

# A full year COVID-19 crisis with interrupted learning and two school closures: The effects on learning growth and inequality in primary education

Citation for published version (APA):

Haelermans, C., Jacobs, M., van Vugt, L., Aarts, B., Abbink, H., Smeets, C., van der Velden, R., & van Wetten, S. (2021). *A full year COVID-19 crisis with interrupted learning and two school closures: The effects on learning growth and inequality in primary education*. ROA. ROA Research Memoranda No. 009 <https://doi.org/10.26481/umaror.2021009>

## Document status and date:

Published: 14/12/2021

## DOI:

[10.26481/umaror.2021009](https://doi.org/10.26481/umaror.2021009)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

Download date: 16 Apr. 2024



# A full year COVID-19 crisis with interrupted learning and two school closures: The effects on learning growth and inequality in primary education

Carla Haelermans  
Madelon Jacobs  
Lynn van Vugt  
Bas Aarts  
Henry Abbink  
Chayenne Smeets  
Rolf van der Velden  
Sanne van Wetten

## ROA Research Memorandum

ROA-RM-2021/9

Researchcentrum voor Onderwijs en Arbeidsmarkt | ROA  
*Research Centre for Education and the Labour Market | ROA*

# **A full year COVID-19 crisis with interrupted learning and two school closures: The effects on learning growth and inequality in primary education**

Carla Haelermans  
Madelon Jacobs  
Lynn van Vugt  
Bas Aarts  
Henry Abbink  
Chayenne Smeets  
Rolf van der Velden  
Sanne van Wetten

ROA-RM-2021/9  
December 2021

**Research Centre for Education and the Labour Market**  
Maastricht University  
P.O. Box 616, 6200 MD Maastricht, The Netherlands  
T +31 43 3883647

secretary-roa-sbe@maastrichtuniversity.nl  
www.roa.nl

ISSN: 2666-8823

## Abstract

### **A full year COVID-19 crisis with interrupted learning and two school closures: The effects on learning growth and inequality in primary education\***

After more than a year of COVID-19 crisis and the school closures that followed all around the world, the concerns about lower learning growth and exacerbated inequalities are larger than ever. In this paper, we use unique data to analyse how one full year of COVID-19 crisis in Dutch primary education has affected learning growth and pre-existing inequalities. We draw on a dataset that includes around 330,000 Dutch primary school students from about 1,600 schools, with standardized test scores for reading, spelling and mathematics, as well as rich (family) background information of the students. The results show a lower learning growth over a full year for all three domains, varying from 0.06 standard deviations for spelling to 0.12 for maths and 0.17 standard deviations for reading. Furthermore, we find that the lower learning growth is (much) larger for vulnerable students with a low socioeconomic background. This implies that pre-existing inequalities between students from different backgrounds have increased. These results are quite alarming and suggest that distance learning could not compensate for classroom teaching, although it prevented some damage that would have occurred if students had not enjoyed any formal education at all.

JEL classification: I20, I21, I24, I26, C90

Keywords: school closures, COVID-19 crisis, learning growth, inequality, socioeconomic status (SES), family situation, Primary Education

Carla Haelermans  
Maastricht University  
ROA  
P.O. Box 616  
NL-6200 MD Maastricht  
The Netherlands  
carla.haelermans@maastrichtuniversity.nl

Madelon Jacobs  
Maastricht University  
ROA  
P.O. Box 616  
NL-6200 MD Maastricht  
The Netherlands  
mce.jacobs@maastrichtuniversity.nl

Lynn van Vugt  
Maastricht University  
ROA  
P.O. Box 616  
NL-6200 MD Maastricht  
The Netherlands  
l.vanvugt@maastrichtuniversity.nl

Bas Aarts  
Maastricht University  
ROA  
P.O. Box 616  
NL-6200 MD Maastricht  
The Netherlands  
bas.aarts@maastrichtuniversity.nl

---

\* This paper has greatly benefited from the feedback from several anonymous reviewers, and participants of ESPE 2021, EEA 2021, ORD 2021, the Learn! & Education Lab conference in June 2021, ResearchEd Netherlands and the Royal Netherlands Academy of Arts and Sciences (KNAW) Webinar in June 2021.

Henry Abbink  
Maastricht University  
ROA  
P.O. Box 616  
NL-6200 MD Maastricht  
The Netherlands  
h.abbink@maastrichtuniversity.nl

Chayenne Smeets  
Maastricht University  
ROA  
P.O. Box 616  
NL-6200 MD Maastricht  
The Netherlands  
csrcm.smeets@maastrichtuniversity.nl

Rolf van der Velden  
Maastricht University  
ROA  
P.O. Box 616  
NL-6200 MD Maastricht  
The Netherlands  
r.vandervelden@maastrichtuniversity.nl

Sanne van Wetten  
Maastricht University  
ROA  
P.O. Box 616  
NL-6200 MD Maastricht  
The Netherlands  
s.vanwetten@maastrichtuniversity.nl

## 1. Introduction

Due to the COVID-19 crisis in 2020 and 2021, many countries had one or more periods of school closures and a year of interrupted learning with distance education. In some countries, schools have even been closed for over a year, causing more than 168 million students to follow only online education or even no education at all (UNICEF, 2021). This fuelled discussions on how this affected the learning growth of students in general, and of students from low socioeconomic backgrounds in particular, as they often do not have a home environment that stimulates learning and often do not have access to a computer for themselves. Theoretically, this relates to the framework of the educational production function, in which (time) investment and ability determine educational outcomes (see e.g. Cunha & Heckman, 2007).

Available empirical studies confirm large challenges related to distance learning, such as access to digital learning devices (e.g., Chetty et al., 2020) and the negative effect of distance schooling on performance and learning growth (e.g., Maldonado and De Witte, 2020; Engzell et al., 2021; Blainey et al., 2020). All these studies show that primary school students encountered serious deficits in their learning growth after a period of school closure(s), as also summarized in the early review study by Patrinos and Donnelly (2021).

Although schools have been striving to reduce the negative impact of the closure on students' learning growth as much as possible, for example by focusing mostly on the basic skills of maths and language, it is unclear to what extent they have succeeded in doing so. Furthermore, the question remains to what extent the school closures worsened pre-existing inequalities. Studies based on the first period of the COVID-19 crisis in 2020 for the UK (Andrew et al., 2020), the Netherlands (Bol, 2020; Engzell et al, 2021), Germany (Grewenig et al., 2021) and Belgium (Maldonado & De Witte 2020) give reason to believe inequalities have increased due to the school closures. For example, British middle-class parents spent more time on home schooling than parents from the working class (Andrew et al., 2020) and in the Netherlands, lower-educated parents felt less capable in helping their students with their schoolwork than high-educated parents (Bol, 2020).

More knowledge on the medium-term development of learning growth is necessary to get a grip on the damage of the global COVID-19 crisis on educational outcomes of students. It is important to know whether the lower learning growth that was observed after the first half year of COVID-19 crisis has worsened in the school year thereafter, in which the crisis continued, and what the state of affairs is after one full year of COVID-19 crisis. Furthermore, we have seen that the inequalities between students have grown in the first period of the

COVID-19 (Haelermans et al., 2021; Engzell et al., 2021) which leaves the question about the current inequalities after a full year of COVID-19 with two school closures. This knowledge is not only valuable for schools, but also for policymakers that must determine future steps to prevent aggravated inequalities. However, only a few studies (published in English) so far look into the effects on student performance after one year of COVID-19 crisis, in the UK (Blainey & Hannay, 2021) and Belgium (Gambi & De Witte, 2021).

We add to this literature by painting a picture of how a full year of COVID-19 crisis in Dutch primary education has affected the learning growth of students, and how it increased pre-existing inequalities<sup>1</sup>. We use unique large-scale data on standardized test scores in primary education, that were recently collected as part of the Netherlands Cohort Study on Education (see Haelermans et al., 2020 for more information on this dataset). This dataset includes about 450,000 students from approximately 1,600 Dutch primary schools (more than 25% of all Dutch primary schools). Test scores are available for national tests taken twice per year in the domains of reading, spelling and mathematics. Test scores resemble the learning growth from grade 1 till grade 6. A major benefit of the Netherlands Cohort Study on Education is that the dataset also contains extensive, rich and yearly information on students' (family) background and school characteristics, based on register data for the full population of Dutch students and schools. Contrary to most previous studies, we are able to investigate inequalities in learning growth between students *at the student level*, by using the detailed background information with respect to their gender and migration background, their socioeconomic status (SES), their household income and wealth, and parental education level and labour market position. We also add to the literature by discussing policy lessons that can be drawn from the results of our study.

Our results show that due to the COVID-19 crisis there was lower learning growth in the domains reading, spelling and maths in the COVID-19 year in comparison with a similar period in the years before. The results vary from -0.06 standard deviation for spelling to -0.12 for maths and -0.17 standard deviation for reading. Furthermore, we find that the drop in learning growth is (much) larger for vulnerable students that have a low socioeconomic status. These findings are confirmed when we look at specific parental outcomes such as education level, household income, household wealth and labour market position. This implies that pre-existing inequalities between students from different backgrounds have further increased.

---

1. In the Netherlands schools closed twice during to the COVID-19-crisis, with the first school closing being in the beginning of the crisis, from the 16<sup>th</sup> of March up to and including the 10<sup>th</sup> of May. The second school closure lasted from the 16<sup>th</sup> of December 2020 up to and including 8<sup>th</sup> of February 2021.

These results are quite alarming and suggest that distance learning could not compensate for classroom teaching, although it prevented some damage that would have occurred if students had not enjoyed any education at all.

The policy lessons that can be drawn from this are the following: 1) targeted interventions for vulnerable groups are necessary to close the widened achievement gaps by student background, 2) available money and resources should be disproportionately allocated to schools with a higher share of vulnerable students, 3) all interventions should be evidence-based, and students' access to these interventions should not depend on their parents' motivation or willingness to opt in, 4) the interventions should take intensity of treatment into account and consider focusing on indirect factors such as socio-emotional wellbeing, and 5) schools should only be closed as a very last resort.

## **2. Theoretical framework**

The theoretical framework underlying this study is the traditional human capital formation setting in which ability and investment are seen as the basic determinants of learning outcomes. When Cunha and Heckman (2007), for example, discuss the 'technology of skill formation', they do not dwell on the nature of the technology, but rather discuss the timing and the level of investments made. The investment part of the educational production function can be divided into an investment part taking place at school and an investment part taking place at home. In the following formula, we relate to those as  $i_s$  and  $i_h$ , respectively and put them alongside ability  $a$  as determinants of the human capital formation function  $f$  which produces a desired educational outcome  $Y$  in a particular schooling process:

$$\Delta Y = f(a; i_s, i_h) \tag{1}$$

In equation (1) we treat ability  $a$  as a compound measure incorporating both innate ability and skills acquired in the past, because we focus on one period of instruction and study the consequences of different investments in education during the period of instruction in times of COVID-19, while taking stock of earlier periods.

Theoretically, one can argue that the COVID-19 crisis and the consequent school closures have first of all directly influenced investment taking place at school, as schools were closed a

considerable time, thereby lowering investment taking place at school during COVID-19 ( $i_{sCOV}$ ).

$$\frac{\delta f}{\delta i_{sCOV}} < \frac{\delta f}{\delta i_s} \quad (2)$$

Because of distance learning and many schools trying to continue to teach online, one can argue that this has indirectly influenced (and increased) investment taking place at home during COVID-19 times ( $i_{hCOV}$ ).

$$\frac{\delta f}{\delta i_{hCOV}} > \frac{\delta f}{\delta i_h} \quad (3)$$

However, it is likely that the lower investment taking place at school is not fully compensated by the higher investment taking place at home, as online learning is considered less efficient, especially for primary school children.

$$\delta i_{sCOV} > \delta i_{hCOV} \quad (4)$$

Although the consequences of the COVID-19 with respect to lower investments taking place at school are the same for all students, this is most likely not the case for the change in investments taking place at home. Even in normal times, we know that time investment at home is unequally distributed over students from different socioeconomic backgrounds. Therefore, it is likely that the increase in time investment taking place at home because of COVID-19 is also unevenly taking place, and is higher for students from a high socioeconomic status (SES) in comparison with student from a low socioeconomic background.

$$\frac{\delta i_{sCOV}}{\delta i_{hCOV}} \text{ low\_SES} > \frac{\delta i_{sCOV}}{\delta i_{hCOV}} \text{ high\_SES} \quad (5)$$

Therefore, it is to be expected that the COVID-19 crisis decreases learning growth for all, but even more so for low-SES students, thereby further increasing the inequality between students from different backgrounds.

### 3. Literature overview

Ever since the start of the COVID-19 crisis, the question arose what the effect is of the crisis and the resulting school closures on students' learning development in primary schools. There are various studies from various countries looking into this question, although they vary in study setting, methodology and in the results they find. Whereas some studies point out that

the development of students is lagging behind by as much as one-third of a year (-0.23 SD) (Kogan & Lavertu, 2021), others argue there is no significant difference in learning growth at all due to lockdowns and school closures (Schwartz, 2021; Chénier et al., 2020; Gonzalez et al., 2020; Gore et al., 2021). Nevertheless, most studies across various countries found a significant negative impact of school closure(s) on learning growth (see e.g. Patrions & Donnelly for an early systematic review).

