Inequality in the Effects of Primary School Closures due to the Covid-19 Pandemic – Evidence from the Netherlands

Carla Haelermans, Madelon Jacobs, Rolf van der Velden, Lynn van Vugt, and Sanne van Wetten

ONLINE APPENDIX

1. Materials and Methods

1.1 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.

In the first 1.5 years of the Covid-19 pandemic, Dutch primary schools had to close twice as a result of the COVID-19 crisis, where the first school closure started on March 16th and lasted up to and including May 10th (school closure of eight weeks, including two weeks of regular holiday). However, after May 10th 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, as they were taken right before the school closure.

From December 16th 2020 up to and including February 8th 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.

Afterwards, schools were open, although occasionally classes were sent home when the teacher was tested positive or had to quarantine, and no replacement could be found. The end of year standardized test took place in the regular testing period.

1.2 Sample

Our sample consists of approximately 500,000 students from ~1,900 primary schools.[1] This means that around 30% 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 2016/2017 up until the end term test in school year 2020/2021.

1.3 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

[1] In our current sample and analyses, students at special primary education are not included as we do not have data about those students.

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 mathematics. 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.

1.4 Standardized test scores

In the Netherlands, most students from grade 1 until grade 6 take a national test that measures the proficiency in key domains. Primary schools are obliged to administer the standardized tests which 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 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 mathematics. 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 mathematics 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 end term 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 the learning growth. We standardize the learning growth per subject, grade level 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. To remove the influence of 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 pandemic.

1.5 Data collection standardized tests

The 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.

1.6 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 before mentioned 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 mathematics generation 2.0 test scores are converted into generation 3.0 test scores following a formula retrieved from CITO.[2] 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).

1.7 Data selection standardized tests

For our main analyses we compare the two cohorts prior to the Covid-19 pandemic with the cohort since the pandemic. This implies that we include the learning growth between midterm test from 2016/2017 and the end term test of 2017/2018 and between 2017/2018 and 2018/2019 as pre-Covid-19 and the learning growth between the midterm test 2019/2020 and the end test 2020/2021 as since-Covid-19. Table A1.1 provides the descriptive statistics of the learning growth by year.

[2] 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 mathematics 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 A1.1 Descriptive statistics learning growth by year

	N	mean	Standard deviation
2016/2017-2017/2018			
Learning growth Reading	91017	23.32	21.21
Learning growth Spelling Lower	82411	89.55	44.75
Learning growth Spelling Upper	49009	40.15	24.75
Learning growth Mathematics	179895	47.58	24.17
2017/2018-2018/2019			
Learning growth Reading	106007	24.59	16.41
Learning growth Spelling Lower	80573	86.07	35.90
Learning growth Spelling Upper	72350	43.86	21.01
Learning growth Mathematics	171243	45.80	19.53
2019/2020-2020/2021			
Learning growth Reading	86587	22.96	16.37
Learning growth Spelling Lower	65000	83.87	35.80
Learning growth Spelling Upper	70891	44.09	21.51
Learning growth Mathematics	143376	44.33	19.96

1.8 Variables of interest

1.8.1 Treatment: one and a half years of Covid-19

Our main explanatory variable of interest is the Covid-19 pandemic. 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 end term test 2020/2021. The period between midterm test 2016/2017 and end term test 2017/2018 and the period between midterm test 2017/2018 and end term test 2018/2019 serve as ‘control’ years and have value 0 in this dummy (pre-Covid-19). Note that we *do not* use the period between the midterm test 2018/2019 and the end term test 2019/2020 because that period comprises of both a pre-Covid-19 and a during-Covid-19 period, which is why we cannot use it as a clean control period.

1.8.2 Student background

To get an idea of the extent to which different student characteristics correlate with the impact of the pandemic 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). 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 A1.2. SES construction

Measures	**Explanation**	**Value**	**Category**
Education	Highest educational attainment of one of the parents	0	Less than high school
		1	High school, vocational or associate degree
		2	Bachelor's degree
		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 in the last year	0	Unemployed
		1	Employed
Total score		*0 - 7*	

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). 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: The income level of the household is divided into three groups: low, medium and high. 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 structure: A dummy variable in which we make a distinction between children living in a household with two parents, and children living in a household with only one parent. Two-parent-household (value 1) are households in which children live with both legal parents, or one of the legal parents with his/her (registered) new partner. One-parent-households (value 0) are households in which only 1 of the legal parents is present, and no other adult resides in the household. Note that we use registered residents from the municipal administrative databases of Statistics Netherlands. New partners of parents already that cohabit with the legal parent of a child, but that are not formally registered at that address will not be taken into account, and these children will be part of the one-parent-household category. Note that the group of children that is not living with either of their legal parents is very small and is therefore left out in this variable.

