Labour Economics 31 (2014) 174–187

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

Causal effects of mathematics

Torberg Falch, Ole Henning Nyhus, Bjarne Strøm *

Department of Economics, Norwegian University of Science and Technology, Norway

HIGHLIGHTS

- Random draw of 16 year olds in Norway to examination in mathematics or languages
- Selection for examination in mathematics affects short and long run outcomes
- Decrease high school dropout
- Increase enrolment in natural science and technology programs in higher education
- The effects are somewhat stronger for males than for females

ARTICLE INFO

Article history:
Received 6 March 2014
Received in revised form 22 July 2014
Accepted 30 July 2014
Available online 7 August 2014

JEL classification:
I21
J24

Keywords:
Mathematical training
Intervention
External exit examination
High school graduation
Higher education

ABSTRACT

This paper exploits that students at age 16 in Norway are randomly selected into one compulsory exit exam in either mathematics or languages. A few days before the actual exam day, the students are notified about exam subject. The students have an intensive preparation period, and examination in mathematics relative to languages is found to decrease dropout from high school, increase enrollment in higher education, and increase enrollment in natural science and technology education programs. Overall, the causal effects seem to be somewhat stronger for males than for females, but the analysis indicates that gender differences interact in complicated ways with prior skills in mathematics. We explore several mechanisms that might contribute to the findings.

1. Introduction

It is a general concern that insufficient student skills in mathematics lead to shortages of key competencies in a time with rapid technological change. A number of empirical studies find that test scores in mathematics are important predictors of future earnings and other individual outcomes, see Murnane et al. (1995) and the literature review in Hanushek (2002). Moreover, the recent cross-country studies suggest that aggregate measures of test scores in mathematics and science are important determinants of economic growth (Hanushek and Woessmann, 2008, 2012). While this evidence suggests an important role for mathematical skills, causal evidence on the impact of mathematics relative to other subjects in school is still scarce.

Our paper is related to two further strands of literature. First, a small, but growing literature initiated by Altonji (1995), investigates the impact of high school curriculum on further school and labor market outcomes. The typical finding is that more mathematics courses in high school increase educational attainment and earnings. The identification issue in this literature is not trivial, however, because the choice of coursework is clearly endogenous. Various instruments for coursework choice are used in the literature, but the identification strategies can be criticized (Altonji et al., 2012). Second, recent evidence from the experimental literature suggests that the effects of rewards and interventions are more pronounced for math tests than for reading tests, see Bettinger (2012) and references therein. The econometric analyses on dropout behavior and returns to education in Oreopoulos (2007) also indicate that students are myopic. Recent experimental evidence by Levitt et al. (2012) support this view and suggest that incentives improve test scores only if rewards are offered immediately after effort is exerted.

While our study does not consider the effect of financial rewards, we study the effect of an intervention which can be interpreted as a

* Corresponding author at: Department of Economics, Norwegian University of Science and Technology, Dragvoll, NO-7491 Trondheim, Norway.
E-mail address: bjarne.strom@svt.ntnu.no (B. Strøm).

http://dx.doi.org/10.1016/j.labeco.2014.07.016
0927-5371/

treatment consisting of two parts; a training and preparation period in ether mathematics or languages followed by a high stake test in either of these subjects. We consider the former part of the treatment to be the most important for the outcomes studied. At the end of compulsory education in Norway, at the age of 16, about 40% of the students are randomly selected to sit for a high stake external exit examination in mathematics, while the rest of the students have an examination in Norwegian or English language. The students are informed of their exam subject a few days in advance, such that there is a period of intensive preparation with extensive support from teachers. The preparation period varies from 2 to 5 working days in our empirical period 2002–2004.

The experimental setting that we exploit provides evidence on whether the observed stronger relationships between skills in mathematics relatively to languages and educational outcomes represent causal effects or merely student sorting. The observed relationships clearly indicate that even a short, but intensive, training period in mathematics immediately prior to a high stake test can have non-negligible treatment effects. This finding makes sense if students are myopic, and is broadly consistent with evidence from the experimental studies suggesting that educational incentives are most effective when the reward comes immediately. We use the population of Norwegian students from administrative registers in the analysis, and find that treatment in mathematics as opposed to languages significantly decreases dropout from high school and increases enrollment in natural science and technology studies in higher education.

The paper is organized as follows: Section 2 reviews related literature and Section 3 presents relevant institutional settings, data and empirical strategy. The empirical results are presented in Section 4, which includes several robustness and heterogeneity analysis, while Section 5 contains concluding remarks.

2. Related literature

A number of papers have investigated the impact of test scores in mathematics and science on earnings and other individual outcomes. For example Bishop (1989), Murnane et al. (1995), and Altonji and Pierret (2001) find that measures of achievement are important determinants of individual earnings for given educational attainment and observed individual and family characteristics. In a recent paper, Koedel and Tyhorst (2012) use a resume-based field experiment and find that stronger mathematical skills improve labor market outcomes.

Another strand of the literature has studied the impact of school curriculum on individual earnings, following the seminal paper by Altonji (1995). These studies typically ask to what extent earnings depend on the number and levels of mathematics and science courses taken in high school. For the US, Altonji (1995), Levine and Zimmerman (1995), and Rose and Betts (2004) generally find a positive impact on earnings of taking more mathematics and science courses. It is a question, however, whether these estimates can be interpreted causally or whether they represent selection effects or omitted variables (Altonji et al., 2012). Given the problem to find credible instruments for students' coursework, or other credible identification strategies, it is not surprising that the results vary somewhat across studies.

Three recent studies apply more credible strategies to identify the impact of curriculum on earnings. Joensen and Nielsen (2009) explore a pilot scheme implemented in some Danish high schools, in which students were allowed to select different combinations of high school courses than students enrolled in other schools. Using this variation as instrument for students' actual choices, they find that taking more advanced mathematics courses has a significant and sizable positive impact on earnings. Their estimates imply that taking one extra course in mathematics increases earnings by 20–25%. The main mechanism seems to be increased likelihood of taking higher education.

Goodman (2012) uses the US state-level changes in high school mathematics requirements as instruments for students' actual coursework and finds that additional mathematics coursework increases earnings, especially for low-skilled students. However it is not entirely clear that the estimated effect reflects only coursework in mathematics since the change in state level math requirements was part of a change toward stricter high school graduation requirements in several subjects.

Cortes et al. (forthcoming) study an algebra policy implemented in Chicago in 2003. Students with achievement below the national median result in an eighth grade exam in mathematics are assigned to algebra courses with double instructional time in ninth grade. Using a regression discontinuity design, they find sizable effects of the double-dosing in algebra on high school graduation rates, college entrance exam scores, and college enrollment rates. The intervention seems to have been most successful for students with relatively low reading skills.

These three studies have different identification approaches, but all find sizable effects of increased coursework in mathematics during the school year. In Joensen and Nielsen (2009) and Cortes et al. (forthcoming), the increased coursework in mathematics is at the expense of coursework in other subjects. Thus, the estimated effects of mathematical coursework are to some extent relative to other coursework. Our study shares this feature, although we do not study coursework per see. The intervention we study differs from the above studies in at least three important ways. First, we study the effect of intensive preparation in a few days without any other school work for the students. Second, the preparation is directly related to a high-stake test very close in time, and third, we are able to estimate average treatment effects because the whole cohort is included in the random assignment of examination subject.

The way the students work on a topic is arguably different in a short intensive preparation period for a high-stake test than during regular teaching. Haeck et al. (2014) investigate the impact of a universal school reform in Quèbec, Canada, that transformed the teaching in mathematics from approaches of memorization and repetition to problem-based and self-directed learning. They find that this change to a socio-constructivist teaching approach reduced student achievement over the whole skill distribution.

Our paper is also related to the growing literature on the impact of instruction time. For example Marcotte and Hemelt (2008) and Hansen (2011) find that reduced instruction time due to more snow-related school day cancelations reduces student performance. In addition, Hansen (2011) finds that variation in the number of instruction days across cohorts implied by state-mandated shifts in test-date is related to student performance. A similar identification strategy is pursued by Carlsson et al. (forthcoming). They exploit the conditionally random variation in the actual date for the test taken by 18 year-old males in Sweden in preparation for military service. They find that 10 days of schooling increases the score on crystallized intelligence by one percent of a standard deviation. Lavy (forthcoming) uses international comparable student tests and exploits variation in instruction time across subjects in a within-student framework. He also finds a positive effect of instruction time on test scores.

Finally, a recent literature is concerned that students have myopic behavior. Levitt et al. (2012) use large field experiments including over 6000 elementary and high school students in Chicago to study student behavior under different incentive schemes. In particular, they study how students perform on a low stake test when rewards are offered immediately after the test compared to when rewards are offered with a delay. They find strong evidence that students have very high discount rates. Incentives increase performance only when the rewards are received shortly after the effort is exerted and is more effective for younger than for older students. Of particular interest is that the effects seem more pronounced for math tests than for reading tests. This is a typical finding in the literature, see Bettinger (2012) and references therein. This evidence combined with myopic behavior motivates us to study whether a short intensive training period followed by a high stake test has different medium- to long-run impact depending on whether the relevant subject is mathematics or a language.

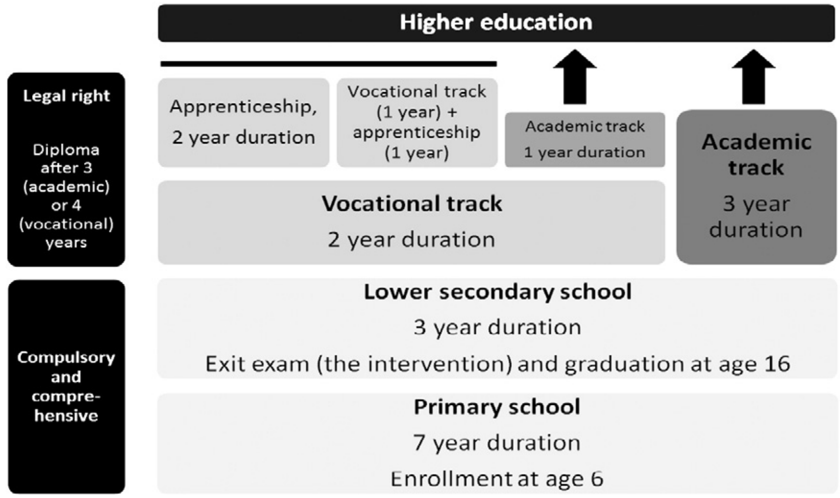

Fig. 1. The Norwegian educational system.

