the causal effects of higher learning intensities on student performance. Due to the fact that German graduates from the most academic track of secondary schooling (*Gymnasium*) are comparatively old[3], almost all German federal states decided to reduce the length of secondary schooling by abolishing the 13th grade without changing the requirements. By leaving the curriculum almost unchanged this major German school reform increased the learning intensity for the treated students considerably. The change in schooling laws was at first enacted in 2003 and realized in 2007 in the federal state of Saxony-Anhalt followed by other German federal states. We use the reform in Saxony-Anhalt as a natural experiment to compare the student achievement of graduates of the double cohort.

[1] In line with that, previous international studies of student performance such as TIMSS and PISA show significant performance differences among students in two adjacent grades in literacy, math and science for most OECD countries (Woessmann, 2003; Fuchs and Woessmann, 2007; OECD, 2002, 2004, 2007). Students in higher grades scored considerably better than students in lower grades.

[2] An increasing body of educational and economical research investigate how school policy, teaching quality and the educational environment effect achievement. See, for example, the educational research reviews by Teddlie and Reynolds (2000) and Creemers and Kyriakides (2006) on school effectiveness, or Hanushek (2005) on the economics of school quality.

[3] Theoretical ages at graduation from secondary schooling are provided by OECD (2005); in Germany, students are aged 19 whereas, e.g., in the Netherlands graduation age is 17-18 years, 18 years in the US, and 17 years in Russia.

The estimated effects of increasing learning intensities on student performance depend on subjects and differ by gender. We find significant negative effects on student performance in math for females and more pronounced negative effects for males. Student performance in foreign language has also decreased due to the reform for both gender, but the effects for males are statistically not significant. In contrast to that, no differences are obtained on achievement scores in literature.

There are only very few studies on the effects of increasing learning intensity on student performance. Pischke (2007) investigates the impact of shortening the instructional time by two short school years 1966-7 in West Germany on grade repetition, higher secondary school track, earnings and employment. He found no negative effects on earnings and employment but an increase in grade repetition and lower track choice. Moreover, it is the only study considering policy-induced variation in schooling time without changing the curriculum. However, as there was no standardized testing system in Germany and no central examinations Pischke (2007) did not estimate the effects on student performance directly. Hence, the scope of deriving insights into the development of human capital is limited for that reason. Moreover, translating the results to today's students may be complicated since the composition of students with regard to the schooling degrees obtained has changed substantially towards higher shares in highest secondary education. Further evidence is provided by Skirbekk (2006) who investigates the effect of variation in school duration on human capital within the highest track using test scores from TIMSS for different Swiss cantons. He finds that differences in the length of the Swiss academic track across regions have no influence on students' math and science performance when school effects are controlled for.

The impact of a considerable lower magnitude of variation in instructional time on student performance is examined for example by Marcotte (2007), Lee and Barro (2001), and Woessmann (2003). Marcotte (2007) uses variation in school days caused by winter weather to identify the impact of increased learning intensity on test scores using Canadian data. His findings are in line with our results. Students with less instructional time perform significantly worse than their peers most notably in math. Lee and Barro (2001) investigate the effects of school resources on student performance measured by internationally comparable test scores across countries. They find significant positive effects of the length of the school term on math and science scores, but significant negative effects for reading. Woessmann (2003) finds significant positive but relatively small effects of instruction time on student performance in math and science. These evidence suggest that learning intensity in some subjects could be increased without negative effects on students' knowledge.

Similar to the German schooling reform also in Canada's province Ontario the length of high school was reduced by one year. However, the major difference with the German reform consists in the more modified curriculum. In Ontario less courses in main subjects like math and English were available to the treated group and, therefore, the impact of the reform on learning intensity is not determinable. Moreover, the 13th grade was not a full-fledged grade like in Germany. Students in Ontario were able to complete their high school after grade 12 also before the reform or complete their degrees by another year to take all of the courses they would like. Morin (2009) estimated the effect of abolishing the highest grade on the performance of high-ability students within their first university year. He found only small effects on student performance. However, Krashinsky (2006) found larger negative impacts on academic performance in university analyzing the impact of the same reform on students with lower high-school averages. In addition to the differences with respect to learning intensity their analysis varies from ours because we control for more student's background information. Another advantage of our study is the fact that we estimate the student performance instantaneously at the end of schooling. Because all students have to take the final exams, we do not have to cope with a potential self-selection as Morin (2009) and Krashinsky (2006) who measure the performance later and only for university students.[4]

Our paper contributes to the existing literature in several respects. We study a policy-induced large scale variation in the length of secondary schooling with only minor changes in curricula which resulted in a considerably increased learning intensity. Identical final written exams for both grades allow us to assess school performance directly. For our empirical analysis we use novel data collected in a survey with students of the double cohort graduated together in 2007. Thus, we can control for a number of student performance influencing factors like family background, student ability and school fixed effects. Based on the comprehensive information, we check the reliability of the assumptions of the natural experiment used to identify the reform effect.

The structure of the rest of the paper is as follows: The next section describes the reform of higher secondary schooling in Germany used to identify the effects of shortening the length of schooling on student performance. The estimation approach and a discussion of the internal and external validation of the estimates is provided in section 3. The data used for the empirical analysis are introduced in section 4 presenting also selected sample descriptives. The empirical estimates of the reform effect on student's achievement are provided in section 5. Section 6

[4]Krashinsky (2009) examined the Ontario reform effect on wages of lower ability students. He found that the wages of the graduates who had the shorter high school instruction time were significantly lower one year after graduation. This effect diminished after another year of labor market experience.

provides a discussion of the implications from these results. The final section concludes.

2 The reform of secondary schooling in Germany

In international comparison, German university graduates are elder on average when they enter the labor market. Apart from the longer duration time of study at university, one fundamental reason is the long secondary schooling time (OECD, 2005). So in the context of the Bologna Process the pressure in Germany to reform the educational system, in particular the academic track, has increased.[5] Education as well as funding of public schools is a matter of the federal states (*Bundeslaender*) in Germany and hence, the educational system differs across the federal states. In the majority of federal states, students are enrolled in school at the age of 6 and attend primary school for four years. Afterwards, they are tracked according to their cognitive skills into three types of secondary schools (basic, intermediate and academic track). The basic track of secondary school (*Hauptschule*) lasts up to grade 9 which is the mandatory schooling duration in Germany. The intermediate track (*Realschule*) ends after grade 10. Afterwards graduates from both of these tracks usually start vocational training in the German apprenticeship system. Until recently, in all federal states (except Saxony and Thuringia) the most academic track of secondary schooling (*Gymnasium*) has led to graduation after 13 years with a school-leaving certificate (*Abitur*) providing the necessary precondition for attending university. Besides the tracked tripartite schooling system, a number of federal states provide an additional type of comprehensive schooling, the so-called integrated comprehensive school (*Integrierte Gesamtschule*). In this type of school, students can graduate after 9, 10 or 13 years, and obtain the corresponding degrees from the three types of secondary schooling. A remarkable difference of this type of school is that students are not tracked ex ante.

The necessity to reform the academic track was also enforced by the East German federal states. As a result of the German unification in 1990, the West German schooling system was adopted by most of the East German federal states in the early 1990s and several further reforms were implemented. The education system of the former German Democratic Republic (GDR) was compulsory without tracking students by abilities until grade 10. Graduates of grade 10 with high cognitive skills and socialist world views were eligible to enter a two-year *Gymnasium*. The university entrance qualification was therefore already obtained after grade 12. After the German unification, the two East German federal states Saxony and Thuringia introduced early

[5] The Bologna Process denotes the process of creating a European Higher Education Area until 2010 including the adoption of the three cycle system (bachelor, master, doctorate) and the introduction of a credit transfer system for higher education courses.

student tracking but retained the 12 grade *Gymnasium* unchanged. Since both federal states performed very successfully in international comparability studies like PISA (see, e.g., PISA-Konsortium Deutschland, 2008), these results supported the decision to abolish the 13th grade of *Gymnasium* in most German states.

The reform of shortening the length of secondary schooling in Germany by one year has been introduced at first in the German state of Saxony-Anhalt. The reform was announced in 2003 and started some months later at the beginning of the school year 2003/2004 in grade 9. Students, who learned in grade 9 at that time, were the first for which the school leaving certificate (*Abitur*) became available after grade 12 at the end of school year 2006/2007. So, in spring of 2007 in Saxony-Anhalt grade 12 (G12) and grade 13 (G13) graduated simultaneously.[6]

How was the reform implemented? For G12 students the 13th grade was abolished whereas the requirements for the school leaving certificate (*Abitur*) remained unchanged. Because in Germany only the two final years of secondary schooling count for the *Abitur*, the curriculum of the former 12th and 13th grades is now teached in the 11th and 12th grade. During the transition period the schools had the opportunity to form new classes and teach students from the double cohort together during the two final years. The majority of schools did not establish new classes but provided common courses for students from both cohorts in some subjects. For G12 students the curriculum of the former 11th grade, the preliminary grade, was distributed to the lower grades. The whole curriculum was brought forward in literature and foreign languages. Only minor reductions were undertaken in math and chemistry whereas in some other subjects like biology and history parts of the curriculum were transformed in additional elective parts. All in all, for G12 students the instructional time during the whole academic track decreased. The drop was mitigated by some extra lessons. Three lessons in grade 9 and three lessons in grade 10, were introduced and the schools were able to decide for which subjects these additional lessons should be provided.

