ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Cattan, Sarah; Kamhöfer, Daniel A.; Karlsson, Martin; Nilsson, Therese

Working Paper

The short- and long-term effects of student absence: Evidence from Sweden

IFS Working Papers, No. W17/21

Provided in Cooperation with:
Institute for Fiscal Studies (IFS), London

Suggested Citation: Cattan, Sarah; Kamhöfer, Daniel A.; Karlsson, Martin; Nilsson, Therese (2017) : The short- and long-term effects of student absence: Evidence from Sweden, IFS Working Papers, No. W17/21, Institute for Fiscal Studies (IFS), London

This Version is available at:
https://hdl.handle.net/10419/200281

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Institute for Fiscal Studies

The short- and long-term effects of student absence: evidence from Sweden

IFS Working Paper W17/21

Sarah Cattan
Daniel A. Kamhöfer
Martin Karlsson
Therese Nilsson

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

The Short- and Long-term Effects of Student Absence: Evidence from Sweden

Sarah Cattan, Daniel A. Kamhöfer, Martin Karlsson and Therese Nilsson*

September 5, 2017

Abstract

Instructional time is seen as an important determinant of school performance, but little is known about the effects of student absence. Combining historical records and administrative data for Swedish individuals born in the 1930s, we examine the impacts of absence in elementary school on short-term academic performance and long-term socio-economic outcomes. Our siblings and individual fixed effects estimates suggest absence has a moderate adverse effect on academic performance. The detrimental effect fades out over time. While absence negatively correlates with final education, income and longevity, we only find robust evidence that it lowers the probability of employment at age 25-30.

Keywords: Absence in school, educational performance, long-term effects, register data

JEL Classifications: *C23, I14, I21, I26*

*Cattan: Institute for Fiscal Studies, London, and IZA, Bonn; Kamhöfer: Paderborn University, University of Duisburg-Essen and CINCH, Essen (Corresponding author: Paderborn University, Warburger Strasse 100, 33098 Paderborn, Germany. E-mail: daniel.kamhoefer@uni-paderbron.de.); Karlsson: CINCH, University of Duisburg-Essen and IZA, Bonn; Nilsson: Lund University and IFN, Stockholm.
For valuable comments we are grateful to Esteban Aucejo, Sonia Bhalotra, Arnaud Chevalier, Paul Devereux, Martin Dribe, Martin Fischer, Petter Lundborg, Erik Plug, Martin Salm, Kjell Salvanes, Hendrik Schmitz, Nina Schwarz, Guido Schwerdt, and Matthias Westphal. We would also like to thank seminar participants at CINCH (Essen), IFN (Stockholm) and Lund University as well as participants of EEA 2015, ESPE 2015, IWAEE 2015, VfS 2015 and the Essen Health Conference 2015. For collecting and digitizing the data used here we are indebted to our colleagues in Essen and Lund as well as a vast team of research assistants. Sarah Cattan gratefully acknowledges financial assistance from the British Academy Postdoctoral Fellowship pf140104. Daniel Kamhöfer is grateful to the Institute for Fiscal Studies, London, and Lund University for hospitality as well as to the German Academic Exchange Service (DAAD) and the University of Duisburg-Essen–ERASMUS+ Mobility Program for financial support.

1 Introduction

Student absence from school is pervasive around the world. In 2015, 19 per cent of fourth-graders in the US were absent from school for three or more days in the last month. Students from low-income backgrounds are more likely to be absent than their more affluent peers, and this is the case for both excused and unexcused absences (Child Trends, 2015). As school attendance correlates with academic achievement and generally viewed as an important input in the education production function, reducing school absences has become a challenging matter for schools and a high priority for local and national governments. Indeed, absence has reached such alarming level in some schools that commentators talk about an "empty-desk epidemic".[1]

Despite the relatively uncontested importance of reducing school absence in the policy arena, there is little causal evidence of the effect of absence on achievement and beyond. Identifying such impact is difficult for several reasons. First, it requires individual-level panel data on school absences, school performance, and any other outcome of interest. Such data rarely exist as many countries only started collecting absence records recently. Second, it requires a credible strategy to identify the causal impact of absence from the vast array of unobserved confounding factors. Students who miss school may be less motivated, in poorer health, or attending schools that fail to promote student engagement, which could lead to spurious correlations between absence and achievement.

The few papers that credibly estimate the causal effect of student absence focus on standardized test scores in the US context (Goodman, 2014, Aucejo and Romano, 2016). To our knowledge, there doesn't exist comparable evidence for a context outside the US, although high rates of student absences are prevalent in many countries. Moreover, these studies focus on the short-term impact of student absence on academic outcomes. Yet, in the presence of dynamic complementarities in the production function of human capital, the adverse effects of absence on the formation of skills could persist and even widen over the long-run (Cunha and Heckman, 2007). Assessing the long-term effects

[1] See Chicago Tribune, 2012.

of individual absences is thus crucial to assess the potential benefits of policies aimed at reducing student absences.

An analysis of such long-run effects will meet several challenges. First, data that link student absence and performance to later labor market outcomes are rare to find. Second, the follow-up period needs to be long enough to allow for a reasonable approximation of life-cycle earnings. This second point is important given that there are several examples in the literature of early-career advantages either fading relatively fast (such as the effect of the business cycle on earnings, cf. Genda et al., 2010; Oreopoulos et al., 2012; Altonji et al., 2016) or becoming more pronounced at higher ages (such as the effects of schooling, cf. Bhuller et al., 2011). An analysis based only on early-career labor market outcomes may thus be biased in an unknown direction.

This paper fills this gap by providing evidence of the short- and long-term impact of student absence using a unique panel following a representative sample of cohorts born between 1930 and 1935 in Sweden. This novel dataset links digitized school records of absence and performance to adult socio-economic outcomes measured up to 60 years later from Census and tax register data. This combination of historical and administrative data allows us to investigate a wide array of outcomes. Specifically, we analyze the effect of student absence in grade 1 and grade 4 (at ages 7-8 and 10-11, respectively) on student performance in these grades, as well as its effect on final education, employment (at ages 25-30 and 35-40), labor market income (at ages 35-40), pensions from past labor market activity (measured at ages 67-72) and mortality.

To deal with the potential endogeneity of absence, we exploit two features of the data. First, the sample includes pairs of siblings, which we use to implement a sibling fixed effect (FE) strategy and control for all time-invariant, family-level characteristics that could simultaneously drive absence and our outcomes of interest. Second, we exploit the fact that absence and achievement were collected for two grades (grade 1 and grade 4) to control for individual FE when looking at short-term impacts. Finally, we also implement two approaches as sensitivity checks: a bounding approach following Altonji

et al. (2005) and an instrumental variable (IV) strategy exploiting local and temporal changes in weather conditions as source of exogenous variation for absence.

In line with the existing literature, we find a negative and significant impact of student absence on academic performance in elementary school equivalent to 3.3 per cent of a standard deviation for ten days of absence (the average number of absences in our sample). To address the arbitrariness of test score scales, we take advantage of our long panel to translate the effect on performance in school into its association with adult earnings. Anchoring the test scores to long-term income this way confirms a moderate effect size.

Our findings for long-term effects suggest that the consequences of absence in elementary school fade out over time. While absence negatively correlate with final educational achievement, employment, income and longevity, only the relationship between employment at ages 25-30 and school absence remains strongly significant when we include sibling fixed effects. In this case, the impact of absence is rather large, as ten days of absence lead to a 4 per cent reduction in employment. Ten years later however, the impact of absence on employment is less precisely estimated and we cannot reject that its effect is no longer distinct from zero.

Our paper makes several contributions to a broad literature examining the impact of instructional time on educational achievement and later socio-economic outcomes. Although school absence is an important determinant of the total individual amount of time spent in school, most existing studies exploit exogenous variation in the length of the school year as source of exogenous variation in instructional time. Among others, such studies use laws and law changes that cause variation in the school year length (e.g., Leuven et al., 2010, Pischke, 2007, Sims, 2008, Agüero and Beleche, 2013, and Fischer et al., 2016)[2]; variation in test dates, where the total amount of education the students receive is eventually the same but some students are tested earlier than others (see, e.g., Carlsson et al., 2015, and Fitzpatrick et al., 2011); and unscheduled school closures re-

[2]Other examples include Battistin and Meroni (2016, evidence for Italy), Huebener and Marcus (2015, Germany) and Bellei (2009, Chile) who use structural reforms that expand instructional time.

sulting from extreme weather events (e.g., Marcotte, 2007, Marcotte and Hemelt, 2008, Marcotte and Hansen, 2010, and Hansen, 2011).

When it comes to school absence, two recent studies, Goodman (2014) and Aucejo and Romano (2016), analyze short-run effects of individual school absence in the US. Using Massachusetts data (school years 2003-2010) for students attending grade 3 onwards and North Carolina data (school years 2006-2010) for grade 3 to 5 students, respectively, they show that school results are negatively affected by absence. Both studies control for institutional heterogeneity using school, teacher and individual fixed effects. To corroborate their results, both studies also implement an IV approach using local variation in snowfall (Goodman, 2014) and infectious diseases (Aucejo and Romano, 2016) to instrument school absence.

Our paper contributes to the above literatures by providing new evidence on the effect of student absence as one determinant of instructional time. Our paper is the first to present estimates of the impact of days of absence on long-term outcomes, including final education, labor market outcomes, and mortality. Moreover, we study individual-level changes in instructional time in a context outside the US. The literature examining the effect of region- or school-level changes in instructional time suggests that the educational system is an important factor for the observed effects, but individual changes in instructional time have not yet been analyzed outside the US.

Our results show that these innovations matter for our understanding of the impact of school absences. In fact, considering effects throughout the life-cycle sheds new light on previous findings regarding the role of school absence. Our short-term point estimates are remarkably close to those of Goodman (2014) and Aucejo and Romano (2016) – even though we analyze the relationship in another country and in another decade. The long-term estimates indicate that the short-term effects – although substantial at the beginning – slowly fade away over time. This highlights the importance of having outcomes measured at different points of the career, as impacts measured early in the career do not reflect impacts found later in the working life. A declining effect of missed

instruction in school throughout the life-cycle is in line with Pischke (2007) who finds a negative effect of school years with reduced instructional time on subsequent schooling but no long-lasting labor market consequences.[3]

The remainder of the paper proceeds as follows. Section 2 provides some background on the schooling system in Sweden in the 1930s. Section 3 describes the data and some descriptive statistics. Section 4 discusses our empirical strategy, while Section 5 presents our results. Section 6 includes our sensitivity analysis and Section 7 concludes.

2 Background

2.1 Elementary education in Sweden in the 1930s and 1940s

The Swedish education system as it appeared in the 1930s has a long history. Compulsory schooling was introduced in 1842 when all parishes of the country had to offer basic education. In the 1930s and 1940s all children went to a common, public, and free school, *Folkskola*[4], and the country was divided into 2,400 school districts responsible for primary education. It was compulsory to enter the first grade at the age of seven and complete at least six years of schooling.

A clear majority of school districts offered six years of compulsory schooling, but a clause introduced in 1921 allowed school districts to introduce seven years of compulsory schooling. The clause was followed by a government decision on July 1, 1936 to increase compulsory schooling by one year over a twelve year period. Accordingly, a mandatory seventh grade was introduced stepwise across districts in the following years. Similarly, the length of the school year corresponded to 34.5 weeks in most districts, but in the period under review, the school year length was increased stepwise to 39 weeks.[5]

[3]More broadly, the absence of long-lasting consequences of the amount of schooling an individual receives is also in line with studies that find zero returns to compulsory schooling, e.g., Stephens and Yang (2014) for the US and Pischke and von Wachter (2008) for Germany.

[4]We use the terms *Folkskola* and elementary school interchangeably.

[5]See Fischer et al. (2013) for an analysis of the seven-year reform, and Fischer et al. (2016) for an examination of the changes in term length. In principle it would be possible to compare the effect of more instructional time due to the school year length increase with the effect of less instructional time because of absence in school. However, we would not expect the former to affect performance in the same grade. Teachers could have adjusted there expectations because all students were affected by the school

Although the responsibility for providing primary education was decentralized, the Ministry of Ecclesiastical Affairs provided clear nationwide standards that applied to all school districts. The most central decree was the 1919 Educational Plan (*Utbildningsplanen*), which included the full curriculum of the *Folkskola*. Students attended elementary school full time, six days a week.[6] Instruction was generally done in classes separated by grade. When the number of students was low, schools were also allowed to pool students in different grades into one classroom, so that a teacher instructed, for instance, students of grade 1 and grade 2 in the same room during the same lesson. The content of the education was grade-specific, however, as stated in the Educational Plan.

The educational system of the 1930s exhibited several features of a modern educational system – like absence of tuition fees and joint instruction of boys and girls at all educational levels (Erikson and Jonsson, 1993) – but education was very selective (Fischer et al., 2016). Students who decided to take more than compulsory education followed a tracking system and generally left *Folkskola* after grade 4 to enter lower secondary school (*Realskola*). All other students remaind in *Folkskola* until they reached the compulsory years of schooling. From 1939 and onwards the admission to *Realskola* was based on grades received in elementary school. The system also offered a second alternative where a student could proceed to lower secondary school after finishing *Folkskola*. After four or five years of lower secondary schooling, students either entered upper secondary school (*Gymnasium*) or finished their educational career. In the birth cohorts that we consider 87.5 per cent of students only have compulsory education[7], and until the 1940s only about 5 per cent of a cohort continued with upper secondary schooling (Fredriksson, 1971a).

year length expansion. Moreover, we do not expect long-term effects because the curriculum remained unchanged.

[6]Instruction ended at noon on Saturdays. Following an exception rule, schools in rural areas had the possibility to offer half-time reading (students went to school every second day or only during certain periods of the year) but this option was very limited in the 1930s and only 0.5 per cent of our sample took half-time reading.

[7]Based on own calculation for the birth cohorts 1930-1935 using the Census 1970, see Table 3 in the next section.

2.2 Historical records of student absence and achievement

As their main organizational tool teachers kept daily records in an exam catalog called *Dagbok med examenskatalog* (see Appendix Figure A1 for a picture). In these catalogs, the teachers recorded students' performance and absences, and noted whether absences were due to sickness, natural obstacles (e.g., heavy snowfall), inappropriate clothes and shoes, other valid reasons for absence, or no valid excuses (that is, truancy). They also included general information about the school and the school year length.

Regarding student performance teachers were encouraged to take notes on the student's performance throughout the entire school year. At the end of the school year, the teachers summarized the days of absence by reason and the final grades by subject in a separate column for end-of-school-year information. Unlike tests that take place on a certain date, the frequent recording of student performance ensures that teaching-to-the-test behavior of teachers and factors on the day of a test did not affect the grades. Moreover, regular record-keeping makes recall bias of the teacher unlikely.

2.3 Grading standards

Three theoretical subjects were taught in *Folkskola*: math; reading and speaking; and writing. Although grades recorded in exam catalogs were not based on standardized tests and hence may partly reflect teachers' subjective impressions of students, a 1940 Royal Commission established precise guidelines for teachers to evaluate and grade their students' performance relative to that of the classmates. For example, to assess a student's math performance, teachers were to take both the ability to solve "standard problems" and more sophisticated ones into account. For reading and speaking, grades were supposed to reflect loud and silent reading and the ability to express a familiar topic in own words. For writing, grades were supposed to assess both the form and content of essays.

The highest possible grade was A ("passed with great distinction") and the poorest grade was C ("not passed"). Teachers were also allowed to add a plus or minus sign in order to express the strength or weakness of the grade. While the grading scheme remained unchanged in the time under review, the grading guidelines changed slightly. From the

school year 1940/41 onwards, teachers were advised to award the grade BA ("passed with credit") for an average performance. One-third of the students in the class should receive a better grade and one-third a poorer grade. Before the school year 1940/41, teachers were more likely to award a student with the grade B for an average performance. The highest grade A was reserved for exceptional students and less than 1 per cent of all students should receive this grade.[8] As we show in the next section, the distribution of grades observed in our sample is remarkably in line with the Royal Commission's guidelines. This gives us confidence that, even though our main measure of academic performance is not a standardized test score, it is a valid measure to compare students' achievement with each other.

3 Data and descriptive statistics

The data we use for analysis combines several historical and administrative data sources.This section provides information about these sources and presents some descriptive statistics on student absence and the main outcomes of interest.

3.1 Data sources

Base data The base foundation for our dataset is individual-level data from administrative church records covering all 30,150 children born between 1930 and 1935 in a representative sample of 133 out of about 2,500 Swedish parishes.[9] Figure 1 presents the spatial distribution of the sample parishes across Sweden. The church records contain individual information on name, gender, date of birth and parish of birth. The records also provide information on the child's parents' birth date, whether the child was born in a hospital (8 per cent of the individuals in our sample), whether the birth was a twin birth (4.2 per cent), and whether the child was born out of wedlock (4.4 per cent). We

[8]In the empirical analysis, the point estimates between specifications with only parish and school fixed effects and with additional teacher fixed effects (that account for subjective grading) do not differ noteworthily. We also change the baseline outcome (performance measured on the 15-point grading scale) to the 7-point grading scale and into a binary indicator that takes the change in the Royal Commission's guidelines into account. The findings do not change our interpretation of the results.