Family size: A dummy variable, based on the number of children that are living with their parents. We make a distinction between a large family (3 or more children still living at home, including the child for whom the indicator is calculated, value 0) and a small family (2 children or less, in other words, having one sibling or being an only child, value 1).

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

In Table A1.3 the descriptive statistics are presented for the background characteristics of the students and their parents.

Table A1.3. Descriptive statistics background characteristics students

	N	share
Low socioeconomic status	77073	12.7%
Med socioeconomic status	309312	51.1%
High socioeconomic status	219657	36.2%
Low educated	63086	11.7%
Med educated	179681	33.3%
High educated	296918	55.0%
Low income	131972	22.0%
Med income	326057	54.2%
High income	143109	23.8%
One-parent household	98391	16.4%
Two-parents household	502838	83.6%
Small family	396812	65.6%
Large family	208246	34.4%
Dutch or western migration	495083	81.7%
Non-western migration background	110825	18.3%

1.9 Empirical strategy: Difference-in-differences

To estimate the effect of the Covid-19 pandemic 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 pandemic. First, we estimate the effect of the Covid-19 pandemic on the learning growth of students. We present standardized learning growth where the Covid-19 learning growth is standardized based on the mean and standard deviation of the two years prior. In this step we estimate the following regression equation, resembling a difference-in-differences design:

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

Where Δy_{ij} stands for the standardized 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 ε_{is} is the school-level clustered standard error. The coefficient of interest is β_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 socioeconomic status, parental education and household income, as presented in equation A2, where C_i stands for the background characteristics of student i:

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

Note that we include the interaction between treatment and socioeconomic status in all other student background analyses. We do this to account for the effect on top of SES. This is shown in equation A3, where S_i stands for the socioeconomic status of student i:

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

1.10 Robustness checks

In addition to the main analyses, we run several robustness checks based on the model presented in equation A1 for the composite score.

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 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 pandemic depends on the assumption that learning growth would have been similar in the absence of the pandemic. 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 Covid-19-cohort and the previous cohorts was caused by the impact of the pandemic.

2. Full Regression and Robustness Results

2.1 Full regression results Table 2

Table A2.1a Full Regression Results Table 2

VARIABLES	SES	Parental Education	Income	Household Structure	Family size	Migration
Treatment	-0.111***	-0.146***	-0.120***	-0.131***	-0.122***	-0.135***
	(0.0106)	(0.0142)	(0.0112)	(0.0120)	(0.0116)	(0.0137)
Medium parental education		-0.0210***				
		(0.00669)				
High parental education		0.0277***				
		(0.00787)				
Treatment # Medium parental educati		0.00806				
		(0.00991)				
Treatment # High parental education		0.0301***				
		(0.0115)				
Medium student SES	-0.0551***	-0.0632***	-0.0448***	-0.0552***	-0.0525***	-0.0188***
	(0.00665)	(0.00732)	(0.00630)	(0.00662)	(0.00654)	(0.00563)
High student SES	-0.0282***	-0.0699***	-0.0280***	-0.0282***	-0.0262***	0.0228***
	(0.00821)	(0.00856)	(0.00752)	(0.00824)	(0.00810)	(0.00663)
Treatment # Medium student SES	0.0238**	0.0427***	0.00766	0.0194**	0.0215**	0.0171**
	(0.00925)	(0.0108)	(0.00931)	(0.00919)	(0.00915)	(0.00832)
Treatment # High student SES	0.0542***	0.0606***	0.0299***	0.0469***	0.0527***	0.0439***
	(0.0111)	(0.0127)	(0.0112)	(0.0111)	(0.0110)	(0.00976)
Medium parental income			-0.0313***			
			(0.00494)			
High parental income			0.0107			
			(0.00651)			
Treatment # Medium parental income			0.0313***			
			(0.00756)			
Treatment # High parental income			0.0341***			
			(0.00944)			
Two-parent family				0.00140		
				(0.00466)		
Treatment # Two-parent family				0.0298***		
				(0.00693)		
Small family (<3 children)					-0.0218***	
					(0.00380)	
Treatment # Small family					0.0200***	
					(0.00545)	
Dutch or western migration backgroun						-0.135***
						(0.00822)
Treatment # Dutch or western migratic						0.0376***
						(0.00999)
Constant	0.0589***	0.0715***	0.0672***	0.0580***	0.0712***	0.133***
	(0.00840)	(0.0111)	(0.00896)	(0.00930)	(0.00913)	(0.0111)
Observations	535,323	476,939	531,808	531,586	534,621	535,218
R-squared	0.003	0.004	0.003	0.003	0.003	0.006
Clusters	1899	1898	1899	1899	1899	1899