This education reform of compressing the secondary schooling period without changing the requirements and comparatively minor changes in the curricula might have affected students in a variety of ways. In this paper, we concentrate on the effects on human capital approximated by the final examination test scores in three different subjects: math, literature, and foreign language (English). Differences could be expected since abolishing one grade resulting in less time for instruction and homework increased the learning intensity. Furthermore, time spent

[6] By now, all federal states except one have decided to cut the last year of secondary schooling in the academic track. The Standing Conference of the Ministers of Education and Cultural Affairs of the Federal States (*Kultusministerkonferenz*, KMK) accentuated the importance "The responsible handling with lifetime and education time of young people is of central concern"(press release, March 2008).

on extracurricular and leisure activities was reduced, which could result in lower non-cognitive skills that have been found to be malleable at least until age 20 (Dahl, 2004) or later (Caspi and Roberts, 1999) and comprise important factors like self-discipline.

3 Estimation approach

3.1 The natural experiment

The policy reform in Saxony-Anhalt provides a natural experiment that allows to identify the effect of shortening the duration of schooling while holding the curriculum almost constant.[7] The main advantage is that given that assignment to the treatment group, i.e. students graduating after 12 years, and to the comparison group, i.e. students graduating after 13 years, could be assumed to be random conditional on the policy reform (as described in section 2) due to the fact, that the time of announcement and the instantaneous implementation, since students have been attending the academic track for several years already, no immediate reaction on tracking resulted. In line with the last point, the education minister of Saxony-Anhalt Jan-Hendrik Olbertz (2009) pointed out that "The introduction of the *Abitur* after 12 grades run surprisingly smoothly." One further advantage consist of centralized final written exams, i.e. the tests, horizons of expections and scoring scales are provided by the State Ministry of Education and are standardized for all students in state. Also, final exams in math and literature are mandatory in Saxony-Anhalt. In addition, students have to choose a foreign language; therefore, examination in English is voluntary only as it could be substituted by for example French. Nevertheless, the vast majority of students take the exam in English. So differences in outcomes could therefore be attributed to be the causal effect of the reform, i.e. the shortening of the schooling duration.

Let D_i denote a binary dummy variable taking the value 1 if the student was assigned to the treatment group, and 0 otherwise. In case the assumptions of the natural experiment hold, we can estimate the treatment effect as follows

$$y_i = \alpha + \beta D_i + \gamma_j + \delta' \mathbf{X} + \varepsilon_i, \tag{1}$$

where y_i denotes the outcome of interest, i.e. the achievement scores in math, literature and foreign language (English) obtained in the final written exams. The achievement scores range in the interval $[0, 15]$, where 0 denotes failure and 15 denotes excellent achievement. α is the constant, γ_j are school-fixed effects, and $\mathbf{X}$ is a matrix of further explanatory variables with

[7] Pischke (2007) analyzes the effects of the German short-school years during the 1960s in a similar way.

δ as the corresponding coefficient vector. The parameter of interest is the average treatment effect (ATE) β which denotes the marginal change in y that is due to the reform.

Schools may influence the average achievement scores of the students in the centralized exam independently of the duration of schooling for a number of reasons, e.g., the quality of teachers, the infrastructure, or the composition of students. Therefore, to allow for an unbiased estimate of the reform effect including the school-fixed effects are crucial as they capture all between-school variation. Moreover, we include a number of further control variables, which were not affected by the reform, to increase the efficiency of the estimates. Examples are gender, age at school enrollment, background characteristics of the family of the student from the time before the reform took place, and the marks in the subject from grade 7. The latter could be interpreted as a proxy for student's ability in the particular subject.[8]

Nevertheless, although the policy reform provides a reasonable natural experiment to analyze the question of interest and to identify the corresponding parameters, there are a number of threats to the validity of the natural experiment that should be regarded carefully. Particularly, if the selection into treatment group and comparison group is not exogenous and members of both groups differ in a systematic way, the outcomes will be affected.

In the case under consideration, this is not very likely since the reform was introduced when students were already in secondary school for a couple of years. If there would be any selection between groups this should be observable when comparing the pre-treatment characteristics of the sample (see below). Reasons for this selection could be the anticipation of the reform, and persons decide to move to a different state within Germany in a very short time span. However, costs of movement would be extraordinarily high particularly for the parents of students as they will have to find a new job and may have to sell their housing property. Therefore, this type of anticipation effect seems to be not very likely. Alternatively, children could be required to commute to a school in a neighboring state. Since the closest border of the next state is far away (about 50 km), commuting of children seems to be less an option as well.

Finally, assuming that the internal validity of our estimates is ensured, there may be concerns with respect to the external validity of the natural experiment, i.e. translating the specific findings to hold in general (see Meyer (1995). A serious threat would be the existence of a general time trend in the accumulation of human capital. Then our estimations will not capture the causal effect of shortening secondary school duration. Whereas a time trend in

[8]Since the students in our sample come from a distinct number of classes within schools, the correlation of in-class outcomes may be interpreted as the treatment effect. For that reason we apply a cluster-robust variance estimator suggested by White (1980).

cognitive achievement is perhaps likely in earlier educational periods of life, it should be not very likely in later periods as the here analyzed periods of final secondary schooling.

3.2 Heterogeneity of effects

The model in equation 1 assumes a homogenous average effect of treatment, i.e. the parameter estimate provides the average change in the outcome of interest that is caused by the reform. Although this is the most important parameter to evaluate the effects, we will also consider heterogeneity in the effects of the reform in the empirical analysis. The first source of heterogeneity are potential gender differences. Besides estimating gender differences in achievement scores that are included in equation 1, a gender specific reaction to the reform could be expected. Since males and females prefer different learning strategies (see, for instance, Green and Oxford, 1995) and there are furthermore differences in biological and mental development (see, for instance, De Bellis et al., 2001), changing the time of education may lead to heterogeneous impacts.[9] To analyze these differences, we will estimate separate models for both genders using the specification from above (and skipping the covariate of gender, respectively).

In addition, we could expect heterogeneity of the effects between schools. Teachers and schools may be differently able to adjust to the new demands of the curriculum and the timing of classes over the day which may lead to school-specific reform effects that differ from the school-fixed effects. To analyze the possible differences between schools, we construct school-specific binary treatment variables D_{s1} to D_{s12} for the 12 schools in our sample. Defined as the interaction between school and the treatment indicator each of the dummy variables takes the value 1 if the student is in the treatment group in the respective school, and 0 if the student is in the comparison grade. The corresponding estimation equation is defined as

$$y_i = \alpha + \sum_{j=1}^{12} \beta_j D_{sj} + \gamma_j + \delta' \mathbf{X} + \varepsilon_i, \tag{2}$$

where the β_j denote the school-specific treatment effects and the remaining variables are defined as before. We estimate the model in equation 2 on the pooled sample and separately by gender.

4 Data

4.1 The students' survey

For the empirical analysis we use novel data collected in a survey of students who graduated from school in 2007. The survey used pen-and-paper questionnaires covering more than 100 questions

[9]See also OECD (2009) for comparison analysis of student's performance of 15-year-old boys and girls.

on different aspects of student's background, experiences, and a number of different outcomes. The questionnaires were sent to all graduates from 12 schools in the city of Magdeburg[10] and the community of Halberstadt[11] in February and March of 2009. Responses were asked to be sent back until the end of April 2009.

Our sample contains all 10 schools from Magdeburg (8 *Gymnasium*, 2 *Integrierte Gesamtschulen*) with double cohorts and the two schools (*Gymnasium*) from Halberstadt. Overall, 1,628 students graduated from these schools in 2007. For 164 graduates, no names or addresses were available. All in all, 1,464 questionnaires were sent to graduates from these schools and 805 responses were received. The response rate of 55% is remarkable.

This provides a reliable number of observations to analyze the effect and effect heterogeneity of the reform in terms of different outcomes. A more detailed description of the available characteristics is provided in appendix A. We impose one limitation on the analysis sample for the sake of homogeneity: only non-grade repeated students who continuously stayed in Germany during their schooling are considered, i.e. students who went abroad for a student exchange and students who repeated a grade are excluded from the analysis.[12] The final sample contains 694 observations.[13]

4.2 Sample description

We will now present some selected descriptive statistics of the sample. Table 1 provides the shares of students from the 12 schools distinguished by grade and gender and corresponding t-tests. The first thing that becomes obvious is that the share of male students is clearly smaller than that of female students. This finding reflects not an imbalance in response rates to the survey but reflects a general tendency in higher secondary school attendance in Germany that has started more than a decade ago (destatis, 2009). Furthermore, schools differ in size. Differences in size of school reflect aspects like location and reputation but also specialization. Specialized schools focus on natural sciences, sports, or have a particular religious emphasis. Overall, the shares of students by grades do not differ significantly; nevertheless, two slight

[10]Magdeburg is located in the center of Saxony-Anhalt and is the federal state capital with about 230,000 inhabitants. For post-secondary education, it hosts a university, a university of applied sciences, and a musical college.

[11]The community of Halberstadt is a rural and mountainous area with about 75,000 inhabitants, living dispersed in villages and smaller cities. Secondary schools are located in towns only and there is only a university of applied sciences in the community.

[12]We excluded these students due to the fact that for the G12 cohort the students who went abroad as well as the students who repeated a grade did the final exams one year later.

[13]The numbers of observations in the estimations presented below may differ due to item non-response in some variables.

differences are observed in the male sample for school 3 and school 6. Both potential imbalances between grades and specialization of schools do not affect the estimates of the reform effect as we take account of school fixed effects, on the one hand, and allow for school-specific reform effects in the extended model considering effect heterogeneity, on the other hand.