[9]The base dataset was originally collected and digitized to evaluate an infant and maternal health program that the Swedish Government introduced between 1931 and 1933. See Bhalotra et al. (forthcoming) for details on the construction and representativeness of the data.

also know the occupation of the parents at the time of birth. For the empirical analysis we generate an indicator for mothers being employed (2.4 per cent) and a set of indicators for the family's socio-economic status based on the main category of the father's occupation according to the first digit of the Historical International Standard Classification of Occupations (HISCO) code[10] (see Table A1 in the Appendix).

Schooling data Individual schooling information was collected in local archives. Specifically, we collected the exam catalogs in which teachers made systematic notes about types of absence and reported grades for each student, for each elementary school of the 133 parishes in our base dataset. As shown in Figure A1 each student is listed with their first name, surname, date of birth and parents' name. Using this information, we merged the schooling information onto the base dataset. We were able to match schooling information for 17,999 out of the 30,150 children in at least grade 1 or grade 4.[11]

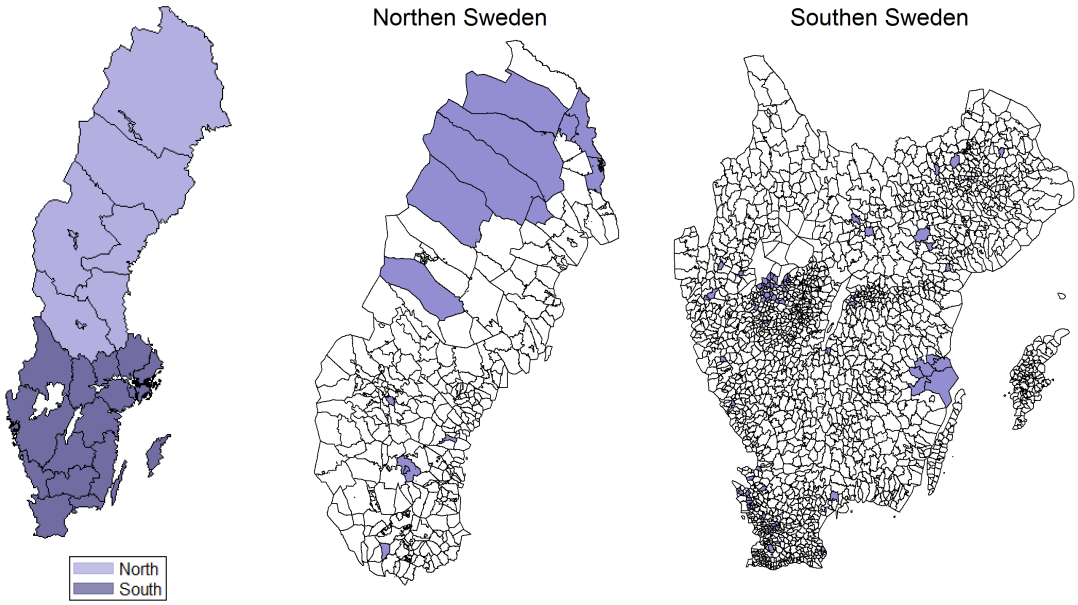
We focus on grade 1 and grade 4 (the last grade in which all students attend *Folkskola*) and digitize the end-of-school-year summary information of the exam catalogs.[12] With

[10]The HICSO code is historical version of today's International Standard Classification of Occupations (ISCO) code, see van Leeuwen et al. (2002). The HISCO occupations coding does not allow to rank jobs according to their prestige or any other criterium. The only group of occupations that can be related to a higher socio-economic status is service workers. If the father had an agricultural occupation, we additionally consider whether he was a farmer, fisherman or hunter (one HISCO category) because this is potentially related to both the family's subsistence as well as the need that children in the household help with reaping the harvest (although we find no evidence that was systematically the case).

[11]The reasons why we are not able to get a perfect match are that (1) exam catalogs were destroyed or cannot be found in the archives, (2) there is insufficient information for identifying an individual, (3) an individual left the sample parish and moved before the age of seven, and/or (4) an individual passed away before reaching school age. The first two reasons are due to the data collection and operationalization and not subject to individual selection. The decision to move and an early death are, however, likely non-random with respect to (sickness) absence and skills. If an early death is health-related, attrition due to mortality may bias the estimates. However, the long-term effect of absence on mortality does not exhibit a noteworthy association between the two factors: see the results section. To address selection due to moving we tried to trace down exam catalogs for individuals who have moved to a different parish before enrolling into *Folkskola* using official registers on movers. For the very few children leaving Sweden before enrolling into *Folkskola* we have no information after they left the country. The assumption we have to make is that the decision to migrate out of Sweden is unrelated to absence in school and educational performance given the socio-economic background. The Online Appendix Table O9 compares the mean value of characteristics in the church records data (that a available of all individuals) between the full sample of all 30,150 individual and the subsample of the individual with schooling information. The results do not indicate a systematic difference in socio-economic factors. Table O10 gives the baseline results separately for individuals that moved between the birth and schooling and individuals that did not move. That coefficients are similar.

[12]Therefore, the data at hand do not allow us to identify the length of absence spells but only the total number of days missed per grade.

Figure 1: Spital distribution of 133 sample parishes within Sweden



Notes: Own illustration. The plot on the left shows the map of Sweden in its regions (*Län*) and the plots in the center and on the right show Northern and Southern Sweden, respectively, in parishes in the time under review. The left plot indicates which regions belong to the Northern and Southern Sweden in the plots in the center and on the right. Parishes belonging to our sample are depicted darker in the plots in the center and on the right.

cohorts born in 1930-1935 the schooling data covers the school years 1936/37 to 1946/47.[13] Table 1 gives an overview over the data structure and corresponding sample sizes by birth cohort. Out of the 30,150 individuals born in the sampled parishes in 1930-1935, we have complete exam catalog records for about 14,000 individuals in either grade 1 or 4. For about 10,000 individuals we have both grades. Reassuringly there is no difference in the matching quality with respect to the birth cohort or the school grade. Using information on the parents, we can also identify sibling pairs born between 1930-1935. Our final sample includes 4,469 siblings for whom we have information on both grades (resulting in 8,938 observations).

As discussed in Section 2, educational performance is measured with the grades teachers assign to students at the end of the school year. Each grade is assigned a numerical value which we refer to as grade point. In our baseline specification we use a scale that takes into account that teachers could assign a plus and a minus sign to a student's grade,

[13]The WWII falls in the time under review. Sweden was neutral in the war and we have not found any historical sources suggesting that the war caused major disruptions in education, nor do the war years reduce the probability that we found exam catalogs in the local archives. In fact, children from Finland were sent to and educated in Sweden because Sweden was less affected by the war, see Santavirta (2012).

Table 1: Number of individuals by birth cohort and sample

# of individuals...	Birth cohort						Total
	1930	1931	1932	1933	1934	1935	
...born in sample	5,355	5,095	5,116	4,743	4,775	5,066	30,150
...in grade 1	2,513	2,438	2,491	2,409	2,360	2,136	14,347
...in grade 4	2,734	2,653	2,647	2,315	2,378	1,864	14,591
...in grades 1 and 4	1,929	1,875	1,887	1,805	1,735	1,448	10,679
...with sibling info.	821	748	818	777	738	567	4,469

Notes: Own calculations based on church records and exam catalog information. For 17,999 out of the 30,150 children born in our sample parishes children we could at least find exam catalog information on one grade. For 10,679 individuals exam catalog information on both grades are available (that is, the individual panel consists 21,358 observations). 4,469 of these individuals have siblings we also observe in both grades (the siblings panel includes 8.938 observations).

Table 2: Grading scale

	Grade														
	Passed...														
	with great distinction		with distinction		with great credit			with credit			without credit			Not passed	
Observed symbols	A	A-	a	a-	AB+	AB	AB-	BA+	BA	BA-	B+	B	B-	BC	C
15-points scale	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
7-points scale	7		6			5			4			3		2	1

Notes: Own illustration based on historical records. The first line states the original grade as denoted in the exam catalog. Lines 2 and 3 give our translation into numerical values on a 15-point and 7-point scale, respectively. The baseline models use the 15-point scale, the Online Appendix includes results for the 7-point scale.

ranging from 1 (poorest grade) to 15 (excellent grade). Table 2 gives the mapping of the potentially ordinal grades into cardinal grade points. To facilitate interpretation we standardize the grade points to have mean 0 and standard deviation 1. In our baseline specification, we measure achievement as the average grade across all subjects. While all students had to take math and reading and speaking, writing was not always graded in the first school year. For the 31.3 per cent of students in our sample with missing writing grade points in the first grade, we calculate the average grade points using the grade points in the other two subjects.

To gain an economically meaningful interpretation of the effect size we "anchor" the raw grades in later-life earnings potential (see Bond and Lang, 2013). Performance anchored in earnings potential should not be confused with the effect of absence on income. The anchored effect of absence still gives the short-term effect on educational performance,

but scaled in units of Swedish krona (SEK, in values of 2002) instead of the somewhat hard-to-interpret numerical grade points. In the analysis we exchange the grade points as dependent variable with the fitted value of the following auxiliary anchoring regression[14]:

$$y_{ig}^{\text{anchor}} = \omega_{0g} + \sum_{j=1}^{13} \omega_{1g,j}\text{math}_{ig} + \sum_{j=1}^{13} \omega_{2g,j}\text{reading}_{ig} + \sum_{j=1}^{13} \omega_{3g,j}\text{writing}_{ig} + \xi_{ig},$$

where y_{ig}^{anchor} is individual i's pension income in 2002, math_{ig}, reading_{ig} and writing_{ig} are her grade points in the particular subject in school grade g, and ξ denotes the estimation error.[15] The anchoring is performed separately for grade-1 grade points and grade-4 grade points. The estimates for the anchoring regressions are reported in the Online Appendix.

Subsequent education data Information on subsequent education beyond elementary school is taken from the highest educational degree as stated in the 1970 Census. Given that individuals are aged 35-40 in 1970, this reflects final education. In our baseline specifications we measure educational attainment with an indicator that takes the value 1 if an individual attains a more advanced track than *Folkskola*, and 0 otherwise.[16]

Labor market data We follow our sample over the life-cycle by using information from the 1960 and 1970 population censuses and tax registers available from 2002 onwards. These sources give long-term information on employment, earnings and pensions. Specifically we identify 11,570 of the individuals in the schooling dataset in the 1960 Census which includes information on individual occupation, and 10,246 individuals in the 1970 Census where we observe employment status and labor market income.[17] While the tax register data include several types of income, we use the income coming from labor

[14]The grade points in each subject enter the regression through full sets of dummy variables. Grade points of 14 and 15 are omitted as these grades are very rare.

[15]Using the labor market income from the Census 1970 for anchoring the performance in school does not change our interpretation of the findings throughout the analysis.

[16]That is, the indicator is 1 if an individual leaves *Folkskola* after grade 4 and attends *Realskola* or if the individual leaves *Folkskola* after the compulsory years of schooling and enrolls into secondary education afterwards. *Realskola* dropouts, *Folkhögskola* or *Folkskola* with 8 or 9 years of compulsory schooling are treated as *Folkskola*, see Fischer et al. (2016).

[17]The effective number of observations used in the final analysis of the long-term effects is lower as we restrict ourselves to empirical strategies that require information on both grades and only consider individuals with siblings in the data.

market activity. As individuals in our sample are well into their retirement ages, this income reflects pensions and, thereby, constitutes a proxy for lifetime earnings.[18] Missing information on labor market outcomes might be due to individuals passing away or migrating from Sweden before 1960. We try to trace down individuals that have migrated (see Appendix) and directly investigate mortality.

Mortality data The exact date of death is taken from the Swedish Death Index of Federation of Swedish Genealogical Societies (see Federation of Swedish Genealogical Societies, 2014). The data includes information on all individuals that passed away between 1901 and 2013.

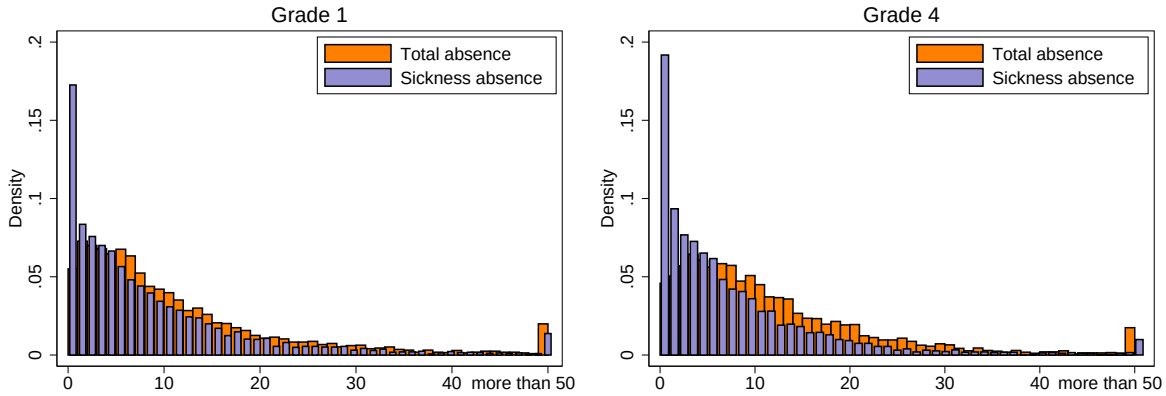
3.2 Descriptive statistics

Our main explanatory variable of interest is the number of missed school days in grade 1 and in grade 4. The data allows us to distinguish between absences due to sickness and absences due to other reasons.[19] Figure 2 shows the distribution of individual days of absence and sickness absence in grade 1 and 4, respectively. In grade 1, 64 per cent of all students miss less than 10 days and 6 per cent of all students have no absence. The average number of missed days in grade 1 is 11 days (median 7 days). In grade 4, students tend to miss slightly more days (mean 11.6 days, median 8.5 days). 59 per cent of all students miss 10 or less days and 5 per cent never miss school. Despite a very different context and time period, the distribution of total days of absence is comparable with that reported in recent US studies (Goodman, 2014; Aucejo and Romano, 2016). We observe a slightly higher density of very high number of absent days than these studies report, but unlike Goodman (2014) who excludes observations with more than 60 days of absence, we do not cap absence days.

[18]For the cohorts considered here, full pensions require thirty years of contributions and the level of the pension is based on the best fifteen years (Sundén, 2006). Widows were in some cases entitled to a certain share of their spouse's earnings after their death and these widow pensions represent the most important deviation from the general rules.

[19]Although the exam catalogs include columns for several reasons for non-health related absence, teachers often only noted other absence without naming the reason.

Figure 2: Distribution of (sickness) absence by grade



Notes: Own calculations based on exam catalog information. 8,938 observations.

Figure 2 illustrates that most absences are sickness absences. Compared to sickness absence, other types of absences only play a minor role – the average number of missed days is 1.6 in grade 1 and 3.3 in grade 4. In grade 1 and 4, 60 per cent and 38 per cent of all students never miss a day for other reasons than sickness, respectively.[20]

Turning to school achievement, Figure 3 shows the distribution of the raw average grade points over math, reading and speaking, and writing by school grade. In line with the suggestion of the Royal Commission, only a few students receive a very low or a very high grade point and the variance of the grade points is higher in grade 4 than in grade 1.[21]

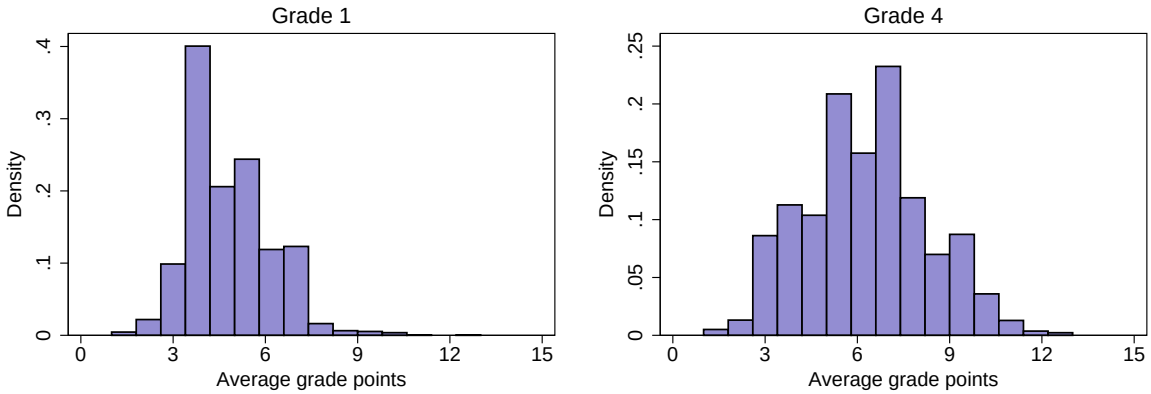
Table 3 presents the long-term outcomes. The highly selective nature of the education system in the time under review, is reflected in only 13 per cent of the individuals in our sample having more than *Folkskola*. Interestingly, this number does not differ by gender. Employment is measured in 1960 and 1970, when our sample is aged 25-30 and 35-40 respectively, and corresponds to a binary indicator equal to 1 if an individual is employed.