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

2.2 Robustness checks

All robustness checks presented below (Tables A2.2-A2.6) show that our results are very stable to the inclusion of inverse probability weights, individual and school level control variables, and the use of a multilevel model or fixed effects instead of clustered standard errors.

Table A2.2 Robustness checks – Inverse probability weights

VARIABLES	Composite score
Treatment	-0.0832***
	(0.00600)
Constant	0.0157***
	(0.00470)
Observations	421,311
R-squared	0.003
Clusters	1855

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

Table A2.3 Robustness checks – Adding individual controls

VARIABLES	Composite score
Treatment	-0.0751***
	(0.00581)
Girl	-0.00502**
	(0.00250)
Medium student SES	-0.0236***
	(0.00683)
High student SES	-0.0116
	(0.00760)
One-parent family	-0.0188***
	(0.00398)
No migration background	0.120***
	(0.00622)
Medium parental education	0.00134
	(0.00516)
High parental education	0.0553***
	(0.00590)
Only the father works	0.0242***
	(0.00396)
Only the mother works	0.0121**
	(0.00613)
Both parents don't work	0.00269
	(0.00840)
Working status parents unknown	-0.0191***
	(0.00295)
Number of children at home	-0.00279
	(0.00174)
Constant	-0.00246
	(0.00944)
Observations	449,673
R-squared	0.006
Clusters	1898

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

Table A2.4 Robustness checks – Adding individual and school level controls

VARIABLES	Composite score
Treatment	-0.0757***
	(0.00620)
Girl	-0.00461*
	(0.00250)
Medium student SES	-0.0205***
	(0.00672)
High student SES	-0.0140*
	(0.00749)
One-parent family	-0.0236***
	(0.00393)
No migration background	0.0908***
	(0.00481)
Medium parental education	0.00260
	(0.00501)
High parental education	0.0487***
	(0.00542)
Only the father works	0.0221***
	(0.00383)
Only the mother works	0.00692
	(0.00605)
Both parents don't work	-0.000663
	(0.00834)
Working status parents unknown	-0.0122***
	(0.00267)
Number of children at home	-0.00343**
	(0.00156)
School level: Mid-sized school board	-0.00156
	(0.0208)
School level: Large school board	-0.0242
	(0.0208)
School level: Mid-sized school	-0.0255***
	(0.00790)
School level: Large school	-0.0262***
	(0.00971)
School level: Low share of disadvantaged students	-0.0175*
	(0.00923)
School level: Medium share of disadvantaged students	-0.0141
	(0.0119)
School level: High share of disadvantaged students	-0.0227
	(0.0156)
School level: Low share of Non-Western students	0.0238**
	(0.0111)
School level: Medium share of Non-Western students	0.0324**
	(0.0137)
School level: High share of Non-Western students	0.0290
	(0.0191)
School level: High share first generation immigrant students	0.00619
	(0.00927)
School level: Medium share second generation immigrant students	-0.00677
	(0.0105)
School level: High share second generation immigrant students	-0.00680
	(0.0164)