Include Table 1 about here

The estimates of the reform effect will be biased in case students in the treatment and the control group differ systematically and these differences affect the outcomes, but are not regarded in the estimation. However, as the reform provides a natural experiment, no systematic differences in students' characteristics that are exogenous to the reform should be expected. Table 2 provides means of selected variables characterizing the students. In the upper panel, the outcomes of interest, i.e. the achievement scores at graduation in math, literature and English that are supposed to be affected by the reform are provided. Below that, the marks from grade 7 in the same subjects are reported. Differences in means have been tested by t-tests and the corresponding p-values are reported in the table. As implied by the assumption and independently of the subject considered, marks obtained at grade 7 do not differ between treatment and control group within gender. However, when regarding the scores at graduation statistically significant differences in math for both genders and additionally in English for female students are revealed. Although the size of these differences could not be completely attributed to the reform since further factors (e.g., school effects) may affect scores, this tends to indicate some impact of the reform on achievement scores.

Include Table 2 about here

With regard to the other variables selected to characterize the students' situation, no clear differences could be obtained with respect to the age at initial school enrollment, the number of siblings, and school choice. The number of books the students possess looks very similar for both grades, too.

Previous empirical literature indicates, that the environment the student lives in matters for educational success (see, e.g., Fuchs and Woessmann, 2007; Todd and Wolpin, 2007). Table 3 provides means of variables characterizing the family background of the graduates. There is a choice of characteristics of mothers, fathers, and items available at home. A common thing to almost all of the variables chosen is that no statistically significant differences between students across grades and the same gender are observed. Hence, the background situation does not differ. Therefore, there is no reason to expect any systematic differences in the outcomes across grades due to the home environment.

With respect to the occupational training of the parents, more than half of the fathers has finished an apprenticeship training after school. A very small percentage of the parents possesses no occupational training, and the shares of university/university of applied sciences graduates and parents with PhD are clearly above the average of society.[14] In addition, parents seem to be quite active in different directions; about 70 percent of the individuals participate in societal events and about half of the parents do regular sports' exercises. Since we consider graduates from higher secondary schooling in the analysis only, these findings are not surprising. However, political and religious engagement is reported for a small fraction of parents only.[15]

Include Table 3 about here

Taking a look at the items available at home shows that the households are well-equipped on average. There are no significant differences in any of the items between grades. Hence, this further supports the picture of the comparability of the treatment group and the comparison group.

5 Empirical results

5.1 Impacts on achievement scores

We have estimated the effects on achievement scores in math, literature, and foreign language (English). The three subjects could be interpreted as proxies for different abilities of the student. Math requires logical thinking and the capacity of abstraction. Literature promotes the linguistic instinct of the students and is therefore useful for the capability of communication of the individuals. English approximates the ability of German students to acquire foreign language skills. Although we can think of a few more subjects taught to develop certain skills of students, the three subjects considered here are able to capture precisely the main requirements for university education, and are therefore of crucial interest when evaluating the effects of a reduction of secondary schooling.

[14] According to destatis (2009), on average, 20.1% of the population possesses a university degree and only 0.5% holds a PhD.

[15] A reason for the limited religious engagement is by far the low rate of people in religious denominations in East Germany.

Math

First we will present the impact of the reform on achievement scores in math. Table 4 provides the corresponding estimates of the homogenous treatment effect for the pooled sample and the gender samples. For each of the samples, we provide two models that differ in the number of additional covariates considered to improve the precision of the estimates. These covariates cover background characteristics of the students that been proved relevant in the empirical literature (see, e.g., Fuchs and Woessmann, 2007). The estimates reveal a significantly negative effect of about 0.7 in the pooled sample, i.e. due to the reform students graduating after 12 years have on average 0.7 points less in the achievement score for math compared to students graduating after 13 years. For the interpretation of this effect, one has to bear in mind that achievement scores range from 0 to 15 points and 0.7 points mean that students graduating after 12 years are loosely speaking almost 5 percent worse off than those receiving a further year of education. Although this is not much, the difference is significant. As changes during the reform occurred purely to the time of teaching, but not to the curriculum, the negative effect indicates a trade-off between the amount and the timing of learning. In the case of math, there tend to be limitations of human ability to accumulate a certain amount of knowledge in a varying time interval.

Include Table 4 about here

Turning to the results for males and females, a gender gap in the effects becomes obvious. Although students graduating after 12 years experience a decrease in achievement scores in math independently of gender, the negative effect tends to be almost twice as large for males compared to females. Males are on average slightly about 1 point worse off, whereas females are worse off about 0.5 points. Nevertheless, the estimated effects for both genders are significant. These findings clarify the different effects of shortening schooling duration for both genders beyond the gender-specific constant regarded in the pooled estimation. Males and females react differently to the applied changes in learning intensity.

Regarding the estimates of the further control variables, the first thing to note is the between-school variation in achievement scores. This variation reflects differences in teaching quality, differences in infrastructure, class sizes, and differences in peer groups. The age at initial enrollment has a negative effect on achievement scores, i.e. students that have started schooling at a younger age have a slightly better achievement score in math at graduation. However, as shown by Muehlenweg and Puhani (2009) this effect should not be interpreted as causal for the age of enrollment, but is more likely to provide a proxy for unobserved ability in the sense that persons with lower unobserved abilities enroll later on average.

As expected, the effect of the mark in math obtained in grade 7 is strong and highly significant. The reversion of the sign is explained by the reversion of the scale of achievement scores after grade 10. Whereas from grade 2 to grade 10 students receive marks defined between 1 (excellent) and 6 (failure), the scale after grade 10 is defined reversely from 15 (excellent) down to 0 (failure). In addition, with regard to the further variables considered to capture details of the background of the student all coefficients show the expected signs.

Literature

In contrast to the effect on the achievement score in math, there is no effect due to shortening the schooling duration on scores in literature. See Table 5 for the corresponding estimation results. Although the point estimates have a negative sign in the pooled sample, this difference is not pronounced enough to differ significantly from zero. The estimates of the further covariates reveal again a large between school variation in achievement scores, but now the effects are better on average among schools compared to the reference school (school 11). Again, persons with good marks in grade 7 are also more likely to make a better exam at the time of graduation.

Include Table 5 about here

In line with the findings for the pooled sample, the estimates do not establish any significant gender differences. Nevertheless, the negative sign of the treatment effect in the pooled sample is determined by the effect for females. For males, even the point estimate is almost zero.

English

Since English is not mandatory, there may be grade-specific self-selection affecting the outcomes. We have checked for differences in the pre-treatment characteristics of students in the treatment and control group, and a slight difference in the marks in grade 7 could be established. As we take account of these differences in the estimation, the parameter estimate for the effect of the reform should not be biased.

The corresponding results of the estimations for the third outcome in analysis, achievement score in English, are provided in Table 6. Similar to the results for literature, the parameter estimate of the treatment effect is negative but not statistically significant in the pooled sample.

Hence, although students experienced the same amount of education in a different time span this has no effect on the written examinations. Nevertheless, when considering gender differences, heterogeneity in the estimates could be revealed. For females, reducing the schooling duration by one year leads to a significant decrease in achievement scores in English of about 0.5 to 0.6 points. In contrast, males are not affected and the point estimates even show a positive (but insignificant) effect of earlier graduation. For this reason, the picture for ability of foreign languages is a bit ambiguous. Compared to math, however, there is a different gender pattern in the effects: Whereas shortening the duration was more harmful for males in terms of mathematical abilities, the picture is reverted for language abilities.

Include Table 6 about here

Regarding the other variables considered in the estimation, marks in English in grade 7 exhibit a larger effect compared to those of math or literature. This indicates that some people are more able to learn foreign languages than others, or that having learned the basics early in the education period is essential for a good final exam. The parameter estimates for the remaining covariates are similar to the other models.

5.2 Effect heterogeneity across schools

We will now turn to the results of effect heterogeneity across schools. This type of effect heterogeneity is likely since schools differ not only with respect to teachers' expertise, peer group, or geographic location, but also differences in the implementation of the reform may have occurred. Schools have managed the adjustment of the schooling years and the double cohort graduating in 2007 differently. To prove the relevance of heterogeneity in the effects across schools we have estimated the effect-heterogeneity models for the pooled sample and the gender-separated samples. The corresponding results are given in Tables 7, 8, and 9.

Include Table 7 about here

Starting with the results for math (Table 7), the estimates establish some heterogeneity in the effects between groups. Almost all of the point estimates are negative in the pooled sample, however, depending on the model specification only between three (spec. 1) and six (spec. 2) parameter estimates are significantly different from zero. Within the significant effects, the estimates are clearly larger than those from Table 4 and point to negative effects of even more than 2 points due to the reform.

The findings derived for both genders separately are in line with the picture from the pooled sample. Independently of gender, a large heterogeneity in the effects of reducing the schooling duration on achievement scores becomes obvious. Nevertheless, males tend to experience more negative effects compared to females when regarding heterogeneity across schools, as well. In some of the schools, students graduating after 12 years obtained even about 3 point lower achievement scores due to the reform.

Include Table 8 about here

In contrast to math, no significant effect of shortening secondary schooling duration on achievement scores in literature was found in Table 5. Possibly, there may be literally no difference between students graduating after 12 or 13 years. However, another reason could be that effects are heterogenous and are averaged out when ignoring school-specific effect heterogeneity. When looking at the corresponding estimation results (Table 9), either the one or the other reason is possible. Although the results establish some heterogeneity in effects across schools, only a few estimates are clearly significantly different from zero in the pooled sample and in the two gender-specific samples. Moreover, the signs of the parameter estimates vary across schools, and in some schools students graduating after 12 years are better off compared to those graduating after 13 years, whereas in other places the picture is reversed.