Income measures are available for 1970 and 2002. The 1970 income measure refers to the labor market income recorded in the 1970 Census when individuals in our sample were in prime working age (35-40 years old). Table 3 states the original values in SEK in the year in which income is measured. With individuals born 1930-1935, the 2002 income measure

[20]The Online Appendix additionally plots the within-family and within-individual distributions of total days of absence and days of sickness absence.

[21]The Online Appendix shows the distributions of grade points by subject and school grade.

Figure 3: Distribution of average grade points across math, reading and speaking, and writing



Notes: Own calculations based on exam catalog information. 8,938 observations.

Table 3: Descriptive statistics on long-term outcomes

		Mean				
	Age range	All	Female	Male	# obs	% female
Education						
More than *Folkskola* (in %)		12.51	12.99	12.02	3,565	50.29
Employment status						
in 1960 (in %)	25-30	65.08	36.51	93.96	4,129	50.28
in 1970 (in %)	35-40	57.93	43.00	72.89	4,469	50.06
Earnings						
in 1970	35-40	23,924	14,989	30,555	2,932	42.60
in 2002	67-72	150,816	128,175	175,138	3,072	51.79
Mortality at age 70						
passed away (in %)		20.70	16.45	24.96	3,072	50.06

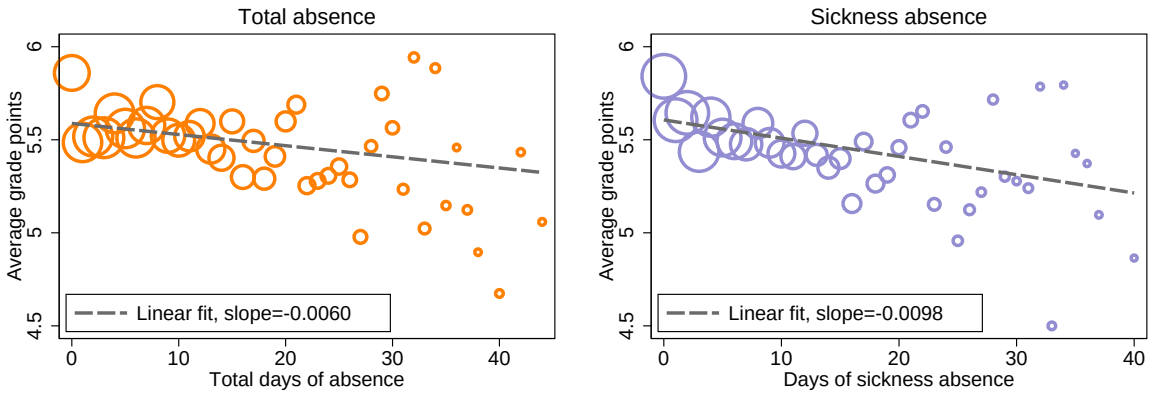
Notes: Own calculations based on the final sample of siblings. Age range gives the individual's age at which the variable is measured. Education is taken from the Census 1970 but is likely to refer to completed schooling for most individuals. Employment in 1960 and 1970 is taken from the Census information in these years. Labor market income 1970 and pensions 2002 are based on Census 1970 and tax registers, respectively, and measured in Swedish krona in the year the information refers to. The mortality information is taken from the Swedish Death Index.

refers to pension income mirroring previous labor market participation (in the baseline specification we do not consider non-labor market income). Looking at longevity, 16 per cent of women and 25 per cent of men in our final sample passed away before reaching the age of 70.

3.3 Correlations between absence, academic and socio-economic outcomes

To set the stage for the empirical analysis we document the associations between the number of days of absences and the outcomes of interest. As expected, the correlation between absence and academic performance is negative, see Figure 4. The linear fits indicate that it is more strongly negative for sickness absence than total absence.

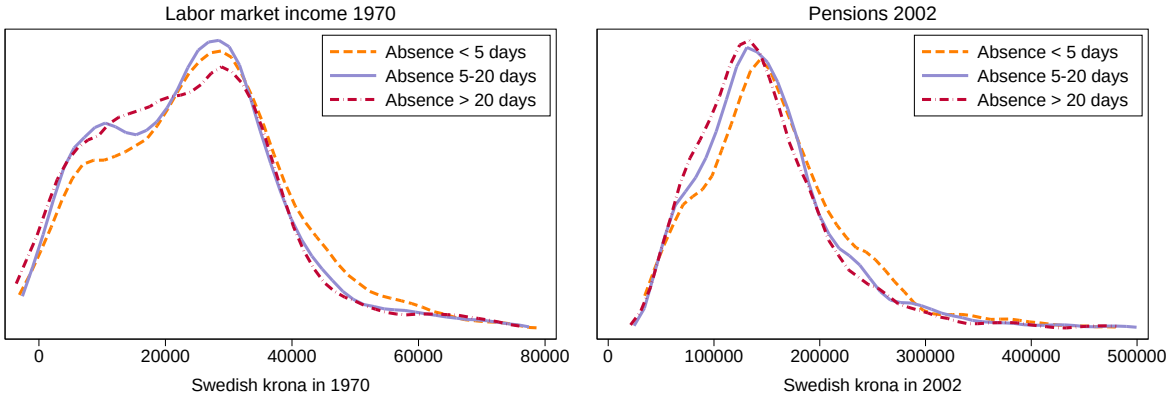
Figure 4: Descriptive relationship between (sickness) absence and performance



Notes: Own calculations based on exam catalog information. 8,938 observations. Grade points are collapsed on the integer of the days of absence. The size of the marker indicates the relative number of observations in the days-of-absence cell. Only cells with 15 or more observations are plotted. The fitted line is taken from a simple linear regression of performance on total absence and sickness absence, respectively, without restricting to the number of observations per cell.

Figure 5 shows the distribution of the income measures by grouped days of absence. While the visual difference between the income distributions for individuals who have missed below 5 days and between 5 and 20 days is rather small, individuals who missed more than 20 days because of sickness seem to earn less later in life. A Kolmogorov-Smirnov test for the equality of the distributions indicates that all conditional distributions but the 5-to-20-days and more-than-20-days distributions for income 1970 differ significantly at the 10 per cent level (see note to the figure). Figure 6 shows the survival rate of individuals who have missed less than 5 days, 5 to 20 days or more than 20 days. The differences between the lines are small (and statistically insignificant), although individuals who missed more than 20 days seem more likely to die younger.

Figure 5: Income distributions by total days of absence



Notes: Own calculation based on exam catalog, Census 1970 and tax register 2002 information. The Census labor market income is limited to values> 0. Using a Kolmogorov-Smirnov test for the equality of the distributions yields that the <5-days distribution of the 1970 income is statistically different from the 5-20-days distribution (corrected p-value 0.025) and the >20-days distribution (p-value 0.006). The 5-20-days and the >20-days distributions do not differ at the conventional levels (p-value 0.223). For 2002 pensions, all three conditional distributions differ statistically significant at the 10 per cent level (p-values for <5 days and 5-20 days: 0.007, <5 days and >20 days: <0.001, 5-20 days and >20 days: 0.086).

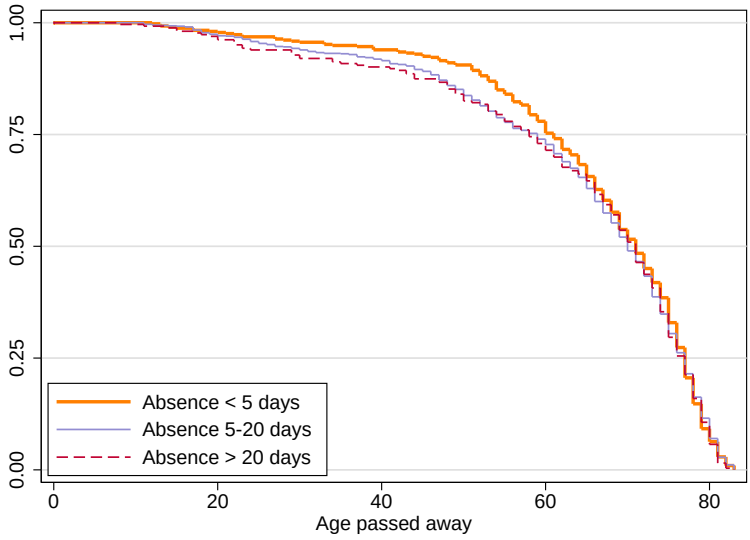
These figures show raw correlations and we should refrain from interpreting them as evidence of a causal link. Indeed, students who are more likely to miss school may also be those of lower ability or those of frailer nature. To start exploring the extent to which such selection may exist, Table 4 reports the average number of days of absence across groups of students defined by observable characteristics. Students whose father is a service worker are more likely to be absent than children whose fathers are agricultural and production workers. Children who have fewer siblings are also more likely to be absent. Based on the available observables it is not obvious whether we should expect students to select positively or negatively into absence. We now turn to the empirical strategy we propose to deal with the potential selection on unobservables.

4 Empirical strategy

4.1 The effect of absence on short- and long-term outcomes

The aim of our analysis is to estimate the causal effect of absence during elementary school on later outcomes, but absence is inherently endogenous as it likely relates to individual unobservable characteristics, including personal health. In addition, our 'treatment

Figure 6: Kaplan-Meier survival function by total days of absence



Notes: Own calculations based on exam catalog and Swedish Death Index information. A Kolmogorov-Smirnov test for the equality of the distributions indicates that the conditional distributions to not differ significantly (corrected p-values for <5 days and 5-20 days: 0.209, <5 days and >20 days: 0.985, 5-20 days and >20 days: >0.999).

variable' – number of days of absence – is a count variable, implying varying treatment intensity. As noted above, most absence days are due to illness, which means that our 'treatment' is in fact typically defined as the combination (*sick*, *absent from school*). Taken together this calls for a careful definition of the treatment effect (and counterfactual treatment) we are seeking to estimate.

In a standard model of a situation with a multi-valued treatment, we would denote potential outcomes under different treatment intensities w by $Y_i(w)$ (cf. Athey and Imbens, 2017), from which we may derive various treatment effects τ_{w_1,w_2} for different levels of treatment w_1 and w_2. Such a specification would require the assumption that the potential outcomes $Y_i(w)$ are insensitive to the source of variation in w. This is a reasonable approximation in many cases, but when most absence is due to illness this assumption may not be warranted. We therefore introduce a second argument, s, in the potential outcome function $Y_i(w, s)$ where s is the number of days of illness during the school year.[22]

[22]The days of illness s should not be confused with the days of sickness *absence*, which also depend on the choice of going to school or not when ill.

Table 4: Summary statistics of absence by type and individual characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Grade 1				Grade 4			
	All absences		Sickness absence		All absences		Sickness absence	
Overall	10.9	(12.0)	9.3	(10.8)	11.6	(11.9)	8.3	(10.3)
Gender								
Female	11.1	(11.9)	9.4	(10.7)	11.9	(12.8)	8.8	(11.3)
Male	10.8	(12.1)	9.2	(11.0)	11.2	(10.9)	7.8	(9.1)
Year of birth								
1930	10.2	(12.5)	8.7	(11.5)	12.0	(12.4)	8.5	(11.2)
1931	9.5	(11.4)	7.8	(9.4)	11.6	(11.6)	7.8	(10.3)
1932	10.9	(10.9)	8.9	(9.6)	12.3	(12.6)	7.9	(10.9)
1933	14.3	(13.2)	12.0	(12.0)	12.8	(12.5)	8.9	(10.1)
1934	11.2	(12.0)	9.2	(10.7)	11.6	(11.5)	8.3	(9.7)
1935	9.2	(11.2)	9.2	(11.2)	8.0	(9.1)	8.0	(9.1)
Occupation of father								
Agricultural worker	10.8	(11.5)	8.7	(10.0)	12.5	(12.0)	8.0	(10.1)
Production worker	10.8	(11.5)	8.7	(10.0)	12.5	(12.0)	8.0	(10.1)
Service worker	12.7	(14.1)	11.2	(13.3)	11.5	(10.7)	9.2	(10.0)
Number of siblings in sample								
0[a]	12.3	(13.8)	11.0	(13.3)	11.1	(11.6)	9.1	(10.7)
1	11.4	(12.2)	9.8	(11.0)	11.6	(11.8)	8.6	(10.1)
2 or more	10.3	(11.8)	8.7	(10.7)	11.5	(12.0)	7.9	(10.5)
Born out of wedlock								
yes	13.7	(13.8)	11.2	(11.6)	14.2	(15.0)	9.7	(11.4)
no	10.8	(11.9)	9.2	(10.8)	11.5	(11.7)	8.2	(10.3)
Born in hospital								
yes	11.3	(12.5)	9.9	(11.7)	11.4	(11.1)	9.0	(9.2)
no	10.9	(12.0)	9.3	(10.8)	11.6	(12.0)	8.2	(10.4)

Notes: Own calculations based on church records and exam catalog information. Observations: 8,938. Columns 1 and 5 give the mean value of the days of absence in total (that is, for all reasons) in grade 1 and 4, respectively. Columns 3 and 7 give the mean value of days of sickness absence in grade 1 and 4, respectively. Standard deviations are given in parentheses in even columns refer to the mean in the odd column on the left.

[a]Information based on the individual panel not restricted to siblings.

Having defined potential outcomes, we may define the causal effect we seek to estimate. In all specifications, we seek to estimate the incremental effect of one additional day of absence from school within a school year. This causal effect corresponds to

$$\tau = \sum_{w=1}^{W} \pi_{w-1} \mathbb{E}\left[Y_i\left(w, s_w\right) - Y_i\left(w-1, s_{w-1}\right)\right], \tag{1}$$

where π_{w-1} represents the empirical frequency of total absence days being equal to $w-1$. τ captures the effect of one additional day of absence averaged over the entire distribution of absence. However, we have not yet made any assumptions regarding s, the number of days the student is ill. Should we keep s constant when comparing different levels of absence, or should we allow it to adjust? When short-term perspectives are concerned, we probably do not want to keep s constant between different levels of w. Doing so would lead to the policy question "Should children go to school when ill?" rather than the seemingly more relevant policy question "Should we try to keep children healthy so that they do not miss school?". Put differently: in the short-term perspective, we may think of s as generating variation in absence days, based on which we can estimate the effects of absence.

For long-term outcomes, it is less clear that we want to allow s to vary in the definition of the treatment effect. Indeed, in the long-run perspective, we are more concerned that a health shock during elementary school may have persistent effects on health, which in turn would affect adult outcomes.[23] In terms of potential outcomes, we would have $Y_i(w,s) \neq Y_i(w,s')$ for $s \neq s'$, and any attempt to use variation in s to identify the effect of absence from school would also pick up an indirect effect operating via the dependence of adult health on childhood health.

Thus, for long-term outcomes, we would prefer to define the incremental effect of a day of absence as $Y_i(w,s) - Y_i(w-1,s)$ for some suitably chosen s. However, if the child's health is the main source of variation in w, it will be difficult to estimate such an effect in the data – because irrespective of the level at which we fix s, some combinations of (w,s) will be very rare in the data. In order to address this potential issue, we try to rule out the possibility that short-term variation in health has an independent effect on outcomes by comparing the estimated effects of absence due to different reasons. If sickness absence has a similar impact on outcomes as other types of absence, it seems safe to conclude that the main component of the treatment is not poor health, but rather the absence. Such

[23]The literature on the dynamics of child health suggests that shocks to a child's health have persistent effects. See for example: Currie and Stabile (2003), Contoyannis and Li (2011), Fletcher and Wolfe (2014) and Conti (2013).

an interpretation is plausible despite persistence in health as long as health persistence is related to unobservables (such as genetic traits or family background) that our empirical strategy adequately controls for.

4.2 Estimation

In order to estimate the incremental effect of absence days, denoted τ in equation (1), a natural starting point is to estimate a model in which the achievement of an individual i in grade g, denoted by y_{ig}, is assumed to depend linearly on the number of days he or she was absent from school in grade g, denoted by W_{ig}, a set of individual-specific controls $X_{1,i}$, a set of of school-specific controls $X_{2,ig}$, and a vector of parish fixed effects P_{ig}:

$$y_{ig} = \beta_0 + \tau W_{ig} + \beta_1 X_{1,i} + \beta_2 X_{2,ig} + P_{ig} + \varepsilon_{ig}, \tag{2}$$

where ε_{ig} captures the unobservables affecting student performance. Given our data, the vector X_1 includes students' characteristics taken from the church records: gender, full sets of year and month of birth dummies as well as interaction terms between the year and the month of birth, age-in-month fixed effects, mother and father's year of birth dummies, father's occupation at the time of birth, and indicators for whether the child was born out of wedlock, whether the child was born in hospital and whether he or she has a twin. The vector X_2 of school-specific factors includes an indicator for grade 4 (in the pooled specification), class size as well as lowest and highest grade taught to students in the same classroom.[24] Finally, the vector of parish fixed effects P_{ig} controls for time-invariant factors that are common to all students going to school in the same parish and that affect their performance in school, for instance compulsory years of education and term length.