For instance, Kogan and Lavertu (2021) analyzed data on Ohio's Third-Grade reading assessments to study the effects of the COVID-19 crisis on students learning development. They show that in the school year 2019/2020, the average achievement on the tests was 0.23 standard deviations lower than before. A Swiss study found that primary school students learned twice as fast when they attended classroom education before the lockdown compared to online education during lockdown (Tomasik et al., 2020). In Germany, Schult, Mahler, Fauth and Lindner (2021) show that the lower learning growth was between 0.03 and 0.09 standard deviations across subjects and age groups. De Witte & Maldonado (2020) also found an overall lower learning growth in Belgium, due to the primary school closure. In another study, the same authors demonstrated a lower learning growth for all subjects with an average of 0.20 standard deviations for maths, 0.25 for Dutch language, 0.25 for French language, 0.27 for sciences, and 0.11 for social sciences, estimated at the school level (Maldonado & De Witte, 2021). Engzell et al. (2021) find that Dutch students also suffered lower learning growth of between 0.05 and 0.09 standard deviations across different age groups and different subjects due to the first school closure. They use data from 350,000 Dutch primary students collected right before and after the school closures, and conclude that on average the lower learning growth due to COVID-19 is about 3 percentile points (corresponding to about 8 weeks of education, which is roughly the duration of the school closures). Based on similar data, but with a smaller dataset of 886 students in grades 3-5 from 13 primary schools in one city in the Netherlands, Henrichs et al. (2020) find that the average lower learning growth is 2-3 months, and that the lower learning growth is worse for grade 3 students than for grade 4 or 5 students. Cohen et al. (2020) paint a similar picture, with larger lower learning growth for grade 3 and 5 based on data from all primary schools in Amsterdam. Lastly, Lek et al. (2020) from the Dutch testing organisation CITO, analyse approximately 20,000 students per grade level for grades 2 to 5 separately, and conclude that scores after the school closures are lower than in the years before, for all grade levels. With respect to special needs children, Bakx et al. (2020) found that Dutch children who have difficulties during physical education, i.e., autism or behavioural

problems, did not perform worse during the lockdowns. Nonetheless, these children did feel less connected with their teachers and classmates during the school closure (Bakx et al., 2020).

The studies mentioned so far were conducted rather at the beginning of the COVID-19 crisis. As mentioned before, only few studies have analysed the period of (over a) year into the pandemic. Blainey and Hannay (2021) show that decreases in learning growth from March 2020 to March 2021 across the UK were twice the size of the decrease seen in the first half year of COVID-19 for reading, maths and spelling. Gambi and De Witte (2021) show that there are serious attainment deficits, with effect sizes of  $-0.14$  SD for Dutch and  $-0.07$  SD for maths, at the student level, since start of the pandemic. One year after the school closures, they report an additional attainment deficit of  $-0.10$  SD for Dutch.

The previous results indicate that the crisis generally seems to have negatively affected learning growth of primary school students. But what about the differential impact on students with different socioeconomic backgrounds? Studies found that the negative effects on learning growth are higher for students from low-income families (Agostinelli et al., 2021; Andrew et al., 2020; Bayrakdar & Guveli, 2020; Bol, 2020; Cunha et al., 2020; Engzell et al., 2021; Maldonado & De Witte, 2021; van de Werfhorst, 2021; Van Lancker, 2021). Andrew et al. (2020) explained that parents from low-income families spent fewer hours a day homeschooling their children compared to parents from higher-income families. Low-income families faced more problems with working from home, as parents' job type demanded them to keep going to work in-person more often. Therefore, these parents had less time to support their children in schoolwork than higher-income parents, who often had office jobs that could be done from home (Agostinelli et al., 2020). These students often did not have a separate study space at home as well (Andrew et al., 2020). This argument is confirmed for Belgian students by Van Lancker (2021). In addition, during school closure low-income students could no longer benefit from the spill-over effects of classmates from higher-income families and were forced to spend more time with children from their neighbourhood who are often "lower-achieving peers" (Agostinelli et al., 2020, p.27).

Not just the parental income, but also their level of education, migration background, and employment status seems to be associated with the decline in learning growth experienced by students. Maldonado and De Witte (2021) concluded that the drop in learning growth of students from low educated mothers was higher compared to students with high-educated mothers. This might be because parents who have obtained higher education feel more capable of helping their children with schoolwork than less-educated parents (Bol, 2020). Bayrakdar

and Guveli (2020) found that the school closure in the UK has been especially hard for students with Pakistani- or Bangladeshi parents or low-income families. Furthermore, another UK study found that, if the father had lost his job due to COVID-19 crisis, the parents were 7.5 percentage points less likely to invest in additional learning resources for their children (Hupkau et al., 2020). This contributed to the growing inequality among students.

#### **4. Study context**

This study is based on standardized test data of students in Dutch primary education. In the Netherlands, primary education starts at age four in kindergarten, with school being compulsory as of age five. At age six, students enter grade 1, in which they formally start to learn how to read, write and do maths. Dutch primary schools are required to have a student administration system and must administer standardized tests every year in the period January/February (midterm test) and the period May/June (end-of-term or end test). Standardized tests at the national level are taken for three main subjects: reading, spelling and maths. These tests are usually administered from the end of grade 1 onwards, up until the midterm of grade 6. However, many schools decide not to take the midterm test in grade 6, since they take a (large) national standardized exit test. The results of this exit test combined with a formal advice from their teachers determines which track in secondary school fits them best.

Dutch primary schools had to close twice as a result of the COVID-19 crisis, where the first school closure started on March 16<sup>th</sup> and lasted up to and including May 10<sup>th</sup> (school closure of eight weeks, including two weeks of regular holiday). However, after May 10<sup>th</sup> students went back to school only half time, so an additional four weeks of disrupted teaching took place. Note that the first school closure had no consequences for the 2020 midterm tests that were taken right before the school closure.

From December 16<sup>th</sup> 2020 up to and including February 8<sup>th</sup> 2021 the schools closed for a second time for 7.5 weeks (including a period of two weeks of regular Christmas holidays). This closure had consequences for the midterm tests of 2021, as schools were still physically closed in the period of January and beginning of February. On average schools decided to delay testing the students by six weeks, to have a few weeks of regular school before the standardized tests were taken. Most students had taken the standardized test by the end of March 2021.

## 5. Materials and Methods

### 5.1 Sample

Our sample consists of approximately 330,000 students from ~1,600 primary schools.<sup>2</sup> This means more than 25% of the total number of primary schools and primary school students in the Netherlands are present in our sample. This is a relatively representative sample for Dutch primary education, with a slight overrepresentation of disadvantaged students and schools. However, correcting for this slight overrepresentation with inverse probability weights hardly changes the coefficients of our analyses (see robustness checks). We use data from the midterm test of the school year 2017/2018 up until the midterm test in school year 2020/2021. A more extensive description of the sample, including a description on how the data was collected, can be found in the Appendix Materials and Methods.

### 5.2 Variables of interest

#### 5.2.1 National test scores

In the Netherlands, most students from grade 1 until grade 6 take a national test that measures the proficiency in key domains. The tests are taken twice a year: the midterm in January/February and the end test in May/June. We use test results from three subject domains: reading, spelling and maths. The reading test assesses the student's ability to understand written texts, including both factual and literary content. The test in spelling asks students to write down a series of words (no verbs), demonstrating that they have learned the spelling rules. The test in maths contains both abstract problems and contextual problems that describe a concrete task. These tests are administered in school and last up to 60 minutes for each subject. Absolute test scores are translated into proficiency scores that are presented on a continuous scale, which shows the learning growth in a certain domain from grade 1 till grade 6. These proficiency scores are used to derive the learning growth.

From these proficiency scores the learning growth can be calculated by extracting the score on the midterm test in the previous year from the midterm test in the current year, for each of the three tested domain separately. The proficiency scores are not comparable across the three domains, therefore we standardize<sup>3</sup> the learning growth. To remove the influence of

- 
2. In our current sample and analyses, students at special primary education are not included as we do not have data about those students.
  3. We standardize the difference in proficiency scores per domain and year for the pre-COVID-19 cohorts and standardize the COVID-19 cohort based on the pooled average and standard deviation of the two previous cohorts.

outliers, the top and bottom 1% in terms of the learning growth scores per year per domain and per grade-year are excluded from the analyses. Note that unlike in some other countries like the UK, test scores were not adjusted for reasons of the COVID-19 crisis.

As described, the midterm test in the school year 2020/2021 was in many cases taken later than in previous years. In order to make a fair comparison with respect to the period over which we calculate learning growth, in comparison with the years before, we corrected<sup>4</sup> for this delay in our calculation of learning growth.

### 5.2.2 Treatment: one year COVID-19

Our main explanatory variable of interest is the COVID-19 crisis. Since we employ a semi-experimental difference-in-difference approach (discussed below) to estimate the effect of COVID-19, we conceptualise the ‘treatment’ through a dummy variable indicating the COVID-19 treatment year, i.e., the period between the midterm test 2019/2020 and the midterm test 2020/2021. The period between midterm test 2017/2018 and midterm test 2018/2019 and the period between midterm test 2018/2019 and midterm test 2019/2020 serve as ‘control’ years and have value 0 in this dummy (pre-COVID-19).

### 5.2.3 Student background

To get an idea of the extent to which different student characteristics correlate with the impact of the crisis on students’ learning growth, we add various student background characteristics to our model. An extensive description of the student background variables and how they are measured can be found in the Appendix Materials and Methods.

*Socioeconomic status (SES):* We use an aggregated socioeconomic status variable which is based upon an index from Berzofsky et al. (2014) and encompasses measures of parental education, household income and parental employment status. The SES measures the position of the student’s family in society. The values range from 0 to 7, the higher the SES, the higher the socioeconomic status of the parents. More information about all student background variables can be found in the Appendix Materials and Methods.

---

4. Although correcting for the number of weeks between tests may have undesired statistical implications (van de Werfhorst, 2021), and underlying assumptions such as linear growth of students during a school year, not correcting at all seems even worse, as six weeks is quite a large part of a school year in which students gain knowledge.

*Parental education level:* The parental education level is derived from both the mother and father's highest educational attainment. This variable reflects the highest attainment of the parents (whether that being the mother/father). We divide it into parents that are low-educated (11.6%), medium-educated (33.8%) and high-educated (54.6%).

*Household income:* The income level of the household is divided into three groups: low, medium and high. In our dataset, 22.8% of students have a low-income indication, 54.2% a medium-income indication, 23% a high-income indication and for 1% it is unknown.

*Household wealth:* The wealth level of the household is also divided into three groups, low (55.4%), medium (10.2%) and high (34.4%).

*Parental labour market position:* The labour market position of parents consists of both parents working (70.9%), only the father works (13.2%), only the mother works (3.8%), both parents don't work (5%) and unknown (7%). However, when looking separately at the father and mother we see that 90.4% of fathers and 78.7% of mothers are actively participating on the labour market, 7.8% of fathers and 12.9% of mothers receive benefits and 1.8% of fathers and 8.4% of mothers are inactive.

*Gender:* We use a dummy variable for gender, which has value 1 for girls (50%) and 0 for boys (50%).

*Migration background:* In our sample we have 72.3% students with a Dutch background, 19.4% have a non-western migration background and 8.2% have a western migration background. Of those students with a migration background, 4.8% is a first-generation migrant and 22.9% is a second-generation migrant.

### *5.3 Empirical strategy: Difference-in-differences*

To estimate the effect of the COVID-19 crisis on students' learning growth, we calculate the learning growth between two midterm tests within a one-year timespan. Resembling a difference-in-differences design, we distinguish between cohorts of students who took the national tests before the start of the COVID-19- crisis. Midterm tests taken in 2017/2018,

2018/2019 are used as pre-COVID-19 tests. The learning growth between the midterm test in 2019/2020 and the midterm test in 2020/2021 is used as the test period affected by COVID-19. We use data for all grades but exclude the midterm test taken in grade 6, as only a few schools expose their students to this test. Thus, we compare the learning growth of the COVID-19 exposed cohort (in the data labelled as the COVID-19 cohort) to the learning growth of students from the two previous cohorts using OLS regressions.

First, we estimate the effect of the COVID-19 crisis on the learning growth of students. Next, we include interaction effects between the treatment and the student-level background characteristics of household income and wealth, parental education level and labour market position, and student's gender and migration background. Note that we include the interaction between treatment and socioeconomic status in all other student background analyses. We do this to draw conclusions on the effect of other background characteristics on top of the effect for socioeconomic status. Lastly, we conduct several robustness checks on the main analysis for the three domains.

## **6. Results**

### *6.1 Main results*

Table 1 shows our main results, in which we look at standardized differences in learning growth and compare the year of the school closures with the previous years. Table 1 shows that the lower learning growth of one full year interrupted learning due to COVID-19 is 0.17 of a standard deviation for reading, 0.12 of a standard deviation for maths and 0.06 of a standard deviation for spelling. These main results are confirmed by several robustness checks, which are presented in Appendix Robustness Checks.

If we want to put that a bit more into perspective, we would like to recalculate that to a number of weeks in a school year. Earlier, Engzell et al. (2021) concluded that the learning loss after the first period of school closures in the Netherlands was around 8 weeks (exactly the number of weeks the schools were closed). They based their calculations on estimates from the literature that the average yearly learning growth in primary education lies between 0.3 and 0.6 standard deviation (Bloom et al., 2009), with a more recent study by the World Bank indicating that this number is around 0.4 SD per year (Azevedo et al., 2020). For fair comparison with the number of weeks estimated by Engzell et al., we base our analyses on the same literature and

assumptions. Given the standardized effects of 0.17, 0.12 and 0.06 SD, this implies a lower learning growth of around 17 weeks for reading, of around 12 weeks for maths and around 6 weeks for spelling for a full school year. If we were to assume that average yearly learning growth is at the lower (upper) end of the estimated range of 0.3 (0.6) SD per year, we would conclude that the lower learning growth was around 22 (11), 16 (8) and 8(4) weeks, for reading, maths and spelling respectively.