Our data also makes it possible to investigate whether the intervention effects are gender specific. While a large and growing literature has investigated gender differences in the response to rewards and competition, no clear cut conclusions emerge, see Lavy (2013) and the references therein. Especially relevant in our context is the small, but growing literature on the gender gap in ability and course work in mathematics and its impact on gender differences in labor market performance. Using the same identification strategy as in Joensen and Nielsen (2009), Joensen and Nielsen (forthcoming) investigate the impact of more advanced math courses by gender and across the ability distribution and find strong earnings effects for high ability females and close to zero effects for marginal males.

3. Institutional setting, data and empirical strategy

3.1. Institutions

Fig. 1 presents an overview of the Norwegian educational system. Students are normally enrolled in primary education the year they turn six. There is no possibility to fail a class neither in primary nor in lower secondary education, which implies that everybody finish compulsory education 10 years after enrollment. The compulsory education is comprehensive with no tracking and a common curriculum for all students.[1]

At graduation the students receive a diploma containing 13 different grades set by the teachers. These grades are determined before the evaluation of the external exit examinations. However, the weakest students do not get a grade in every subject (see below). The grading scale is from one to six, where six is the best grade. In addition, the diploma includes the result from the external exit examination.

After the end of compulsory education, students can choose to leave school or to enroll in high school education. About 95% enrolls in high school the year they finish compulsory education. The students could apply for 15 different study tracks in the empirical period. Three of the study tracks qualify for higher education (academic tracks) and 12 tracks give a certificate for work in a broad amount of occupations (vocational tracks). The academic tracks consist of three years, while the vocational study tracks normally consist of two years in school plus two years as apprentice. In addition, as shown in Fig. 1, there is a possibility to qualify for higher education by choosing a demanding academic year after two years in a vocational study track.

[1] Few students enroll in private schools. About two and five percent of a cohort enroll in private compulsory schools and non-compulsory high schools, respectively.

Students have to rank three different study tracks when applying for enrollment. All students have a legal right to be enrolled in one of these three tracks, but which track and school they actually enroll depends on achievement in compulsory education measured by their teacher grades and the result on the exit examination. The application deadline is in the winter/early spring, many weeks before the external exit examination. At this stage the students are well informed about their achievements in the different subjects. The diploma with the final grades is forwarded after the end of the school year. There are some possibilities to reverse the priority ranking of study tracks after the diploma is ready, and there are some possibilities to change study track early in the fall after the enrollment in high school. To our knowledge such changes are rare, but application data has not been available for this project. Thus, this dimension cannot be used as a completely clean placebo test.

The 430 municipalities are responsible for compulsory education, while the 19 counties are responsible for high school education. Enrollment into compulsory schools is based on catchment areas, while the counties decide the admission system for the high schools. They determine the capacity of the individual schools and study tracks according to local needs and student demand. Some counties use catchment areas for the individual study tracks, other counties have school choice within certain regions, while some do not have any restrictions on school choice.

3.2. The treatment

All students have to sit for a written external exit examination at the end of compulsory education either in Norwegian language, English language, or mathematics. The exam is a test that lasts five hours. The Norwegian Directorate for Education and Training prepares the exams, decides external examiners (teachers from other schools than the students they evaluate), and gives clear instructions about the randomization of students. Local authorities, including the county governor who is appointed by the central government, are responsible for the assignment of examination subjects to schools and students. For this purpose they mainly randomize schools, but at some schools different students are examined in different subjects. The randomization process is not clearly spelled out, probably in order to keep the assignment hard to predict. Anecdotal evidence clearly suggests that students have no clue about their examination subject, and there is a lot of anxiety among students before the announcement of the subject assignment.

The examination subject is unknown up to a few days before the examination day. The exam is carried out on the same day in all subjects in the empirical period of this paper, and the students are well informed in advance about the procedure.[2] The length of the period from when the students are informed about their exam subject to the examination day varies across the years. In 2002, 2003 and 2004 the students were informed 2, 7 and 5 days prior to the exam, respectively. In the latter two years the preparation period included a weekend and the national day, which implies that the preparation period ranges from 2 to 5 working-days.[3] During this period, the students have no other obligations at school than the preparation for the exam. How the preparation is done depends to a large extent of the individual student. Their teachers offer extensive support during school hours, but the students are expected to work on the examination subject also in the evenings.

The treatment starts with an intensive training period and ends with the examination. It is reasonable to consider the examination situation as relatively equal in mathematics and languages. The students perform high effort over five hours in the examination independent of subject. Training and preparation for the examination are, however, very different in mathematics and languages. While the preparation in mathematics mainly includes repetition and drilling, the preparation in writing essays is more extensive. The relevant competency for producing essays in language examination arguably requires longer time to mature. The training period is most likely more intense and effort-demanding in the former than in the latter case. The findings by Haeck et al. (2014) indicate that this is an important distinction. Nevertheless, our empirical analysis cannot distinguish between these two parts of the treatment.

About 40% of the students are randomly selected to sit an examination in mathematics, while the other students sit an examination in either the Norwegian language or the English language. However, the share of students with the examination in mathematics varies across the counties. For example, this share varied from 0.41 to 0.30 in 2002–2004 in the smallest of the 19 counties in the country. In six percent of the schools no student had the examination in mathematics during the period 2002–2004, while two percent of the schools had students with examination in mathematics each year.

This written external exit examination is the final written test in compulsory education. Most students also have an oral external examination after the written external examination. The oral examination is organized by the individual schools, and initial analyses suggest that the allocation of students is not random.[4]

3.3. Data

We use register data from Statistics Norway covering all students that finished compulsory education in the years 2002–2004. To make the sample more homogeneous we only include students that turn 16 years of age the year they finish compulsory education.[5] In addition, we only include students with teacher grading information on the three examination subjects and information on which compulsory school they graduated from. Teacher set grades are available for the whole sample and not just for the subject drawn at random, but a small number of students have missing grade information. We do not know the exact reasons for missing teacher set grades, but the main reason is probably that teachers have too little information to evaluate performance due to students not participating sufficiently in lectures and tests.

Table 1
Data reduction.

	Observations	Percent
Finish compulsory education in 2002–2004	174,067	100.0
Not turning 16 years the year finishing compulsory education	10,059	5.8
Missing information about teacher assessed grade for at least one of the subjects mathematics, Norwegian language and English language	6,878	4.0
Missing compulsory school identifier	704	0.4
Central exit examination in mathematics and one of the languages	578	0.3
Have exemption from central exit exam	146	0.1
Analytical sample	155,702	89.4

We also exclude a few students registered with exam grades in both mathematics and one of the languages, and students with exemption from the external examination. Details of the data reduction are presented in Table 1. The analytical sample consists of 89.4% of the population, amounting to 155,702 observations.

Table 2 contains descriptive statistics of the main variables used in the analysis. Panel A describes the intervention variable. Close to 40% of students were examined in mathematics, which is the treatment group in our analysis. Another 38% were examined in the English language, while 21% were examined in the Norwegian language. About two percent of the students did not appear on the examination day, where illness might be one explanation. These shares are the same for each cohort.

Panel B in Table 2 presents descriptive statistics for our main outcome variables. About 46% of the cohort enrolls high school in an academic study track the year they finish compulsory education. The majority of the sample enrolls in a vocational study track (about 51%), while some do not enroll this year (about 3%). The dropout from high school is, however, relatively large. Only about 71% of the sample graduates from high school education within 5 years after the end of compulsory education. About 44% of the sample enrolls in higher education, defined as enrollment within six years after the end of compulsory education.[6] Higher education programs in science and technology are the most demanding in terms of mathematical skills. About 6% of the sample enrolls in higher education programs in science or technology within six years after the end of compulsory education.[7] All the outcomes are relatively stable in the empirical period.

In the empirical analysis below we perform separate analyses for females and males. We also distinguish between students with low prior skills in mathematics, defined as a teacher set grade in mathematics of 3 or lower (51.4% of the sample), and students with high prior skills in mathematics, defined as a teacher set grade of 4 or higher (48.6% of the sample). This grade is set by the teachers prior to the examination.

[2] There are two formal written Norwegian languages. Students that are drawn to have their exam in Norwegian have two exam days, one in each language. The first exam day is the same as for the students drawn for exam in the other subjects.

[3] In 2002 the students were informed about their exam subject in May 22 and sat the exam in May 24. The relevant dates were May 15 and May 22 of 2003 and May 14 and May 19 of 2004.

[4] About two thirds of the students have an oral examination in one of the subjects on the curriculum. The oral examination is organized by the school district in cooperation with the individual schools, without any influence by the Directorate for Education and Training. Inspection of the data indicates that in particular students without any oral examination tend to have low teacher set grades.

[5] Since no students fail any grade in the Norwegian compulsory education, one could expect that all students turn 16 years of age the year they finish compulsory education. However, there are some exceptions. If a child is not considered to be mature enough, the parents together with the school and psychologists can postpone enrollment one year. It is also possible to start one year ahead the birth cohort. In addition, some older students return to improve their grades, and immigrants are often over-aged.

[6] The fact that about the same number of students enrolls in higher education as the number of students enrolling an academic study track in high schools arise from students choosing an academic year after a minimum of two years in a vocational track (see Fig. 1), and that some students change study track in high school. 12% of the students not enrolling the academic study track right after compulsory education graduate with an academic certificate within five years, while 3% of the students enrolling the academic study track graduate with a vocational certificate within five years.