The key problem with interpreting the OLS estimates of τ in equation (2) as the causal effect of days of absence is that days of absence likely correlate with the unobservables ε_{ig} and the exogeneity assumption $\mathbb{E}(\varepsilon|W, X_1, X_2, P) = 0$, which is necessary to interpret $\hat{\tau}$

[24]Class size is taken into account through spline variables. That is, we include variables that group the number of classmates in bins of five, where the bins for more than five classmates only give the marginal number relative to the previous bin.

as a causal effect, is likely to be violated. For example, days of absence may be correlated with unmeasured school factors, such as school resources and teacher quality, which we are not able to control. These factors are presumably positively correlated with performance and negatively correlated with absence as they determine students' engagement in school. Neglecting them would therefore overestimate the impact of absence in the OLS model.

A common approach to address this concern is to augment the above equation with school and teacher fixed effects, thus effectively relating the absences and performance of students attending the same school and taught by the same teacher. This is one of the strategies implemented by Goodman (2014) and Aucejo and Romano (2016). While this approach controls for all school-specific and teacher-specific time-invariant factors that may be confounding the effect of absence on performance, there may well be other individual-specific unobservable characteristics that distinguish students who are more frequently absent than others. If, conditional on the observables included in the model, these characteristics are correlated with performance in school, the effect of days of absence on performance will still be biased. Students who are less able, less motivated or whose parents place less emphasis on education may be absent more frequently. If these students also perform worse in school, then this unobserved difference result in an downward bias in the effect of days of absence on performance.

To address this further concern, we take advantage of two key features of our dataset: that we observe sibling pairs and that we observe students' absence and performance twice. Exploiting the fact that we observe sibling pairs we augment equation (2) not only with school and teacher fixed effects, but also with family fixed effects. That is, our estimating equation becomes the following siblings fixed effect model:

$$y_{i(f),g} = \beta_0 + \tau W_{i(f),g} + \beta_1 X_{1,i(f),} + \beta_2 X_{2,i(f),g} + P_{i(f),g} + S_{i(f),g} + T_{i(f),g} + \lambda_f + \varepsilon_{i(f),g}, \quad (3)$$

where $S_{i(f),g}$ is the school fixed effect, $T_{i(f),g}$ the teacher fixed effect, and λ_f the family fixed effect for individual i in family f. This design controls for any unobserved individual characteristics that has the same additive effect on outcomes of both siblings. While siblings fixed effects remove innate genes and other family-constant factors with certainty,

parental involvement could per se differ between siblings. Given that siblings in our sample are born in a relatively tight time span of five years (1930 to 1935) the underlying parenting style is less likely to differ across offspring compared to siblings born farther apart. Moreover, in the time period we study, parental involvement in their offspring's education was quite low in Sweden (Fredriksson, 1971b).

When analyzing the impact of absences on short-term attainment, we strengthen the strategy even further by exploiting the fact that we observe student's absence and performance twice. We pool observations on grade 1 and grade 4 for each individual and include individual fixed effects in the estimating equation:

$$y_{i(f),g} = \beta_0 + \tau W_{i(f),g} + \beta_1 X_{1,i} + \beta_2 X_{2,i(f),g} + P_{i(f),g} + S_{i(f),g} + T_{i(f),g} + \alpha_{i(f)} + \varepsilon_{i(f),g}, \quad (4)$$

where α_i is an individual fixed effect. This design controls for any unobserved individual characteristic that has the same linear effect on achievement in grade 1 and grade 4. Even if unobserved ability, motivation or parental taste for education differs between siblings, α_i will absorb this as long as the difference is constant between grades 1 and 4.

To implement this equation we effectively estimate the following within-student model:

$$\Delta y_{i(f)} = \tau^{FE} \Delta W_{i(f)} + \beta_2^{FE} \Delta X_{2i(f)} + \Delta P_{i(f)}^{FE} + \Delta S_{i(f)}^{FE} + \Delta T_{i(f)}^{FE} + \Delta \varepsilon_{i(f)},$$

with $\Delta y_{i(f)} \equiv y_{i(f),4} - y_{i(f),1}$, $\Delta W_{i(f)} \equiv W_{i(f),4} - W_{i(f),1}$, etc., and $\Delta \varepsilon_{i(f)} \equiv \varepsilon_{i(f),4} - \varepsilon_{i(f),1}$. The intercept β_0, the vector of time-constant observables X_1 as well as the time-constant unobservables α_i will be removed from the estimation. The parish, school and teacher fixed effects will only be identified from students that move to another parish, switch schools and/or are assigned to a new teacher between grades 1 and 4. As an individual always belongs to the same family, the individual fixed effects model nests siblings fixed effects at the same time. As long as $\mathbb{E}(\Delta \varepsilon_i | \Delta (W, X_2, P, S, T)) = 0$, $\hat{\tau}$ will be unbiased.

While the individual fixed effect strategy is arguably more valid than a strategy only relying on within-school or within-teacher variation, it is not without limitations. First, it requires us to assume that the effect of absence on performance is the same in grade

1 and grade 4. *A priori*, it is unclear if these effects are the same, but in Section 5 we present suggestive evidence supporting this assumption.

Second, the individual (siblings) FE estimates do not recover τ if there are individual-specific (family-specific) factors of student achievement that vary over time (across siblings) and are correlated with the student's absence. An example of such a threat to the identification would be changes in class size – which may lead to increased absence and to changes in student performance. In order to address this potential issue, we include class size as a control variable. Another issue is dynamic parental investments: parents may adjust their own inputs in response to a child's absence which can lead to biased estimates. In our case, however, spillover effects of this kind can only occur in the rare cases that two siblings are observed in the same school year. Nevertheless, we cannot rule out all threats to identification. To address this, we provide in Section 6 a comprehensive sensitivity analysis where we bound our estimates against unobservable confounders and complement our main analysis with an instrumental variables (IV) strategy.

5 Estimation results

5.1 Short-term effects of absence in school

Table 5 reports the pooled estimates of the effect of days of absence in grade 1 and grade 4 on performance in the same grade.[25] The rows of the table indicate the different ways of measuring average performance. In column 1, we regress average performance on days of absence and control variables (including parish FE, but not family or individual FE). These estimates show that one additional day of absence is significantly associated with a 0.35 per cent of a standard deviation (SD) decrease in average performance. Once school, teacher and siblings fixed effects are added in the model (column 2), the negative effect of days of absence becomes slightly larger in magnitude, 0.40 per cent of a SD. This is in line with the hypothesis that students whose parents invest less in their children's education may also be more likely to miss school, but the fact that these estimates are so

[25] Appendix Table A2 reports the coefficient estimates associated with all the control variables included in these specifications

close to the OLS estimates suggests that selection on family unobservables may not be very important in this context.[26] When including individual fixed effects to the model (column 3), the point estimate returns to the magnitude of 0.33 per cent of a SD.[27]

The effect of absence is of moderate size and statistically significant at the 1 per cent level in all specifications. Assuming linearity, the effect of 10 days of absence – about the average in our sample – corresponds to around 3 per cent of a SD in student performance. Interestingly, despite analyzing absence in a very different context and literally in another century, our results measured in SD units are comparable to those in Goodman (2014) and Aucejo and Romano (2016). Using a similar identification strategy with recent US data Goodman (2014) finds an effect 0.8 per cent of a SD in math and English and Aucejo and Romano (2016) find effects of 0.55 per cent of a SD in math and 0.29 per cent in reading in their preferred specifications.

This effect size is also comparable to what has been found in the literature examining the effect of teacher quality on performance. For instance, Chetty et al. (2014) find that a one SD increase in teacher Value Added improves pupils' math test scores by 0.14 SD and English test scores by 0.1 SD. Rivkin et al. (2005) find that a one-year increase in teacher experience increases pupil performance by up to 0.13 SD in math and 0.06 SD in reading.

An advantage of our study is that we can anchor student performance, which is measured on a somewhat arbitrary scale, to adult outcomes. In other words, we can translate the short-term effect of absence on school performance into its effect on earnings potential. This still measures the short-term effect of absence, but in a unit (SEK) that is economically more meaningful than standardized grade points. In the individual FE specification the impact of ten additional days of absence on school performance translates into a

[26]The same seems to hold true if we include school and teacher FE stepwise in Table A2 in the Appendix. Given the conditioning variables in the OLS model, particularly the full sets of year-of-birth, age and parish indicators, the coefficient of absence does not change noteworthily across the school, teacher and siblings FE specifications.

[27]We restrict the individual FE sample to contain individuals with siblings in the dataset. The Online Appendix shows the effect when all observations are used. The point estimates are quite similar, suggesting that siblings and singletons react in the same way to an additional day of absence.

Table 5: Baseline results

	(1) OLS	(2) Sibl. FE	(3) Indi. FE
Average grade points in units of SD (mean: 0, SD 1)			
Days of absence	-0.0035^{***}	-0.0040^{***}	-0.0033^{***}
	(0.0009)	(0.0011)	(0.0012)
Average grade points in units of pension 2002 (mean 2002 pension in sample: 150,816 SEK)			
Days of absence	-56.1501^{***}	-39.7389^{*}	-45.1403^{**}
	(14.7496)	(20.2815)	(20.5733)
Conventional controls			
Time-variant	✓	✓	✓
Time-invariant	✓	✓	
Fixed effects			
Socio-economics	✓	✓	
Parish	✓	✓	✓
School		✓	✓
Teacher		✓	✓
Siblings		✓	
Individual			✓
# observations	8,938	8,938	8,938
# families/individuals		1,988	4,469

Notes: Each cell states the coefficient of days of absence for a separate regression. The rows give different measures of the dependent variable average grade points. In the first row average performance over math, reading and speaking, and writing is standardized with mean 0 and standard deviation 1. The second row measures average grade points in units of pensions 2002, see the data description in the text for details. Time-variant conditional variables: grade, range of grades instructed in the same classroom, length of the school year in weeks. Time-invariant conditional variables: female, born out of wedlock, twin birth, mother employed at the time of birth, born in hospital. Socio-economics fixed effects include full sets of fixed effects for the year and month of birth, year and month interactions, age, parent's year of birth, and the family's socio-economic status based on the first-digit HISCO code of the father. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

decrease in earnings potential of 451 SEK (in values of 2002). Given that the average pension is about 150,000 SEK, this effect seems rather humble.

5.2 Long-term effects of absence in school

Table 6 reports the effects on our long-term outcomes of the average number of absences across grades 1 and 4 in columns 1 (OLS) and 2 (siblings FE) as well as the effects of

the number of absences in each grade by grade in columns 3 (OLS) and 4 (siblings FE), respectively.

Our estimates suggest there is a robust negative effect of absence on secondary school enrollment, but we are unable to attribute it to a certain school grade. The point estimates are negative for both grades and of at most 0.1 percentage points – which can be compared to a baseline probability of 12.5 per cent. For comparison, our estimates of the effect of school performance, as measured by average grade points, on secondary schooling enrollment range between 0.05 and 0.13 (results available upon request). Multiplying this estimate with the estimated effect of absence on school performance of around -0.003 (cf. Table 5 above), we would expect an effect of absence on enrollment of between -0.0003 and -0.0005. Our estimates of the impact of absence on enrollment are in general larger and typically twice as large as this indirect estimate, even though the differences between the two are insignificant.

Turning to employment, the results are mixed. When it comes to early-career employment (as measured in the 1960 Census when individuals are between 25 and 30 years old), our results suggest that there is a negative relationship driven by absences in grade 1. Indeed, when siblings FE are included in the model, ten days of absence in grade 1 leads to a decrease in the probability of being employed at ages 25-30 by 2.6 percentage points, and this estimate is strongly significant. Relative to the average employment at the time (65 per cent), this corresponds to a 4 per cent reduction, which is a rather large effect. On the other hand, for employment in 1970 (at ages 35-40) the estimates again have the expected negative sign, but the point estimates are smaller and noisier. The 95 per cent confidence interval for the effect of ten days of absence (across both grades) on employment in 1970 ranges between -0.017 and -0.002, suggesting that there may still be a negative effect of absences on employment at 35-40, but this effect is unlikely to be large and is rather getting smaller over time.

With respect to labor market earnings, the association between income and absence is negative for most specifications. For 1970 labor market income the estimates are not

Table 6: Long-term effect of absence in school

	(1)	(2)	(3)	(4)
	Pooled effect		Separate effects	
	OLS	Sibl. FE	OLS	Sibl. FE
More than Folkskola *(1=yes)*				
Total abs. (avg. both grades)	−0.0012	−0.0013		
	(0.0010)	(0.0009)		
Total absence in grade 1			−0.0003	−0.0011*
			(0.0006)	(0.0007)
Total absence in grade 4			−0.0010*	−0.0002
			(0.0006)	(0.0005)
Employment 1960 (1=yes)				
Total abs. (avg. both grades)	−0.0016**	−0.0016*		
	(0.0007)	(0.0009)		
Total absence in grade 1			−0.0014**	−0.0026***
			(0.0005)	(0.0006)
Total absence in grade 4			−0.0003	0.0008
			(0.0007)	(0.0008)
Employment 1970 (1=yes)				
Total abs. (avg. both grades)	−0.0010*	−0.0020		
	(0.0006)	(0.0015)		
Total absence in grade 1			−0.0006	−0.0008
			(0.0005)	(0.0009)
Total absence in grade 4			−0.0004	−0.0012
			(0.0006)	(0.0009)
Labor market income 1970				
Total abs. (avg. both grades)	−54.4802	−13.9569		
	(36.9846)	(57.2241)		
Total absence in grade 1			−18.5468	−1.3469
			(21.2133)	(40.3369)
Total absence in grade 4			−36.0149	−12.2763
			(23.2125)	(28.5027)
Pensions 2002				
Total abs. (avg. both grades)	−124.6255	−186.0209		
	(126.1594)	(264.0324)		
Total absence in grade 1			0.1536	46.2308
			(125.6788)	(169.5014)
Total absence in grade 4			−120.9218*	−215.0359
			(72.6683)	(147.7769)
Passed away before the age of 70 (1=yes)				
Total abs. (avg. both grades)	−0.0001	0.0011		
	(0.0008)	(0.0008)		
Total absence in grade 1			0.0002	0.0008
			(0.0005)	(0.0006)
Total absence in grade 4			−0.0003	0.0003
			(0.0006)	(0.0006)

Notes: Number of observations: More than *Folkskola* 3,019 (in 1,373 families), employment 1960 3,902 (1,750), employment 1970 4,469 (1,988), income 1970 2,137 (985), pensions 2002 2,363 (1,080), passed away before age 70 4,469 (1,988). Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

statistically significant and rather small in size – in most specifications ten days of absence correspond to less than 1 per cent of the average income. Besides the long time horizon (27 to 32 years after grade 1 and 24 to 29 years after grade 4), the lack of a relationship between absence and income at ages 35-40 may also be due to the Swedish wage structure being extremely compressed at this time, so that individual productivity had a very limited impact on earnings (Bhalotra et al., 2016). For 2002 pensions the pooled and fourth-grade estimations are negative (while the coefficients for grade 1 absences are close to zero). The OLS estimate of the effect of grade 4 absence on pensions 2002 is significant at the 10 per cent level. When adding siblings fixed effects, the point estimate increases in magnitude, but becomes less precise.[28] Finally, for mortality the coefficients alternate around zero and do not exceed 0.01 percentage points. This indicates that there is no effect on mortality.[29]

Overall, the results presented so far suggest that absence in elementary school has a robust negative impact on short-term performance of small but non-negligible magnitude. Absences in elementary school also have a detrimental impact on early-career employment, but this effect fades out with increased labor market experience. Broadly speaking, the findings are very much in line with those of Pischke (2007) and Dustmann et al. (forthcoming) who find that initial differences in the quantity and quality of schooling have no long-lasting labor market effects. The fact that the effect of absence on labor market outcomes are significant early in the career but not later on underscores the value of having access to data on outcomes at different points of the life-cycle in order to get an unbiased perspective of the full impacts of absence.

That being said, it is possible that the above results mask some heterogeneity between certain groups or for specific causes of absences (e.g., sickness absences) or that absence only has adverse effects if it occurs with large frequency (non-linearities). We now explore these different margins.

[28] In the baseline specification we only consider income and pension values >0. The Online Appendix gives the results for alternative income measures. The alternative measures do not change our interpretation of the baseline results.

[29] The missing association between absence and longevity indicates that sample attrition due to selective mortality is not a major concern when interpreting the other long-term effects.

5.3 Heterogeneity

Subgroup analysis Table 7 reports the estimates for academic performance in elementary school where we allow the impact of the total number of days of absence to vary between males and females, as well as between children of agricultural and non-agricultural workers. For each panel, the first row reports the main effect of the number of days of absence, while the second row reports the coefficient on the interaction. Overall, we find little evidence of heterogeneous impacts. With respect to gender heterogeneity, the individual FE model estimates suggest that the effect of absences may be worse for men than for women's achievement, but the difference between the two groups is statistically not different from zero.[30] While absence is more strongly negatively correlated with the performance of children of agricultural workers than with the that of other children, differences in the impact of absences wash away once we account for unobserved heterogeneity at the family level. This lack of a clear effect heterogeneity along the socio-economic status is in line with Goodman (2014), while Aucejo and Romano (2016) find evidence of some heterogeneous impacts between students of different abilities. This result also underlines the educational and societal context of the analysis. In the setting we investigate, textbooks were, for instance, provided by the parish if families could not afford them otherwise.