School level: Medium share low income household students	0.00423
	(0.0103)
School level: High share low income household students	-0.00577
	(0.0187)
School level: Medium share low wealth household students	-0.000781
	(0.00947)
School level: High share low wealth household students	-0.0225*
	(0.0116)
School level: Medium share one-parent household students	0.0126
	(0.00939)
School level: High share one-parent household students	0.0114
	(0.0145)
School level: Low share large family students	0.00969
	(0.00983)
School level: Medium share large family students	0.00188
	(0.0114)
School level: High share large family students	0.0148
	(0.0141)
School level: medium share students father employed	-0.0196
	(0.0172)
School level: high share students father employed	-0.0413**
	(0.0201)
School level: medium share students mother employed	0.0190**
	(0.00950)
School level: high share students mother employed	0.0168
	(0.0126)
School level: medium share low SES students	-0.00844
	(0.00910)
School level: high share low SES students	-0.0115
	(0.0146)
School level: Province 2	0.0511*
	(0.0282)
School level: Province 3	0.0530
	(0.0474)
School level: Province 4	0.0343
	(0.0286)
School level: Province 5	0.0374
	(0.0341)
School level: Province 6	0.0316
	(0.0272)
School level: Province 7	0.0702**
	(0.0295)
School level: Province 8	0.0550*
	(0.0295)
School level: Province 9	0.0700**
	(0.0288)
School level: Province 10	0.0151
	(0.0407)
School level: Province 11	-0.00530
	(0.0282)
School level: Province 12	-0.0166
	(0.0285)

School level: Public schools	-0.00315
	(0.0180)
School level: Schools based upon educational or pedagogical concepts	0.000181
	(0.00840)
School level: Schools with mixed denominations	0.186***
	(0.0287)
School level: very low level urbanization	0.0113
	(0.0127)
School level: low level urbanization	0.0153
	(0.0143)
School level: medium level urbanization	0.0249*
	(0.0148)
School level: high level urbanization	0.0378*
	(0.0205)
School level: very high level urbanization	0.0349*
	(0.0183)
Constant	-0.0152
	(0.0440)
Observations	443,107
R-squared	0.011
Clusters	1845

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

Table A2.5 Robustness checks – Using school fixed effects

VARIABLES	Composite score
Treatment	-0.0801***
	(0.00233)
Constant	0.374***
	(0.0286)
Observations	535,323
R-squared	0.041

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A2.6 Robustness checks – Using a multilevel model

VARIABLES	Composite score	lns1_1_1	lnsig_e
Treatment	-0.0798***		
	(0.00232)		
Constant	0.0166***	-1.859***	-0.271***
	(0.00386)	(0.0186)	(0.000968)
Observations	535,323	535,323	535,323
Number of groups	1,899	1,899	1,899

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figures A2.1, A2.2 and A2.3 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 pandemic. 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. 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 A2.1 Trends in learning growth – reading