Include Table 9 about here

Finally, Table 9 provides the estimates of between-school heterogeneity in the effects on achievement scores in English. In the pooled sample as well as in the females' sample, the picture reveals heterogeneity in the effects between schools similar to that for the other outcomes. In contrast to math, the effects are a bit less pronounced which is expected from the estimates from Table 6 already. Compared to literature, on the other hand, the overall tendency of the results is more negative indicating that graduates with 12 years of schooling are slightly worse off. However, this finding is not supported by the results for males: Here, despite the heterogeneity in the effects significant estimates indicate benefits for the treatment group from the reform. Nevertheless, as only two (spec. 1) or three (spec. 2) of the point estimates are statistically significant this finding is not at all that conclusive for the improved achievement scores in English.

5.3 Robustness checks

In addition to the estimation results presented so far, we have estimated a number of specifications to check the robustness of the empirical findings. For the sake of brevity, we discuss these checks for the models of the pooled sample analyzing the effects on achievement scores in math and literature.

The figures in Tables 10 and 11 provide parameter estimates of four models. Models (1) and (2) provide separate regressions for the two regions sampled, i.e. model (1) refers to the city of Magdeburg and model (2) to the community of Halberstadt. The results indicate that the reform effects for achievement scores in math (Table 10) differ slightly between regions, and the negative effect for graduates after 12 years is a bit more pronounced in the city of Magdeburg. Nevertheless, despite this small difference the parameter estimates are qualitatively very similar indicating no strong regional variation in the treatment effects (not captured by school-fixed effects already). This finding is supported by the results for the effects on achievement scores in literature (Table 11). The separate models for the two regions indicate a negative point estimate that is not statistically significant. Similar to the result for math, the level of the point estimate is a bit higher for the city of Magdeburg than for the community of Halberstadt.

Include Tables 10 and 11 about here

Models (3) and (4) provide additional sample selections. As noted above, besides *Gymnasium* the so-called integrated comprehensive school allows higher secondary school graduation (*Abitur*) without tracking of students. We have skipped the latter type of schools from the estimation sample in model (3)and consider only *Gymnasium* schools located in the city of Magdeburg. Regarding the estimate of the effect on achievement scores in math, Table 10 shows an even slightly higher effect compared to the overall effect and that for all schools in the city of Magdeburg. This points out that the implementation of the reform in this type of schools exhibits a larger trade-off between duration of schooling and achievement scores at graduation compared to integrated comprehensive schools, where the treatment effect has to be smaller with respect to the overall effect. Although statistically not significant, the point estimate of model (3) for literature (Table 11) is more negative than in the other models supporting this finding.

Finally, model (4) is estimated on large schools only, i.e. schools with more than 80 students graduating in 2007. The corresponding estimates show that the treatment effect on achievement scores in math (Table 10) is further pronounced compared to the other models. This might be an indicator for larger schools being less effective than smaller schools in providing the teaching quality for the graduates after 12 years. Nevertheless, this argumentation holds only for math. Considering the estimate for literature (Table 11), the point estimate remains statistically insignificant but the sign is reverted to positive.

6 Discussion

Shortening the secondary schooling time with an unchanged curriculum might affect students' behavior in a number of ways. Instructional time is strongly reduced in almost all subjects and only a small number of additional hours of instruction is given in core subjects. Thus, the total time for instruction, homework, learning and repeating the relevant matter has declined, while the requirements per year have increased. As shown by the empirical estimates above, this acceleration in pace affects student performance in different ways. In addition, some students may not have been able to cope with the increasing requirements per year. Hence, leaving school with a lower grade of completion, changing to special schools where graduation is still possible after 13 years, or repeating a grade might be likely. Although these patterns have existed before the reform already, there may also be effects of the reform on these responses.

To analyze these reform effects, the numbers of all leaving, repeating, and changing students for all considered schools are required. Unfortunately, the access to these numbers is restricted. We therefore have to rely on data providing the fractions of students who complete the highest academic track after regular schooling time in the period of study.[16] Despite the tracked system, mandatory schooling ends after 9 years. Hence, students can drop off afterwards to start a vocational training for example. The drop-off behavior could be influenced by the reform, too. Therefore, we take the number of grade 9 students attending the highest academic track (*Gymnasium*) and compare the fractions, which complete final secondary school examinations in regular time. In the 13-grade cohort 78.3 percent of the students graduated and in the 12-grade cohort 69.7 percent. A fraction of this difference rest on the fact that the rules for spending a year abroad are different depending on cohorts. Students of G12 cohort who studied a year abroad had to graduated one year later in 2008 whereas students of the G13 cohort were allowed to return to their own cohort. The remaining difference incorporates students who were not able to cope with the increased learning intensity.

If especially low ability students disappeared from the grade 12 cohort, the average ability of grade 12 students would probably be higher than the average ability of grade 13 students at the time of final examinations. Our estimators of the effects on student performance in the analyzed subjects would have been biased and the results should have been interpreted as the lower bound of the reform effects. However, nothing indicates biased estimators. For the results presented above the difference is relevant if drop-off rates were non-random across grades, since the data used in the empirical analysis comprise retrospective information surveyed from graduates. As

[16] These data have been provided uniquely for the purpose at hand by the Statistical Office of Saxony-Anhalt.

shown above, comparing the pre-reform characteristics of the students, no observable differences with regard to ability or background variables could be established, and there is no indication that estimates are biased due to possible self-selection.

The empirical results suggest substantial heterogeneous effects of shortening secondary schooling by one year on human capital accumulation. The impacts are significant negative on student performance in math for both genders and in foreign language for females, however, insignificant in literature. One possible explanation for these findings can be different requirements in higher grades for the subjects considered here. Whereas the curriculum in math requires to open up new mathematical fields (e.g., stochastic) accompanied by learning new methods and understanding the underlying concepts, the curricula in literature and foreign language focus on the refinement of familiar concepts and on the appliance of these concepts on a broader matter. A similar argument is made by Eren and Henderson (2009) regarding effects of additional homework on test scores. They find evidence of positive and significant effects of homework on math test scores (see also Aksoy and Link, 2000; Eren and Henderson, 2007), but little or no impact on test scores in other subjects like literature. Despite there is no reform effect in literature, this result does not indicate whether the level of education is satisfying. All in all, however, the marginal contribution of the 13th grade to native language skills is negligible.

On the contrary, the marginal contribution to the mathematical skills is significant and may have serious consequences on labor demand and labor supply. To fill the shortage of engineers and graduates from natural science society needs people with excellent mathematical skills. The poor impacts of the reform indicate that, on the one hand, even more responsibility for education will be transferred to universities. But in light of the limitations of instructional time in universities a change of the curriculum will become necessary. The effects of this trade-off could not be evaluated yet. On the other hand, because graduates from grade 12 perform comparatively worse in math, probably less of them may enroll in engineering or natural sciences and the shortage will tighten. This requires a change in the instructional time at schools by transferring the latter from subjects with low or zero marginal contribution to subjects with higher contribution to skills.

7 Conclusion

Despite the important role of time as the most basic input in human capital accumulation, understanding the process of how the instruction time at school affects human capital is still in its infancy. Human capital gain per instruction hour respectively per grade depends on a number of factors such as students' ability and effort, teachers' expertise and educational skills, peers, school resources, institutional aspects as well as their interactions. Among them, there is little evidence on the role of curriculum as an important institutional factor in the process of human capital accumulation. The curriculum affects human capital accumulation by its impact on the learning intensity. Leaving the curriculum constant, less time at school increases learning intensity, whereas the effect of less schooling time combined with a restricted curriculum is ambiguous. Hence, it would be useful to focus on the role of learning intensity on student performance. The current lack of knowledge has been caused by the difficulty of collecting appropriate data.

This paper has studied a very rare policy reform of shortening time at school holding the curriculum almost constant, which increased learning intensities. Thus, the empirical results close the gap. There are some direct impacts in terms of human capital accumulation, but the magnitude of these impacts varies across subjects. Our findings suggest inefficient learning intensity in the acquisition of linguistic skills. However, increased learning intensity in math cannot compensate for shortening schooling time. These results suggest that a human capital model in which each hour of instruction time or each grade increases human capital represented by test scores by a constant amount is inappropriate to explain the knowledge gains of additional schooling (c.f. Pischke (2007), p. 1240).[17]

Moreover, some students are not able to cope with the increased learning intensity. Perhaps, lowering the learning intensity in such demanding subjects like math by additional instructional time at the expense of subjects with low learning intensity is a reasonable recommendation. In addition, a revision of the curricula could be an appropriate response.

More research is required to study the role of schooling time on the skill formation process. Our results suggest that institutional features, such as learning intensity, matter. Policy makers should turn their attention from raising the quantity of education to raising the quality. Valuable time at school is wasted. Consequently, educational expenditures are also wasted, which means it could have yielded much higher returns from human capital while invested in earlier periods of life (see Cunha, Heckman, Lochner, and Masterov, 2006).

[17] In contrast to rather scarce reforms of shortening schooling, opposed reforms of increasing schooling by compulsory school laws or minimum school leaving laws were more common and their impacts have been widely investigated. Among the extensive literature, there are some studies which find zero returns of additional schooling (Oosterbeek and Webbink, 2007; Pischke and von Wachter, 2008; Grenet, 2009). These findings suggest that inefficiencies in time using for education also exist at other levels of education.