Table 8 reports the results of a similar subgroup analysis for our long-term outcomes. As the difference in labor market participation and earnings between men and women is quite substantial in the time under review (see Table 3), interacting pooled days of absence with the female indicator yields rather remarkable findings. Although not statistically significant, ten days of absence associate with a decrease in probability of being employed in 1960 for women by 6.6 per cent, while the corresponding number males is less than one-tenth.[31] The long-term effect of absence on pensions in 2002 seems to be stronger for females as well. Looking at the social gradient along the father's occupation indicates

[30] One should keep in mind the gender difference is only identified through a variation in days of absence in the individual FE model. Similarly, the interaction term between father's occupation and absence in the individual FE model is only identified through variation in absence.

[31] For women the effect is $\big((10 \text{ days} \times -0.0004) + (10 \text{ days} \times -0.0020)\big)/0.3651$ baseline probability $= -0.0657 \approx -6.6$ per cent and for men $(10 \times -0.0004)/0.9396 = -0.0043 \approx -0.4$ per cent.

Table 7: Heterogeneity in the short-term effects by subgroup

	(1) OLS	(2) Sibl. FE	(3) Indi. FE
Gender			
Absence	-0.0028^{*}	-0.0037^{**}	-0.0043^{***}
	(0.0014)	(0.0016)	(0.0013)
Absence×female	-0.0013	-0.0004	0.0017
	(0.0016)	(0.0018)	(0.0016)
Father's occupation			
Absence	-0.0024^{**}	-0.0041^{***}	-0.0031^{**}
	(0.0011)	(0.0012)	(0.0016)
Absence×agri. worker	-0.0027^{*}	0.0002	-0.0005
	(0.0015)	(0.0017)	(0.0026)
Grade			
Absence	-0.0040^{***}	-0.0047^{***}	
	(0.0014)	(0.0014)	
Absence×grade 1	0.0009	0.0014	
	(0.0016)	(0.0017)	

Notes: Each panel states the coefficient of total days of absence as well as of an interaction between total days of absence and the subgroup indicator. 8,938 observations. Control variables as in the baseline specification. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

some heterogeneity on this dimension, where the effect of absences on 1970 income is significantly larger for children of agricultural workers than other children.

Grade-specific effects An assumption underlying our individual FE strategy is that the impact of absence in grade 1 and in grade 4 is the same. While we cannot test for this assumption in the context of the individual FE model, we present suggestive evidence that this assumption holds by estimating a model where we allow for different effects of absence in grade 1 and grade 4 and control for teacher, school and siblings FE. The results presented in the bottom panel of Table 7 indicate that the effect of absence in grade 1 cannot be statistically distinguished from the effect of absence in grade 4 on any academic performance measure.

Table 8: Heterogeneity in the long-term effects by subgroup

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable					
	> *Folk-skola*	Empl. 1960	Empl. 1970	Income 1970	Pensions 2002	Passed away≤70
Gender						
Absence	−0.0023	−0.0004	−0.0021	−8.9994	−28.2702	0.0011
	(0.0016)	(0.0011)	(0.0020)	(56.4558)	(327.9810)	(0.0016)
Absence×female	0.0016	−0.0020	0.0002	−10.0480	−277.5613	0.0001
	(0.0017)	(0.0017)	(0.0018)	(64.3459)	(287.1122)	(0.0017)
Father's occupation						
Absence	−0.0008	−0.0021**	−0.0022	70.2782	−315.6315	0.0015
	(0.0013)	(0.0010)	(0.0016)	(60.2351)	(280.5521)	(0.0012)
Abs.×agri. worker	−0.0011	0.0011	0.0005	−192.8888***	328.3218	−0.0010
	(0.0014)	(0.0017)	(0.0033)	(54.5382)	(357.2552)	(0.0019)

Notes: Each panel states the coefficient of total days of absence (average over grades 1 and 4) as well as of an interaction between total days of absence and the subgroup indicator. Number of observations: More than *Folkskola* 3,087 (in 1,396 families), employment 1960 3,904 (1,751), employment 1970 4,471 (1,989), income 1970 2,139 (986), pensions 2002 2,365 (1,081), passed away before age 70 4,471 (1,989). Dependent variables defined as in the baseline long-term results. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

5.4 Sickness vs. non-sickness absences

Our data allows us to distinguish between absences due to sickness and absences due to other reasons. Table 9 reports the estimates of the short-term effects of days of absences on attainment, when we allow the effect to be different for absences due to sickness and absences due to other reasons. At the bottom of the table, we report in brackets the p-value of a F-test that the two coefficients of interest are equal to each other. Comparing the effect of sickness absence with non-sickness absence across the different specifications (including school and teacher FE) reveals an interesting pattern. In the OLS model and the individual FE model (columns 1 and 5 in Table 9), the estimated effects of both types of absence are similar in magnitude and statistically undistinguishable from each other. If we only compare students who have the same teacher (and, thereby, are in the same school and generally in the same class) to each other in column 3, the association between non-sickness absence and performance is more than twice as strong as the association between

sickness absence and performance (the coefficients differ at the 10 per cent level). This pattern suggests that the association between non-sickness absence and performance is driven by family-level or individual factors.

A candidate for such a factor may be behavioral issues that cause truancy (which is reported as non-sickness absence). Accounting for unobserved behavioral problems either through within-family or within-individual comparison yields rather similar results – at least when compared to the effect in the teacher FE specification. If there are no time-varying behavioral problems (or other time-varying unobservable confounders of non-sickness absence), the coefficient of days of non-sickness absence in column 5 may as well be a reasonable approximation for the effect absence days w while holding the number of days in illness s constant. Taking up the discussion of what an ideal experiment may look like in Section 4.1, the similar coefficients of sickness and non-sickness absence in column 5 (the p-value is close to 1) thus lends support to the notion that the reduced performance associated with absence is driven by the absence in itself, and not by the student's health. The fact that only the impact of sickness absence is significantly different from zero might be driven by the fact that there is a lot more variation in this variable than there is in days of non-sickness absence.[32]

A similar exercise for our long-term outcomes reveals very comparable patterns. Table 10 reports the OLS (first panel) and siblings FE (second panel) coefficients of sickness absence and non-sickness absence for grades 1 and 4. Interestingly, the negative association between total absence and having more than *Folkskola* education seems driven by non-sickness absence – even though sickness is the main cause for overall absence. This supports the view that non-sickness absence is driven by behavioral problems, which teachers account for when recommending students for *Realskola* enrollment. Once behavioral problems on family level are partialled out (through the inclusion of siblings FE), the negative association between absences with education vanishes.

[32] Aucejo and Romano (2016) distinguish between excused and unexcused absence. They do not compare the effect size across different FE models, but in line with our findings, their estimated effect of unexcused absence exceeds the effect of excused absence.

Table 9: Short-term effects – total absence vs. sickness absence

	(1) OLS	(2) School FE	(3) Teacher FE	(4) Sibl. FE	(5) Indi. FE
Average grade points in units of SD					
Days of sickness absence	-0.0037^{***}	-0.0042^{***}	-0.0036^{***}	-0.0044^{***}	-0.0034^{**}
	(0.0009)	(0.0011)	(0.0010)	(0.0011)	(0.0014)
Days of non-sickness absence	-0.0027	-0.0039^{**}	-0.0075^{***}	-0.0017	-0.0031
	(0.0018)	(0.0017)	(0.0022)	(0.0024)	(0.0025)
	[0.5789]	[0.8444]	[0.0793]	[0.2651]	[0.9319]
# observations	8,938	8,938	8,938	8,938	8,938
# individuals/families		749	1,259	1,988	4,469

Notes: See note to the baseline results table. The brackets at the bottom of the table give the p-value of a F-test of equality of the coefficients of sickness and non-sickness absence. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

Both the OLS and siblings FE specifications point to the fact that both types of absences significantly decrease early employment. Although the effect of non-sickness absences is more negative than that of sickness absences, we cannot distinguish the two from each other at conventional levels of significance. Regarding employment in 1970, the absence coefficients are negative but neither statistically different from each other nor from zero. The income measures are only negatively associated with non-sickness absence in the OLS estimates. To the extent that non-sickness absence reflects a bolder behaviour, these results that suggest there might be a wage premium for being "pushy".

5.5 Non-linearities

While a student may be able to compensate a few days of absence, this may not be possible for a longer period of absence. This would result in a non-linear relationship between absence and educational performance. To investigate the presence of non-linearities we run the individual FE specification (similar to column 3 in Table 5) where we define the treatment as an indicator for whether the number of days of absence exceeds a certain threshold, where we vary this threshold between 1 and 50. The dots in Figure 7 give the coefficients associated with such indicator along the different threshold values on the

Table 10: Long-term effects – total absence vs. sickness absence

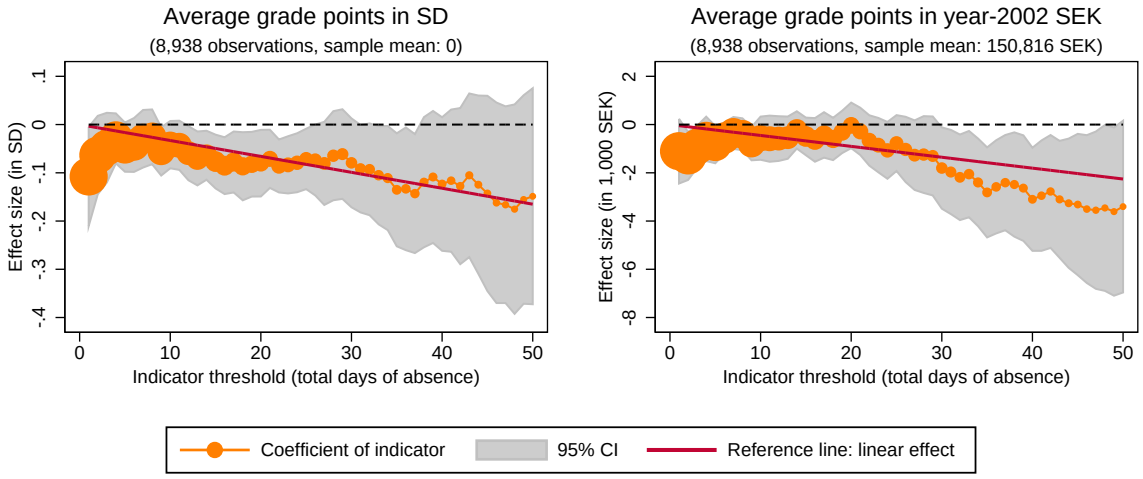
	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable					
	> *Folk-skola*	Empl. 1960	Empl. 1970	Income 1970	Pensions 2002	Passed away≤70
OLS estimation						
Sick. abs. grade 1	0.0002	−0.0011*	−0.0005	−16.0216	−24.3926	−0.0000
	(0.0007)	(0.0007)	(0.0006)	(17.1981)	(139.5187)	(0.0005)
Non-sick. abs. grade 1	−0.0025**	−0.0028*	−0.0013	−33.0191	158.5347	0.0014
	(0.0010)	(0.0014)	(0.0015)	(78.3599)	(380.7130)	(0.0013)
	[0.0289]	[0.3363]	[0.6589]	[0.8195]	[0.6619]	[0.3175]
Sick. abs. grade 4	−0.0003	0.0002	−0.0005	−36.6376	−88.3797	0.0004
	(0.0007)	(0.0007)	(0.0005)	(27.8351)	(91.9464)	(0.0006)
Non-sick. abs. grade 4	−0.0035***	−0.0020	−0.0002	−32.5383	−255.0109	−0.0031***
	(0.0011)	(0.0013)	(0.0015)	(50.2889)	(238.1689)	(0.0009)
	[0.0253]	[0.1076]	[0.8157]	[0.9462]	[0.5525]	[0.0004]
	{0.0280}	{0.1316}	{0.9465}	{0.8493}	{0.8239}	{0.0012}
Siblings FE						
Sick. abs. grade 1	−0.0013	−0.0022***	−0.0008	−25.7318	−15.8283	0.0004
	(0.0008)	(0.0008)	(0.0010)	(42.5099)	(198.1435)	(0.0006)
Non-sick. abs. grade 1	−0.0004	−0.0055**	−0.0009	138.0634	443.4685	0.0032**
	(0.0010)	(0.0022)	(0.0024)	(101.3751)	(397.9748)	(0.0016)
	[0.5476]	[0.2390]	[0.9675]	[0.1317]	[0.3465]	[0.0775]
Sick. abs. grade 4	−0.0006	0.0011	−0.0011	−39.8570	−272.8657	0.0008
	(0.0007)	(0.0008)	(0.0009)	(25.4238)	(174.1033)	(0.0007)
Non-sick. abs. grade 4	0.0016*	−0.0009	−0.0017	81.6146	56.3521	−0.0018
	(0.0010)	(0.0027)	(0.0027)	(65.2114)	(249.8072)	(0.0015)
	[0.1149]	[0.4702]	[0.8370]	[0.0692]	[0.3066]	[0.1160]
	{0.1381}	{0.0003}	{0.9703}	{0.1736}	{0.2434}	{0.2317}

Notes: Number of observations: More than *Folkskola* 3,087 (in 1,396 families), employment 1960 3,904 (1,751), employment 1970 4,471 (1,989), income 1970 2,139 (986), pensions 2002 2,365 (1,081), passed away before age 70 4,471 (1,989). The brackets give the p-value of a F-test of equality of the coefficients of sickness and non-sickness absence in the respective grade. The braces state the p-value of a F-test of equality of all four coefficients. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

x-axis.[33] The size of the dots indicates the relative number of observations for which the indicator is 1. While 95 per cent out of the c.a. 9,000 student-grade observations exhibit

[33]Because the number of students that miss a large number of days in one school year is often rather low (e.g., only 12 students miss exactly 45 days), it is not meaningful to regress performance on a full set of binary indicators for each number of days in a single regression. The Online Appendix gives the results of a regression using indicator variables that bin days of absence. The coefficients of the indicator variables lie around the linear effect reference line.

Figure 7: Non-linearities in the short-term effect of absence for different threshold values

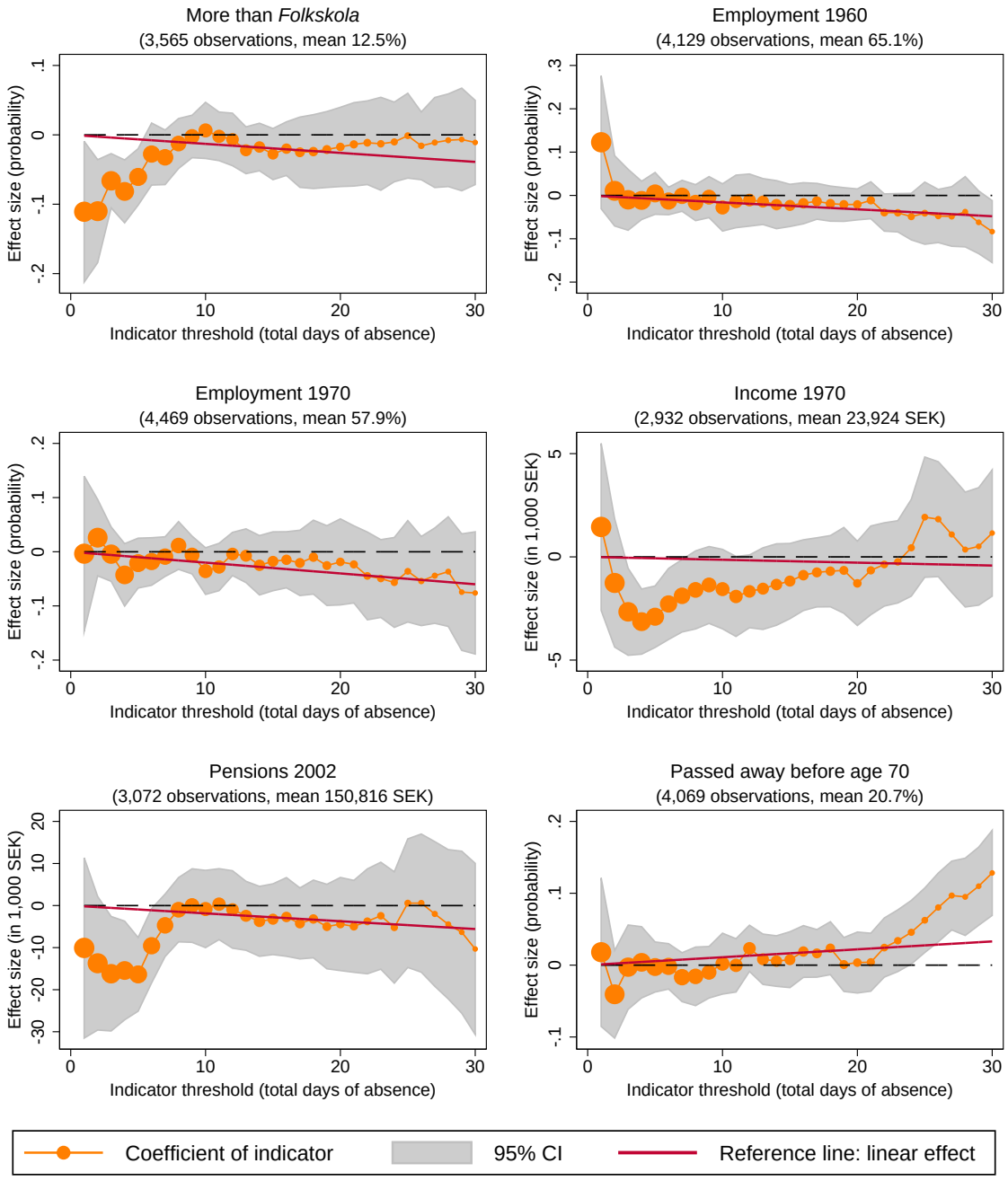


Notes: This graph plots the coefficient of a regression of performance in school on a binary indicator for total days of absence. In the left plot performance is measured in units of standard deviations and in the right plot in year-2002 SEK. The indicator threshold is given on the x-axis. The size of the coefficient markers depicted through orange dots is proportional to the number of observations for that the indicator is 1. Out of the 8,942 student-grade observations 95 per cent have at least one day of absence (leftmost orange dot in the plots) and 2 per cent missed 50 or more days in one school year (rightmost orange dot in the plots). The gray area indicates the 95 per cent significance band of the coefficient estimates. The red line depicts the linear effect of an additional day of absence taken from the baseline model.

at least one day of absence, less than 2 per cent of the student-grade observations have 50 or more days. Naturally, the 95 per cent confidence interval of the point estimators – depicted in gray – increases with the threshold value. The red reference line depicts the linear effect of the baseline model as of Table 5 multiplied by the number of absent days.