[7] We use the Norwegian classification of education. At the level "first stage tertiary education, undergraduate level", the educational programs are divided into 10 different areas. The area that requires the most in terms of mathematical and cognitive skills is denoted "Natural science, vocational and technical education", which we denote higher education programs in science and technology. This area is sub-divided into 9 different fields. 50% of the students enroll in "Information and computer technology" and "Electrical, electronic, mechanical and machine subjects". For some study programs there are explicit requirements of advanced mathematics in the high school diploma, but that is not the case for most of the study programs. While 40% of the students who achieve the academic certificate have second year mathematics in high school, that is the case for 75% of the students that enroll in a higher education program in science and technology.

Table 2
Descriptive statistics for the treatment and educational attainment, percent.

	All	2002	2003	2004	Females	Males	Low prior math skills	High prior math skills
Panel A: central exit examination								
Examination in mathematics	38.6	39.7	38.5	37.9	38.6	38.7	38.7	38.6
Examination in Norwegian	21.1	20.5	21.5	21.4	21.1	21.1	21.0	21.2
Examination in English	38.1	38.3	37.5	38.6	38.1	38.2	38.0	38.2
Students not appearing on the examination day	2.1	1.5	2.7	2.2	2.3	2.0	2.3	2.0
Panel B: educational attainment								
Enrolling academic study track in high school	46.5	46.9	46.0	46.5	50.7	42.4	26.2	67.9
Graduating high school within 5 years	70.6	70.3	70.3	71.0	75.1	66.1	52.4	89.7
Enrollment in higher education	44.4	44.0	44.1	45.2	54.6	34.5	21.5	68.7
Enrollment in higher education, science or technology	6.0	5.9	6.1	6.0	4.2	7.8	1.5	10.8
Observations	155,702	49,534	51,185	54,983	76,770	78,932	80,038	75,664

The descriptive statistics presented in Table 2 shows that the probability of examination subject is equal across these subsamples, but that the mean values of the outcome variables vary as expected.

We include in some of the models a range of socioeconomic characteristics, including immigration, birth quartile, parental education, parental income, parental employment, and parental marital status. Descriptive statistics for the variables are presented in Appendix Table A1. Parental education is classified into four levels (only compulsory education; graduated from high school; bachelor degree; master or PhD degree) and is based on the education category of the parent with the highest education. Parental income is measured by taxable income and is included as quartile indicators. For marital status we use indicators of whether the parents are married when the student finishes compulsory education and whether the parents are divorced at that time. 61.5% of the parents were registered as married, 12.5% were registered as divorced, and 26% had never been married. The latter includes cohabiting parents, which is much less common when the age of the child is 16 (as in our case) than when the child is newly born.

3.4. Educational outcomes and subject specific achievement in compulsory education

The empirical literature generally finds a positive relationship between educational attainment and test scores in mathematics. In this section we show that this is the case also in Norway. While the evidence is descriptive, it suggests some hypotheses regarding the effects of the intensive training and testing in mathematics that we investigate below.

We run regression models with grades set by the teachers in the potential exit examination subjects (mathematics, English, and Norwegian language) as explanatory variables. In addition, the model includes the grade point average (GPA) covering the 13 teacher set grades on the diploma from compulsory school. Since GPA is included in the models, the coefficients for the specific subjects should be interpreted as whether the subject contributes more or less than the other subjects. The models include the rich set of socioeconomic characteristics described above. To account for possible differences in grading practices between schools, the models also include cohort times compulsory school fixed effects. Models are estimated for the full sample as well as for subsamples by gender, and below and above average skills in mathematics.

The results are presented in the Appendix Table A2. Overall, the results suggest that the association between educational outcomes and compulsory school grades is stronger for mathematics than for languages and this is more pronounced for males than for females. Based on these findings and the results in the literature, a natural hypothesis to test is that students randomly selected to the mathematics examination are exposed to a treatment more important for educational attainment than students selected to language examination. In addition, the treatment effects are expected to be most pronounced for males, while heterogeneity related to prior mathematical skills is expected to depend on the specific outcome variable. As to the choice of study track in high school, which is mainly made prior to the external examination in compulsory education, we expect no treatment effect.

3.5. Empirical strategy

We investigate the effect of treatment in terms of intensive training and testing in mathematics relative to languages by exploring that each student is randomly selected to external exit examination in only one of the subjects. In the empirical analysis we take treatment intensity into account as the number of treatment days varies from 2 to 5 working-days during the empirical period. Obviously, since the treatment consists of training in a short period of time, effect sizes should be much lower than the impact of one standard deviation in the teacher set grades in mathematics.

We estimate variants of the following model

$$Y_{ic} = \alpha + \beta NTD_i + X_i'\delta + \gamma_c + \varepsilon_{ic}, \tag{1}$$

where Y_{ic} represents the outcome for individual i in cohort c, NTD is the number of treatment days, X is a vector of socioeconomic characteristics, γ_c is cohort specific effects, and ε_{ic} is the random error term. β can be interpreted as the average treatment effect. If treatment is random, the estimated treatment effect would be independent of whether the model condition on X and γ_c or not. To gauge the plausibility of the randomness assumption, we present results both for the model in Eq. (1) and for models without any controls included.

In his study of the Tennessee STAR experiment, Krueger (1999) includes school fixed effects to take account of the fact that randomization was done within schools. In our case the central government gives clear instruction about randomization, while actual implementation is done at the local level under the inspection of the county governor. Thus, as specification checks we also present model versions with fixed effects for the counties, the municipalities, and the schools, respectively.

To further investigate the issue of randomness, Table 3 presents descriptive evidence on the relationship between treatment and student characteristics. Columns (1)–(3) show that the mean values of the teacher set grades, the socioeconomic characteristics, and the cohort dummy variables are similar in the treatment and the control group. In particular, the mean values of the teacher set grade in mathematics are identical in the two groups. Out of the 26 variables in the table, three of the differences are significant at 10% level and none at 5% level.[8]

[8] In the analysis in the present paper, we use the highest educational level of the parents and the sum of their incomes. The analysis can be done in more detail by exploiting information on each single parent. Then we have twice as many parental education levels and parental income quartiles related to each student than in Table 3. We find no significant differences across the treatment and control group for the income variables. For educational level, we find that the indicator for the mother having high school education as the highest educational level is significantly related to treatment at 10% level (the difference between the groups is about 0.6 percentage points), while the other relationships are clearly insignificant. Notice that the relationship between treatment and parents with high school educational level has a p-value of 0.162 in Table 3.

Table 3
Treatment in mathematics, balancing tests.

	(1) Mean value, math exam	(2) Mean value, no math exam	(3) Difference	(4) Partial OLS on number of treatment days	(5) OLS on number of treatment days	(6) OLS on number of treatment days
Grade in mathematics	−0.001	0.000	−0.0009	0.0040	−0.0195	–
			(0.0093)	(0.0081)	(0.0178)	
Grade in Norwegian language	0.006	−0.004	0.0095	0.0151*	0.0142	–
			(0.0094)	(0.0080)	(0.0195)	
Grade in English language	0.002	−0.002	0.0041	0.0065	−0.0077	–
			(0.0093)	(0.0082)	(0.0147)	
Grade point average (GPA)	0.005	−0.003	0.0077	0.0127	0.0187	–
			(0.0093)	(0.00783)	(0.0282)	
Female	0.493	0.493	−0.0008	0.0078	−0.0117	0.0016
			(0.0026)	(0.0091)	(0.0129)	(0.0088)
First generation immigrant	0.033	0.034	−0.0005	−0.0134	0.0062	0.0073
			(0.0017)	(0.0423)	(0.0468)	(0.0468)
Second generation immigrant	0.021	0.020	0.0004	0.0221	0.0211	0.022
			(0.0020)	(0.0821)	(0.0821)	(0.0822)
Parents' highest educational level is high school education	0.470	0.464	0.0061	0.0178	0.0294*	0.031**
			(0.0043)	(0.0149)	(0.0157)	(0.016)
Parents' highest educational level is bachelor degree	0.288	0.290	−0.0015	0.0059	0.0153	0.0197
			(0.0028)	(0.0118)	(0.0196)	(0.0194)
Parents' highest educational level is master or PhD	0.102	0.105	−0.0031	−0.0287	−0.0079	−0.0029
			(0.0033)	(0.0308)	(0.0323)	(0.0329)
Benefits due to disease before the age of 18	0.019	0.019	−0.0006	−0.0404	−0.0535	−0.0533
			(0.0007)	(0.0331)	(0.0378)	(0.0378)
Benefits due to disabilities before the age of 18	0.024	0.025	−0.0003	0.0058	0.0246	0.0224
			(0.0008)	(0.0294)	(0.0329)	(0.0328)
One parent employed	0.239	0.243	−0.004	−0.0151	0.0046	0.0052
			(0.0028)	(0.0130)	(0.0242)	(0.0242)
Both parents employed	0.710	0.704	0.0058	0.0161	0.0304	0.0316
			(0.0035)	(0.0147)	(0.0276)	(0.0276)
2nd quartile parental income	0.250	0.250	0.0008	−0.0237*	−0.0224	−0.022
			(0.0031)	(0.0143)	(0.0152)	(0.0151)
3rd quartile parental income	0.253	0.248	0.0047*	0.0199	−0.0046	−0.0039
			(0.0028)	(0.0131)	(0.0200)	(0.0199)
4th quartile parental income	0.246	0.252	−0.0060	0.0090	−0.0203	−0.0192
			(0.0049)	(0.0232)	(0.0275)	(0.0274)
Married parents	0.613	0.608	0.0056*	0.0066	0.0071	0.0088
			(0.0032)	(0.0123)	(0.0146)	(0.0146)
Divorced parents	0.124	0.127	−0.0031*	−0.0172	−0.0160	−0.0166
			(0.0018)	(0.0146)	(0.0166)	(0.0165)
Mobility	0.112	0.111	0.0012	0.0137	0.0207	0.0199
			(0.0021)	(0.0181)	(0.0188)	(0.0188)
Mobility unknown	0.021	0.022	−0.0011	−0.0371	−0.0176	−0.0208
			(0.0009)	(0.0381)	(0.0377)	(0.0377)
Born second quartile	0.265	0.268	−0.0028	−0.0096	−0.0065	−0.0069
			(0.0024)	(0.0107)	(0.0126)	(0.0125)
Born third quartile	0.258	0.259	−0.0009	−0.0053	−0.0029	−0.0038
			(0.0024)	(0.0107)	(0.0125)	(0.0124)
Born fourth quartile	0.230	0.227	0.0029	0.0213*	0.0132	0.0117
			(0.0022)	(0.0115)	(0.0127)	(0.0126)
Cohort 2003	0.327	0.330	−0.0025	0.949***	1.130***	1.130***
			(0.0254)	(0.113)	(0.114)	(0.114)
Cohort 2004	0.346	0.357	−0.0112	−0.231***	0.342***	0.340***
			(0.0233)	(0.082)	(0.070)	(0.070)
Test of joint significance, excluding cohort specific effects, p-value					0.383	0.377

Note. 155,702 observations. Standard errors in parentheses are clustered at the compulsory school level.
*** Denotes significance at the1% level.
** Denotes significance at the 5% level.
* Denotes significance at the 10% level.