References

Aksoy, T., and C. Link (2000): "A Panel Analysis of Student Math Achievement in the US in the 1990s: Does Increasing the Amount of Time in Learning Activities Affect Math Achievement?," *Economics of Education Review*, 19, 261–277.

Caspi, A., and B. W. Roberts (1999): "Personality Continuity and Change across the Life Course," in *Handbook of Personality: Theory and Research*, ed. by L. A. Pervin, and O. P. John, pp. 300–327. Guilford Press, New York.

Creemers, B., and L. Kyriakides (2006): "Critical Analysis of the Current Approaches to Modelling Educational Effectiveness: The Importance of Establishment a Dynamic Model," *School Effectiveness and School Improvement*, 17, 347–366.

Cunha, F., J. J. Heckman, L. Lochner, and D. Masterov (2006): "Interpreting the Evidence on Life Cycle Skill Formation," in *Handbook of the Economics of Education*, ed. by A. Hanushek, and F. Welch, chap. 12, pp. 697–812. Elsevier B.V.

Dahl, R. E. (2004): "Adolescent Brain Development: A Period of Vulnerabilities and Opportunities.," in *Annals of the New York Academy of Sciences*. New York Academy of Sciences.

DeBellis, M., M. Keshavan, S. Beers, J. Hall, K. Frustaci, A. Masalehdan, J. Noll, and A. Boring (2001): "Sex Differences in Brain Maturation during Childhood and Adolescence," *Celebral Cortex*, 11(6), 552–557.

destatis (2009): *Allgemeine Schulen-Schuljahr 2008/09, Fachserie 11 Reihe 1-2008/09*. Statistisches Bundesamt Deutschland, Wiesbaden.

Eren, O., and D. Henderson (2007): "The Impact of Homework on Student Achievement," *Econometrics Journal*, 11, 326–348.

——— (2009): "Are we Wasting our Children's Time by Giving them More Homework?," Working paper, University of Nevada.

Fuchs, T., and L. Woessmann (2007): "What Accounts for International Differences in Student Performance? A Re-Examination Using PISA Data," *Empirical Economics*, 32(2-3), 433–464.

Green, J., and R. Oxford (1995): "A Closer Look at Learning Strategies, L2 Proficiency, and Gender," *TESOL Quarterly*, 29(2), 261–297.

GRENET, J. (2009): "Is it Enough to Increase Compulsory Education to Raise Earnings? Evidence from French and British Compulsory Schooling Laws," Discussion paper, LSE.

HANUSHEK, E. A. (2005): "The Economics of School Quality," *German Economic Review*, 6, 269–286.

KRASHINSKY, H. (2006): "How Would One Extra Year of High School Affect Academic Performance in University? Evidence from an Unique Policy Change," Discussion paper, University of Toronto.

——— (2009): "How Would One Extra Year of High School Affect Wages? Evidence from an Unique Policy Change," Working paper no. 20, Canadian Labour Market and Skills Researcher Network.

LEE, J.-W., AND R. BARRO (2001): "School Quality in a Cross-Section of Countries," *Economica*, 68, 465–488.

MARCOTTE, D. (2007): "Schooling and Test Scores: A Mother-Natural Experiment," *Economics of Education Review*, 95(26), 629–640.

MEYER, B. (1995): "Natural and Quasi-Experiments in Economics," *Journal of Business & Economic Statistics*, 13(2), 151–161.

MORIN, L. P. (2009): "Estimating the Benefit of High School for College-Bound Students," Discussion paper, University of Ottawa.

MUEHLENWEG, A. M., AND P. A. PUHANI (2009): "The Evolution of the School Entry Age Effect in a School Tracking System," *Journal of Human Resources*, forthcoming.

OECD (2002): *Reading for Change: Perfomance and Engagement Across Countries*. OECD, Paris.

——— (2004): *PISA: Learning for Tomorrow's World*. OECD, Paris.

——— (2005): *Education at a Glance*. OECD, Paris.

——— (2007): *PISA 2006: Science Competencies for Tomorrow's World*. OECD, Paris.

——— (2009): *Equally Prepared for Life? How 15-Year-Old Boys and Girls Perform in School*. OECD, Paris.

OOSTERBEEK, H., AND D. WEBBINK (2007): "Wage Effects of an Extra Year of Basic Vocational Education," *Economics of Education Review*, 26, 408–419.

PISA-Konsortium Deutschland (2008): *PISA '06: PISA 2006 in Deutschland - Die Kompetenzen der Jugendlichen im dritten Laendervergleich.* Waxmann, Muenster, New York, Muenchen, Berlin.

Pischke, J., and T. von Wachter (2008): "Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation," *The Review of Economics and Statistics*, 90(3), 592–598.

Pischke, J. S. (2007): "The Impact of Length of the School Year on Student Performance and Earnings: Evidence From the German Short School Years," *Economic Journal*, 117(October), 1216–1242.

Skirbekk, V. (2006): "Does School Duration Affect Student Performance? Findings from Canton-Based Variation in Swiss Educational Length," *Schweizerische Zeitschrift fuer Volkswirtschaft und Statistik*, 142(1), 115–145.

Teddlie, C., and D. Reynolds (2000): *The International Handbook of School Effectiveness Research.* Falmer Press, London.

Todd, P., and K. Wolpin (2007): "The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps," *Journal of Human Capital*, 1, 91–136.

White, H. (1980): "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test of Heteroskedasticity," *Econometrica*, 48, 817–838.

Woessmann, L. (2003): "Schooling Resources, Educational Institutions and Student Performance: The International Evidence," *Oxford Bulletin of Economics and Statistics*, 65(2), 117–170.

A Appendix

To provide an overview on the available information in the survey, the 101 questions could be categorized into the following ten categories:

1. *Personal information:* The first set of questions covers personal characteristics of the student, like date of birth, gender, place of residence, place of residence during schooling education, movements, nationality, number of close friends, etc.

2. *Family background:* The second category provides information on the family background of the graduate. These variables comprise quite detailed information about the parents and the siblings of the students. Characteristics of the parents are collected separately for father and mother and cover the age, the time of cohabitation with each parent, divorce, dead, changing partners of the parent, education, occupational degree, unemployment, and several types of engagements, e.g., cultural, political, religious, or sports. Besides that, the family background includes details about the number of books of the parents and further items available in the parental household. For these items the use by the graduate is also recorded. Examples are access to the internet, dictionaries, newspapers, or reference books. With regard to the siblings of the graduate the survey provides information on the row of siblings, and for each sibling further information on gender, age, education, and some others.

3. *Schooling, general information:* Details of schooling of the graduate could be categorized into two classes. The first covers general information on schooling, i.e. the time of pre-school, primary and early secondary schooling, changes of residence during that time, grade repetition, and so on.

4. *Schooling, detailed information:* The second category of schooling variables contains details of the curriculum of the student. An example are the sequence of enrollment, the grades when courses where started, the duration, and the number of foreign languages learned at school. Moreover, information on natural sciences (biology, chemistry, physics) is provided in this category and there are questions covering details of additional in-school education the students attained. Furthermore, a number of questions is devoted to assess the stress and burden of schooling of the students, an assessment of the skills learned at school and the valuation of teaching these skills at school.

5. *Education outside school:* Classes at school provide a relevant part of individual's education but many students participate in a number of educational activities outside school.

These activities comprise, e.g., musical classes, sports, journalistic activities (students' newspapers), political engagement etc. Information on different activities and the number of years of these activities are given by this category.

6. *Last year of school and graduation:* Questions describing the last year of school and graduation cover the class size, the types of the main courses (basic courses and intensified courses taught with more hours per week), the achievement score in each of these courses, the overall achievement score, activities outside school (working, homework etc.), the state of health during the last year of school, spending of leisure time and leisure activities (dating friends, reading, chatting, etc.), and consumption of alcoholic beverages and smoking behavior.

7. *Support from parents, teachers and other persons:* This category comprises the incidence and amount of support with schooling tasks and homework from close relatives, particularly the parents, teachers, and other persons like friends, siblings and peers.

8. *Education after graduation:* Since students in the survey have graduated in 2007, about 18 months have passed between graduation and the date of interview. The activities that took place during that time are reported in a retrospective monthly calendar covering various states of employment, civil and military service, education, and times spent abroad. In addition, information on the financing of living today, the type of education (apprenticeship, university or university of applied sciences studies), the subject, the aspired degree (e.g., bachelor, master, PhD), and on reasons for the choice of education is provided.

9. *Assessment of school:* In this category, students had to assess the value of school for different skills, like the ability of logical thinking, independence, ability to accept/reflect criticism, cooperation in teams, practical skills, technical skills etc. In addition, several items evaluating the relationship between teachers and students were collected.

10. *Attitudes and non-cognitive skills:* In the final set of questions information on various items were collected in order to identify aspects of personality of the individual. The set of items could be used to derive measures of non-cognitive skills of the students.

Tables

Table 1: Distribution of Survey Respondents by Schools, Gender and Grades

	Men			Women		
	Grade 13	Grade 12	p-value[a]	Grade 13	Grade 12	p-value[a]
School 1	0.035	0.034	0.964	0.016	0.021	0.676
School 2	0.085	0.083	0.943	0.033	0.039	0.719
School 3	0.05	0.103	0.088	0.061	0.056	0.81
School 4	0.057	0.069	0.672	0.081	0.112	0.262
School 5	0.057	0.103	0.147	0.073	0.094	0.402
School 6	0.149	0.062	0.016	0.098	0.099	0.966
School 7	0.099	0.124	0.507	0.11	0.137	0.36
School 8	0.142	0.103	0.324	0.13	0.155	0.445
School 9	0.05	0.034	0.524	0.057	0.026	0.089
School 10	0.057	0.097	0.208	0.069	0.06	0.689
School 11	0.092	0.083	0.778	0.15	0.124	0.411
School 12	0.128	0.103	0.523	0.122	0.077	0.104
N	141	145		246	233	

[a] p-value from t-test on equality of shares.