A comparison of the orange and red lines indicates that the per-day effect in SD (left plot) of the non-linear estimations using the binary indicators does not substantially differ from the linear effect. If short-term performance in school is measured in SEK in the year 2002, the non-linearly estimated effect exceeds the linear effect for more than 30 days of sickness absence. That is, only if an individual is absent more than 30 days in one school year, the effect of absence increases disproportionately. Regardless of how we measure the outcome variable, the deviation from the linear trend is never significantly different. Given that less than 6 per cent of all student-grade observations have 30 or more days of absence, non-linearities do not seem to play an important role. This finding is in line with Aucejo and Romano (2016) who neither find evidence of non-linearities.

Figure 8: Non-linearities in the effect of average days of absence in both grades on long-term outcomes using individual fixed effects



Notes: This graph plots the coefficient of a regression of the long-term on a binary indicator for average days of total absence over grades 1 and 4. The indicator threshold is given on the x-axis. The size of the orange coefficient plot is proportional to the number of observations for that the indicator is 1. The gray area indicates the significance band of the coefficient estimates. The red line depicts the linear effect of an additional day of absence in the baseline specification.

Figure 8 reports the results of a similar exercise for our long-term outcomes.[34] Overall, there is no strong evidence that absence has non-linear impacts on final educational achievement, adult employment, income or mortality.

6 Sensitivity analysis

The estimates in Tables 5 and 9 are remarkably stable across specifications, suggesting that selection on unobservables into absence may not be a salient feature. However, the FE estimates could still be biased if there are unobservable factors that correlate with individual absence and student outcomes *and* that vary over time or across siblings. To address this caveat and show the robustness of our results, we implement a bounding approach where we bound the influence of omitted variables. We also present results from an instrumental variables approach using local weather shocks as instrument for absences.

6.1 Bounds

We employ the bounding approach suggested by Oster (forthcoming), building on the idea of Altonji et al. (2005). Our goal is to bound the effect of absence assuming that the selection on unobservables is as strong as the selection on observables. We consider the case where the selection on unobservables is in the same or the opposite direction as the selection on observables, thus allowing the true effect to be overestimated or underestimated. The exercise is only helpful if the observables are informative with respect to the selection, so we control for a large array of control variables and the full set of siblings fixed effects. This removes factors such as constant family resources, parental preferences and genetic endowment that are likely negatively correlated with absence and positively correlated with performance. As omitting these factors would likely cause an upward bias that challenges our implications, the bounding approach seems particulary useful in the application at hand.

[34] Because of the fewer observations for the long-term outcomes, we only run the absence indicator up to the threshold of 30 or more days of absence.

The starting point is to compare the coefficient of absence in the baseline model ($\tilde{\beta}$), with the coefficient of absence in a simple linear regression of the dependent variable on absence and an intercept ($\dot{\beta}$). Formally, the bound around the coefficient of absence β^* is:[35]

$$\beta^* \approx \tilde{\beta} - \delta(\dot{\beta} - \tilde{\beta})\frac{R_{max} - \tilde{R}}{\tilde{R} - \dot{R}},$$

where the degree of proportionality of selection on observables to selection on unobservables, δ, is either set to 1 (unobservable selection goes into the same direction) or -1 (unobservable selection is into the adverse direction). In a second step, the movement in the coefficient of absence, $\dot{\beta} - \tilde{\beta}$, is re-scaled by the movement in the R^2 relative to the potential change in the R^2 (where $\tilde{R}$ and $\dot{R}$ denote the R^2 of the baseline model and the simple regression, respectively, and R_{max} denotes the highest possible value of the R^2).[36]

Table 11 shows the bound estimates for the short- and long-term effects of days of absence. For the baseline model estimate $\tilde{\beta}$ we take the siblings FE specification in order to be able to calculate the bounds separately for each grade.[37]

For the short-term effect on performance in grade 1 the estimated β^* for $\delta = 1$ is -0.0078. Thus, if unobservable selection is proportional to and goes in the same direction as the observable selection an additional day of absence associates with a decrease in performance by 0.78 per cent of a SD. If instead assuming the same amount of selection but in the opposite direction ($\delta = -1$), the upper bound is -0.0017. Even if our baseline model fails to account for selection on unobservables that affect the outcome as strongly as family-level time-invariant characteristics, the effect of ten days of absence would be a decrease performance by 1.7 per cent of a SD. For the short-term effects for absence

[35]This expression is only an approximation, see Oster (forthcoming) for the exact calculation. To calculate the bounds we use the Stata ado-file `psacalc` provided online by Emily Oster. All errors are our own responsibility.

[36]Following Hener et al. (2016) we consider as $R_{max} = \min(2.2 \times \tilde{R}, 1)$.

[37]To estimate $\tilde{\beta}$ we stratify the sample by grade and run separate regressions. For the short-term estimates for grade heterogeneity in Table 7 we use instead an interaction term, for the long-term estimates by grade in Table 6 we regress the outcome on absence in both grades in the same regression. Thereby, the $\tilde{\beta}$ estimates used here are not exactly the same as in the analysis before, but they are quite similar, see column 2 of Table 11.

in the fourth grade the bounds have the same sign for $\delta = 1$ and $\delta = -1$ and both bounds are barely distinguishable from the baseline estimate. This strongly supports our interpretation that the baseline results do not indicate an omitted variable bias.

For all long-term outcome variables all bounds for the adverse selection case point towards the same direction as in the baseline model (absence in school is harmful). For $\delta = 1$, some of the bound estimates point to a positive effect of absence. However, these are the cases when conditioning on the observables reduces the absence coefficient in absolute terms and omitting the variables causes an underestimation. The bounds, again, reinforce our belief that serious omitted variable bias is unlikely.

Table 11: Coefficient bounds for $\delta = 1$ and $\delta = -1$ selection

Dependent variable	Absence	(1) Restricted model $\dot{\beta}$	(2) Baseline model $\tilde{\beta}$	(3) Bound β^* for $\delta = 1$	(4) Bound β^* for $\delta = -1$
Short-term outcome					
Average performance	Grade 1	−0.0032	−0.0028	−0.00777	−0.00191
		(0.0007)	(0.0015)		
		[0.0029]	[0.3721]		
	Grade 4	−0.0032	−0.0046	−0.00496	−0.00456
		(0.0019)	(0.0018)		
		[0.0031]	[0.3940]		
Long-term outcomes					
More than *Folkskola*	Grade 1	−0.0004	−0.0012	−0.0012	−0.0011
		(0.0005)	(0.0007)		
		[0.0002]	[0.1562]		
	Grade 4	−0.0010	−0.0003	−0.0000	−0.0004
		(0.0006)	(0.0006)		
		[0.0013]	[0.1548]		
Employment 1960	Grade 1	−0.0014	−0.0026	−0.0017	−0.0032
		(0.0006)	(0.0006)		
		[0.0012]	[0.4712]		
	Grade 4	−0.0011	0.0007	0.0022	−0.0003
		(0.0008)	(0.0008)		
		[0.0007]	[0.4688]		
Employment 1970	Grade 1	−0.0008	−0.0008	−0.0004	−0.0011
		(0.0006)	(0.0010)		
		[0.0004]	[0.2104]		
	Grade 4	−0.0007	−0.0012	−0.0010	−0.0014
		(0.0007)	(0.0009)		

Continued on next page

Table 11 – *continued*

Dependent variable	Absence	(1) Restricted model $\dot{\beta}$	(2) Baseline model $\tilde{\beta}$	(3) Bound β^* for $\delta = 1$	(4) Bound β^* for $\delta = -1$
		[0.0003]	[0.2108]		
Income 1970	Grade 1	−22.6056	−1.6248	55.9162	−28.4313
		(16.4742)	(39.6379)		
		[0.0004]	[0.4697]		
	Grade 4	−48.0319	−12.3049	95.0982	−60.4634
		(23.0625)	(28.2772)		
		[0.0015]	[0.4698]		
Pensions 2002	Grade 1	−191.7764	35.9554	171.1624	−20.6282
		(103.5173)	(170.6840)		
		[0.0012]	[0.3100]		
	Grade 4	−258.6889	−213.1113	171.4723	−401.7276
		(79.8756)	(148.5325)		
		[0.0022]	[0.3111]		
Passed away ≤ age 70	Grade 1	0.0004	0.0008	0.0002	0.0013
		(0.0004)	(0.0006)		
		[0.0001]	[0.1456]		
	Grade 4	0.0001	0.0004	−0.0011	0.0014
		(0.0006)	(0.0006)		
		[0.0000]	[0.1453]		

Notes: Column 1 gives the coefficient of absence in the restricted model where the outcome variable is regressed on absence and an intercept. The unrestricted model in column 2 is similar to the baseline results for the short- and long-term effects. Column 3 states the bound for proportional unobservable selection that goes in the same direction as observable selection. In column 4 the unobserved selection is proportional but in an adverse direction. Parish-clustered standard errors for the regression coefficients in parentheses. The resulting R^2s are given in brackets.

6.2 Instrumental variables estimates

Short-term IV estimates Next, we follow Goodman (2014) and Aucejo and Romano (2016) and employ an instrumental variables strategy. In line with Goodman (2014), Marcotte (2007) and Marcotte and Hemelt (2008), we exploit changes in weather conditions using local meteorological time series data on the temperature collected from Matsuura and Willmott (2012).[38] In the simplest feasible specification we instrument

[38]The data includes monthly temperature information interpolated to a 0.5 degree by 0.5 degree latitude/longitude grid and we assign each school parish to the closest grid node using latitude and longitude information on the parish center as well as the grid node centroid.

yearly absence using a count variable that gives the number of "benign" months in the school year. Following Bruckner et al. (2014), we define a month in a certain parish as benign if the average temperature in this parish is within the upper quartile of the temperatures measured in all parishes and all years for this month. The Online Appendix Figure O4 shows the average temperature for the school years 1936/37-1947/48 across regions and the spatial variation in the average number of benign months.[39]

The exclusion restriction of the instrument is that the number of benign months only affects an individual's performance in school through days of absence. As pointed out by Goodman (2014) weather may affect both individual absence and school closures. If the number of benign months correlates with weather-related school closures, this would violate the exclusion restriction. While we have no school-level information on closures, the context of our analysis makes weather-related school closures unlikely. The northern part of Sweden is often covered with snow from October to April, and it seems fair to assume that schools, parents and students were well-adapted to the situation. Moreover, schools were usually at walking distance from a student's home and historical sources do not mention snow-related school closures. Another concern is that weather-related teacher and classmate absence affects performance. There is no direct information on teacher absence, but teachers generally lived in one part of the school building, and students were provided with a substitute teacher if the teacher was sick.

We consider the two-stage least square (2SLS) estimator. To identify the impact of weather on days of absence, we run a first-stage regression of days of absence W_{ig} on the number of benign months experienced by student i when he or she is in grade g. In the second stage, we regress the short-term outcome measure on fitted days of absence:

$$\begin{aligned} W_{ig} &= b_0 + b_1 \text{ \#benign months}_{ig} + b_2 X_{1i} + b_3 X_{2ig} + P_{ig} + v_{ig} \quad \text{and} \\ y_{ig} &= \alpha + \tau \widehat{W}_{ig} + \beta_1 X_{1i} + \beta_2 X_{2ig} + P_{ig} + \varepsilon_{ig} \quad \text{for } g = 1, 4, \end{aligned}$$

[39]Naturally the parishes in the north with very low average temperatures are also the ones that barely experience months with benign temperatures. This has limited consequences for our analysis. Even in the north there is variation in the number of benign months and most parishes in our sample are located in the more populated southern part, where the variation in benign months is higher.

where b_1 is the effect of the number of benign months on days of absence.[40] The instrument needs to satisfy the independence assumption that $Y_i(w,s) \perp\!\!\!\perp$ #benign months$_{ig}$ | X_{1i}, X_{2ig}, P_{ig}. Given the inclusion of parish fixed effects that account for more educated families living into certain regions (e.g., warmer, more industrial regions), we belive this assumption is likely to hold.

Additionally we consider a two-sample two-stage least square (TS2SLS) approach.[41] The universe of the weather and absence information across grades constitute the first-stage sample that allows a pooled estimation of the first-stage effect. We use this first-stage sample to estimate the effect of the number of benign months on days of absence in either grade. In the reduced-form sample, we then predict the days of absence based on the effect of benign months in the first-stage sample. Instead of using the effect of weather in grade 1 or grade 4 on absence in the same grade to predict absence in this grade, we thus use the coefficient of the pooled estimation to fit absence in either grade.[42] To get standard errors of the second-stage TS2SLS estimates we apply the Delta method.

Regardless of the IV estimator, two remarks should be kept in mind. First, our weather instrument is more plausibly providing exogenous variation in absences due to sickness than to other reasons such as truancy. To the extent that the effect of absence is heterogeneous across individuals missing school due to sickness or non-sickness reasons, the IV estimates of τ will be the LATE for individuals whose absence is affected by weather. Second, as remarked by Goodman: "estimates of the impact of weather on student achievement can be quite sensitive to the chosen definition of the instrument" (see Goodman, 2014, p.15). The effect of temperature on days of absence seems to be non-linear, and mapping

[40] In order to keep as much variation as possible in the first stage we only rely on parish fixed effects. Moreover, we do not restrict our sample to siblings observed in both grades as neither individual nor siblings FE seem necessary for the exclusion restriction to hold. This allows us to use all individuals for that we have grade-1 and/or grade-4 information.

[41] When implementing 2SLS, we can pool grade-1 and grade-4 information to receive average first- and second-stage effects or conduct the analysis separately by grade. Pooling has the advantage that we use more information at once and receive more precise estimates. Yet, a separate estimation allows the first- and second-stage effects to differ across grades. To circumvent this trade-off we consider the TS2SLS approach.

[42] The underlying assumption is that the first-stage effect does not differ across grades. Comparing the separate 2SLS first-stage estimates with the jointly estimated TS2SLS first-stage estimate allows us to investigate the plausibility of this assumption.

Table 12: Short-term IV results for total days of absence

	(1)	(2)	(3)	(4)
	Specification			
	2SLS		TS2SLS	
	grade 1	grade 4	grade 1	grade 4
First-stage effect of weather on absence				
# benign months	-0.3610^{***}	-0.3553^{***}	-0.3494^{***}	-0.3494^{***}
	(0.0080)	(0.1632)	(0.0701)	(0.0701)
Second-stage effect of fitted absence on performance				
Absence	-0.0038	-0.0220	-0.0039	-0.0252
	(0.0222)	(0.0252)	(0.0653)	(0.0654)
# observations first stage	13,884	14,152	28,036	28,036
# observations second stage	13,884	14,152	13,884	14,152
F-statistic instrument	7.91	4.74	24.86	24.86

Notes: Each cell (but the first stage in column 3 and 4) states a separate regression. In the two-sample two-stage least square (TS2SLS) specification the first stage is estimated jointly over both grades. All standard errors in parentheses. The standard errors of the second stage in columns 1 and 2 are Stata's default standard errors clustered on parish level. In columns 3 and 4 the standard errors of the second stage are calculated using the Delta method and clustered on parish level. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

monthly information on weather conditions in days of absence per school year without losing the variation that triggers absence is not trivial. For this reason, we think our IV estimates provide complementary (rather than superior) evidence to our individual fixed effects estimates.