Column (4) in Table 3 replaces the dummy variable for treatment with the number of treatment days and presents partial regressions. Since the number of treatment days only varies across cohorts, they are clearly related to the cohort dummy variables. For the other variables, the relationship is significant at 10% level in only two cases.

Column (5) in Table 3 presents results from a multivariate regression with the number of treatment days as the dependent variable. This model includes the socioeconomic characteristics that are used as control variables in the analyses below, in addition to the teacher set grades. The indicator for whether the parents' highest educational level is high school education is the only variable that turns out as significant at 10% level. Using an F-test we cannot reject the hypothesis that all explanatory variables have jointly zero effects (p-value of 0.38).

The last column in Table 3 presents a regression only including the variables that are used as control variables in analysis below, i.e., we only condition on socioeconomic characteristics and not on prior grades. As expected, excluding the grades from the equation does not alter the results, since each of them are unrelated to the treatment. As a further check on treatment randomness, Appendix Table A3 presents results from this regression for the subsamples we use in the analysis below. The socioeconomic characteristics are jointly unrelated to the

Table 4
Effects of treatment days on educational attainment.

	(1)	(2)	(3)	(4)	(5)
Panel A: enrolling academic study track in high school					
Number of treatment days	0.00052	0.00118	0.00119	0.00080	0.00071
	(0.00126)	(0.00103)	(0.00097)	(0.00092)	(0.00088)
Panel B: graduating high school within 5 years					
Number of treatment days	0.00199**	0.00198***	0.00185**	0.00200***	0.00188***
	(0.00088)	(0.00073)	(0.00072)	(0.00071)	(0.00072)
Panel C: enrollment in higher education					
Number of treatment days	0.00153	0.00171**	0.00164**	0.00190**	0.00153*
	(0.00112)	(0.00083)	(0.00082)	(0.00079)	(0.00081)
Panel D: enrollment in higher education, science or technology					
Number of treatment days	0.00105***	0.00107***	0.00105***	0.00106***	0.00093**
	(0.00038)	(0.00039)	(0.00038)	(0.00040)	(0.00042)
Socioeconomic characteristics	No	Yes	Yes	Yes	Yes
Cohort specific effects	No	Yes	Yes	Yes	Yes
County fixed effects	No	No	Yes	Yes	Yes
School district fixed effects	No	No	No	Yes	Yes
School fixed effects	No	No	No	No	Yes
Observations	155,702	155,702	155,702	155,702	155,702

Note. The socioeconomic characteristics included in columns (2)–(5) are described in Section 3.3 and presented in Appendix Table A1. Standard errors in parentheses are clustered at the compulsory school level.
*** Denotes significance at the 1% level.
** Denotes significance at the 5% level.
* Denotes significance at the 10% level.

treatment in each year, for both genders, and both for students with low and high prior skills in mathematics.[9]

Overall, the empirical evidence clearly indicates that the treatment is random, as it should be according to the institutions.

4. Empirical results

We first estimate effects of the treatment in mathematics on the educational outcomes presented above. Thereafter we investigate in more detail the possible channels behind these effects. In addition to presenting average effects for the population, we split the sample with regard to gender and prior skills in mathematics. Finally, we provide some analyses on the robustness of our specification of the treatment.

4.1. Educational attainment

Table 4 presents the results. The models in column (1) include only the treatment variable. As discussed above, the treatment effect on the probability to enroll in an academic study track in high school is expected to be zero because this choice is basically made several weeks before treatment takes place. The results in Panel A confirm this hypothesis. The effect of the treatment is very close to zero and highly insignificant.

Panel B presents results for the probability to graduate from high school within 5 years after the end of compulsory education. One day of intensive mathematical training, combined with the high-stake testing, increases the probability to graduate from high school with 0.2 percentage points. The effect is significant at 5% level, and it is a non-trivial average treatment effect. For the cohort in 2003, which had a treatment of five working days, the estimate indicates an effect of 1 percentage point.

The result in Panel C in Table 4 implies that the treatment of one day increases the probability to enroll in higher education by 0.15 percentage points, while Panel D shows that the effect on enrollment in study programs in science and technology in higher education is about 0.11. These are the educational fields that are the most demanding in terms of mathematical skills. The effect of treatment of one day is 1.7% of the mean value, and is significant at one percent level.

The models in column (2) in Table 4 include socioeconomic characteristics and cohort fixed effects, similar to Eq. (1) above. This does not change the estimated treatment effects, but increases the precision somewhat. In particular, the effect on enrollment in higher education is significant at 5% level in this model. Since the county governors control the randomization process, the models in column (3) include county fixed effects for the 19 counties. This does neither affect the estimated effects. The models in column (4) include fixed effects for the 440 municipalities. The municipalities are responsible for compulsory education and are involved in the assignment of exam subjects. In this model specification, the effect on enrollment in higher education increases to 0.19 percentage points. Notice that in particular in small municipalities, there will typically be some clustering of exam subjects for a given cohort, although that is less likely across the 3 cohorts in our sample. Finally, column (5) includes school fixed effects. Also in this case the estimated treatment effects mainly remain unchanged.[10]

Taking the point estimates in Table 4 at face value, most of the gain in high school graduation maturates in enrollment in higher education, and about 2/3 of the latter turns up in the fields of science and technology. This does not necessarily imply that the marginal students induced to enroll higher education by the treatment enroll in these fields. Rather, it is plausible that some students enrolling in less demanding fields in terms of mathematical skills in the absence of treatment switch to science and technology because of the treatment.

A possible interpretation is that intensive training in a short period of time is more productive for mathematics than for languages as competency in languages requires longer time to mature. While our natural experiment does not allow for a rigorous test, this interpretation does not easily fit with the descriptive evidence suggesting a stronger positive relationship between longer-term outcomes and grades in mathematics than of grades in languages.

4.2. Heterogeneous treatment effects

The models in Table 5 split the sample according to gender and mathematical skills prior to the exam. We present results for models without any control variables and for models including socioeconomic characteristics and cohort fixed effects. Panels A and B present gender specific models. Again, it turns out that choice of study track in high school education is unrelated to treatment. For the other three outcomes, the treatment has a significant effect only for males. This is in accordance with the findings in Levitt et al. (2012) that male students are more responsive to incentives than females. In addition, a larger effect on males than on females is in accordance with the associations between these outcomes and compulsory school grades in mathematics reported in Appendix Table A2. Further, the relative size of the impact on the different outcomes is similar for the males as for the population in Table 4. Notice, however, that the differences across gender are not statistically significant (p-value 0.12).

[9] The treatment is also unrelated to the teacher set grades in mathematics and English. For the full sample, the correlation with the teacher set grade in the Norwegian language is significant at 10% level as shown in Table 3. This is driven by males and students with low mathematical skills.

[10] Throughout the paper, we report results from linear probability models. A well-known problem is the possibility that such models may predict probabilities outside the [0,1] interval. In the specifications in column (2) in Table 4 (the model without fixed effects), this happens in 0.1%, 0.7%, 0.6% and 3.7% for the outcomes in Panels A–D, respectively. However, the results are very robust to the choice of other estimation methods. For example, using the logit model on the same specifications, the estimated effects are significant at the same level as for the linear probability model. The marginal effects changes from 0.00198 to 0.00214, from 0.00171 to 0.00212, and from 0.00107 to 0.00090 for the outcomes graduating high school, enrolment in higher education, and enrolment in a science or technology, respectively. Using the probit model gives similar results.

Table 5
Effects of treatment days on educational attainment, subsamples.

	(1) Enrolling academic study track in high school	(2) Graduating high school within 5 years	(3) Enrollment in higher education	(4) Enrollment in higher education, science or technology
Panel A: females				
Without control variables	0.00118	0.00092	0.00090	0.00062
	(0.00148)	(0.00106)	(0.00132)	(0.00045)
With control variables	0.00179	0.00126	0.00109	0.00063
	(0.00132)	(0.00094)	(0.00108)	(0.00046)
Observations	76,770	76,770	76,770	76,770
Panel B: males				
Without control variables	−0.00023	0.00293**	0.00191	0.00150***
	(0.00157)	(0.00114)	(0.00139)	(0.00059)
With control variables	0.00054	0.00260***	0.00229**	0.00147**
	(0.00131)	(0.00098)	(0.00112)	(0.00061)
Observations	78,932	78,932	78,932	78,932
Panel C: low prior math skills				
Without control variables	0.00026	0.00320***	0.00214**	0.00042
	(0.00121)	(0.00116)	(0.00104)	(0.00027)
With control variables	0.00103	0.0031***	0.00225**	0.00039
	(0.00115)	(0.00112)	(0.00096)	(0.00027)
Observations	80,038	80,038	80,038	80,038
Panel D: high prior math skills				
Without control variables	0.00060	0.00062	0.00069	0.00167**
	(0.00143)	(0.00073)	(0.00120)	(0.00069)
With control variables	0.00160	0.00100	0.00151	0.00178**
	(0.00129)	(0.00070)	(0.00107)	(0.00071)
Observations	75,664	75,664	75,664	75,664

Note. Each cell represents an independent regression with the number of treatment days as independent variable. The control variables include the socioeconomic characteristics described in Section 3.3 and presented in Appendix Table A1 and cohort fixed effects. Standard errors in parentheses are clustered at the compulsory school level.