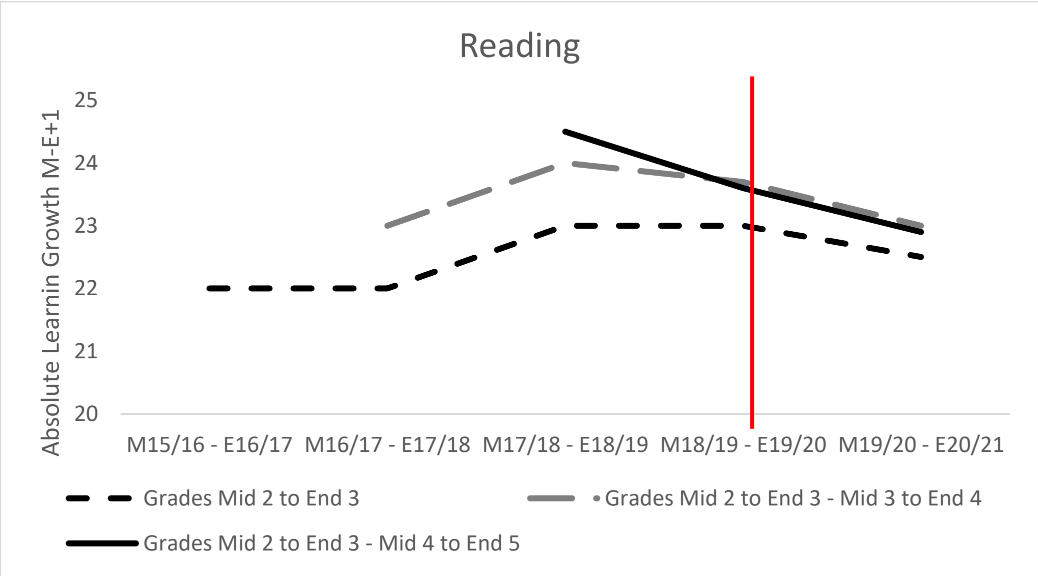


Figure A2.2 Trends in learning growth – spelling

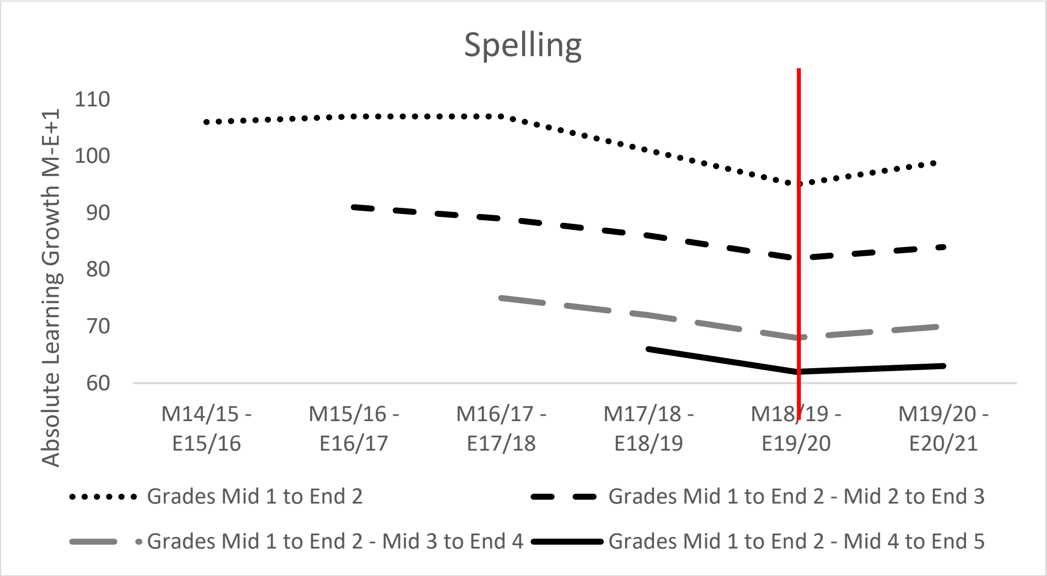
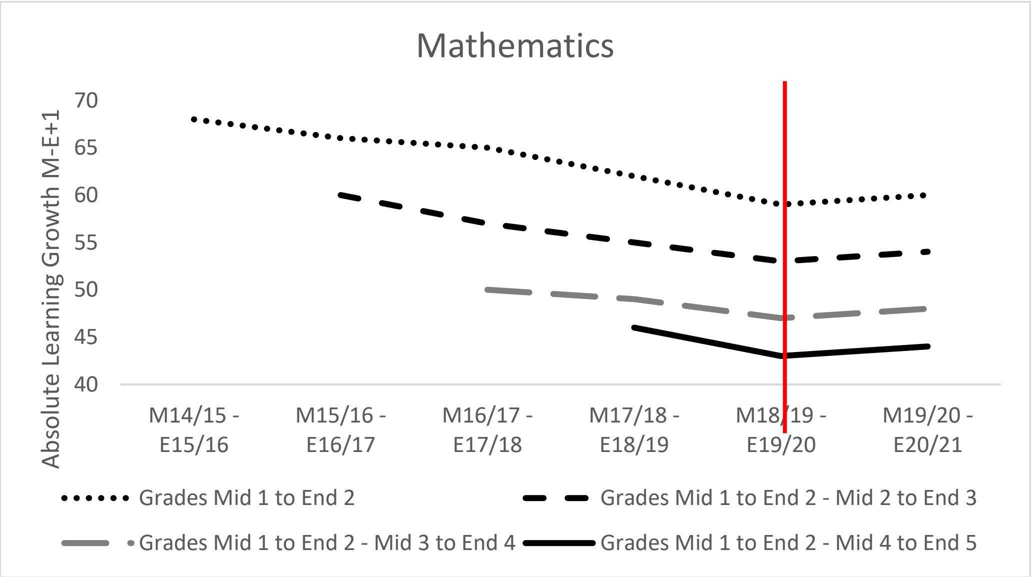


Figure A2.3 Trends in learning growth - mathematics



3. References Online Appendix

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

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/jcaa027/5871552