Table 12 shows the IV results for performance. Columns 1 and 2 give the 2SLS effects for grade 1 and 4, and columns 3 and 4 state the corresponding TS2SLS effects when the first stage is estimated jointly. Regardless how the first stage is estimated, an additional month with benign temperatures reduces, on average, the number of days of absence by about 0.35. The assumption that the first-stage relationship is the same across grade 1 and grade 4 seems justified. Noteworthily, only the F-statistic of the pooled specification exceeds the Staiger and Stock (1997) rule-of-thumb value of 10. Turning to the structural estimates, all coefficients indicate a negative relationship between absence and performance in school.

While the second-stage coefficients exhibit a large difference across grades, the decision of the IV estimator (2SLS vs. TS2SLS) makes no noteworthy difference. All second-stage estimates have rather large standard errors and we cannot reject that the coefficient is zero. Still, the point estimates are negative and lower than the baseline result. In fourth grade, an additional day of absence decreases performance by 2 per cent of a SD.[43] Using a similar strategy, the IV results of Goodman (2014) are also up to 2 per cent of a SD.

All in all, we interpret the short-term IV results as confirmation of the baseline estimates. Given the limited external validity of the LATE estimates and the assignment mechanism, the IV approach is not a silver bullet to address identification. Still, an additional day of absence seems negatively related to academic performance.

Long-term IV estimates For the long-term outcomes we extend the TS2SLS approach to all seven grades of compulsory education.[44] As before, we estimate the first-stage effect of benign months on absence in grades 1 and 4:

$$W_{ig} \quad = \quad b_0 + b_1 \text{ \#benign months}_{ig} + b_2 X_{1i} + P_i + v_{ig} \quad \text{for } g = 1, 4,$$

but we calculate the fitted days of absence for all seven grades. This is possible due to two features of the data. First, weather information is available for all years. Second, knowing when students were supposed to visit a certain grade based on the year of birth allows us to assign the instrument value. The first-stage sample of the TS2SLS approach is the same as before (weather and absence information for grades 1 and 4). The reduced-form sample is now the universe of the weather information for all grades and long-term outcomes. In the second stage we regress the long-term outcome variable $y_i^{\text{long-term}}$ on the predicted number of days of absence for all seven grades:

[43] A reason for the big effect size may be that the IV strategy yields LATE and not the population average treatment effect.

[44] It is not possible to employ the 2SLS implementation of the IV approach to the long-term outcomes. The following example illustrates the problem: individual A is hit by a weather shock in grade 4 in 1938, individual B is in grade 4 in 1939 and his or her absence in grade 4 is not affected by the weather shock in 1938. However, individual B's absence in grade 3 in 1938 is affected by the weather shock. This is problematic as absences in grades 3 and 4 probably affect the long-term outcomes similarly, but we do not observe absence in grade 3 (or any other grade but grades 1 and 4).

$$y_i^{\text{long-term}} = \alpha + \sum_{g=1}^{7} \tau_g \widehat{W}_{ig} + \beta_1 X_{1i} + P_i + \varepsilon_i.$$

Socio-economic controls X_1 come from the church records while the information on the parish of residence during childhood P_i comes from the exam catalog. We restrict the analysis to individuals that did not move between grades 1 and 4 and assume that the school parish was the same in all grades.

This strategy relies on the assumption that weather conditions affect absence in the same way in all seven grades. As indicated in Table 12, at least for the observed grades 1 and 4, this assumption seems justified. We estimate the first stage using all absence information wherefore the first-stage estimates for the long-term TS2SLS approach are (nearly) those stated in Table 12.[45] Table 13 presents the second-stage estimates for long-term outcomes. The point estimates have the expected sign, but despite the strong first stage, the estimated structural effects of absence in school fail to reach statistical significance. For employment in 1960 the point estimates clearly exceed the baseline estimates and the grade-1 effect is, again, bigger than the one in grade 4. A similar pattern holds for employment in 1970 but the grade-4 effect is close to zero with a positive sign. All in all, as for the short-term effects, the long-term IV estimates exceed the baseline FE estimates but do not indicate that omitted variables cause an overestimation in the baseline model.

To overcome the assumption that weather conditions have the same effect on absence in all grades, we additionally consider a model where days of absence in grades 1 and 4 are treated as two endogenous variables and weather conditions in grades 1 and 4 as two instruments. Controlling for the number of benign months in the other five grades, solves the problem of weather shocks in unobserved grades while, at the same time, does not rely on assuming equal first-stage effects. The Online Appendix shows the first- and

[45] Because we limit the first stage to individuals that did not move between grade 1 and 4, we lose about 100 observations for the first stage. The point estimates barely change compared to Table 12. The Online Appendix reports the first-stage estimates for the final long-term sample. Moreover, we present first-stage estimates by grade when the first-stage sample is restricted to individuals with available second-stage information. This does not affect our results and underlines similar first-stage effect across both grades.

Table 13: Long-term IV using the TS2SLS approach – Second-stage results for total days of absence

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable					
	> *Folk-skola*	Empl. 1960	Empl. 1970	Income 1970	Pensions 2002	Passed away≤70
Fitted days of sickness absence in						
grade 1	−0.0033	−0.0096	−0.0217	−39.36	−248.18	0.0091
	(0.0238)	(0.0254)	(0.0256)	(623.10)	(4127.05)	(0.0201)
grade 4	−0.0054	−0.0055	0.0006	154.13	−379.79	0.0039
	(0.0269)	(0.0149)	(0.0213)	(941.11)	(4325.53)	(0.0158)
# observations	22231	25657	27913	18459	19009	27913
# individuals	13814	15980	15980	11483	11848	17422
F-statistic instr.	19.72	19.22	25.59	18.38	13.54	25.59

Notes: Each cell states a separate regression. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

second-stage estimations using sickness absences as endogenous variables. The first-stage F-statistics are well-above zero, but below 10. The reason for this is that we use two variables containing of number of benign months in grade 1 and in grade 4, respectively, in each first-stage regression. Given the high serial correlation of weather conditions in both grades and that we would only expect contemporary weather to affect absence, we may not overstate the importance of the F-statistics here. In fact, the results indicate, that weather conditions in the same grade have an effect on days of sickness absence (as also indicated by the short-term IV and long-term TS2SLS results). The second-stage results do not exhibit a different pattern than for the TS2SLS estimates.

7 Conclusions

Student absence from school is an important but often overlooked determinant of instructional time. To date, little is known about the long-run impact of students missing school, and the only studies providing causal evidence of the impact of student absence on academic performance focus on the US. The major contribution of this paper is to estimate the impact of student absence in elementary school on short- and long-term

outcomes for a non-US context by using a unique combination of historical records and administrative datasets from Sweden.

Our analysis shows that absence in elementary school has a significant impact on student performance: ten days of absence over a school year leads to a reduction in grade point average of 3.3 per cent of a standard deviation. The estimated effect is very robust across empirical strategies and comparable in magnitude to results found for the US. This immediate impact on school performance spills over into secondary school admissions, which are based on elementary school performance. Our estimated effect of absence on secondary schooling admissions is at least as large as one would expect based on the effect of absence on performance – even though we are unable to attribute it to a certain school grade. For the other long-term outcomes, the effect of student absence in elementary school is only pronounced for early-career employment. For employment and income at age 35-40, pension income, and mortality, our sibling fixed effect estimates indicate that the negative effect of absence is undistinguishable from zero.

Our findings for the short-term effects of absence on school performance deliver very robust results and consistently suggest that the existence of an omitted variable bias is rather unlikely. Nevertheless, we are careful in interpreting the causality of our long-term effects. When considering long-term outcomes, it becomes more difficult to define what the alternative to the 'treatment' is. A large majority of absence days are due to illness, and we cannot rule out *a priori* that a persistent health shock from elementary school has an independent effect on long-term outcomes. As a result, we exploit the fact that our data has information on reasons for absence to compare the long-term effects of absence due to sickness with those of absence due to other reasons. We find that there are no important differences between the two, which we interpret as evidence that our long-term effects most likely capture the impact of reduced instructional time.

Our findings are obviously specific to a particular context, but as the first study providing evidence of the long-term effects of student absence, we believe that it is highly relevant to both academics and policy-makers concerned with high rates of absence around the

world. Our results suggest that although student absence leads to worse performance at the end of the school year, the associated penalty on the labor market eventually fades out. Thus, the reduction in instructional time resulting from individual absence does not prevent those missing school from acquiring the skills that determine their long-run productivity. The fact that absence only affects employment at age 25-30, but no outcome later on is consistent with the labor market using educational performance as a signal of ability early on and progressively learning about workers' true productivity (Altonji and Pierret, 2001).

Our findings are strikingly in the same vein as those of a small but growing number of studies interested in the long-run consequences of variation in instructional time through changes in the school year length (Pischke, 2007) or in number of years of compulsory schooling (Stephens and Yang, 2014 and Pischke and von Wachter, 2008). One possible explanation for the patterns we find is that students are able to compensate for the educational content they miss over the next few years in school and/or that teachers are effective at helping students catch up the skills that have the most return in the labor market (though not being able to help them catch up on the whole curriculum, as reflected by the negative effect of absence on grade point average and secondary schooling enrollment). At a time when policy-makers around the world are paying increasing attention to school absence, our findings indicate that policies aimed at reducing student absence may not be particularly effective at increasing productivity in the economy. At least in the context that we study, students – perhaps with the support of their teachers and/or parents – seem able to compensate for any shortfall in learning associated with their absence in a way that does not affect their long-run productivity.

References

Agüero, J. and T. Beleche (2013). Test-Mex: Estimating the effects of school year length on student performance in Mexico. *Journal of Development Economics 103*(C), 353–361.

Altonji, J., T. Elder, and C. Taber (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy 113*(1), 151–184.

Altonji, J. G., L. B. Kahn, and J. D. Speer (2016). Cashier or consultant? entry labor market conditions, field of study, and career success. *Journal of Labor Economics 34*(S1), S361–S401.

Altonji, J. G. and C. R. Pierret (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics 116*(1), 313–350.

Athey, S. and G. W. Imbens (2017, May). The State of Applied Econometrics: Causality and Policy Evaluation. *Journal of Economic Perspectives 31*(2), 3–32.

Aucejo, E. M. and T. F. Romano (2016). Assessing the effect of school days and absences on test score performance. *Economics of Education Review 55*, 70 – 87.

Battistin, E. and E. Meroni (2016). Should we increase instruction time in low achieving schools? Evidence from Southern Italy. *Economics of Education Review 55*, 39 – 56.

Bellei, C. (2009). Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile. *Economics of Education Review 28*(5), 629–640.

Bhalotra, S., M. Karlsson, and T. Nilsson (forthcoming). Infant Health and Longevity: Evidence from a Historical Intervention in Sweden. *Journal of the European Economic Association.*

Bhalotra, S. R., M. Karlsson, T. Nilsson, and N. Schwarz (2016, November). Infant Health, Cognitive Performance and Earnings: Evidence from Inception of the Welfare State in Sweden. IZA Discussion Papers 10339, Institute for the Study of Labor (IZA).

Bhuller, M., M. Mogstad, and K. G. Salvanes (2011). Life-cycle bias and the returns to schooling in current and lifetime earnings.

Bond, T. N. and K. Lang (2013, December). The Evolution of the Black-White Test Score Gap in Grades K-3: The Fragility of Results. *The Review of Economics and Statistics 95*(5), 1468–1479.

Bruckner, T., G. van den Berg, K. Smith, and R. Catalano (2014). Ambient temperature during gestation and cold-related adult mortality in a Swedish cohort, 1915-2002. *Social Science & Medicine 119*(0), 191–197.

Carlsson, M., G. Dahl, B. Öckert, and D.-O. Rooth (2015). The Effect of Schooling on Cognitive Skills. *The Review of Economics and Statistics 97*(3), 533–547.

Chetty, R., J. Friedman, and J. Rockoff (2014, September). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review 104*(9), 2593–2632.

Chicago Tribune (2012). "An emptydesk epidemic" by David Jackson, Gary Marx and Alex Richards. *Chicago Tribune,* November 11, 2012 (http://www.chicagotribune.com/ct-met-truancy-mainbar-20121111-story.html, last assessed June 29, 2017).

Child Trends (2015). "Student Absenteeism". *Child Trends Databank* (https://www.childtrends.org/?indicators=student-absenteeism, last assessed August 7 2017).

Conti, G. (2013). The developmental origins of health inequality. In *Health and Inequality*, pp. 285–309. Emerald Group Publishing Limited.

Contoyannis, P. and J. Li (2011). The evolution of health outcomes from childhood to adolescence. *Journal of Health Economics 30*(1), 11–32.

Cunha, F. and J. Heckman (2007). The Technology of Skill Formation. *American Economic Review 97*(2), 31–47.

Currie, J. and M. Stabile (2003). Socioeconomic status and child health: Why is the relationship stronger for older children? *The American Economic Review 93*(5), 1813–1823.

Dustmann, C., P. A. Puhani, and U. Schönberg (forthcoming). The Long-term Effects of Early Track Choice. *The Economic Journal*.

Erikson, R. and J. O. Jonsson (1993). *Ursprung och Utbildning – Social Snedrekrytering till Högre Studier*. Stockholm, Sweden: Utbildningsdepartementet.

Federation of Swedish Genealogical Societies (2014). *Swedish Death Index 1901-2013*. Farsta, Sweden: Federation of Swedish Genealogical Societies.

Fischer, M., M. Karlsson, and T. Nilsson (2013). Effects of Compulsory Schooling on Mortality: Evidence from Sweden. *International Journal of Environmental Research and Public Health 10*(8), 3596–3618.

Fischer, M., M. Karlsson, T. Nilsson, and N. Schwarz (2016, December). The Sooner the Better? Compulsory Schooling Reforms in Sweden. IZA Discussion Papers 10430, Institute for the Study of Labor (IZA).

Fitzpatrick, M., D. Grissmer, and S. Hastedt (2011). What a difference a day makes: Estimating daily learning gains during kindergarten and first grade using a natural experiment. *Economics of Education Review 30*(2), 269–279.

Fletcher, J. and B. Wolfe (2014). Increasing our understanding of the health-income gradient in children. *Health economics 23*(4), 473–486.

Fredriksson, V. A. (1971a). *Svenska Folkskolans Historia*. Stockholm, Sweden: Albert Bonniers Förlag.

Fredriksson, V. A. (1971b). *Svenska folkskolans historia*, Volume 6. Albert Bonniers förlag.

Genda, Y., A. Kondo, and S. Ohta (2010). Long-term effects of a recession at labor market entry in japan and the united states. *Journal of Human Resources 45*(1), 157–196.

Goodman, J. (2014). Flaking Out: Student Absences and Snow Days as Disruptions of Instructional Time. NBER Working Papers 20221, National Bureau of Economic Research, Inc.

Hansen, B. (2011). School Year Length and Student Performance: Quasi-Experimental Evidence. mimeo.

Hener, T., H. Rainer, and T. Siedler (2016). Political socialization in flux? Linking family non-intactness during childhood to adult civic engagement. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 179*(3), 633–656.

Huebener, M. and J. Marcus (2015). Moving up a Gear: The Impact of Compressing Instructional Time into Fewer Years of Schooling. Technical report.

Leuven, E., M. Lindahl, H. Oosterbeek, and D. Webbink (2010). Expanding schooling opportunities for 4-year-olds. *Economics of Education Review 29*(3), 319–328.

Marcotte, D. (2007). Schooling and test scores: A mother-natural experiment. *Economics of Education Review 26*(5), 629–640.

Marcotte, D. and B. Hansen (2010). Time for School? *Education Next 10*(1), 53–59.

Marcotte, D. and S. Hemelt (2008). Unscheduled School Closings and Student Performance. *Education Finance and Policy 3*(3), 316–338.

Matsuura, K. and C. Willmott (2012). Terrestrial Air Temperature: 1900-2010 Gridded Monthly Time Series (Version 3.01). Technical report, Department of Geography, University of Delaware. `http://climate.geog.udel.edu/~climate/html_pages/Global2011/README.GlobalTsT2011.html`.

Oreopoulos, P., T. Von Wachter, and A. Heisz (2012). The short-and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics 4*(1), 1–29.

Oster, E. (forthcoming). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics*.

Pischke, J.-S. (2007). The Impact of Length of the School Year on Student Performance and Earnings: Evidence From the German Short School Years. *The Economic Journal 117*(523), 1216–1242.

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

Rivkin, S., E. Hanushek, and J. Kain (2005, 03). Teachers, Schools, and Academic Achievement. *Econometrica 73*(2), 417–458.

Santavirta, T. (2012). How large are the effects from temporary changes in family environment: evidence from a child-evacuation program during world war ii. *American Economic Journal: Applied Economics 4*(3), 28–42.

Sims, D. (2008). Strategic responses to school accountability measures: It's all in the timing. *Economics of Education Review 27*(1), 58–68.