*** Denotes significance at the 1% level.
** Denotes significance at the 5% level.
* Denotes significance at the 10% level.

The last part of Table 5 splits the sample by compulsory school grades in mathematics set by teachers before treatment takes place. Again, both for students with low and high prior skills in mathematics, there is no treatment effect on the choice of study track in high school. However, for the other outcomes, some interesting patterns emerge.

First, the treatment effect on high school graduation only appears for students with mathematical skills below the mean. This result makes sense since 70% of the students graduate, and the students on the margin of graduation is likely to be in the group with low prior mathematical skills. For this group, the treatment effect is equal to 0.6% of the average graduation rate. We find the same pattern for the probability to enroll in higher education, where the treatment effect is equal to 1.0% of the average value for students with prior mathematical skills below the mean.

Second, for the enrollment in higher education studies in science or technology, the treatment effect is present only for students with prior mathematical skills above the mean. Again, this makes sense when we take into account that these are the students for which such studies are the most likely alternative. 10.8% of the students with above average prior skills in mathematics enroll in such study programs, while that is the case for only 1.5% of the students with prior skills below average (see Table 2). For the students with prior mathematical skills above mean, the estimated treatment effect is 1.6% of the mean value. The effects are significantly different across the skill groups at 10% significance level.

A student's perceived probability to obtain a low or high grade on the exam grade depends on examination subject, and since the result

Table 6
Effects of treatment days on educational attainment, detailed subsamples.

	(1) Enrolling academic study track in high school	(2) Graduating high school within 5 years	(3) Enrollment in higher education	(4) Enrollment in higher education, science or technology
Panel A: females with low prior math skill				
Number of treatment days	0.00147	0.00323**	0.00349**	0.00009
	(0.00164)	(0.00154)	(0.00148)	(0.00033)
Observations	37,894	37,894	37,894	37,894
Panel B: males with low prior math skill				
Number of treatment days	0.00064	0.00300**	0.00116	0.00065
	(0.00145)	(0.00145)	(0.00107)	(0.00041)
Observations	42,144	42,144	42,144	42,144
Panel C: females with high prior math skills				
Number of treatment days	0.00300*	0.00026	−0.00019	0.00128
	(0.00160)	(0.00085)	(0.00124)	(0.00082)
Observations	38,876	38,876	38,876	38,876
Panel D: males with high prior math skills				
Number of treatment days	0.00006	0.00180*	0.00324**	0.00233**
	(0.00173)	(0.00104)	(0.00163)	(0.00116)
Observations	36,788	36,788	36,788	36,788

Note. Each cell represents an independent regression with the number of treatment days as independent variable. The control variables include the socioeconomic characteristics described in Section 3.3 and presented in Appendix Table A1 and cohort fixed effects. Standard errors in parentheses are clustered at the compulsory school level.

*** Denotes significance at the 1% level.
** Denotes significance at the 5% level.
* Denotes significance at the 10% level.

on the external exit examination matters for high school enrollment this may potentially affect student effort. Suppose students' perception is that a good exam grade is less likely in mathematics than in languages. A student with weak prior performance in mathematics relatively to languages could then have an incentive to exert more effort if he/she is randomly assigned to an exam in mathematics as opposed to languages.[11] To investigate whether such mechanisms can be driving our results, we have restricted the sample to the students with at least the same teacher set grade in mathematics as in both language subjects. However, when using this restricted sample, we get qualitatively the same results as reported in Table 4.[12]

A related concern is that the response to prior skills in mathematics and the confidence with, and perception of performance in mathematics and languages may differ between genders as the evidence in Niederle and Vesterlund (2010) suggests. Joensen and Nielsen (forthcoming) suggests that succeeding with more advanced mathematics in high school courses may affect the preferences, self-confidence or self-perception of females. Although our treatment variable is very different from theirs, relating the treatment effect differences across gender to the differences across skill groups might potentially increase the understanding of the underlying mechanisms.

Table 6 therefore interacts gender and skill group in order to provide a richer description of the heterogeneity in the treatment effects. The differences across gender are small for students with low prior skills in mathematics. The gender difference is related to the high ability students. Treatment effects are found for male students with high ability, while this is not the case for females. Accordingly, the differences across skill groups regarding high school graduation and enrollment in higher

[11] In our sample, the average grade on the exam is 3.30, 3.67, and 3.58 in mathematics, Norwegian, and English, respectively. The respective averages for the teacher set grades are 3.48, 3.85, and 3.73.

[12] 48% of the students have at least the same grade in mathematics as in both Norwegian and English. Restricting the sample to these students, the estimated effects are −0.0001, 0.0016, 0.0009, and 0.0018 for enrolling academic study track, graduating from high school, enrolling higher education, and enrolling higher education program in science or technology, respectively, in the models without controls.

Table 7
Effects of treatment days on progression in high school education.

	(1) Enrolling third year on-time	(2) Enrolling third year on-time, academic study track	(3) Enrolling third year on-time, not academic study track	(4) Enrolling final semester within 5 years, not academic study track	(5) Has been apprentice, not academic study track
Panel A: all					
Without control variables	0.00188**	0.00066	0.00267**	0.00280***	0.00134
	(0.00088)	(0.00080)	(0.00122)	(0.00100)	(0.00130)
With control variables	0.00151**	0.00032	0.00204*	0.00228**	0.00075
	(0.00074)	(0.00076)	(0.00113)	(0.00092)	(0.00118)
Observations	155,702	72,352	83,350	83,350	83,350
Mean dependent variable	0.752	0.900	0.623	0.754	0.409
Panel B: females					
Without control variables	0.00098	0.00059	0.00076	0.00222*	0.00071
	(0.00103)	(0.00096)	(0.00161)	(0.00130)	(0.00141)
With control variables	0.00073	0.00029	0.00047	0.00225*	0.00058
	(0.00093)	(0.00093)	(0.00152)	(0.00123)	(0.00142)
Observations	76,770	38,918	37,852	37,852	37,852
Mean dependent variable	0.786	0.913	0.656	0.780	0.236
Panel C: males					
Without control variables	0.00268**	0.00068	0.00430***	0.00331**	0.00181
	(0.00115)	(0.00115)	(0.00154)	(0.00131)	(0.00170)
With control variables	0.00220**	0.00034	0.00327**	0.00225*	0.00078
	(0.00101)	(0.00113)	(0.00146)	(0.00124)	(0.00160)
Observations	78,932	33,434	45,498	45,498	45,498
Mean dependent variable	0.719	0.886	0.595	0.732	0.552
Panel D: low prior math skills					
Without control variables	0.00253**	0.00072	0.00310**	0.0033***	0.00095
	(0.00119)	(0.00191)	(0.00136)	(0.00119)	(0.00138)
With control variables	0.00233**	0.00082	0.00257*	0.00272**	0.00071
	(0.00113)	(0.00189)	(0.00132)	(0.00115)	(0.00128)
Observations	80,038	20,983	59,055	59,055	59,055
Mean dependent variable	0.603	0.779	0.540	0.685	0.397
Panel E: high prior math skills					
Without control variables	0.00105	0.00061	0.00177	0.00182*	0.00231
	(0.00066)	(0.00059)	(0.00149)	(0.00100)	(0.00218)
With control variables	0.00084	0.00033	0.00147	0.00162	0.00036
	(0.00065)	(0.00060)	(0.00147)	(0.00102)	(0.00187)
Observations	75,664	51,369	24,295	24,295	24,295
Mean dependent variable	0.909	0.950	0.824	0.923	0.438

Note. Each cell represents an independent regression with the number of treatment days as independent variable. The control variables include the socioeconomic characteristics described in Section 3.3 and presented in Appendix Table A1 and cohort fixed effects. Standard errors in parentheses are clustered at the compulsory school level.
*** Denotes significance at the 1% level.
** Denotes significance at the 5% level.
* Denotes significance at the 10% level.

education are related to female students. High ability female students are not affected by the treatment. Likewise, the effect on enrollment in science and technology in higher education is driven by high ability male students. Thus, while the overall impression from simple models is that the treatment effects are stronger for males than for females, allowing for detailed heterogeneous effects in several dimensions suggest that gender differences are more complicated and related to the interaction with prior skills. Unfortunately, the results are too fragile to allow distinguishing between different underlying mechanisms.

Taken together, the results indicate that the treatment affects students across the whole ability distribution, but at different margins. The effects seem to be mediated through different channels for students located at different points in the ability distribution. In the next sections we investigate treatment effects on outcomes during high school education in order to shed some more light on the potential channels through which the effects on high school graduation and higher education enrollment might spell out.

4.3. Student progression in high school

Table 7 investigates treatment effects on the progression in high school education. The mean values of the dependent variables are presented in the table. Column (1) shows that one day of treatment implies a statistically significant 0.15 percentage points increase in the probability to enroll on-time in the third year in high school. In the academic study track this is the final year, while in vocational study tracks it depends on whether an apprentice part is included in the program. The normal progression in the apprentice system is to start as an apprentice in the beginning of the third school year and graduate with a craft certificate two years later. The treatment effect on progression is entirely driven by males and students with prior mathematical skills below the mean, and the coefficient sizes are close to the results for graduation within five years in Table 4.