**Table 1.** Main results effects (in pre-COVID SDs) of one full year COVID-19 crisis in primary education on standardized learning growth

	Reading	Spelling	Maths
Treatment	-0.171*** (0.00616)	-0.0573*** (0.00527)	-0.124*** (0.00624)
Observations	425,461	400,293	450,602
R-squared	0.006	0.001	0.004
Clusters of schools	1665	1656	1667

Robust standard errors in parentheses. Coefficients are presented in terms of standard deviations.

\*\*\* p<0.01. \*\* p<0.05. \* p<0.1

### 6.2 Inequality in the lower learning growth

In Table 2, we show the results by socioeconomic status (SES) of the student. In this table, the constant resembles the learning growth of low-SES students in the pre-COVID-19 years. In the pre-COVID-19 years, high-SES students had higher learning growth than medium-SES and low-SES students in the domains of reading and maths. The treatment indicates the difference in learning growth before and during COVID-19 for low-SES students. We observe a 0.17 SD lower learning growth for low-SES students for reading, a 0.18 SD lower for maths, and around 0.08 SD for spelling. In the pre-COVID-19 years, medium-SES students had a little lower learning growth than low-SES students (but this difference is only significant for reading), whereas high-SES students had more learning growth than both the other groups, except for spelling. For maths and spelling the negative effects of the full year of COVID-19 crisis are largest for low-SES students. Medium- and high-SES students have a lower learning growth of around 0.06 and 0.05 SD for spelling, and around 0.12 and 0.11 SD for maths, respectively. There are no significant differences in the interaction effect for reading. The total effects are

best observed in Figure 1. The lower learning growth was lowest for the high-SES group, implying that the already existing inequality has increased.

**Table 2.** Differential effects (in pre-COVID SDs) of one full year COVID-19 crisis in primary education on standardized learning growth – by socioeconomic status (SES)

	Reading	Spelling	Maths
Treatment	-0.175*** (0.0117)	-0.0830*** (0.0105)	-0.182*** (0.0119)
Medium SES	0.00936 (0.00638)	-0.0169*** (0.00559)	-0.00595 (0.00680)
High SES	0.0581*** (0.00750)	-0.000550 (0.00639)	0.0180** (0.00791)
Treatment x Medium SES	-0.00514 (0.0115)	0.0256*** (0.00982)	0.0600*** (0.0112)
Treatment x High SES	0.0217 (0.0134)	0.0347*** (0.0111)	0.0746*** (0.0127)
Constant	-0.0244*** (0.00672)	-0.260*** (0.00616)	-0.177*** (0.00772)
Observations	425,461	400,293	450,602
R-squared	0.007	0.002	0.005
Clusters of schools	1665	1656	1667

Robust standard errors in parentheses. Coefficients are presented in terms of standard deviations.

\*\*\* p<0.01. \*\* p<0.05. \* p<0.1

**Figure 1.** Effects of one full year COVID-19 crisis in primary education on learning growth – by socioeconomic status (SES)

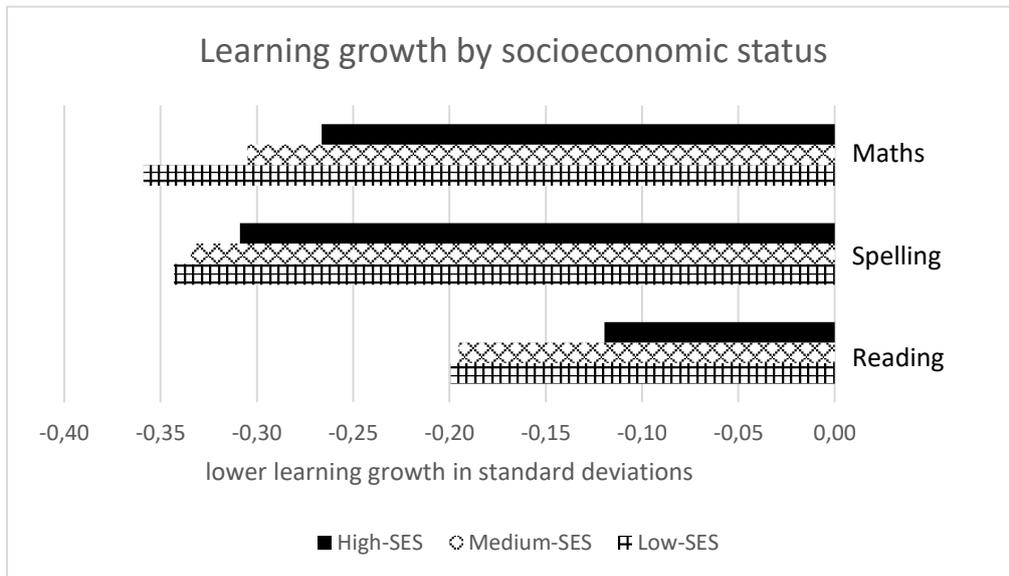
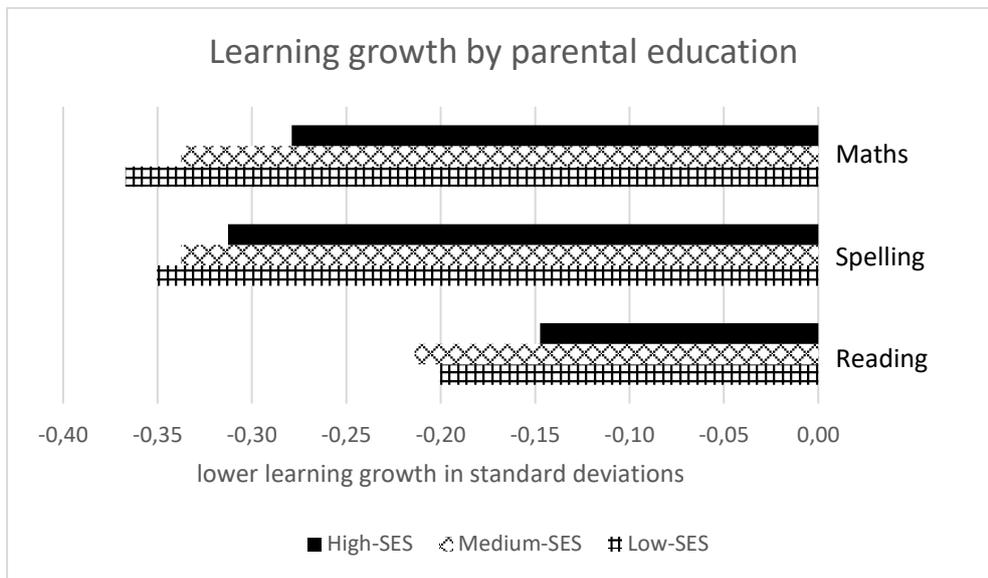
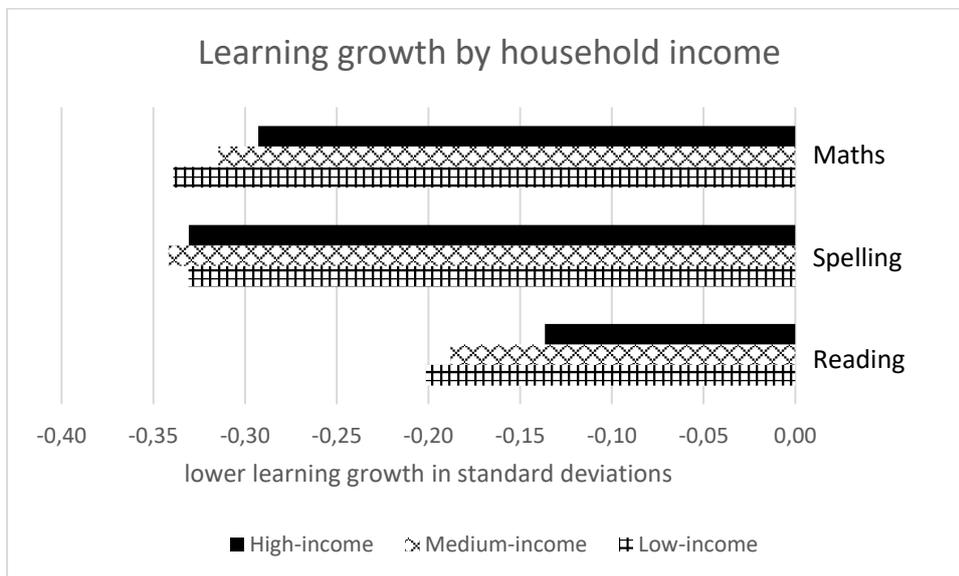


Figure 2 to 7 show the effects of one full year COVID-19 crisis on learning growth, differentiated by parental education, household income, household wealth, gender and migration background. The results show that disadvantaged students are more severely affected than their peers. Figure 2 shows that students from low-educated parents have the largest lower learning growth for maths and spelling, whereas for reading students from middle-educated parents have the lowest learning growth (although for spelling and reading the effects are not significantly different from the lower-educated parents group). Figures 3 and 4 show a similar picture for household income and household wealth. Figure 5 shows the effects by parental labour market position. Here we see that for maths and reading the lower learning growth is the largest for students of which neither of the parents have a job and students where only mother has a job. For spelling we do not see a clear pattern. Furthermore, in Figure 6 we see that the effects by gender significantly differ by domain. We see only a small difference for reading, while for spelling for girls the difference in learning growth is larger, and for maths boys have a larger difference. Figure 7 shows migration background, where we see that rather the students without migration background have lower learning growth due to the COVID-19 crisis, although the interaction term in the regression is not significant of any of the domains. The underlying regression results of these figures, as well as significance levels for each coefficient, can be found in the Appendix Results.

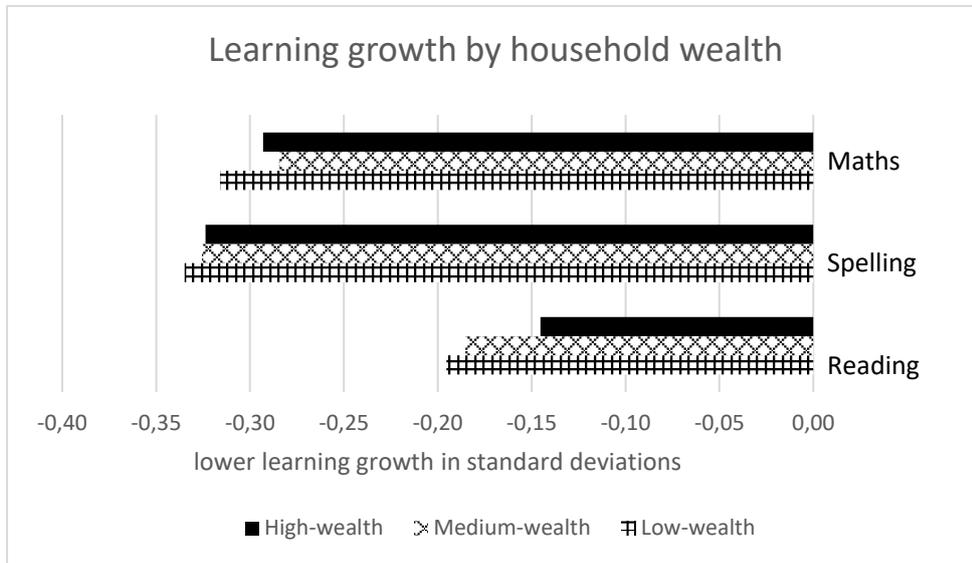
**Figure 2.** Effects of one full year COVID-19 crisis in primary education on learning growth – by parental education level



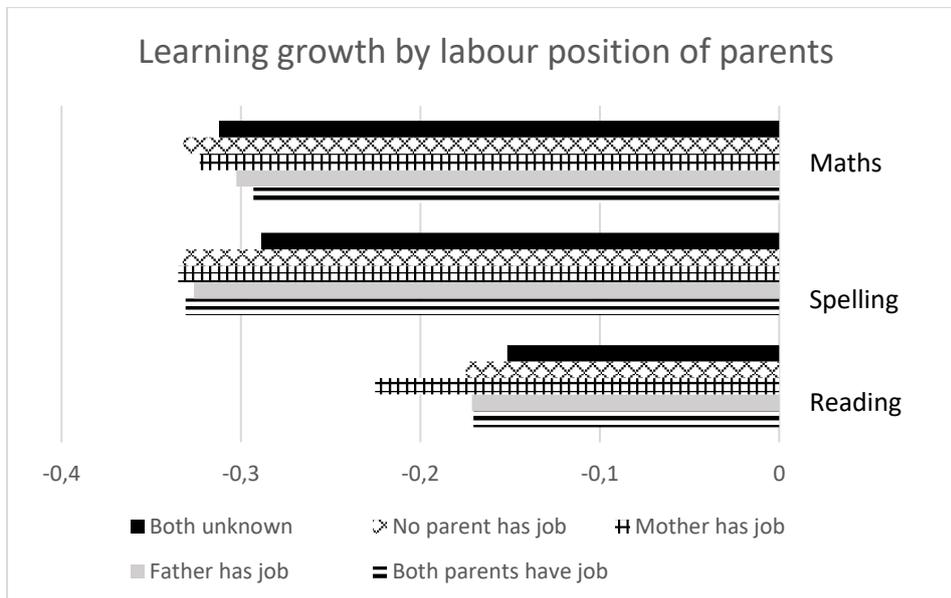
**Figure 3.** Effects of one full year COVID-19 crisis in primary education on learning growth – by household income