Table 2: Means of Selected Characteristics by Grade and Gender

	Men			Women		
	Grade 13	Grade 12	p-value[a]	Grade 13	Grade 12	p-value[a]
Scores at Graduation[b]						
Literature	8.007	8.120	0.745	8.806	8.858	0.851
Math	7.721	6.845	0.016	7.519	6.982	0.062
English	7.418	7.819	0.235	8.554	8.315	0.439
Scores in Grade 7[b]						
Math	2.221	2.140	0.383	2.306	2.343	0.56
English	2.514	2.371	0.124	2.290	2.152	0.051
Literature	2.421	2.287	0.091	2.058	2.000	0.309
Grade repeated	0.093	0.021	0.009	0.053	0.052	0.948
Grade skipped	0.000	0.021	0.087	0.004	0.013	0.291
Age at school enrollment	6.234	6.234	0.994	6.191	6.126	0.085
No. of siblings	0.901	1.021	0.259	0.947	0.940	0.924
Choice of School for Reason						
Close distance	0.579	0.662	0.147	0.622	0.652	0.490
Reputation	0.679	0.669	0.863	0.618	0.738	0.005
No. of own books						
0 to 50	0.397	0.421	0.687	0.232	0.253	0.584
51 to 100	0.340	0.372	0.574	0.431	0.335	0.031
101 to 200	0.135	0.124	0.790	0.220	0.266	0.235
201 to 500	0.113	0.062	0.125	0.093	0.133	0.172
More than 500	0.014	0.021	0.676	0.024	0.013	0.355
N	141	145		246	233	

[a] p-value from t-test on equality of means.
[b] Scores until grade 10 range from 1 (excellent) to 6 (failed) and are reverted from scores in grades 12/13 ranging from 0 (failed) to 15 (excellent).

Table 3: Means of Selected Background Characteristics by Grade and Gender

	Men			Women		
	Grade 13	Grade 12	p-value[a]	Grade 13	Grade 12	p-value[a]
Characteristics of father						
Age	49.109	48.243	0.181	49.238	47.743	0.002
Unemployment	0.255	0.228	0.585	0.301	0.232	0.088
Occupational degree						
No occupational training	0.000	0.000		0.004	0.009	0.533
Apprenticeship training	0.546	0.580	0.582	0.602	0.589	0.787
University/University of Applied Sciences	0.392	0.362	0.614	0.339	0.348	0.835
PhD	0.062	0.058	0.902	0.055	0.054	0.943
Activities						
Cultural	0.213	0.254	0.447	0.201	0.229	0.488
Sports	0.541	0.458	0.189	0.413	0.495	0.096
Societal	0.738	0.773	0.518	0.737	0.827	0.026
Politics	0.050	0.032	0.478	0.038	0.047	0.662
Religious	0.083	0.083	0.984	0.028	0.024	0.821
Characteristics of mother						
Age	46.640	46.210	0.347	46.764	46.264	0.243
Unemployment	0.284	0.324	0.459	0.293	0.270	0.589
Occupational degree						
No occupational training	0.007	0.000	0.313	0.016	0.009	0.448
Apprenticeship training	0.489	0.627	0.020	0.533	0.556	0.612
University/University of Applied Sciences	0.460	0.338	0.036	0.426	0.388	0.396
PhD	0.043	0.035	0.732	0.025	0.047	0.181
Activities						
Cultural	0.371	0.374	0.963	0.325	0.336	0.798
Sports	0.449	0.547	0.105	0.483	0.525	0.371
Societal	0.674	0.803	0.015	0.752	0.852	0.007
Politics	0.029	0.008	0.188	0.021	0.033	0.459
Religious	0.073	0.120	0.188	0.041	0.070	0.162
Items at home						
Desk	0.879	0.890	0.788	0.907	0.927	0.418
Place for handicraft	0.199	0.083	0.005	0.337	0.339	0.970
Experiment kit	0.369	0.359	0.859	0.215	0.245	0.449
Cell phone	0.879	0.903	0.515	0.951	0.936	0.460
Computer	0.752	0.703	0.361	0.545	0.472	0.113
Internet access	0.908	0.903	0.900	0.915	0.893	0.417
Classical literature	0.426	0.345	0.162	0.508	0.489	0.681
Poetry	0.149	0.110	0.333	0.195	0.236	0.277
Reference book	0.887	0.869	0.652	0.939	0.936	0.878
Dictionary	0.929	0.924	0.873	0.951	0.953	0.936
Newspaper (regional)	0.660	0.572	0.131	0.565	0.584	0.681
Newspaper (national)	0.163	0.179	0.718	0.122	0.099	0.419
N	141	145		246	233	

[a] p-value from t-test on equality of shares.

Table 4: Achievement Score Math (Regression Estimates)[a]

	Pooled Sample		Female Sample		Male Sample	
	(1)	(2)	(1)	(2)	(1)	(2)
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
D	-0.704***	-0.698***	-0.515**	-0.501**	-0.999***	-1.047***
School fixed effects						
School 1	-2.605***	-2.329***	-2.710***	-2.364***	-2.318***	-2.120**
School 2	-0.471	-0.635	-1.025	-0.986	0.076	-0.394
School 3	-2.041***	-1.566***	-2.077***	-1.588**	-1.918***	-1.111
School 4	-1.462***	-1.277***	-1.702***	-1.472***	-0.930*	-0.416
School 5	-1.085**	-1.032**	-1.498**	-1.286**	-0.380	-0.389
School 6	-0.803**	-1.114***	-0.847	-1.377**	-0.823	-0.915
School 7	-1.462***	-1.367***	-1.951***	-1.749***	-0.584	-0.732
School 8	-1.151**	-1.160**	-1.281**	-1.122**	-0.889	-1.296
School 9	-3.140***	-2.873***	-3.020***	-2.555***	-3.124***	-3.501***
School 10	-0.398	-0.193	-0.663	-0.492	0.042	0.233
School 12	-1.396***	-1.196***	-1.795***	-1.570***	-0.729	-0.542
Sociodemographic variables						
Age (enrollment at school)	-0.643***	-0.674***	-0.526	-0.550	-0.854**	-0.824*
Male	-0.118	-0.260	–	–	–	–
Score in Math (grade 7)	-1.676***	-1.741***	-1.938***	-1.900***	-1.365***	-1.465***
Father unemployed	–	-0.037	–	0.279	–	-0.459
Occupational degree of father (reference: apprenticeship or no occupational training)						
Grad. from Univ./Univ. of Appl. Sciences	–	0.664***	–	0.985***	–	0.067
PhD	–	0.516	–	1.699**	–	-1.624**
Help with homework from father (reference: no support)						
Frequently	–	-0.426	–	-0.113	–	-1.264**
Infrequently	–	-0.560*	–	-0.530	–	-0.554
Rarely	–	0.061	–	0.194	–	-0.092
No. of books of parents (reference: less than 50 books)						
51 to 100	–	0.260	–	0.073	–	0.884
101 to 250	–	0.709*	–	0.463	–	1.555***
251 to 500	–	0.803**	–	0.593	–	1.418**
501 to 2,000	–	0.813**	–	0.603	–	1.502**
More than 2,000	–	1.605***	–	1.283*	–	2.573***
Constant	16.597***	16.204***	16.604***	15.725***	16.872***	16.100***
Statistics						
R^2	0.242	0.281	0.263	0.304	0.238	0.316
N	728	671	452	421	276	250
No. of clusters	93	93	87	87	84	82

[a] All standard errors are clustering-robust based on class as the sampling unit. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Achievement Score Literature (Regression Estimates)[a]

	Pooled Sample		Female Sample		Male Sample	
	(1)	(2)	(1)	(2)	(1)	(2)
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
D	-0.117	-0.156	-0.124	-0.148	0.019	-0.025
School fixed effects						
School 1	1.056*	1.831***	1.677*	2.974**	0.209	0.171
School 2	1.462**	1.241**	0.627	0.916*	1.815*	1.290
School 3	1.011**	1.258**	0.951**	1.135**	0.751	0.817
School 4	0.853**	1.278***	1.501***	2.051***	-0.650	-0.481
School 5	0.559	0.623	1.086*	1.401**	-0.597	-0.795
School 6	0.739**	0.559	0.549	0.193	0.777	0.608
School 7	0.427	0.609	0.676	1.006**	-0.120	-0.151
School 8	1.017***	1.104***	0.909**	1.198***	1.098	0.857
School 9	0.277	0.619	0.413	1.069	-0.055	-0.383
School 10	1.918***	2.074***	2.284***	2.454***	1.045	0.866
School 12	0.455	0.603	0.521	0.745	0.183	0.185
Sociodemographic variables						
Age (enrollment at school)	-0.033	0.044	0.345	0.385	-0.565*	-0.615*
Male	-0.292	-0.276	–	–	–	–
Score in Literature (grade 7)	-1.681***	-1.713***	-1.696***	-1.763***	-1.649***	-1.749***
Father unemployed	–	0.126	–	0.310	–	-0.129
Occupational degree of father (reference: apprenticeship or no occupational training)						
Grad. from Univ./Univ. of Appl. Sciences	–	0.438*	–	0.450	–	0.250
PhD	–	0.327	–	1.124	–	-0.365
Help with homework from father (reference: no support)						
Frequently	–	-0.696*	–	-1.169**	–	0.406
Infrequently	–	-0.210	–	-0.229	–	0.002
Rarely	–	0.005	–	0.096	–	0.194
No. of books of parents (reference: less than 50 books)						
51 to 100	–	0.250	–	0.353	–	-0.293
101 to 250	–	0.735**	–	0.646	–	0.593
251 to 500	–	1.127***	–	1.260***	–	0.611
501 to 2,000	–	0.876**	–	0.974*	–	0.546
More than 2,000	–	1.599***	–	1.712**	–	1.375
Constant	11.789***	10.460***	9.376***	8.127***	15.036***	15.161***
Statistics						
R^2	0.170	0.210	0.163	0.228	0.209	0.240
N	724	667	448	417	276	250
No. of clusters	93	93	87	87	84	82