Columns (2) and (3) in Table 7 investigate whether the treatment effect on progression differs between students in the academic and vocational study tracks. We split the sample according to the study track in the first year in high school since it is relatively common to change study track during high school education,[13] in particular from a vocational study track to the main academic study track. While the division of the sample by outcome variables might in general introduce selection

[13] Three percent of the sample are not registered in high school education the fall in the year they finish compulsory education, but a few of those students nevertheless graduate from high school within five years. They are included in the sample of students not enrolling an academic track in Table 4 in order to keep the population of students in the regressions. The qualitative results are not altered by excluding these students.

problems, this is likely not a problem here since we found no treatment effect on the initial choice of study track. The results in columns (2) and (3) show that the positive treatment effect on progression is entirely driven by the students enrolling a vocational study track. Further, the effect is strongest and highly significant for males (Panel C) and for students with low prior mathematical skills (Panel D), similar to the findings above.

Since several vocational study tracks are stipulated to four years study, we also estimate the effect on enrolling the final semester within five years for the relevant students. Notice that in this case we allow the students to be delayed compared to the normal progression. In the sample of students enrolling a vocational study track immediately after compulsory education, 75.4% enrolls the final semester within five years.[14] The treatment effect on this outcome (column (4)) is very close to the effect on on-time progression. For completeness, we also present results for the probability of becoming an apprentice (column (5)). No significant treatment effects are found for this outcome.[15]

Overall, the results in Table 7 clearly indicate that the positive treatment effect on high school graduation is largely a result of decreased dropout of students initially enrolled in vocational study tracks. However, these results cannot explain the positive treatment effect on the probability to enroll in higher education studies in science and technology. First, decreased dropout rate is found for students initially enrolled in vocational study tracks and most of them achieve a vocational certificate that does not qualify for higher education. Second, decreased dropout rate occurs for students with prior mathematical skills below mean, while the treatment effect on enrollment in higher education in science and technology was found for students with prior skills above mean.

4.4. Student achievement in high school

In order to identify possible channels for the treatment effect on enrollment in science and technology higher education programs, we exploit information from high school diplomas and estimate treatment effects on high school grades in mathematics for students initially enrolling in academic study tracks. The results are presented in Table 8.[16]

In the first year of the academic study track, all students take a two-semester course in mathematics. While the students in our sample had the same curriculum in the first semester (fall), the students could choose between an advanced course and a "practical" course in the second semester (spring). Column (1) in Table 8 investigates this choice. The results indicate that the treatment slightly increases the probability to choose the advanced course, but the effect is clearly insignificant. The effect is largest, however, for students with prior mathematics skills above mean.

Column (2) in Table 8 uses the value added in grade in mathematics from compulsory education to the end of the first year in high school as the outcome.[17] The value added is calculated using standardized values,

Table 8
Effects of treatment days on high school achievement, academic study track.

	(1)	(2)	(3)	(4)
	Choosing first year advanced mathematics	Value added in first year mathematics	Choosing second year mathematics	Value added in second year mathematics
Panel A: all				
Without control variables	0.00059	0.0048	−0.00020	0.0013
	(0.00167)	(0.0030)	(0.00158)	(0.0041)
With control variables	0.00127	0.0080**	0.00136	0.0038
	(0.00163)	(0.0031)	(0.00143)	(0.0042)
Observations	58,012	58,012	58,012	29,392
Mean dependent variable	0.618	−0.665	0.505	−1.007
Panel B: females				
Without control variables	0.00129	0.0056	0.00088	0.0027
	(0.00193)	(0.0035)	(0.00195)	(0.0051)
With control variables	0.00134	0.0083**	0.00230	0.0052
	(0.00197)	(0.0035)	(0.00185)	(0.0053)
Observations	32,500	32,500	32,500	14,048
Mean dependent variable	0.556	−0.654	0.431	−0.912
Panel C: males				
Without control variables	−0.00001	0.0037	−0.00124	−0.0008
	(0.00216)	(0.0037)	(0.00211)	(0.0051)
With control variables	0.00118	0.0076**	0.00016	0.0027
	(0.00208)	(0.0037)	(0.00186)	(0.0051)
Observations	25,512	25,512	25,512	15,344
Mean dependent variable	0.698	−0.679	0.600	−1.095
Panel D: low prior math skills				
Without control variables	−0.00258	−0.0004	−0.00287	0.0020
	(0.00286)	(0.0040)	(0.00227)	(0.0101)
With control variables	−0.00050	0.0028	−0.00062	−0.0017
	(0.00289)	(0.0041)	(0.00223)	(0.0105)
Observations	12,051	12,051	12,051	2,058
Mean dependent variable	0.222	−0.457	0.169	−0.527
Panel E: high prior math skills				
Without control variables	0.00117	0.0062*	0.00028	0.0019
	(0.00169)	(0.0033)	(0.00170)	(0.0041)
With control variables	0.00176	0.0090*	0.00173	0.0046
	(0.00163)	(0.0032)	(0.00152)	(0.0042)
Observations	45,961	45,961	45,961	27,334
Mean dependent variable	0.722	−0.720	0.593	−1.044

Note. Each cell represents an independent regression with the number of treatment days as independent variable. The control variables include the socioeconomic characteristics described in Section 3.3 and presented in Appendix Table A1 and cohort fixed effects. Standard errors in parentheses are clustered at the compulsory school level.
*** Denotes significance at the 1% level.
** Denotes significance at the 5% level.
* Denotes significance at the 10% level.

and the mean value are negative because there is a selection of students with relatively high grades into the academic study track. The treatment effect is positive for the full sample of students, and significant at 5% level in the model including socioeconomic characteristics. The estimated effect implies that one day of intensive training increases value added in mathematics by 0.8% of a standard deviation. This effect is positive for both females and males, but not for student with low prior skills. However, the latter sample is very small due to the selection into study tracks.

Mathematics was not a compulsory subject in the second year in high school in the empirical period. Students studying further mathematics could choose between an advanced course and a course in "business" mathematics. Column (3) in Table 8 shows estimated treatment effects on the probability to study further mathematics (either of the two courses). Column (4) estimates treatment effects on the

[14] This number can be decomposed into 58.3% who graduate, 10.9% who fail the final examination, and 6.2% who drop out before the final examination.

[15] As explained in Section 3.1 and Fig. 1, there is a possibility for students initially enrolled in a vocational study track to qualify for higher education studies by choosing a demanding academic year after two years in a vocational study track. Regression models using an indicator variable for this outcome as the dependent variable shows a positive point estimate for the treatment variable, but far from significant at conventional levels for the total sample and for males and females separately.

[16] This sub-sample is not significantly related to the treatment. Relating the sample to the number of treatment days, the p-value is equal to 0.46 and 0.16 with and without control variables, respectively. The average grade in mathematics from compulsory education in this sample is equal to 4.22 and 4.23 for the students with and without treatment, respectively.

[17] Both compulsory and high school grades are set by students' teachers. Notice that the high school grades are from two different courses and that some students are enrolled in two minor academic tracks. The models that condition on socioeconomic characteristics include a dummy variable for the advanced mathematics course and dummy variables for study track. Excluding these dummy variables from the models do not affect the qualitative results.

Table 9
Effects of treatment dummy on educational attainment.

	(1) Enrolling academic study track in high school	(2) Graduating high school within 5 years	(3) Enrollment in higher education	(4) Enrollment in higher education, science or technology
Panel A: all				
Without control variables	0.00019	0.00686**	0.00287	0.00377***
	(0.00482)	(0.00333)	(0.00413)	(0.00145)
With control variables	0.00198	0.00671**	0.00436	0.00397***
	(0.00366)	(0.00264)	(0.00295)	(0.00140)
Observations	155,702	155,702	155,702	155,702
Panel B: females				
Without control variables	0.00053	0.00249	0.00045	0.00275
	(0.00563)	(0.00386)	(0.00483)	(0.00168)
With control variables	0.00323	0.00328	0.00271	0.00307*
	(0.00470)	(0.00333)	(0.00385)	(0.00165)
Observations	76,770	76,770	76,770	76,770
Panel C: males				
Without control variables	−0.00000	0.01126***	0.00554	0.00470**
	(0.00590)	(0.00433)	(0.00518)	(0.00216)
With control variables	0.00093	0.00997***	0.00631	0.00473**
	(0.00462)	(0.00355)	(0.00395)	(0.00215)
Observations	78,932	78,932	78,932	78,932
Panel D: low prior math skills				
Without control variables	0.00019	0.01144***	0.00577	0.00110
	(0.00468)	(0.00436)	(0.00385)	(0.00098)
With control variables	0.00060	0.01006**	0.00563*	0.00119
	(0.00404)	(0.00395)	(0.00337)	(0.00096)
Observations	80,038	80,038	80,038	80,038
Panel E: high prior math skills				
Without control variables	0.00164	0.00330	0.00142	0.00692***
	(0.00540)	(0.00265)	(0.00443)	(0.00262)
With control variables	0.00399	0.00351	0.00387	0.00682***
	(0.00462)	(0.00247)	(0.00379)	(0.00259)
Observations	75,664	75,664	75,664	75,664

Note. Each cell represents an independent regression with the treatment dummy as independent variable. The control variables include the socioeconomic characteristics described in Section 3.3 and presented in Appendix Table A1 and cohort fixed effects. Standard errors in parentheses are clustered at the compulsory school level.
*** Denotes significance at the 1% level.
** Denotes significance at the 5% level.
* Denotes significance at the 10% level.

value added in the mathematics grade from compulsory education to the end of the second year in high school, given that mathematics courses are chosen.[18] For both of these outcomes, the estimated treatment effects are small and insignificant. Even though the treatment seems to increase mathematics performance in the first year, that does not seem to carry on to the second year for those students who choose mathematics the second year.

Overall, the treatment seems to have increased graduation from high school by increasing the progression of male students and students with low prior mathematical skills, who choose to enroll in a vocational study track. On the other hand, the treatment might have increased the enrollment in higher education study programs in science and technology by increasing the mathematical skills of students with high prior mathematical skills. This is a group of students who are not on the margin of dropping out of high school.