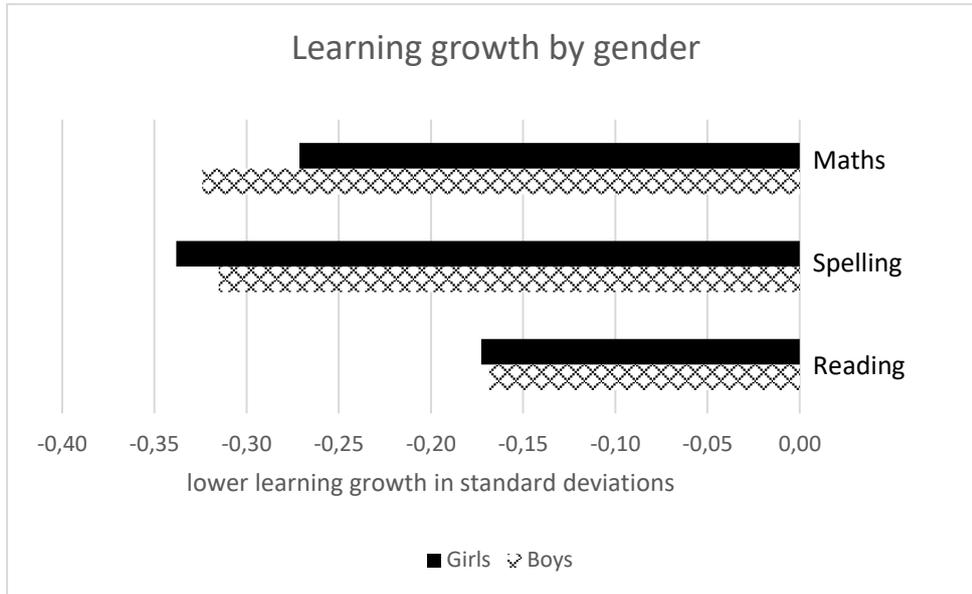
**Figure 4.** Effects of one full year COVID-19 crisis in primary education on learning growth – by household wealth



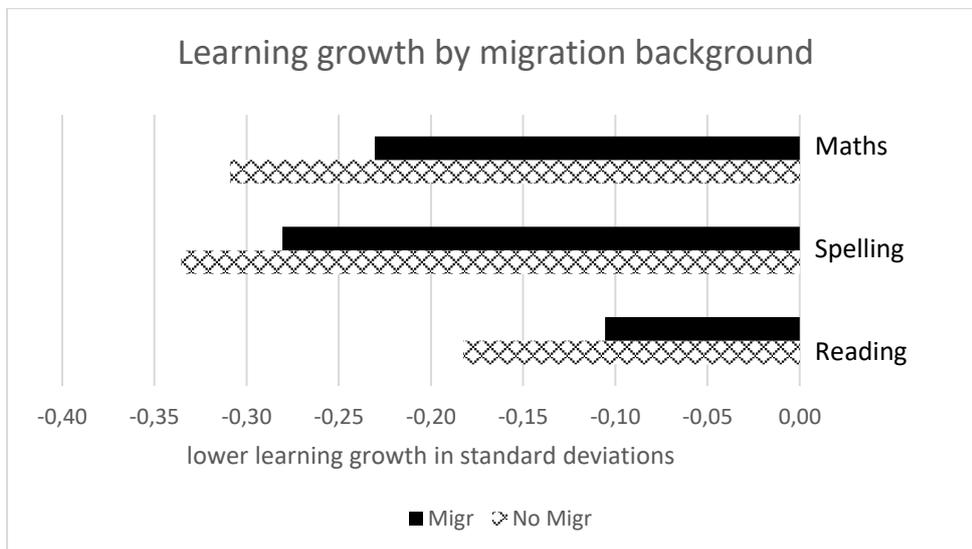
**Figure 5.** Effects of one full year COVID-19 crisis in primary education on learning growth – by labour position parents



**Figure 6.** Effects of one full year COVID-19 crisis in primary education on learning growth – by gender



**Figure 7.** Effects of one full year COVID-19 crisis in primary education on learning growth – by migration background



## 7. Conclusion & discussion

After more than a year of COVID-19 crisis and the school closures that followed, the concerns about decreased learning growth and increased inequalities are larger than ever. To develop targeted policies, it is highly relevant to know what the total lower learning growth was, and

whether this lower learning growth differentially affected various social groups. In this paper we have analysed how one full year of COVID-19 crisis in education (and two school closures) has affected the learning growth of primary school students in the Netherlands. We compared the learning growth of the COVID-19-exposed cohort to the learning growth of students from the two previous cohorts. Using interaction terms, we investigated the inequality in the effects between students with respect to household income, household wealth, parental education level and labour market position, and students' gender and migration background.

The results show that the lower learning growth in reading, spelling and maths over the full COVID-19 year varies between 0.06 and 0.17 of a standard deviation. These lower learning growth is (much) larger for disadvantaged students than for other students, implying that the pre-existing inequalities between students from different backgrounds have increased. For gender, however, the inequality has decreased and for migration background the results are exactly the other way around than expected.

The overall effect-sizes found in this study suggest a lower learning growth of around 17 weeks for reading, of around 12 weeks for maths and around 6 weeks for spelling, based on a full school year. And for disadvantaged students the lower learning growth is even larger. The overall numbers imply that students' learning growth during the 15.5 weeks of school closures has taken place at a rate of 61 percent at best (for spelling), but also that students hardly learned anything for reading in the periods the schools were closed and the weeks of interrupted learning. On average this is slightly higher than what Engzell et al. (2021) found, which could be explained by the second school closure. Engzell et al. (2021) concluded that the learning loss in weeks was about the same amount of time as the schools had been closed (including two holiday weeks). In our analyses, the average number of weeks delay in learning growth are lower than the number of weeks the schools have been closed (15.5 weeks including holiday) in this first full year of COVID-19 crisis.

In relation to previous literature, the effects we find are somewhat in the middle range of standardized effect scores that authors report, and quite similar to the results found in Belgium, when estimated at the individual level by Gambi & De Witte (2021). Studies that find (much) larger standardized effects are performed in the United States (Kogan & Lavertu, 2021), the United Kingdom (Blainey & Hannay, 2021) and in Belgium (Maldonado & De Witte, 2021). Studies that find (much) smaller effects are based in Germany (Schult et al., 2021) and the Netherlands (Engzell et al., 2021). However, it is important to note that most studies that

we now relate our findings to analysed only the first period of the COVID-19 crisis and not a whole year. This seems especially relevant with respect to the effect reported by Engzell et al. (2021) as they use the same test scores from the Netherlands. However, it is important to notice is that they estimate their effects in percentile points, and recalculate their effects from percentile points to standard deviations using a conversion method. They do not standardize their data based on the pre-COVID period, such as we do in our analyses.

Furthermore, whereas one could argue that the educational (and testing) system in the United States and also in the United Kingdom is very different from the Netherlands, this argument does not hold for Belgium and Germany. There are a couple of potential reasons why the effects are much larger in the studies of Maldonado and De Witte (2021). First of all the age of the students: they look at grade 6 students, whereas we analyse up until grade 5 and take the average over the grade levels. However, in additional analyses of our data in our (Dutch) factsheets), we also see that the negative effects are largest for students in grade 5, especially for math (Haelermans et al., 2021). Also, the study of Maldonado & De Witte estimates the effects at the school level. In the follow-up study by De Witte & Gambi (2021) they also estimate effects at the individual level for Belgium, and they are indeed a lot smaller than at the school level. Another potential reason is that Maldonado & De Witte use absolute test scores, whereas we focus on learning growth between two periods. The latter also holds for Schult et al. (2021), who estimate a much smaller effect across subjects and age groups in Germany. Furthermore, another potential reason as to why Schultz et al. find smaller results could be due to the fact that they only analyse the tests after the summer break, instead of directly after re-opening the schools.

The results presented in this study are quite alarming and show that in-class education really matters, specifically for the most vulnerable groups. Although distance learning may have prevented further damage, it could not compensate for classroom teaching. These findings therefore call for an (inter)national focus on reducing the lower learning growth of students, in general, but most importantly for students with lower-educated parents, a lower household income and from a single-parent household. It is worrisome, and unfortunately not unlikely, that the increased inequalities in learning loss due to the pandemic may lead to long-lasting inequalities (Kautz et al., 2014; Psacharopoulos et al., 2021), deepening the gap in adult outcomes between these groups in the population. This very much stresses the need for targeted interventions to reduce the current inequalities in learning loss caused by the pandemic.

However, targeted interventions are not enough, as there were already some interventions in place in the second part of the Covid-19 crisis (the academic year 2020/21). These interventions were mostly remedial teaching programs that were specifically targeted at the lower-performing and more vulnerable children. Since inequalities have remained the same since the beginning of the crisis, this seems to not have helped reducing these inequalities. On the other hand, we cannot rule out that these targeted interventions are part of the reason that more vulnerable students did not continue to grow worse relative to their less vulnerable peers over the course of the pandemic. This question is currently being studied by Jacobs et al. (2021).

Therefore, there are several policy implications, both at the country and the school level, that can be drawn based on this: 1) targeted interventions for vulnerable groups are necessary to close the widened achievement gaps by student background; 2) available money and resources should be disproportionately allocated to schools with a higher share of vulnerable students (low socioeconomic status, low-educated parents, and/or low household income); 3) all targeted interventions should be evidence-based to avoid spending a lot of money on educational programs that may not yield results, and should be group specific for the vulnerable groups. Examples of successful evidence-based interventions are High Dosage Tutoring (daily intense sessions in very small groups of students with a tutor) (Nickow et al., 2020) and summer schools (Xie et al., 2021). Furthermore, students' access to these interventions should not depend on their parents' motivation or willingness to opt in, as this crisis shows the large differences between parents abilities and time (e.g. Bol, 2020); 4) the interventions should take intensity of treatment into account; i.e. when schools already spend a considerable amount of time on maths and language, only one more hour per week may not make a significant difference; 5) consider focusing more on indirect factors such as the socio-emotional wellbeing of children. Given the large share of time already spent on basic skills as well as the fact that socio-emotional skills in general explain a substantial part of the differences in direct outcomes such as maths and language (e.g. Cunha & Heckman, 2007; Borghans et al., 2016), it may be well worth the investment; and 6) in the event of another crisis, or in case the current Covid-19 pandemic continues, schools should only be closed as a very last resort to avoid further inequalities.

Although this study presents important results for the current levels of reading, spelling and maths skills in primary education, some caveats are in place. It is important to realize that in this study we did not have any information about other relevant skills, such as other cognitive

skills and socio-emotional skills. Other domains may have suffered more if school focused all attention to the skills in the aforementioned key domains. And the effect of the school closure may have been stronger for students who are less emotionally stable or who have less self-regulatory skills. These are important questions that should be addressed in further research.

Overall, this study is an important contribution to the literature on the short-term effects of the COVID-19 crisis on learning growth and inequality in primary education.

## References

Agostinelli, F., Doepke, M., Sorrenti, G., & Zilibotti, F. (2020). *When the Great Equalizer Shuts Down: Schools, Peers, and Parents in Pandemic Times*. Insitute of Labor Economics. <http://ftp.iza.org/dp13965.pdf>

Andrew, A., Cattan, S., Costa-Dias, M., Farquharson, C., Kraftman, L., Krutikova, S., Phimister, A., & Sevilla, A. (2020). *Learning during the lockdown: real-time data on children's experiences during home learning*. Institute for Fiscal Studies. [https://ifs.org.uk/uploads/Edited\\_Final-BN288%20Learning%20during%20the%20lockdown.pdf](https://ifs.org.uk/uploads/Edited_Final-BN288%20Learning%20during%20the%20lockdown.pdf)

Azevedo, J. P., Hasan, A., Goldemberg, D., Iqbal, S. A., & Geven, K. (2020) Simulating the potential impacts of COVID-19 school closures on schooling and learning outcomes: A set of global estimates. *World Bank Policy Research Working Paper* (9284). <https://openknowledge.worldbank.org/handle/10986/33945>. Accessed 3 November 2021.

Bakx, A., Samsen-Bronsveld, E., Hornstra, L., Diepstraten, I., & Speetjens, P. (2020). *Het welbevinden en motivatie van kinderen in het basisonderwijs tijdens onderwijs op school en afstandsonderwijs: De visie van ouders*. 36. <https://www.point013.nl/wp-content/uploads/2020/09/PO-Rapportage-Afstandsonderwijs.pdf>

Bayrakdar, S., & Guveli, A. (2020). *Inequalities in home learning and schools' provision of distance teaching during school closure of COVID-19 lockdown in the UK* (Working Paper Nr. 2020–09). ISER Working Paper Series. <https://www.econstor.eu/handle/10419/227790>

Berzofsky, M., Smiley-McDonald, H., Moore, A., & Krebs, C. (2014). Measuring socioeconomic status (SES) in the NCVS: Background, options, and recommendations. Retrieved from Bureau of Justice Statistics website: [https://www.bjs.gov/content/pub/pdf/Measuring\\_SES-Paper\\_authorship\\_corrected.pdf](https://www.bjs.gov/content/pub/pdf/Measuring_SES-Paper_authorship_corrected.pdf)

Blainey, K., Hiorns, C., Hannay, T. (2020) *The Impact of Lockdown on Children's Education: a Nationwide Analysis*. RS Assessment from Hodder Education | SchoolDash.

Blainey, K. & T. Hannay (2021). The impact of school closures on spring 2021 attainment. White paper. [https://www.risingstarsuk.com/rsassessment/whitepapers?utm\\_source=organic&utm\\_medium=sharing&utm\\_campaign=whitepaper\\_21](https://www.risingstarsuk.com/rsassessment/whitepapers?utm_source=organic&utm_medium=sharing&utm_campaign=whitepaper_21).

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Res. Edu. Effect*, 1, 289–328.

Bol, T. (2020). Inequality in home schooling during the Corona crisis in the Netherlands. First results from the LISS Panel. SocArXiv. <https://doi.org/10.31235/osf.io/hf32q>

Borghans, L., Duckworth, A., Heckman, J., Weel, B. ter (2008). The economics and psychology of personality traits. *Journal of Human Resources* 43(4), 972-1059.