[a] All standard errors are clustering-robust based on class as the sampling unit. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Achievement Score English (Regression Estimates)[a]

	Pooled Sample		Female Sample		Male Sample	
	(1)	(2)	(1)	(2)	(1)	(2)
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
D	-0.287	-0.277	-0.636**	-0.519*	0.278	0.152
School fixed effects						
School 1	-1.552*	-0.764	-0.938	-0.559	-1.966***	-1.023
School 2	-1.773***	-1.887***	-2.945***	-2.775***	-1.081	-1.315*
School 3	-1.861***	-1.564**	-2.170***	-1.855**	-1.726***	-1.217**
School 4	-1.369**	-1.332*	-1.270	-1.250	-1.501	-1.488
School 5	-0.635	-0.587	-0.551	-0.490	-0.829	-0.809
School 6	-0.070	-0.239	-0.114	-0.498	0.012	-0.057
School 7	-0.926	-0.684	-0.693	-0.395	-1.326	-1.417*
School 8	-1.127*	-1.031*	-1.003	-0.874	-1.219*	-1.396**
School 9	-2.973***	-2.757***	-3.644***	-3.229***	-2.029***	-1.836***
School 10	-0.581	-0.406	-0.338	-0.063	-1.274*	-1.248*
School 12	-0.389	-0.076	-0.584	-0.473	-0.171	0.267
Sociodemographic variables						
Age (enrollment at school)	0.004	0.187	0.241	0.525	-0.444	-0.383
Male	-0.259	-0.288	–	–	–	–
Score in English (grade 7)	-1.856***	-1.832***	-1.982***	-1.980***	-1.697***	-1.648***
Father unemployed	–	-0.076	–	-0.061	–	-0.147
Occupational degree of father (reference: apprenticeship or no occupational training)						
Grad. from Univ./Univ. of Appl. Sciences	–	0.556**	–	0.749**	–	0.342
PhD	–	0.552	–	1.262*	–	-0.572
Help with homework from father (reference: no support)						
Frequently	–	-0.444	–	-0.048	–	-1.359**
Infrequently	–	-0.413	–	-0.524	–	-0.153
Rarely	–	-0.055	–	-0.142	–	0.049
No. of books of parents (reference: less than 50 books)						
51 to 100	–	0.059	–	-0.169	–	0.278
101 to 250	–	0.146	–	0.569	–	-0.457
251 to 500	–	0.805*	–	0.887**	–	0.643
501 to 2,000	–	0.586	–	0.535	–	0.664
More than 2,000	–	0.619	–	0.613	–	0.498
Constant	13.545***	11.866***	12.563***	10.105***	15.407***	14.720***
Statistics						
R^2	0.345	0.367	0.379	0.409	0.315	0.381
N	602	553	355	328	247	225
No. of clusters	92	92	85	85	82	80

[a] All standard errors are clustering-robust based on class as the sampling unit. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Achievement Score Math (Regression Estimates, Effect Het.)[a]

	Pooled Sample		Female Sample		Male Sample	
	(1)	(2)	(1)	(2)	(1)	(2)
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
D_{s1}	-1.938***	-1.979**	-1.205	-1.549	-2.315***	-2.459**
D_{s2}	-0.941	-1.427*	-1.897	-1.585	-0.309	-1.002
D_{s3}	-0.501	-0.629	1.508	0.939	-2.993***	-3.065***
D_{s4}	-0.931**	-1.320***	-1.038**	-1.226***	-0.885***	-1.773***
D_{s5}	0.524	0.533	-0.723	-0.578	2.648*	2.279**
D_{s6}	-0.675	-0.441	-0.530	-0.381	-1.007	-0.558
D_{s7}	-0.576	-0.470	-0.205	-0.188	-1.280	-1.419*
D_{s8}	-0.578	-0.703**	-0.558	-0.780**	-0.632	-0.713
D_{s9}	0.342	0.813**	1.080**	1.542***	-0.400	-0.846
D_{s10}	-0.538	-0.358	-0.167	-0.023	-0.973*	-0.377
D_{s11}	-0.413	-0.231	-0.258	-0.201	-0.681	0.119
D_{s12}	-2.287***	-2.251***	-1.880***	-1.534***	-2.816***	-3.325***
School fixed effects						
School 1	-1.832***	-1.323*	-2.216**	-1.575**	-1.512**	-0.678
School 2	-0.230	-0.083	-0.222	-0.303	-0.107	0.176
School 3	-2.023***	-1.402**	-2.865***	-2.133*	-0.411	0.754
School 4	-1.208**	-0.710*	-1.312*	-0.910	-0.842	0.543
School 5	-1.725**	-1.564**	-1.263	-1.117	-2.717*	-2.043
School 6	-0.683	-0.994*	-0.729	-1.276	-0.673	-0.466
School 7	-1.398*	-1.281*	-2.007**	-1.787**	-0.278	0.043
School 8	-1.079*	-0.959	-1.148	-0.858	-0.899	-0.897
School 9	-3.428***	-3.154***	-3.507***	-3.119***	-3.227***	-3.022***
School 10	-0.355	-0.169	-0.715	-0.588	0.174	0.354
School 12	-0.629	-0.365	-1.190	-1.075	0.268	1.159
Sociodemographic variables						
Age (enrollment at school)	-0.637***	-0.677***	-0.527	-0.550	-0.860**	-0.796*
Male	-0.115	-0.248	–	–	–	–
Score in Math (grade 7)	-1.655***	-1.731***	-1.938***	-1.896***	-1.356***	-1.491***
Father unemployed	–	-0.023	–	0.263	–	-0.403
Occupational degree of father (reference: apprenticeship or no occupational training)						
Grad. from Univ./Univ. of Appl. Sciences	–	0.654***	–	0.951***	–	-0.049
PhD	–	0.522	–	1.616**	–	-1.583**
Help with homework from father (reference: no support)						
Frequently	–	-0.499	–	-0.113	–	-1.377**
Infrequently	–	-0.578**	–	-0.473	–	-0.661
Rarely	–	0.055	–	0.250	–	-0.278
No. of books of parents (reference: less than 50 books)						
51 to 100	–	0.177	–	0.018	–	0.958
101 to 250	–	0.708*	–	0.497	–	1.768***
251 to 500	–	0.774**	–	0.608	–	1.451**
501 to 2,000	–	0.739*	–	0.594	–	1.498**
More than 2,000	–	1.525***	–	1.253	–	2.687***
Constant	16.382***	16.045***	16.502***	15.569***	16.746***	15.485***
Statistics						
R^2	0.254	0.296	0.280	0.318	0.284	0.368
N	728	671	452	421	276	250
No. of clusters	93	93	87	87	84	82

[a] All standard errors are clustering-robust based on class as the sampling unit. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Achievement Score Literature (Regression Estimates, Effect Het.)[a]

	Pooled Sample		Female Sample		Male Sample	
	(1)	(2)	(1)	(2)	(1)	(2)
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
D_{s1}	-1.538	-2.898***	-2.197	-4.498***	-1.032	-0.784
D_{s2}	0.282	-0.425	0.070	-0.034	0.269	-0.862
D_{s3}	-1.413**	-1.483**	-1.737***	-2.139***	-1.346*	-1.332
D_{s4}	-0.149	-0.563	0.172	-0.101	-1.065	-2.229
D_{s5}	-1.272*	-1.389**	-1.696**	-1.664*	0.009	-0.066
D_{s6}	-0.145	-0.191	0.051	0.001	0.037	0.127
D_{s7}	-0.891*	-0.797	-1.043	-0.940	-0.600	-0.620
D_{s8}	0.398	0.350	0.103	-0.186	1.101	1.460
D_{s9}	1.536*	2.811***	1.997*	3.477***	0.951	1.830*
D_{s10}	0.444	0.535	1.031	0.970	-0.019	-0.013
D_{s11}	0.258	0.562	-0.265	0.079	1.699**	1.941**
D_{s12}	0.527	0.589	1.410*	1.535*	-0.536	-0.458
School fixed effects						
School 1	1.965***	3.857***	2.781**	5.810***	1.530**	1.558
School 2	1.438*	1.694**	0.454	0.936**	2.476	2.486
School 3	1.883***	2.276***	1.620***	2.147***	2.488***	2.660***
School 4	1.036***	1.830***	1.281**	2.122**	0.747	1.664
School 5	1.441*	1.677**	1.959**	2.351***	0.213	0.140
School 6	0.920*	0.887**	0.416	0.205	1.565	1.453
School 7	1.013*	1.262**	1.103	1.491**	1.005	1.034
School 8	0.938***	1.174***	0.759**	1.346***	1.433	1.055
School 9	-0.200	0.015	-0.424	-0.012	0.346	-0.047
School 10	1.786***	2.006***	1.704**	2.000**	1.869***	1.743*
School 12	0.362	0.593	-0.080	0.198	1.241	1.364
Sociodemographic variables						
Age (enrollment at school)	-0.043	0.027	0.280	0.247	-0.546	-0.589
Male	-0.261	-0.236	–	–	–	–
Score in Literature (grade 7)	-1.734***	-1.761***	-1.820***	-1.891***	-1.615***	-1.629***
Father unemployed	–	0.164	–	0.468**	–	-0.152
Occupational degree of father (reference: apprenticeship or no occupational training)						
Grad. from Univ./Univ. of Appl. Sciences	–	0.443*	–	0.539*	–	0.180
PhD	–	0.348	–	1.218*	–	-0.406
Help with homework from father (reference: no support)						
Frequently	–	-0.688*	–	-1.026**	–	0.326
Infrequently	–	-0.230	–	-0.202	–	-0.025
Rarely	–	-0.043	–	0.027	–	0.187
No. of books of parents (reference: less than 50 books)						
51 to 100	–	0.058	–	0.021	–	-0.198
101 to 250	–	0.700**	–	0.609	–	0.797
251 to 500	–	1.103***	–	1.142***	–	0.970
501 to 2,000	–	0.889**	–	0.932	–	0.818
More than 2,000	–	1.404***	–	1.333*	–	1.427
Constant	11.782***	10.401***	10.084***	9.175***	14.041***	13.672***
Statistics						
R^2	0.186	0.236	0.191	0.268	0.232	0.277
N	724	667	448	417	276	250
No. of clusters	93	93	87	87	84	82