4.5. Alternative treatment variable

The treatment effects above are estimated using information on both treatment status and variation in the number of treatment days – the period between students receive notification on exam subject and the exam day – across cohorts. We investigate in this section whether the treatment effect varies across cohorts in accordance with this model formulation. To some extent, using a pure treatment indicator can be used to distinguish between being exposed to the test (the dummy) or being exposed to extra training (the continuous variable).

Firstly, we estimate models where the number of treatment days is replaced by a treatment indicator. Since the treated have on average 3.3 working days of training, we expect the effect of the dummy variable for treatment to be around 3.3 times as large as the effect of the number of treatment days as reported above. Table 9 presents the results. As before, there are no significant effects on the choice of study track. Column (2) shows the same pattern for high school graduation as in Table 4. In fact, the coefficients that are significantly different from zero are very close to expectation; they are 3.2–3.8 times larger than the effects reported in Table 4. This might indicate that the formulation with the number of treatment days represents the data generating process quite well. The findings for enrolling a study program in science or technology in higher education are similar (column (4)), while for enrollment in higher education (column (3)), the treatment effects tend to be less precisely estimated than in the model formulation with the number of treatment days.

In order to discriminate between the model formulations in Tables 5 and 9, we also used an encompassing approach by estimating a general model including both the treatment indicator and the number of treatment days as explanatory variables. The effect of the number of treatment days does not change much when the treatment dummy variable is included in the model, while the effect of the dummy variable gets much smaller than in the models reported in Table 9. For example, in the model for enrollment in higher education, the effect of the treatment dummy variable turns negative in all cases.[19] However, limited variation in the number of treatment days makes it impossible to statistically discriminate between the formulations.

Secondly, we estimate models with cohort specific treatment effects. Table 10 shows that for graduating high school within five years, the average effects are about 0.5, 1.0, and 0.5 percentage points in 2002, 2003, and 2004, respectively. These differences in effect sizes are very similar to the differences in the training period. The effect is clearly largest in 2003 when the training period was longest. Notice, however, that because the treatment effect is small, we cannot formally reject that it is equal across the cohorts.[20] Indeed, we cannot formally distinguish between the two parts of the intervention, the intensive training period and the high-stake test itself.

The effect is largest in 2003 also for the other outcomes. While the effect on enrollment in higher education is estimated to be close to zero in the other years, the effect for enrollment in a study program in science or technology is estimated to be of about the same size in 2002 as in 2003.[21] Taken together, the results seem to indicate that it

[18] The sub-sample of students with mathematics in the second year in high school does not seem to be a selected sample related to the treatment. Relating the sample to the number of treatment days, the p-value is equal to 0.89 and 0.72 with and without control variables, respectively.

[19] In the model for high school graduation including both the number of training days and the dummy variable for the intervention, the effects are 0.0017 (0.0025) and 0.0013 (0.0092), respectively, in the model specification without controls and using all observations (standard errors in parentheses). Including controls changes the estimates to 0.0017 (0.0026) and 0.0011 (0.0095), respectively. In the model for enrollment in a study program in science or technology, the effects are 0.0007 (0.0045) and 0.0014 (0.0035), respectively, in the model specification without controls, and 0.0002 (0.0011) and 0.0034 (0.0041), respectively, in the models including controls.

[20] For example, the p-value on a test of equal treatment effects on high school graduation across cohorts is equal to 0.59 in the model without controls. Likewise, the p-value on a test on whether the effect in 2003 is larger than the other years is equal to 0.31.

[21] The p-values of a test of equal treatment effects across cohorts are equal to 0.30 and 0.19 for enrollment in higher education and enrollment in science and technology, respectively, in the model specifications without controls.

Table 10
Effects of treatment dummy on educational attainment by cohorts.

	(1) Enrolling academic study track in high school	(2) Graduating high school within 5 years	(3) Enrollment in higher education	(4) Enrollment in higher education, science or technology
Panel A: 2002				
Without control variables	−0.0078	0.0040	−0.0053	0.0058**
	(0.0102)	(0.0071)	(0.0090)	(0.0027)
With control variables	−0.0045	0.0052	−0.0015	0.0058**
	(0.0069)	(0.0053)	(0.0056)	(0.0025)
Observations	49,534	49,534	49,534	49,534
Panel B: 2003				
Without control variables	0.0160	0.0137**	0.0171*	0.0058**
	(0.0108)	(0.0069)	(0.0095)	(0.0026)
With control variables	0.0127*	0.0099**	0.0130**	0.0053**
	(0.0072)	(0.0051)	(0.0058)	(0.0024)
Observations	51,185	51,185	51,185	51,185
Panel C: 2004				
Without control variables	−0.0074	0.0033	−0.0026	0.0001
	(0.0110)	(0.0069)	(0.0093)	(0.0024)
With control variables	−0.0028	0.0053	0.0010	0.0010
	(0.0075)	(0.0050)	(0.0058)	(0.0023)
Observations	54,983	54,983	54,983	54,983

Note. Each cell represents an independent regression with the treatment dummy as independent variable. The control variables include the socioeconomic characteristics described in Section 3.3 and presented in Appendix Table A1 and cohort fixed effects. Standard errors in parentheses are clustered at the compulsory school level.
*** Denotes significance at the 1% level.
** Denotes significance at the 5% level.
* Denotes significance at the 10% level.

is the intensive training part of the intervention that drives the findings, and not the high-stake testing itself.

5. Concluding remarks

We estimate the causal effect of training and preparation, followed immediately by a high-stake test, in mathematics relative to languages by exploiting random selection of students into external exit examination in different subjects in the Norwegian compulsory education. We find that treatment in terms of an intensive preparation period of 2–5 days in mathematics instead of languages increases the probabilities to graduate from high school, to enroll in higher education, and to enroll a study program in natural sciences or technology in higher education. Five days of intensive training is estimated to increase these probabilities by about 1.0, 0.8, and 0.5 percentage points, respectively. For all outcomes, the overall picture is that males appear to benefit somewhat more from the treatment than females, but the gender differences interact in complicated ways with students prior skills.

The treatment generally affects students across the whole ability distribution although at different margins. The positive effect on high school graduation is mostly related to improved progression for students initially enrolling in vocational study tracks in high school. These students have typically relative low prior skills. On the other hand, the positive effect on enrollment in science and technology programs in higher education seems to be restricted to students with relatively high skills in mathematics prior to the treatment. For these students, the treatment seems to have a positive short-term effect on grades in mathematics in high school.

The causal evidence is in accordance with simple descriptive associations between skills in different subjects measured by teacher set grades and the relevant outcomes. Taken together, the results suggest that mathematical skills are more important for broad measures of educational success than skills in languages.

One policy implication of these findings might be that all students should have an external exit examination in mathematics. If everybody knew that they should have their external examination in mathematics, they would, however, most likely prepare for that examination during the whole school year. The effect of such an institutional change is not possible to predict from the present study since student effort incentives for such a long period are likely to differ from that occurring in a short-term intensive preparation period just before the high-stake test. The present study indicates, however, that an intensive preparation period prior to mathematics exams is beneficial.

Taken together with the existing literature, our findings suggest that it would be beneficial for the students to be exposed to more training and testing in mathematics in school. That can be achieved by extending the mathematics courses or by increasing the incentives by making mathematical skills more high-stake. For example, putting more weight on mathematical skills in enrollment procedures of students in high schools and higher education institutions would make mathematics a more high-stake subject for the students.

Acknowledgments

We thank two anonymous referees, Flavio Cunha, Joshua Goodman, Erik Plug, Helena Skyt Nielsen, seminar and conference participants at European Economic Association in Gothenburg, European Society of Population Economics in Bern, European Association of Labor Economists in Bonn, Area Conference at CESifo, Norwegian University of Science and Technology, Statistics Norway, and a workshop in Oslo for insightful comments on earlier versions of this paper. Financial support from the Norwegian Research Council under The Programme for Educational Research (UTDANNING 2020) is gratefully acknowledged.

Appendix A

Appendix Table A1
Descriptive statistics for independent variables.

	Mean value
Female	0.493
First generation immigrant	0.034
Second generation immigrant	0.020
Parents' highest educational level is high school education	0.466
Parents' highest educational level is bachelor degree	0.289
Parents' highest educational level is master or PhD	0.103
Benefits due to disease before the age of 18	0.019
Benefits due to disabilities before the age of 18	0.025
One parent employed	0.241
Both parents employed	0.706
Parental income 2nd quartile	0.250
Parental income 3rd quartile	0.250
Parental income 4th quartile	0.250
Married parents	0.610
Divorced parents	0.126
Mobility	0.111
Mobility unknown	0.022
Born second quartile	0.267
Born third quartile	0.259
Born fourth quartile	0.228
Observations	155,702

Appendix Table A2
The relationship between teachers assessed grades in compulsory education and educational attainment.