Chénier, M. A., Maldonado, J. E., & De Witte, K. (2020). *De invloed van extra schoolsluitingsdagen op gestandaardiseerde testcores in het zesde leerjaar*. <https://feb.kuleuven.be/drc/LEER/in-the-press/invloed-van-extra-schoolsluitingsdagen-26-01-2020.pdf>

Chetty, R., The Opportunity Insights Team, Friedman, J. N., Hendren, N., & Stepner, M. (2020, June). *The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data*. NBER. <https://doi.org/10.3386/w27431>

Cohen, L. van Spijker, F., T. Brils, L. de Graaf, M. van der Wouden, S. Ahamiane, W. Boterman, Z. Attar (2020). *De Staat van het Amsterdamse primair onderwijs 2020*.

Cunha, F., & J.J. Heckman (2007). “The Technology of Skill Formation,” *American Economic Review*, 97(2):31- 47.

Cunha, J., Silva, C., Guimarães, A., Sousa, P., Vieira, C., Lopes, D., & Rosário, P. (2021). No Children Should Be Left Behind During COVID-19 Pandemic: Description, Potential Reach, and Participants' Perspectives of a Project Through Radio and Letters to Promote Self-Regulatory Competences in Elementary School. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.647708>

De Witte, K., & Maldonado, J. (2020). *De effecten van de COVID-19 crisis en het sluiten van scholen op leerlingprestaties en onderwijsongelijkheid* (Vol. 2020, Issue 181, pp. 1–6). KU Leuven. Faculteit Economie en Bedrijfswetenschappen. Departement Economie. <https://lirias.kuleuven.be/3189076?limo=0>

Engzell, P., Frey, A., & Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 crisis. *Proceedings of the National Academy of Sciences*, 118(17). <https://doi.org/10.1073/pnas.2022376118>

Gambi, L. & De Witte, K. (2021) The resiliency of school outcomes after the COVID-19 pandemic. Standardised test scores and inequality one year after long term school closures. KU Leuven Department of Economics Discussion Paper Series 21.12

Grevenig, E., Lergetporer, P., Werner, K., Woessmann, L., & Zierow, L. (2021). COVID-19 and educational inequality: How school closures affect low- and high-achieving students. *European Economic Review*, 140(1) 10390. <https://doi.org/10.1016/j.euroecorev.2021.103920>

Gonzalez, T., De La Rubia, M. A., Hincz, K. P., Comas-Lopez, M., Subirats, L., Fort, S., & Sacha, G. M. (2020). Influence of COVID-19 confinement on students' performance in higher education. *PloS one*, 15(10), <https://doi.org/10.1371/journal.pone.0239490>

Gore, J., Fray, L., Miller, A., Harris, J., & Taggart, W. (2021). The impact of COVID-19 on student learning in New South Wales primary schools: An empirical study. *The Australian Educational Researcher*. <https://doi.org/10.1007/s13384-021-00436-w>

Haelermans, C., Huijgen, T., Jacobs, M., Levels, M., van der Velden, R., van Vugt, L., & van Wetten, S. (2020). Using Data to Advance Educational Research, Policy, and Practice: Design, Content, and Research Potential of the Netherlands Cohort Study on Education. *European Sociological Review*, 36(4), 643–662. <https://doi.org/10.1093/esr/jcaa027>

Haelermans, C., Korhals, R., Jacobs, M., de Leeuw, S., Vermeulen, S., van Vugt, L., Aarts, B., Prokic-Breuer, T., van der Velden, R., van Wetten, S., & de Wolf, I. (2021). Sharp increase in inequality in times of the COVID-19 pandemic. *Working paper*.

Haelermans, C., van der Velden, R., Aarts, B., Bijlsma, I., Jacobs, M., Smeets, C., van Vugt, L., & van Wetten, S. (2021d). Leergroei deels hersteld, behalve bij begrijpend lezen. Nationaal Regieorgaan Onderwijsonderzoek. NCO Factsheet No. 5.

Henrichs, L., L. Hornstra, S. Polderdijk & T. Schuurman (2020). *Lessen uit de lockdown*. Didactief, 31 December 2020.

Hupkau, C., Isphording, I., Machin, S., & Ruiz-Valenzuela, J. (2020). *Labour Market Shocks during the COVID-19 Pandemic, Inequalities and Child Outcomes* (Nr. 15). Centre for Economic Performance. <https://cep.lse.ac.uk/pubs/download/cepcovid-19-015.pdf>

Jacobs, M., Haelermans, C., & Meeter, M. (2021). Effecten van IOP op leergroei in het PO. Factsheet.

Kautz, T., Heckman, J.J., Diris, R., ter Weel, B. & Borghans, L. (2014), Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success National Bureau of Economic Research.

Kogan, V., & Lavertu, S. (2021). *The COVID-19 Pandemic and Student Achievement on Ohio's Third-Grade English Language Arts Assessment*. John Glenn College of Public Affairs. [http://glenn.osu.edu/educational-governance/reports/reports-attributes/ODE\\_ThirdGradeELA\\_KL\\_1-27-2021.pdf](http://glenn.osu.edu/educational-governance/reports/reports-attributes/ODE_ThirdGradeELA_KL_1-27-2021.pdf)

Lancker, W. van (2020). *Schoolachterstand tijdens de coronastop*. SAMPOL Samenleving & Politiek. <https://sampol.be/2020/04/school-achterstand-tijdens-de-corona-stop>

Lek, K. R. Feskens, & J. Koning (2020). *Het effect van afstandsonderwijs op leerresultaten in het PO*. CITO, November 2020.

Nickow, A., Oreopoulos, P., & Quan, V. (2020). The impressive effects of tutoring on pre-K-12 learning: A systematic review and meta-analysis of the experimental evidence. NBER Working Paper no. 27476.

Notenboom, A., Aartsen, A., & Lit, S. (2017). *Tussendoelen rekenen-wiskunde voor het primair onderwijs*. SLO.

Maldonado, J. E., & De Witte, K. (2021). The effect of school closures on standardised student test outcomes. *British Educational Research Journal* (in print). <https://doi.org/10.1002/berj.3754>

Patrinos, H. & Donnelly, R. (2021). Learning Loss During COVID-19: An Early Systematic Review.

Report. ISSN: 2693-5015.

Psacharopoulos, G., Collis, V., Patrinos, H. A., & Vegas, E. (2020). Lost Wages: The COVID-19 Cost of School Closures. SSRN Electronic Journal.

Schult, J., Mahler, N., Fauth, B., & Lindner, M. A. (2021). *Did Students Learn Less During the COVID-19 Pandemic? Reading and Mathematics Competencies Before and After the First Pandemic Wave*. <https://doi.org/10.31234/osf.io/pqtgf>

Schwartz, S. (2021). *"Learning Loss, in General, Is a Misnomer": Study Shows Kids Made Progress During COVID-19*. Education Week. <https://www.edweek.org/leadership/learning-loss-in-general-is-a-misnomer-study-shows-kids-made-progress-during-covid-19/2021/04>

Tomasik, M. J., Helbling, L. A., & Moser, U. (2020). Educational gains of in-person vs. distance learning in primary and secondary schools: A natural experiment during the COVID-19 pandemic school closures in Switzerland. *International Journal of Psychology*. <https://doi.org/10.1002/ijop.12728>

UNICEF. (2021). *COVID-19: Schools for more than 168 million children globally have been completely closed for almost a full year, says UNICEF*. <https://www.unicef.org/press-releases/schools-more-168-million-children-globally-have-been-completely-closed>

Werfhorst, H.G. van de (2021). Inequality in learning is a major concern after school closures. *PNAS May 18, 2021 118* (20) e2105243118 <https://doi.org/10.1073/pnas.2105243118>

Xie, C., Neitzel, A., Cheung, A., & Slavin, R. E. (2021). The effects of summer programs on K-12 students' reading and mathematics achievement: A meta-analysis. Baltimore, MD: Center for Research and Reform in Education, Johns Hopkins University.

## **8. Appendix Materials and Methods**

### *8.1 Dataset*

The underlying dataset on the full population of students in all schools in Dutch primary education is obtained through the Netherlands Cohort Study on Education (in Dutch abbreviated as NCO [Nationaal Cohortonderzoek Onderwijs] (Haelermans et al., 2020)). The Netherlands Cohort Study on Education uses longitudinal register data on track placement of cohorts of students in primary and secondary education. The dataset is based at Statistics Netherlands, where it is combined with school administrative data on students' performance, such as the standardized tests scores used in our study. The Netherlands Cohort Study on Education currently consists of three pillars. The first pillar maps students' pathways through education and their trajectory into tertiary education and combines this with very rich and extensive information on their (social and family) background using register data from Statistics Netherlands. For each student, information on age, gender, country of origin, marital status of the parents, household information, socioeconomic status (SES) of both student and his/her parents, and regional variables are available. Parental variables include variables on their highest obtained educational level, SES, working status, income, and wealth. The funding of Dutch primary schools is partly based on the socioeconomic background of the school's population. The NCO data set therefore also holds information on whether or not and to what extent a school population consists of students with disadvantageous socioeconomic status.

The second pillar consists of additional information at the school level. This information is available through the Dutch Ministry of Education and the Dutch Inspectorate of Education. It consists of data on e.g. school size, the level of urbanization of the location of the school and its (religious/secular) denomination. In the future this data will be complemented with more detailed information on the educational process and school quality.

A third pillar consists of microdata on student performance from school administrative systems. Primary schools in the Netherlands are required to monitor their students' progress in domains such as reading, spelling and maths. Most schools use a national standardized test for this, which makes it possible to have information on the development of students' performance between the age of 8 and 12. These primary school standardized tests are used for the paper at hand.

## 8.2 Data collection standardized tests

In the Netherlands, students take standardized tests throughout grades 1 to grade 6 in primary education. Primary schools are obliged to administer the standardized tests students take every year. These standardized tests come from different test suppliers, with the largest supplier in the Netherlands being CITO, with which we collaborated for this paper. Schools use administration systems to store the information about the standardized test scores, through which the data are collected. Three of these administration systems, called CITO-LOVS, ParnasSys and ESIS exported the data on standardized test scores from school year 2013/2014 onwards as part of the Netherlands Cohort Study on Education project. With permission of the schools, the administration system exports the data on the standardized test scores to Statistics Netherlands, who pseudonymized the student-id and school-id. Before any data was exported, parents were given the opportunity to object against export of their child(ren)'s data. Data was not exported from those students whose parents objected.

The raw data consist of ~450,000 unique students per school year from ~1,600 primary schools (approximately 25% of all Dutch primary schools) from the period 2013/2014 until 2020/2021 (see Table A8.1). In these data, we have information available from standardized tests written before and after the first and second lockdown and the corresponding first and second school-closure in the Netherlands.

**Table A8.1** Available information on standardized tests per year and per grade.

	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6
School year 2013/2014	✓	✓	✓	✓	✓	✓
School year 2014/2015	✓	✓	✓	✓	✓	✓
School year 2015/2016	✓	✓	✓	✓	✓	✓
School year 2016/2017	✓	✓	✓	✓	✓	✓
School year 2017/2018	✓	✓	✓	✓	✓	✓
School year 2018/2019	✓	✓	✓	✓	✓	✓
School year 2019/2020	✓	✓	✓	✓	✓	✓
School year 2020/2021	✓	✓	✓	✓	✓	✓

### *8.3 Data cleaning standardized tests*

The data collected via the administration systems is shared by the school through a ‘click-button’ in the system which approves the data to be transferred to Statistics Netherlands by the beforementioned administrations systems. Statistics Netherlands uploads this information in their Remote Access location; a secured infrastructure where the data is stored, and where researchers can work with the data (after permission). During the cleaning process a variety of points received our attention: (1) not all students are in the final dataset, only students with officially registered in the municipality. (2) students who switch schools usually have their personal data transferred to the new schools, giving rise to duplicate records. To correct for this, test records are valid if the date on which they are taken falls in between the registration and deregistration on the specific school. (3) test records without valid scores, test coding or dates were deleted. (4) duplicate student observations were removed if they were identical based on school id, student id, gender, registration and deregistration date, grade, class, test score, test coding, postal code and date of birth. (5) only test records from CITO are considered valid; test records from other test suppliers are removed from the data. (6) test scores outside of the CITO-specified acceptable range are recoded into missing. (7) reading and maths generation 2.0 test scores are converted into generation 3.0 test scores following a formula retrieved from CITO.<sup>5</sup> Spelling generation 2.0 test scores cannot be converted into generation 3.0 test scores. After the data cleaning, the data is merged with the dataset of the Netherlands Cohort Study on Education (Haelermans et al., 2020).

### *8.4 Data selection standardized tests*

For our main analyses we compare the two cohorts prior to the COVID-19 crisis with the cohort since the crisis. This implies that we include the learning growth between midterm test from 2017/2018 and 2018/2019 and between 2018/2019 and 2019/2020 as pre-COVID-19 and the learning growth between the midterm tests 2019/2020 and 2020/2021 as since-COVID. Next, the selections as listed in Table A8.2 were applied.

---

5. In earlier years of the generation 2.0 tests were taken, in more recent years the generation 3.0 test is used. The reading and maths generation 2.0 test scores can be converted into generation 3.0 test scores following a formula retrieved from CITO, spelling generation 2.0 test scores cannot be converted into generation 3.0 test scores and therefore cannot be used for the analyses.