[a] All standard errors are clustering-robust based on class as the sampling unit. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Achievement Score English (Regression Estimates, Effect Het.)[a]

	Pooled Sample		Female Sample		Male Sample	
	(1)	(2)	(1)	(2)	(1)	(2)
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
D_{s1}	-0.375	-2.120	-1.553	-3.089	0.654	-1.291*
D_{s2}	0.659	0.353	-1.331	-0.483	1.553*	1.201
D_{s3}	-1.264***	-1.394***	-2.497***	-2.775***	-0.152	-0.162
D_{s4}	-1.402**	-1.558**	-1.909***	-1.861**	-0.611	-1.317
D_{s5}	0.355	0.213	-1.550**	-1.839**	3.410***	2.806***
D_{s6}	0.315	0.281	-0.003	0.265	1.124*	0.811
D_{s7}	-0.722*	-0.441	-0.831**	-0.752**	-0.577	-0.606
D_{s8}	-0.512	-0.442	-0.333	-0.319	-0.790	-0.560
D_{s9}	0.171	0.358	-0.067	0.505	0.428	0.390
D_{s10}	0.094	0.251	0.246	0.399	0.822	2.306**
D_{s11}	-1.370	-1.178	-1.414	-1.222	-1.159	-0.825
D_{s12}	1.160**	1.244**	2.180***	2.502***	-0.114	-0.626
School fixed effects						
School 1	-1.965**	0.042	-0.820	0.677	-2.751***	-0.362
School 2	-2.632***	-2.530***	-3.023**	-3.177**	-2.200**	-2.094**
School 3	-1.771***	-1.376***	-1.672***	-1.110	-1.999***	-1.485*
School 4	-1.245**	-1.018	-0.966	-0.798	-1.585	-1.157
School 5	-1.492	-1.278	-0.382	-0.097	-3.496**	-2.901**
School 6	-0.747	-0.848*	-0.749	-1.221*	-0.764	-0.600
School 7	-1.159*	-1.002	-0.951	-0.636	-1.444	-1.378
School 8	-1.484**	-1.346**	-1.469**	-1.268*	-1.303	-1.452
School 9	-3.609***	-3.346***	-4.177***	-3.838***	-2.666***	-2.350**
School 10	-1.258**	-1.127**	-1.083	-0.854	-2.242	-3.219***
School 12	-1.471**	-1.151**	-2.036***	-2.019***	-0.557	0.263
Sociodemographic variables						
Age (enrollment at school)	-0.003	0.155	0.227	0.444	-0.435	-0.351
Male	-0.253	-0.283	–	–	–	–
Score in German (grade 7)	-1.881***	-1.859***	-2.057***	-2.092***	-1.706***	-1.674***
Father unemployed	–	-0.029	–	0.119	–	-0.150
Occupational degree of father (reference: apprenticeship or no occupational training)						
Grad. from Univ./Univ. of Appl. Sciences	–	0.585**	–	0.866***	–	0.334
PhD	–	0.587	–	1.303*	–	-0.369
Help with homework from father (reference: no support)						
Frequently	–	-0.267	–	0.107	–	-1.266**
Infrequently	–	-0.432	–	-0.585	–	-0.073
Rarely	–	-0.102	–	-0.228	–	-0.092
No. of books of parents (reference: less than 50 books)						
51 to 100	–	-0.004	–	-0.562	–	0.431
101 to 250	–	0.147	–	0.443	–	-0.344
251 to 500	–	0.766*	–	0.648	–	0.535
501 to 2,000	–	0.567	–	0.502	–	0.537
More than 2,000	–	0.613	–	0.368	–	0.590
Constant	14.107***	12.538***	13.162***	11.308***	15.947***	14.986***
Statistics						
R^2	0.366	0.389	0.425	0.462	0.364	0.427
N	602	553	355	328	247	225
No. of clusters	92	92	85	85	82	80

[a] All standard errors are clustering-robust based on class as the sampling unit. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Robustness Checks of Estimation: Math[a]

	(1) Coeff.	(2) Coeff.	(3) Coeff.	(4) Coeff.
D	-0.721***	-0.625**	-0.750***	-0.879***
School fixed effects				
School 1	-2.361***	–	–	–
School 2	-0.671	–	-0.687	–
School 3	-1.547***	–	-1.532***	–
School 4	-1.209***	–	-1.212***	–
School 5	-1.102**	–	-1.098**	–
School 6	-1.015**	–	-1.110***	–
School 7	–	-0.191	–	-1.310**
School 8	–	–	–	-1.116**
School 9	-2.961***	–	–	–
School 10	-0.190	–	-0.144	–
School 12	-1.191***	–	-1.197***	-1.187**
Sociodemographic variables				
Age (enrollment at school)	-0.428	-1.432***	-0.408	-1.122***
Male	-0.339	-0.111	-0.288	-0.088
Score in German (grade 7)	-1.637***	-1.902***	-1.674***	-1.782***
Father unemployed	0.026	-0.153	0.021	-0.121
Occupational degree of father (reference: apprenticeship or no occupational training)				
Grad. from Univ./Univ. of Appl. Sciences	0.706***	0.533	0.864***	0.855***
PhD	0.484	4.361***	0.587	1.867
Help with homework from father (reference: no support)				
Frequently	-0.062	-1.333**	-0.060	-0.616
Infrequently	-0.109	-1.519***	-0.170	-0.996***
Rarely	0.709**	-1.437**	0.625*	-0.296
No. of books of parents (reference: less than 50 books)				
51 to 100	0.175	-0.055	0.052	0.261
101 to 250	0.322	1.031	0.284	1.236**
251 to 500	0.402	1.002	0.493	1.196**
501 to 2,000	0.460	1.333	0.423	0.815
More than 2,000	1.175**	2.283***	1.478***	1.633**
Constant	14.343***	20.706***	14.287***	19.054***
Statistics				
R^2	0.273	0.395	0.248	0.322
N	502	169	461	325
No. of clusters	73	20	63	39

[a] All standard errors are clustering-robust based on class as the sampling unit. See text for details. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Robustness Checks of Estimation: Literature[a]

	(1) Coeff.	(2) Coeff.	(3) Coeff.	(4) Coeff.
D	-0.140	-0.104	-0.190	0.138
School fixed effects				
School 1	1.711**	–	–	–
School 2	1.223**	–	1.218**	–
School 3	1.186**	–	1.146**	–
School 4	1.226***	–	1.241***	–
School 5	0.598	–	0.598	–
School 6	0.547	–	0.505	–
School 7	–	-0.431	–	0.723*
School 8	–	–	–	1.177***
School 9	0.520	–	–	–
School 10	2.031***	–	2.011***	–
School 12	0.552	–	0.514	0.595
Sociodemographic variables				
Age (enrollment at school)	0.173	-0.357	0.111	-0.409
Male	-0.335	-0.214	-0.213	0.073
Score in Literature (grade 7)	-1.638***	-1.899***	-1.708***	-2.070***
Father unemployed	0.071	0.120	0.253	0.173
Occupational degree of father (reference: apprenticeship or no occupational training)				
Grad. from Univ./Univ. of Appl. Sciences	0.304	0.778	0.399	0.826**
PhD	0.120	2.826***	0.143	1.414
Help with homework from father (reference: no support)				
Frequently	-0.381	-2.020***	-0.128	-1.691***
Infrequently	-0.154	-0.262	-0.120	-0.065
Rarely	-0.009	0.170	-0.014	0.084
No. of books of parents (reference: less than 50 books)				
51 to 100	0.092	0.235	-0.068	0.294
101 to 250	0.418	1.178*	0.219	1.079**
251 to 500	0.819*	1.600**	0.718	1.078**
501 to 2,000	0.577	1.386*	0.436	0.600
More than 2,000	1.758***	0.698	1.785***	1.019
Constant	9.809***	14.065***	10.333***	13.546***
Statistics				
R^2	0.188	0.326	0.200	0.278
N	496	171	455	323
No. of clusters	73	20	63	39

[a] All standard errors are clustering-robust based on class as the sampling unit. See text for details. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$