Sample	(1)	(2)	(3)	(4)	(5)
	All	Females	Males	Low prior math skills	High prior math skills
Panel A: enrolling academic study track in high school					
Grade in mathematics	0.0236***	0.0274***	0.0178***	−0.0235***	−0.0097***
	(0.0022)	(0.0032)	(0.0030)	(0.0038)	(0.0037)
Grade in Norwegian language	0.0066***	−0.0091***	0.0232***	0.0123***	−0.0098***
	(0.0024)	(0.0033)	(0.0033)	(0.0030)	(0.0035)
Grade in English language	0.0251***	0.0035	0.0420***	0.0335***	0.0104***
	(0.0021)	(0.0031)	(0.0028)	(0.0028)	(0.0031)
Grade point average (GPA)	0.2018***	0.2394***	0.1697***	0.1755***	0.2707***
	(0.0038)	(0.0054)	(0.0050)	(0.0046)	(0.0059)
Panel B: graduating high school within five years					
Grade in mathematics	0.0159***	−0.0016	0.0306***	0.0720***	−0.0250***
	(0.0020)	(0.0027)	(0.0028)	(0.0044)	(0.0024)
Grade in Norwegian language	−0.0281***	−0.0351***	−0.0241***	−0.0252***	−0.0133***
	(0.0021)	(0.0029)	(0.0030)	(0.0032)	(0.0025)
Grade in English language	−0.0548***	−0.0568***	−0.0535***	−0.0593***	−0.0408***
	(0.0018)	(0.0026)	(0.0026)	(0.0029)	(0.0023)
Grade point average (GPA)	0.2915***	0.3101***	0.2803***	0.3155***	0.1876***
	(0.0032)	(0.0045)	(0.0043)	(0.0045)	(0.0045)
Panel C: enrollment in higher education					
Grade in mathematics	0.0501***	0.0400***	0.0570***	0.0031	0.0083**
	(0.0021)	(0.0030)	(0.0027)	(0.0033)	(0.0036)
Grade in Norwegian language	0.0044**	−0.0104 ***	0.0207***	0.0059**	−0.0063*
	(0.0022)	(0.0032)	(0.0028)	(0.0027)	(0.0033)
Grade in English language	−0.0033*	−0.0298***	0.0185***	−0.0029	−0.0082***
	(0.0019)	(0.0028)	(0.0025)	(0.0024)	(0.0030)
Grade point average (GPA)	0.2261***	0.2916***	0.1694***	0.2057***	0.2882***
	(0.0035)	(0.0047)	(0.0046)	(0.0042)	(0.0056)
Panel D: enrollment in higher education, science or technology					
Grade in mathematics	0.0500***	0.0400***	0.0600***	0.0087***	0.0694***
	(0.0013)	(0.0015)	(0.0020)	(0.0011)	(0.0029)
Grade in Norwegian language	−0.0051***	−0.0043***	−0.0076***	−0.0009	−0.0143***
	(0.0012)	(0.0015)	(0.0020)	(0.0009)	(0.0024)
Grade in English language,	0.0074***	0.0018	0.0116***	0.0022***	0.0119***
	(0.0011)	(0.0013)	(0.0017)	(0.0008)	(0.0021)
Grade point average (GPA)	−0.0014	−0.0017	0.0005	0.0089***	−0.0023
	(0.0018)	(0.0021)	(0.0028)	(0.0014)	(0.0041)
Socioeconomic characteristics	Yes	Yes	Yes	Yes	Yes
Cohort times school fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	155,702	76,770	78,932	80,038	75,664

Note. The socioeconomic characteristics included in the models as described in Section 3.3 and presented in Appendix Table A1 are included in all models. Standard errors in parentheses are clustered at the compulsory school level.
*** Denotes significance at the 1% level.
** Denotes significance at the 5% level.
* Denotes significance at the 10% level.

Appendix Table A3
Number of treatment days, balancing tests for subsamples.

Sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2002	2003	2004	Females	Males	Low prior math skills	High prior math skills
Female	−0.0168*	0.0099	0.0103	–	–	0.0153	−0.0128
First generation immigrant	−0.0275	−0.0361	0.0844	−0.0204	0.0336	0.0403	−0.0465
Second generation immigrant	−0.0280	−0.0454	0.1070	0.0470	−0.0046	0.0072	0.0424
Parents' highest educational level is high school education	−0.0164	0.1240***	−0.0142	0.0123	0.0494**	0.0333*	0.0272
Parents' highest educational level is bachelor degree	−0.0269	0.1110**	−0.0248	0.0041	0.0351	0.0333	0.0088
Parents' highest educational level is master or PhD	−0.0007	0.0691	−0.0723	−0.0253	0.0188	−0.0139	−0.00415
Benefits due to disease before the age of 18	0.0098	−0.1980**	0.0227	−0.0339	−0.0685	−0.0078	−0.1220**
Benefits due to disabilities before the age of 18	−0.0342	0.1000	−0.0020	0.0440	0.0111	−0.0187	0.0934*
One parent employed	0.0517**	−0.0218	−0.0071	0.0041	0.0066	0.0113	−0.0035
Both parents employed	0.0569**	−0.0110	0.0521	0.0310	0.0324	0.0264	0.0388
Parental income 2nd quartile	0.0009	−0.0255	−0.0410*	−0.0288	−0.0153	−0.0219	−0.0202
Parental income 3rd quartile	−0.0060	0.0179	−0.0256	0.0149	−0.0221	−0.0201	0.0168
Parental income 4th quartile	−0.0348	0.0525	−0.0735	−0.0274	−0.0112	−0.0259	−0.0082
Married parents	0.0188	0.0130	−0.0023	0.0012	0.0157	0.0227	−0.0108
Divorced parents	0.0094	−0.0234	−0.0331	−0.0044	−0.0290	0.0060	−0.0533**
Mobility	0.0270	0.0650	−0.0279	−0.0068	0.0477*	0.0229	0.0136
Mobility unknown	−0.0168	0.0073	−0.0564	0.0061	−0.0460	−0.0731	0.0481

Appendix Table A3 (*continued*)

Sample	(1) 2002	(2) 2003	(3) 2004	(4) Females	(5) Males	(6) Low prior math skills	(7) High prior math skills
Born second quartile	−0.0185	0.0013	−0.0036	0.0036	−0.0169	−0.0240	0.0102
Born third quartile	−0.0208	−0.0119	0.0198	0.0076	−0.0145	−0.0005	−0.0080
Born fourth quartile	−0.0088	0.0287	0.0139	0.0275	−0.0035	0.0139	0.0078
Cohort specific effects	–	–	–	Yes	Yes	Yes	Yes
Observations	49,534	51,185	54,983	76,770	78,932	80,038	75,664
Test of joint significance of the socioeconomic characteristics, p-value	0.403	0.262	0.470	0.589	0.305	0.446	0.160

Note. Standard errors in parentheses are clustered at the compulsory school level.
*** Denotes significance at the 1% level.
** Denotes significance at the 5% level.
* Denotes significance at the 10% level.

References

Altonji, J., 1995. The effect of high school curriculum on education and labor market outcomes. J. Hum. Resour. 30, 409–438.

Altonji, J., Pierret, C.R., 2001. Employer learning and statistical discrimination. Q. J. Econ. 116, 313–350.

Altonji, J., Blom, E., Meghir, C., 2012. Heterogeneity in human capital investments: high school curriculum college major, and careers. Annu. Rev. Econ. 4, 185–223.

Bettinger, E., 2012. Paying to learn: the effect of financial incentives on elementary school test scores. Rev. Econ. Statist. 94, 686–698.

Bishop, J.H., 1989. Is the test score decline responsible for the productivity growth decline? Am. Econ. Rev. 79, 178–197.

Carlsson, M., Dahl, G.B., Öckert, B., Rooth, D.-O., 2012. The Effect of Schooling on Cognitive Skills. Rev. Econ. Stat. (forthcoming).

Cortes, K.E., Goodman, J., Nomi, T., 2013. Intensive Math Instruction and Educational Attainment: Long-Run Impacts of Double-Dose Algebra. J. Hum. Resour. (forthcoming).

Goodman, J., 2012. The Labor of Division: Returns to Compulsory Math Coursework. Working Paper. Harvard University.

Haeck, C., Lefebvre, P., Merrigan, P., 2014. The distributional impacts of a universal school reform on mathematical achievements: a natural experiment from Canada. Econ. Educ. Rev. 41, 137–160.

Hansen, B., 2011. School Year Length and Student Performance: Quasi Experimental Evidence'. University of Oregon, (mimeo).

Hanushek, E.A., 2002. Publicly Provided Education'. In: Auerbach, A.J., Feldstein, M. (Eds.), Handbook of Public Economics. vol. 4. Elsevier, Amsterdam, pp. 2045–2141.

Hanushek, E.A., Woessmann, L., 2008. The role of cognitive skills in economic development. J. Econ. Lit. 46, 607–668.

Hanushek, E.A., Woessmann, L., 2012. Do better schools lead to more growth? cognitive skills, economic outcomes, and causation. J. Econ. Growth 17, 267–321.

Joensen, J.S., Nielsen, H.S., 2009. Is there a causal effect of high school math on labor market outcomes. J. Hum. Resour. 44, 171–198.

Joensen, J.S., Nielsen, H.S., 2014. Mathematics and Gender: Heterogeneity in Causes and Consequences. Econ J. (forthcoming).

Koedel, C., Tyhorst, E., 2012. Math skills and labor-market outcomes: evidence from a resume-based field experiment. Econ. Educ. Rev. 33, 131–140.

Krueger, A.B., 1999. Experimental estimates of education production functions. Q. J. Econ. 114, 497–532.

Lavy, V., 2010. Do Differences in School's Instruction Time Explain International Achievement Gaps in Math, Science, and Reading? Evidence from Developed and Developing Countries,. Econ. J. (forthcoming).

Lavy, V., 2013. Gender differences in market competitiveness in a real workplace: evidence from performance-based pay tournaments among teachers. Econ. J. 123, 540–573.

Levine, P.B., Zimmerman, D.J., 1995. The benefit of additional high-school math and science classes for young men and women. J. Bus. Econ. Stat. 13, 137–149.

Levitt, S.D., List, J.A., Neckermann, S., Sadoff, S., 2012. The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance'. Working Paper. NBER.

Marcotte, D.E., Hemelt, S.W., 2008. Unscheduled school closings and student performance. Educ. Finance Pol. 3, 316–338.

Murnane, R.J., Willett, J.B., Levy, F., 1995. The growing importance of cognitive skills. Rev. Econ. Statist. 77, 251–266.

Niederle, M., Vesterlund, L., 2010. Explaining the gender gap in math test scores: the role of competition. J. Econ. Perspect. 24, 129–144.

Oreopoulos, P., 2007. Do dropouts drop out too soon? wealth, health and happiness from compulsory schooling. J. Public Econ. 91, 2213–2229.

Rose, H., Betts, J.R., 2004. The effect of high school courses on earnings. Rev. Econ. Statist. 86, 497–513.