**Table A8.2** Data selection

	<b>Number of schools</b>	<b>Number of students<sup>i</sup></b>	<b>Number of unique students</b>
1. Starting number of observations for school years 2017/2018 2018/2019 and 2019/2020 after merge with NCO-data <sup>ii</sup>	1 687	755 919	332 490
2. Only keeping students from grades 1-5 (grades 2-5 for reading)	1 686	676 845	372 197
3. Removing special education primary schools	1 680	676 823	327 876
4. Dropping schools with 5 or less students in the data	1 676	676 794	327 855
5. Proficiency score equal to or below zero are set to missing, students with four or more missings are deleted from the data <sup>iii</sup>	1 676	676 794	327 538
6. Background information available			
7. Final dataset	<b>1 676</b>	<b>676 794</b>	<b>327 538</b>

Notes:

i. Students can appear in the data a maximum of three time, when they were in primary education in all three school years (2017/2018, 2018/2019 and 2019/2020).

ii. After merging with the NCO-data only students that are known in the admin data of Statistics Netherlands are left in the data.

iii. We have a total of three tests per domain (reading, spelling, maths) per testing moment, leading to a total of maximum 6 proficiency scores within a school year. When a student misses five or six proficiency scores (in other words, has only one or zero valid scores), we are unable to calculate learning growth, which is why these students are removed.

## 8.5 Measures

### 8.5.1 Standardized test scores and calculation of learning growth

Within the Netherlands students in the same grade level of primary education take the same tests within a given year. These tests are administered in school and last up to 60 minutes for each subject. Absolute test scores are translated into proficiency scores that are presented on a continuous scale for all grade levels together.

The reading test assesses the student's ability to understand written texts, including both factual and literary content. The test in spelling asks students to write down a series of words (no verbs), demonstrating that they have learned the spelling rules. The test in maths contains both abstract problems and contextual problems that describe a concrete task.

The learning growth is calculated based upon the proficiency scores of the standardized tests, by extracting the score of the midterm from previous year from the midterm test of the current year for of each domain, with the condition that the student must have taken two midterm tests within the same school. To remove the influence of outliers, the top and bottom 1% of the

proficiency scores per year per domain are excluded from the analyses. As our outcome measure, we standardize the learning growth and look at the learning growth for each domain separately.

Because midterm tests are on average taken 6 weeks later than usual, but are not corrected by test supplier CITO, we apply a correction to the learning growth between the midterm test in 2019/2020 and 2020/2021. Based on long term school level average number of weeks between the two midterm tests (which on average over all schools is 52 weeks, as expected) we correct at the school level for the number of weeks the test was taken later in 2021 than in the years before by multiplying the learning growth by  $(\frac{52}{52-\#weeks\ later\ in\ 2021})$ .

Table A8.3 shows the descriptive statistics of the absolute (corrected) learning growth by year, presenting a first indication of a lower learning growth in the year of the COVID-19 crisis.

**Table A8.3** Descriptive statistics learning growth by year

**2017/2018-2018/2019**

	N	mean	sd	min	max
Learning growth Reading	150475	18.404	20.69	-151.32	173.436
Learning growth Spelling	124412	33.538	28.486	-137	247
Learning growth Maths	156615	27.36	18.308	-188	181.74

**2018/2019-2019/2020**

Learning growth Reading	152209	17.987	19.785	-126	185
Learning growth Spelling	148407	33.012	27.781	-228	264
Learning growth Maths	159727	26.759	18.042	-177	208

**2019/2020-2020/2021**

Corrected learning growth Reading	122777	14.732	18.086	-122.828	168.772
Corrected learning growth Spelling	127474	31.088	25.815	-124.621	235.857
Corrected learning growth Maths	134260	24.489	16.906	-159.586	151.439

*8.5.2 Independent variables*

In this paper, we use the following independent variables, next to the treatment that was already described in full above:

*Socioeconomic status (SES)*: The parental socioeconomic status is based upon the educational attainment of parents, the parental income as a percentage of the baseline of poverty income and whether or not the parents work in the previous year. The scoring of the variable can be found in Table A8.4, in the end the maximum score on the SES indicator is 7. Based upon this scoring range from 0-7 a categorical variable is computed as well, scores from 0 to 3 are categorized as ‘low SES’, scores 4 and 5 are categorized as ‘medium SES’ and scores 6 and 7 are categorized as ‘high SES’.

**Table A8.4.** SES construction

Measures	Explanation	Value	Category
Education	Highest	0	Less than high school
	educational	1	High school, vocational or associate degree
	attainment of one	2	Bachelor’s degree
	of the parents	3	Master’s degree, doctorate or professional degree
Income	Parental income as percentage of poverty income	0	0-100% of poverty income
		1	100-200% of poverty income
		2	200-400% of poverty income
		3	400% of poverty income
Employment status	Employment status	0	Unemployed
	in the last year	1	Employed
<i>Total score</i>		<i>0 - 7</i>	

*Parental education level*: Parents’ highest educational attainment is divided into three categories. Parental education is defined as low when the highest obtained degree of (at least one of) the parents is in pre-vocational secondary education (vmbo b/k), or a degree in upper secondary vocational education (mbo 1), or grades 7 to 9 in pre-vocational secondary education (vmbo gl/tl) or senior general secondary education or university preparatory education (1). It is defined as middle with a degree in upper secondary vocational education level 2, 3 or 4, or when they completed senior general secondary education or university preparatory education (2). The indicator is given the value ‘high’ for students whose parents that have a university of applied sciences degree or higher (3).

*Household income:* Highest household income is defined as low when the highest income of one of the parents is below the Dutch minimum income level (1), middle when the income is higher than minimal level but below twice the minimum income level (2) and high when the income of one of the parents is higher than twice the minimum income (3).

*Household wealth:* Highest household wealth is defined as low when the highest wealth of one of the parents is below the Dutch minimum wealth level (1), middle when the wealth is higher than minimal level but below twice the minimum wealth level (2) and high when the wealth of one of the parents is higher than twice the minimum wealth (3).

*Labour market position of parents:* The parental labour market status is based upon the socioeconomic category of the previous year of the parents, where the original categories are that the parent is either working; based upon employee, director, entrepreneur or working family member, receiving benefit; based upon receiving unemployment benefits, social assistance benefits, social services benefits, sickness benefits or pension benefits, or being inactive; which is based upon either being a student without income, or not receiving an income at all. We have transformed this information into a categorical variable with five categories: (1) both parents working, (2) only the father works, (3) only the mother works, (4) both parents don't work and (5) unknown for both parents.

*Gender:* A dummy variable with value 1 for girls and 0 for boys.

*Migration background:* A dummy variable for having either Dutch background or western-migration background (0) or having a non-western migration background (1).

In Table A8.5 the descriptive statistics are presented for the background characteristics of the students and their parents.

**Table A8.5.** Descriptive statistics background characteristics students

	N	share
Low socioeconomic status	58,398	10.8%
Med socioeconomic status	308,193	57.0%
High socioeconomic status	174,423	32.2%
Low educated	56,839	11.8%
Med educated	160,403	33.4%
High educated	262,816	54.7%
Low income	120,769	22.5%
Med income	289,357	54.0%
High income	125,513	23.4%
Low wealth	296,085	54.7%
Med wealth	55,282	10.2%
High wealth	189,647	35.1%
Both parents have job	383,481	70.9%
Only father has job	71,122	13.1%
Only mother has job	20,933	3.9%
None of parents have job	26,940	5.0%
Unknown for both parents	38,538	7.1%
Girls	270,182	49.9%
Boys	270,832	50.1%
Dutch or western migration	435,571	81.5%
Non-western migration background	105,324	19.5%

### 8.6 Method of analysis

In order to estimate the effect of the COVID-19 crisis on students' learning growth, we compare the learning growth between the midterm of the one year and the midterm of the next school year. Learning growth is determined by the difference in overall knowledge and skills in a certain year compared to the year prior (Notenboom, Aarts & Lit, 2017). The Dutch government has set out yearly goals and reference levels. Therefore, we can research the learning growth by comparing to which extent the level and goals are met in one year compared to the next year. We distinguish between cohorts of students who took standardized tests before the start of the COVID-19 crisis (up until the midterm test in school year 2019/2020 = pre-COVID-19) and students who took standardized tests during the COVID-19 crisis (from the end test of school year 2019/2020 onwards). We compare the learning growth of the COVID-19-exposed-cohort to the learning growth of students from the two previous cohorts using OLS regressions.

First, we estimate the effect of the COVID-19 crisis on the learning growth of students. We present standardized learning growth where 2019 standardized based on the mean and standard

deviation of 2017 and 2018. In this step we estimate the following regression equation, resembling a difference-in-differences design:

$$\Delta y_i = \beta_0 + \beta_1 T_i + \varepsilon_{is} \quad (A1)$$

Where  $\Delta y_{ij}$  stands for the learning growth in proficiency score between two test moments for student  $i$ .  $T_i$  is an indicator for treatment, which is the COVID-19 exposed-cohort, and  $\varepsilon_{is}$  is the school-level clustered standard error. The coefficient of interest is  $\beta_1$ , which captures the difference in average learning growth between the COVID-19-exposed-cohort and the average learning growth of the (pooled) preceding two cohorts.

In the second step, we include interaction effects between the treatment and the student level background characteristics such as household income, household wealth, parental education level, parental labour market position, parental socioeconomic status, migration background and gender of the student, as presented in equation 2, where  $C_i$  stands for the background characteristics of student  $i$ :

$$\Delta y_i = \beta_0 + \beta_1 T_i + \beta_2 C_i + \beta_3 T_i C_i + \varepsilon_{is} \quad (A2)$$

Note that we include the interaction between treatment and parental education level and household income in all other student background analyses. We do this to account for the effect on top of parental education level or household income. This is shown in equation 3, where  $S_i$  stands for the socioeconomic status of student  $i$ :

$$\Delta y_i = \beta_0 + \beta_1 T_i + \beta_2 C_i + \beta_3 T_i C_i + \beta_4 S_i + \beta_5 T_i S_i + \varepsilon_{is} \quad (A3)$$

### 8.7 Robustness checks

In addition to the main analyses, we run several robustness checks based on the model presented in equation 1. The robustness checks refer to the standardized results as presented in Appendix 9.3.

1. We used inverse probability weights to limit the impact of selectivity and over-representation of certain students and schools in our data. In calculating weights, we use population data on all students enrolled in Dutch primary education and calculate the probability that they are in the standardized test data separately per academic year and test subject domain. The weight is a function of the following student observable characteristics: parental education, household income, gender, share of students with

low educated parents at their school, number of students at the school, urbanisation level (based on location of the school), province (based on location of the school) and school denomination.

2. We ran our main regressions including additional background characteristics at a) the student level, and b) the student and the school level.
3. We added school fixed effects instead of clustering standard errors.
4. We use a multilevel model instead of clustering standard errors.
5. We did a placebo analyses in which we pretend that 2018/2019 is the COVID-19-exposed cohort and compare them with the two years prior. Note that including the school year 2016/2017 here means that we have to restrict our analysis to grades 1-4 (as the 2016/2017 data is not available for grade 5). Therefore, we run our main analyses again based on grades 1-4 for reasons of comparison.
6. We present trend figures of the development of learning growth over time. The interpretation of the differences in learning growth being due the impact of the COVID-19 crisis depends on the assumption that learning growth would have been similar in the absence of the crisis. While this assumption is untestable, we can provide supporting evidence for it by looking at the variability of learning growth over time. If these trends are relatively stable, we can be reasonably sure that the difference between the 2019/2020 cohort and the previous cohorts was caused by the impact of the crisis.

## 9. Appendix Results

### 9.1 Main unstandardized results

**Table A9.1a** Unstandardized main results

	Reading	Spelling	Maths
Treatment	-3.462*** (0.125)	-2.164*** (0.199)	-2.567*** (0.129)
Constant	18.19*** (0.0807)	33.25*** (0.136)	27.06*** (0.0869)
Observations	425,461	400,293	450,602
R-squared	0.006	0.001	0.004
Clusters	1665	1656	1667

Robust standard errors, clustered at the school level, in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

### 9.2 Main results standardized interaction with student characteristics

**Table A9.2** Standardized main results – Parental socioeconomic status

	Reading	Spelling	Mathematics
Treatment	-0.175*** (0.0117)	-0.0830*** (0.0105)	-0.182*** (0.0119)
Medium SES	0.00936 (0.00638)	-0.0169*** (0.00559)	-0.00595 (0.00680)
High SES	0.0581*** (0.00750)	-0.000550 (0.00639)	0.0180** (0.00791)
Treatment x Medium SES	-0.00514 (0.0115)	0.0256*** (0.00982)	0.0600*** (0.0112)
Treatment x High SES	0.0217 (0.0134)	0.0347*** (0.0111)	0.0746*** (0.0127)
Constant	-0.0244*** (0.00672)	-0.260*** (0.00616)	-0.177*** (0.00772)
Observations	425461	400293	450602
R-squared	0.007	0.002	0.005
Clusters	1665	1656	1667

Robust standard errors in parentheses

\*\*\* p<0.01. \*\* p<0.05. \* p<0.1

**Table A9.3** Standardized main results – Parental education level

	Reading	Spelling	Mathematics
Treatment	-0.172*** (0.0127)	-0.0833*** (0.0111)	-0.191*** (0.0127)
Medium edu	-0.00517 (0.00684)	0.0140** (0.00550)	0.00608 (0.00654)
High edu	0.0522*** (0.00767)	0.0188*** (0.00623)	0.0410*** (0.00729)
Treatment x medium edu	-0.00826 (0.0135)	-0.00118 (0.0105)	0.0234* (0.0130)
Treatment x high edu	0.000859 (0.0148)	0.0188 (0.0115)	0.0470*** (0.0139)
Medium SES	-0.0152** (0.00739)	-0.0279*** (0.00616)	-0.0266*** (0.00734)
High SES	0.0234** (0.00926)	-0.00824 (0.00737)	-0.00124 (0.00878)
Treatment x Medium SES	-0.00571 (0.0135)	0.0168 (0.0113)	0.0341** (0.0132)
Treatment x High SES	0.0196 (0.0174)	0.0204 (0.0137)	0.0536*** (0.0157)
Constant	-0.0232*** (0.00716)	-0.264*** (0.00649)	-0.179*** (0.00826)
Observations	378,232	356,201	400,562
R-squared	0.008	0.002	0.006
Clusters	1664	1656	1667

Robust standard errors in parentheses

\*\*\* p&lt;0.01. \*\* p&lt;0.05. \* p&lt;0.1

**Table A9.4** Standardized main results – Household income

	Reading	Spelling	Mathematics
Treatment	-0.181*** (0.0120)	-0.0879*** (0.0109)	-0.192*** (0.0123)
Medium income	-0.0172*** (0.00519)	-0.0352*** (0.00432)	-0.0283*** (0.00476)
High income	0.0302*** (0.00662)	-0.0327*** (0.00554)	-0.00630 (0.00639)
Treatment x medium income	0.0305*** (0.00981)	0.0243*** (0.00762)	0.0527*** (0.00914)
Treatment x high income	0.0347*** (0.0124)	0.0329*** (0.00961)	0.0526*** (0.0115)
Medium SES	0.0127* (0.00652)	0.00419 (0.00543)	0.00806 (0.00653)
High SES	0.0454*** (0.00734)	0.0189*** (0.00626)	0.0249*** (0.00743)
Treatment x Medium SES	-0.0240* (0.0123)	0.00964 (0.00983)	0.0277** (0.0116)
Treatment x High SES	-0.00453 (0.0142)	0.0103 (0.0111)	0.0378*** (0.0132)
Constant	-0.0210*** (0.00699)	-0.253*** (0.00643)	-0.171*** (0.00801)
Observations	422.231	397.111	446.988
R-squared	0.008	0.002	0.005
Clusters	1665	1656	1667

Robust standard errors in parentheses

\*\*\* p&lt;0.01. \*\* p&lt;0.05. \* p&lt;0.1

**Table A9.5** Standardized main results – Household wealth

	Reading	Spelling	Mathematics
Treatment	-0.177*** (0.0118)	-0.0867*** (0.0107)	-0.185*** (0.0121)
Medium wealth	0.000440 (0.00629)	-0.0180*** (0.00487)	-0.0174*** (0.00560)
High wealth	0.0188*** (0.00473)	-0.0273*** (0.00371)	-0.00284 (0.00447)
Treatment x medium wealth	0.00965 (0.0110)	0.0272*** (0.00835)	0.0489*** (0.00972)
Treatment x high wealth	0.0312*** (0.00849)	0.0384*** (0.00676)	0.0258*** (0.00765)
Medium SES	0.00473 (0.00627)	-0.00896* (0.00544)	-0.00394 (0.00666)
High SES	0.0491*** (0.00704)	0.0136** (0.00608)	0.0205*** (0.00743)
Treatment x Medium SES	-0.0119 (0.0114)	0.0145 (0.00966)	0.0502*** (0.0111)
Treatment x High SES	0.00737 (0.0132)	0.0148 (0.0109)	0.0592*** (0.0124)
Constant	-0.0258*** (0.00684)	-0.257*** (0.00625)	-0.176*** (0.00783)
Observations	425.461	400.293	450.602
R-squared	0.007	0.002	0.005
Clusters	1665	1656	1667

Robust standard errors in parentheses

\*\*\* p&lt;0.01. \*\* p&lt;0.05. \* p&lt;0.1

**Table A9.6** Standardized main results – Parental labour market position

	Reading	Spelling	Mathematics
Treatment	-0.153*** (0.0138)	-0.0529*** (0.0112)	-0.145*** (0.0135)
Father has job	0.0151*** (0.00527)	0.0262*** (0.00433)	0.0345*** (0.00517)
Mother has job	0.0126 (0.00886)	0.0122* (0.00681)	0.000280 (0.00786)
No parent has job	0.0745*** (0.00961)	0.0831*** (0.00824)	0.0702*** (0.00896)
Both unknown	0.0218*** (0.00794)	0.0711*** (0.00758)	0.0311*** (0.00852)
Treatment x Father has job	-0.00318 (0.0102)	-0.0144* (0.00816)	-0.0309*** (0.0100)
Treatment x Mother has job	-0.0522*** (0.0174)	-0.00921 (0.0135)	-0.0171 (0.0154)
Treatment x No parent has job	-0.0576*** (0.0175)	-0.0650*** (0.0140)	-0.0680*** (0.0168)
treatment x both unknown	0.00408 (0.0147)	-0.0204* (0.0113)	-0.0318** (0.0139)
Medium SES	0.0383*** (0.00674)	0.0255*** (0.00582)	0.0245*** (0.00680)
High SES	0.0873*** (0.00761)	0.0414*** (0.00662)	0.0492*** (0.00778)
Treatment x Medium SES	-0.0231* (0.0127)	-0.000358 (0.0102)	0.0302** (0.0123)
Treatment x High SES	0.00195 (0.0140)	0.00763 (0.0113)	0.0439*** (0.0135)
Constant	-0.0579*** (0.00740)	-0.310*** (0.00653)	-0.215*** (0.00771)
Observations	425.461	400.293	450.602
R-squared	0.007	0.002	0.005
Clusters	1665	1656	1667

Robust standard errors in parentheses

\*\*\* p&lt;0.01. \*\* p&lt;0.05. \* p&lt;0.1

**Table A9.7** Standardized main results – Gender

	Reading	Spelling	Mathematics
Treatment	-0.181*** (0.0123)	-0.0881*** (0.0109)	-0.197*** (0.0123)
Girls	-0.0166*** (0.00328)	-0.0327*** (0.00254)	0.0219*** (0.00299)
Treatment x Girls	0.0123* (0.00677)	0.00964* (0.00504)	0.0307*** (0.00602)
Medium SES	0.00905 (0.00638)	-0.0175*** (0.00559)	-0.00562 (0.00679)
High SES	0.0576*** (0.00750)	-0.00143 (0.00639)	0.0185** (0.00791)
Treatment x Medium SES	-0.00489 (0.0115)	0.0259*** (0.00983)	0.0601*** (0.0112)
Treatment x High SES	0.0221 (0.0134)	0.0352*** (0.0111)	0.0747*** (0.0127)
Constant	-0.0158** (0.00708)	-0.243*** (0.00633)	-0.188*** (0.00778)
Observations	425461	400293	450602
R-squared	0.007	0.002	0.005
Clusters	1665	1656	1667

Robust standard errors in parentheses

\*\*\* p&lt;0.01. \*\* p&lt;0.05. \* p&lt;0.1

**Table A9.8** Standardized main results – Migration background

	Reading	Spelling	Mathematics
Treatment	-0.168*** (0.0122)	-0.0797*** (0.0100)	-0.173*** (0.0119)
Migration background	0.0881*** (0.00574)	0.0609*** (0.00526)	0.0929*** (0.00670)
Treatment x Migration background	-0.0111 (0.00976)	-0.00600 (0.00892)	-0.0144 (0.0103)
Medium SES	0.0402*** (0.00609)	0.00414 (0.00511)	0.0268*** (0.00617)
High SES	0.0937*** (0.00714)	0.0235*** (0.00601)	0.0554*** (0.00726)
Treatment x Medium SES	-0.00889 (0.0115)	0.0235** (0.00926)	0.0545*** (0.0108)
Treatment x High SES	0.0152 (0.0134)	0.0310*** (0.0106)	0.0663*** (0.0123)
Constant	-0.0705*** (0.00646)	-0.291*** (0.00574)	-0.226*** (0.00690)
Observations	425.383	400.218	450.514
R-squared	0.008	0.002	0.006
Clusters	1665	1656	1667

Robust standard errors in parentheses

\*\*\* p&lt;0.01. \*\* p&lt;0.05. \* p&lt;0.1

### 9.3 Robustness checks (Standardized results)

All robustness checks presented below (Tables A.8.9-A.8.14b) show that our results are very stable to the inclusion of inverse probability weights, dummy for end test after the 2020 summer holiday, individual and school level control variables and fixed effects. Furthermore, the use of a multilevel model and placebo analyses also show the robustness of our results.

**Table A9.9** Robustness checks standardized – Inverse probability weights

	Reading	Spelling	Maths
Treatment	-0.175*** (0.00754)	-0.0588*** (0.00590)	-0.120*** (0.00681)
Constant	-0.00249 (0.00442)	-0.270*** (0.00381)	-0.176*** (0.00437)
Observations	345.292	325.811	365.931
R-squared	0.008	0.002	0.005
Clusters of schools	1624	1616	1626

Robust standard errors. clustered at the school level. in parentheses

\*\*\* p<0.01. \*\* p<0.05. \* p<0.1.

**Table A9.10** Robustness checks standardized – Adding individual controls

	Reading	Spelling	Maths
Treatment	-0.173*** (0.00634)	-0.0573*** (0.00539)	-0.122*** (0.00636)
Gender	-0.0132*** (0.00275)	-0.0304*** (0.00206)	0.0320*** (0.00245)
Disadvantaged student 0.3	-0.0438*** (0.00763)	-0.0318*** (0.00596)	-0.0654*** (0.00721)
Disadvantaged student 1.2	-0.000991 (0.00828)	-0.00937 (0.00672)	-0.0124 (0.00821)
Household structure migration background	-0.668*** (0.0802)	-1.032*** (0.122)	-1.103*** (0.0754)
Average SES	0.0882*** (0.00493)	0.0470*** (0.00406)	0.0917*** (0.00523)
High SES	-0.00437 (0.00452)	-0.0132*** (0.00379)	-0.0158*** (0.00417)
Only the father works	0.0400*** (0.00597)	-0.00693 (0.00475)	-0.00386 (0.00562)
Only the mother works	0.0108** (0.00442)	0.0105*** (0.00340)	0.0186*** (0.00405)
Both parents don't work	6.73e-06 (0.00747)	-0.00131 (0.00545)	-0.00960 (0.00637)
Working status parents unknown	0.0251*** (0.00809)	0.0148** (0.00601)	0.0102 (0.00732)
Low income household	0.0197*** (0.00740)	0.0264*** (0.00548)	0.0259*** (0.00712)
Number of children at home	-0.0239*** (0.00338)	0.0115*** (0.00268)	-0.00880*** (0.00325)
Constant	-0.00318* (0.00170)	-0.00938*** (0.00140)	-0.0140*** (0.00154)
	-0.0351*** (0.00865)	-0.259*** (0.00720)	-0.195*** (0.00804)
Observations	377.109	355.173	399.390
R-squared	0.010	0.003	0.008
Clusters of schools	1664	1656	1667

Robust standard errors. clustered at the school level. in parentheses

\*\*\* p<0.01. \*\* p<0.05. \* p<0.1.

**Table A9.11** Robustness checks standardized – Adding individual and school level controls

	Reading	Spelling	Maths
Treatment	-0.173*** (0.00643)	-0.0586*** (0.00550)	-0.124*** (0.00649)
Gender	-0.0132*** (0.00278)	-0.0304*** (0.00206)	0.0317*** (0.00246)
Disadvantaged student 0.3	-0.0394*** (0.00755)	-0.0266*** (0.00568)	-0.0597*** (0.00714)
Disadvantaged student 1.2	-0.00188 (0.00804)	-0.0112* (0.00663)	-0.0172** (0.00795)
Household structure	-0.714*** (0.0800)	-1.130*** (0.119)	-1.167*** (0.0730)
Migration background	0.0657*** (0.00453)	0.0247*** (0.00346)	0.0689*** (0.00418)
Average SES	-0.00450 (0.00453)	-0.0114*** (0.00374)	-0.0138*** (0.00412)
High SES	0.0275*** (0.00569)	-0.0156*** (0.00441)	-0.0142*** (0.00510)
Only the father works	0.00803* (0.00440)	0.00937*** (0.00331)	0.0153*** (0.00391)
Only the mother works	-0.00708 (0.00742)	-0.00621 (0.00542)	-0.0156** (0.00630)
Both parents don't work	0.0190** (0.00816)	0.0105* (0.00599)	0.00440 (0.00722)
Working status parents unknown	0.0104 (0.00718)	0.0187*** (0.00538)	0.0162** (0.00683)
Low income household	-0.474*** (0.0597)	0.564*** (0.0878)	-0.0613 (0.0578)
Number of children at home	-0.000894 (0.00161)	-0.00780*** (0.00128)	-0.0122*** (0.00143)
School level: Mid-sized school board	0.0134 (0.0246)	0.0531*** (0.0143)	-0.0290 (0.0247)
School level: Large school board	0.0129 (0.0246)	0.0450*** (0.0142)	-0.0336 (0.0241)
School level: Mid-sized school	0.000797 (0.00710)	-0.0163** (0.00659)	0.00717 (0.00773)
School level: Large school	0.0243*** (0.00887)	-0.0181** (0.00832)	0.00495 (0.0101)
School level: Low share of disadvantaged students	-0.0108 (0.00906)	-0.0182** (0.00798)	-0.0157* (0.00919)
School level: Medium share of disadvantaged students	-0.0188* (0.00906)	-0.0178* (0.00906)	-0.0258** (0.00906)

	(0.0105)	(0.0102)	(0.0110)
School level: High share of disadvantaged students	-0.0218 (0.0148)	-0.0167 (0.0137)	-0.0315* (0.0161)
School level: Low share of Non-Western students	0.0113 (0.0109)	0.0137 (0.00877)	0.0158 (0.0108)
School level: Medium share of Non-Western students	0.0129 (0.0137)	0.0229* (0.0120)	0.000573 (0.0138)
School level: High share of Non-Western students	0.0267 (0.0183)	0.0267 (0.0171)	0.00992 (0.0198)
School level: High share first generation immigrant students	0.00794 (0.00887)	-0.000871 (0.00774)	0.00930 (0.00955)
School level: Medium share second generation immigrant students	-0.00154 (0.0103)	0.00838 (0.00891)	0.00346 (0.0103)
School level: High share second generation immigrant students	0.00314 (0.0149)	0.00509 (0.0134)	0.0143 (0.0164)
School level: Medium share low income household students	-0.00902 (0.0109)	0.00829 (0.00901)	-0.00834 (0.0102)
School level: High share low income household students	-0.00484 (0.0185)	0.0342** (0.0163)	0.0127 (0.0186)
School level: Medium share low wealth household students	-0.00240 (0.00909)	0.00503 (0.00806)	0.00620 (0.00920)
School level: High share low wealth household students	0.00564 (0.0125)	0.00109 (0.0109)	-0.000260 (0.0129)
School level: Medium share one-parent household students	0.00243 (0.00952)	-0.00366 (0.00808)	-0.00505 (0.00977)
School level: High share one-parent household students	-0.00338 (0.0148)	-0.0187 (0.0126)	-0.00549 (0.0157)
School level: Low share large family students	-0.0146 (0.00938)	-0.00994 (0.00852)	-0.000989 (0.00972)
School level: Medium share large family students	-0.000702 (0.0108)	-0.0124 (0.00961)	0.00117 (0.0113)

School level: High share large family students	-0.0193 (0.0128)	-0.0335*** (0.0114)	-0.0109 (0.0142)
School level: medium share students father employed	-0.0230* (0.0134)	-0.0202 (0.0134)	-0.0143 (0.0159)
School level: high share students father employed	-0.0227 (0.0173)	-0.0412*** (0.0159)	-0.0182 (0.0186)
School level: medium share students mother employed	0.00685 (0.0104)	0.0336*** (0.00967)	0.0191* (0.0113)
School level: high share students mother employed	-0.00171 (0.0142)	0.0410*** (0.0127)	0.0230 (0.0152)
School level: medium share low SES students	-0.00422 (0.00934)	-0.00498 (0.00824)	0.0114 (0.00927)
School level: high share low SES students	-0.0128 (0.0145)	-0.0111 (0.0135)	0.0135 (0.0148)
School level: Province 2	0.00754 (0.0261)	0.0360* (0.0209)	0.00812 (0.0224)
School level: Province 3	-0.00822 (0.0371)	-0.0202 (0.0303)	-0.0302 (0.0347)
School level: Province 4	0.000822 (0.0260)	-0.0207 (0.0198)	0.00255 (0.0214)
School level: Province 5	0.0286 (0.0285)	0.00663 (0.0220)	0.0377 (0.0240)
School level: Province 6	0.0152 (0.0250)	0.00401 (0.0188)	0.0200 (0.0204)
School level: Province 7	0.0436 (0.0275)	0.0180 (0.0211)	0.0467** (0.0233)
School level: Province 8	0.0276 (0.0265)	-0.00410 (0.0209)	0.0213 (0.0235)
School level: Province 9	0.0198 (0.0254)	0.0190 (0.0199)	0.0307 (0.0216)
School level: Province 10	-0.00365 (0.0331)	0.0151 (0.0275)	0.0240 (0.0346)
School level: Province 11	0.0171 (0.0255)	-0.0314 (0.0194)	0.0216 (0.0211)
School level: Province 12	0.0276 (0.0261)	-0.0816*** (0.0194)	0.0319 (0.0212)
School level: Public schools	0.0222 (0.0184)	0.0479** (0.0192)	0.0163 (0.0221)

School level: Schools based upon educational or pedagogical concepts	0.00366 (0.00760)	0.00250 (0.00719)	0.00593 (0.00847)
School level: Schools with mixed denominations	-0.204*** (0.0289)	-0.0622*** (0.0196)	-0.0658** (0.0295)
School level: very low level urbanization	0.00982 (0.0134)	0.00381 (0.0110)	-0.0171 (0.0144)
School level: low level urbanization	0.0223 (0.0148)	0.0201 (0.0122)	-0.00658 (0.0152)
School level: medium level urbanization	0.0151 (0.0148)	0.0239* (0.0126)	-3.38e-05 (0.0159)
School level: high level urbanization	0.0159 (0.0218)	0.0115 (0.0193)	0.00787 (0.0238)
School level: very high level urbanization	0.0403** (0.0185)	0.0458*** (0.0170)	0.0439** (0.0209)
Constant	-0.0697* (0.0409)	-0.298*** (0.0323)	-0.200*** (0.0408)
Observations	371.058	349.866	392.876
R-squared	0.011	0.007	0.010
Clusters of schools	1622	1614	1624

Robust standard errors. clustered at the school level. in parentheses

\*\*\* p<0.01. \*\* p<0.05. \* p<0.1.

**Table A9.12** Robustness checks standardized – Using school fixed effects

	Reading	Spelling	Maths
Treatment	-0.174*** (0.00335)	-0.0625*** (0.00251)	-0.127*** (0.00283)
Constant	0.255*** (0.0354)	-0.0499 (0.0356)	0.105*** (0.0315)
School level FE	Yes	Yes	Yes
Observations	425.461	400.293	450.602
R-squared	0.023	0.028	0.028

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table A9.13** Robustness checks standardized – Using a multilevel model

	Reading	Spelling	Maths
Treatment	-0.173*** (0.003)	-0.062*** (0.002)	-0.126*** (0.003)
Constant	-0.011** (0.003)	-0.264*** (0.003)	-0.182*** (0.004)
School variance	0.013*** (0.001)	0.014*** (0.001)	0.016*** (0.001)
Error-term variance	0.930*** (0.002)	0.515*** (0.001)	0.722*** (0.002)
Observations	425,461	400,293	450,602
Number of schools	1,665	1,656	1,667

Robust standard errors in parentheses  
\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table A9.14a** Robustness checks standardized – Placebo model - Original analyses without grade 5 (2017/2018-2019/2020 grades 1-4)

	Reading	Spelling	Maths
Treatment	-0.137*** (0.00657)	-0.0507*** (0.0057)	-0.106*** (0.00674)
Constant	-0.0111*** (0.00395)	-0.209*** (0.00377)	-0.121*** (0.00443)
Observations	333,741	335,472	357,960
R-squared	0.004	0.001	0.003
Clusters of schools	1664	1656	1667

Robust standard errors, clustered at the school level, in parentheses  
\*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

**Table A9.14b** Robustness checks standardized – Placebo model - Placebo analyses 2016/2017-2018/2019 grades 1-4

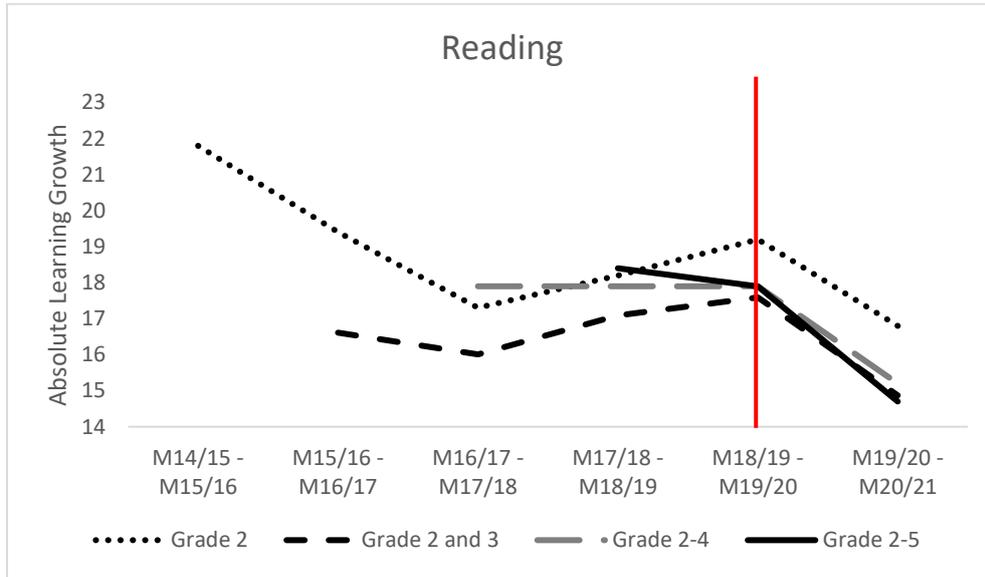
	Reading	Spelling	Maths
Treatment	-0.0003 (0.006)	0.008 (0.0007)	-0.004*** (0.007)
Constant	-0.000*** (0.0041)	0.000*** (0.0055)	-0.000*** (0.052)
Observations	351.836	312.298	370.920
R-squared	0.000	0.000	0.000
Clusters of schools	1666	1657	1669

Robust standard errors, clustered at the school level, in parentheses

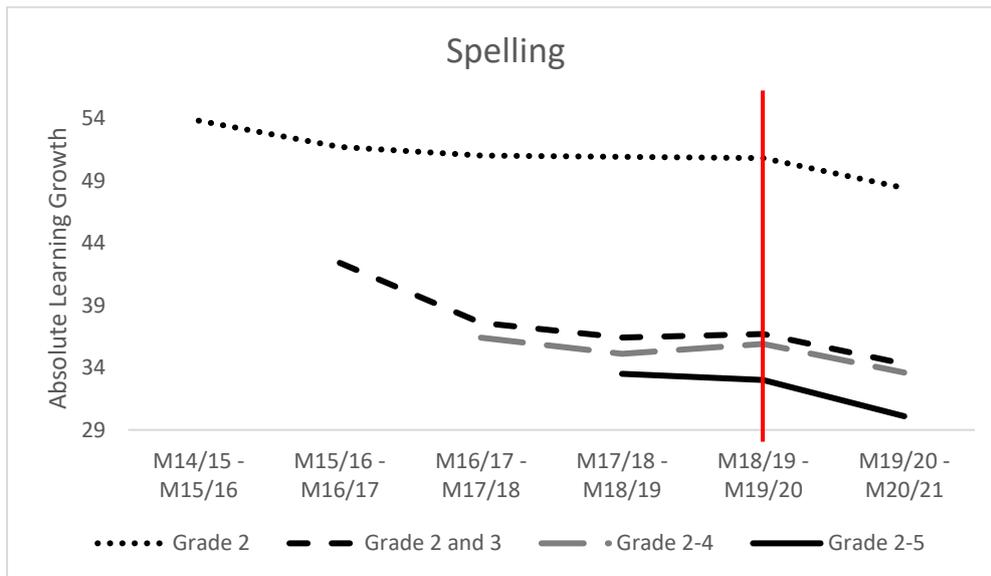
\*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Figures A9.1, A9.2 and A9.3 show the mean learning growth for all available grades for all available cohorts in reading, spelling, and maths respectively. As noted earlier, we do not have information on grade 1 learning growth for the reading domain, as grade 1 students do not take a midterm test for this domain. For higher grades, we also have fewer available cohorts due to the manner in which data was collected. The red line indicates the start of the COVID-19 crisis. The figures clearly show a marked lower learning growth between the COVID-19 affected cohort of the school year 2019/2020 relative to the prior cohorts in all domains. As could be seen in Table A9.14b there is some year-to-year variation for the domain of maths. Furthermore, the year to-year variation might seem bigger as there are not that many series available per grade about the learning growth in the domains

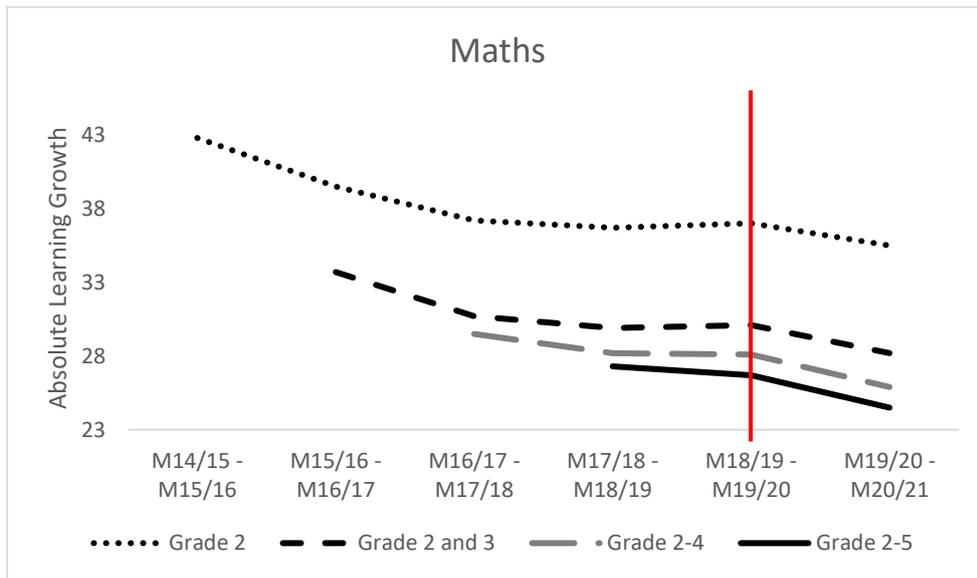
**Figure A9.1** Trends in learning growth – reading



**Figure A9.2** Trends in learning growth – spelling



**Figure A9.3** Trends in learning growth - maths





## Colophon

© Research Centre for Education and the Labour Market  
Nothing from this publication may be reproduced in any way  
without the prior written permission of the ROA director.

**Research Centre for Education and the Labour Market**  
Maastricht University  
School of Business and Economics  
[secretary-roa-sbe@maastrichtuniversity.nl](mailto:secretary-roa-sbe@maastrichtuniversity.nl)  
[www.roa.nl](http://www.roa.nl)