Staiger, D. and J. Stock (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica 65*(3), 557–586.

Stephens, Melvin, J. and D.-Y. Yang (2014, June). Compulsory Education and the Benefits of Schooling. *American Economic Review 104*(6), 1777–92.

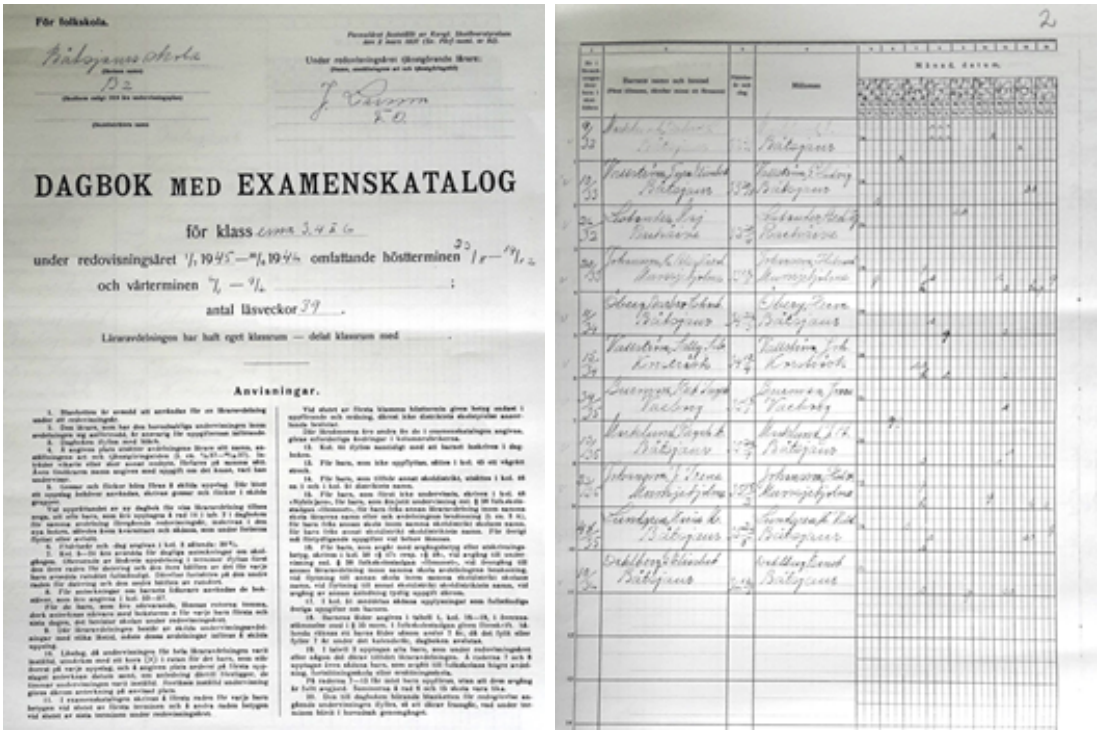
Sundén, A. (2006). The Swedish experience with pension reform. *Oxford Review of Economic Policy 22*(1), 133–148.

van Leeuwen, M., I. Maas, and A. Miles (2002). *HISCO: Historical International Standards Classification of Occupations*. Leuven, Belgium: Leuven University Press.

Appendix

Appendix figures

Figure A1: Example of an exam catalog



Notes: Pictures of an exam catalog taken in an archive in Sweden.

Appendix tables

Table A1: Summary statistics for control variables

Time-invariant variables	Mean
Female (in %)	50.06
Number of siblings	1.56
Year of birth (in %)	
1930	18.37
1931	16.74
1932	18.30
1933	17.39
1934	16.51
1935	12.69
(we additionally control for the month of birth and interaction terms between the year and the month of birth)	
Occupation of the parents at the time of birth (in %)	
Father: farmer, fisherman, hunter	42.13
Father: agricultural worker	34.59
Father: service and sales worker	7.21
Father: production workers	49.18
Mother: employed (binary)	2.39
Living at the time of time and birth conditions (in %)	
Born out of wedlock (in %)	4.36
Born in hospital (in %)	7.97
Twin birth (in %)	4.16
Mother's year of birth	1902
Father's year of birth	1897
(mother's and father's year is controlled for by using 10-year dummies)	

	Mean grade	
Time-variant variables	1	4
Age (in years)	8.13	11.27
(included through age-in-months fixed effects)		
Class characteristics		
All classmates in same grade (in %)	34.93	29.38
Some classmates in lower grade (in %)	0.00	64.85
Some classmates in higher grade (in %)	65.07	30.81
Class size (number of students)	13.84	15.64
(measured through 5-day splines from 0 to 25)		

Notes: Own calculation based on church records and exam catalog information. Sample restricted to individuals with available sibling information when all siblings are observed in grade 1 and grade 4. Observations: 8,938. Mutually exclusive indicators may not add up to 100 per cent because of missing information. For the estimations, missing information are coded as separate category taken into account that the reason for the missing information might be meaningful in its own right (e.g., the father's occupation is missing because the father is unknown).

Table A2: Full estimation output for all fixed effects specifications

	(1) OLS	(2) School FE	(3) Teacher FE	(4) Sibl. FE	(5) Indi. FE
Total days of absence	−0.0035***	−0.0043***	−0.0043***	−0.0041***	−0.0035***
	(0.0009)	(0.0010)	(0.0010)	(0.0011)	(0.0013)
Female	0.2761***	0.2794***	0.2863***	0.3256***	
	(0.0318)	(0.0313)	(0.0328)	(0.0397)	
Birth year: 1931	−0.6899***	−0.9647***	−0.7462***	−0.2439	
	(0.2360)	(0.1333)	(0.2010)	(0.1929)	
Birth year: 1932	−0.8483***	−0.5598***	−0.6575***	−0.7286	
	(0.1765)	(0.1308)	(0.1894)	(0.4565)	
Birth year: 1933	2.4848***	−0.0635	−0.0588	0.0116	
	(0.4723)	(0.1455)	(0.1706)	(0.2730)	
Birth year: 1934	−0.6706**	−0.9025***	−0.6238***	−0.4792	
	(0.2738)	(0.1762)	(0.2176)	(0.6530)	
Birth year: 1935	2.4016***	2.6556***	3.4850***	−0.0263	
	(0.3734)	(0.3098)	(0.6624)	(0.3861)	
Occup. father: agriculture	0.0616	0.0051	0.0053	0.0491	
	(0.0622)	(0.0609)	(0.0638)	(0.0994)	
Occup. father: services	0.1276*	0.1637**	0.1675**	0.0057	
	(0.0723)	(0.0802)	(0.0818)	(0.0911)	
Occup. father: farmer	0.0071	0.0533	0.0554	−0.0920	
	(0.0595)	(0.0607)	(0.0586)	(0.1011)	
Occup. father: unknown	0.0411	0.1140**	0.1393**	0.1571**	
	(0.0576)	(0.0573)	(0.0578)	(0.0664)	
Mother employed	−0.1685	−0.1324	−0.1503	−0.0723	
	(0.1094)	(0.1021)	(0.1026)	(0.0902)	
Wedlock	0.1362	0.1202	0.0684	0.0647	
	(0.0946)	(0.0848)	(0.0757)	(0.0980)	
Born in hospital	0.1455**	0.1322**	0.1134*	0.1111*	
	(0.0648)	(0.0662)	(0.0641)	(0.0646)	
Twin	−0.2138***	−0.2091***	−0.2080**	−0.1821	
	(0.0721)	(0.0742)	(0.0831)	(0.1362)	
Grade 4	1.2066***	0.9332***	1.0493***	0.7943***	0.5899***
	(0.2107)	(0.1712)	(0.2313)	(0.2286)	(0.0344)
Classmates in lower grade	0.0451	0.0757	0.1033*	0.1355***	0.1445**
	(0.0505)	(0.0461)	(0.0552)	(0.0440)	(0.0621)
Classmates in higher grade	0.0556	−0.0411	−0.0792	−0.0588	−0.0574
	(0.0378)	(0.0480)	(0.0702)	(0.0664)	(0.0553)
Class size 1-5	−0.0546**	−0.0176	−0.0015	−0.0151	0.0258
	(0.0259)	(0.0297)	(0.0351)	(0.0371)	(0.0414)
Class size 6-10	0.0169	0.0148	0.0142	0.0030	0.0024
	(0.0129)	(0.0157)	(0.0164)	(0.0118)	(0.0233)
Class size 11-15	−0.0062	−0.0043	0.0032	0.0123	0.0204
	(0.0092)	(0.0098)	(0.0098)	(0.0119)	(0.0177)
Class size 16-20	−0.0009	0.0132	−0.0052	−0.0018	0.0051
	(0.0086)	(0.0097)	(0.0192)	(0.0154)	(0.0172)
Class size 21-25	0.0083	−0.0024	0.0137	0.0100	0.0149
	(0.0122)	(0.0137)	(0.0166)	(0.0166)	(0.0233)
Class size >25	0.0008	0.0137***	0.0142***	0.0121**	0.0118*
	(0.0041)	(0.0040)	(0.0047)	(0.0054)	(0.0068)
Class size (missing)	0.0029	0.4751***	0.5963***	0.5285***	0.7646***
	(0.1297)	(0.1591)	(0.1686)	(0.1791)	(0.1854)
# observations	8,942	8,942	8,942	8,942	8,942
# units		748	1,259	1,989	4,471

Notes: See note to the baseline results table. Fixed effects are suppressed. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

The Short- and Long-term Effects of Student Absence: Evidence from Sweden

—Online Appendix—

Sarah Cattan* Daniel A. Kamhöfer† Martin Karlsson‡ Therese Nilsson§

September 5, 2017

*Institute for Fiscal Studies, London, and IZA, Bonn

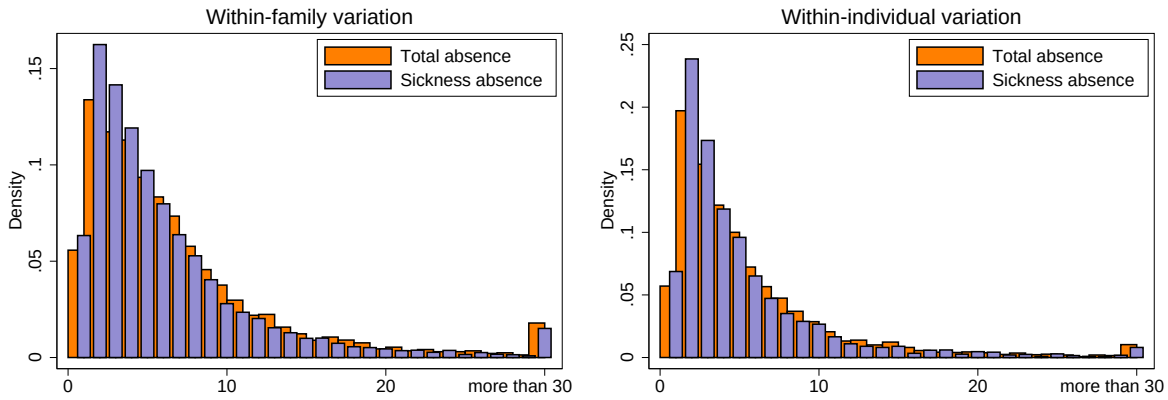
†Paderborn University, University of Duisburg-Essen and CINCH, Essen (Corresponding author: Paderborn University, Warburger Strasse 100, 33098 Paderborn, Germany. E-mail: daniel.kamhoefer@uni-paderbron.de.)

‡CINCH, University of Duisburg-Essen and IZA, Bonn

§Lund University and IFN, Stockholm

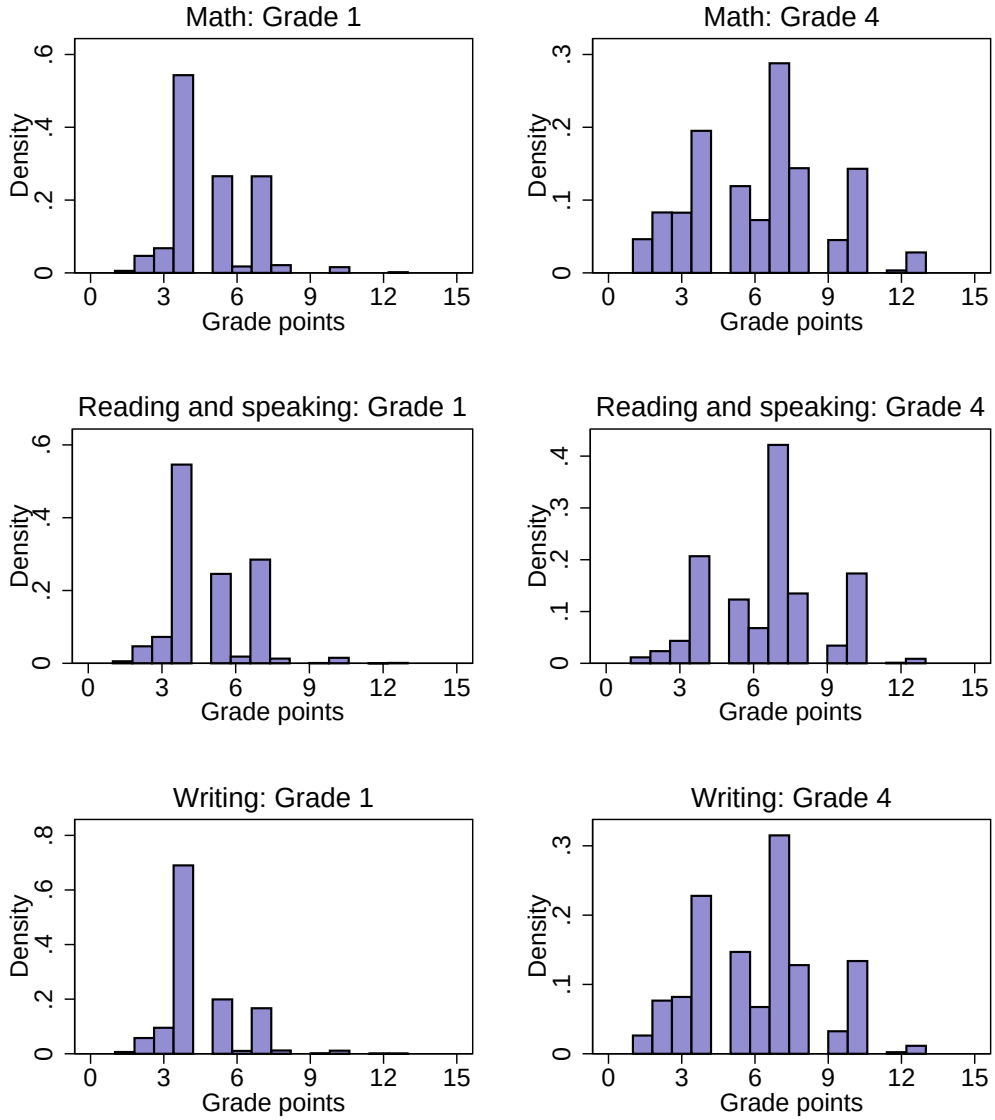
Online Appendix: figures

Figure O1: Distribution of the within-family and within-individual variation in (sickness) absence



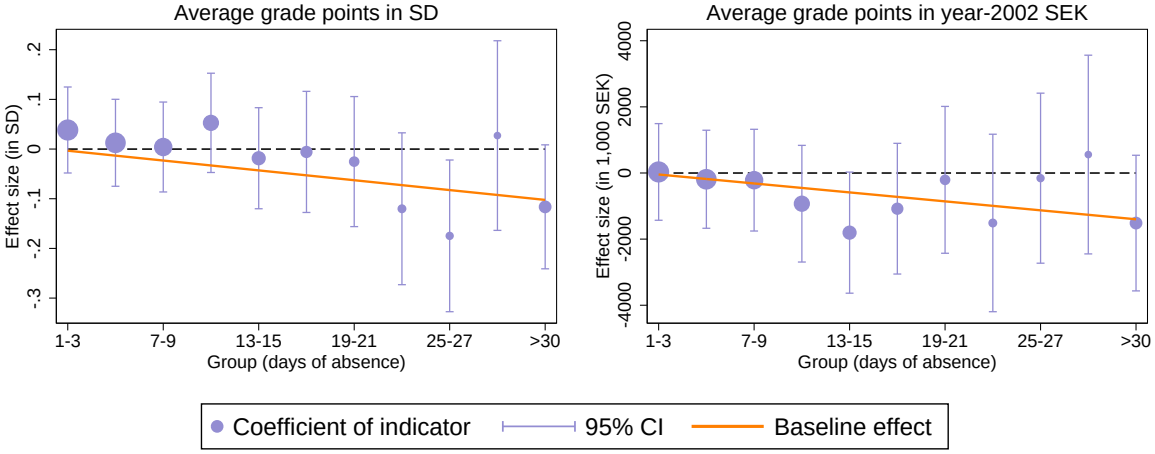
Notes: Own calculations based on exam catalog information. 8,938 observations.

Figure O2: Distribution of grades by subject



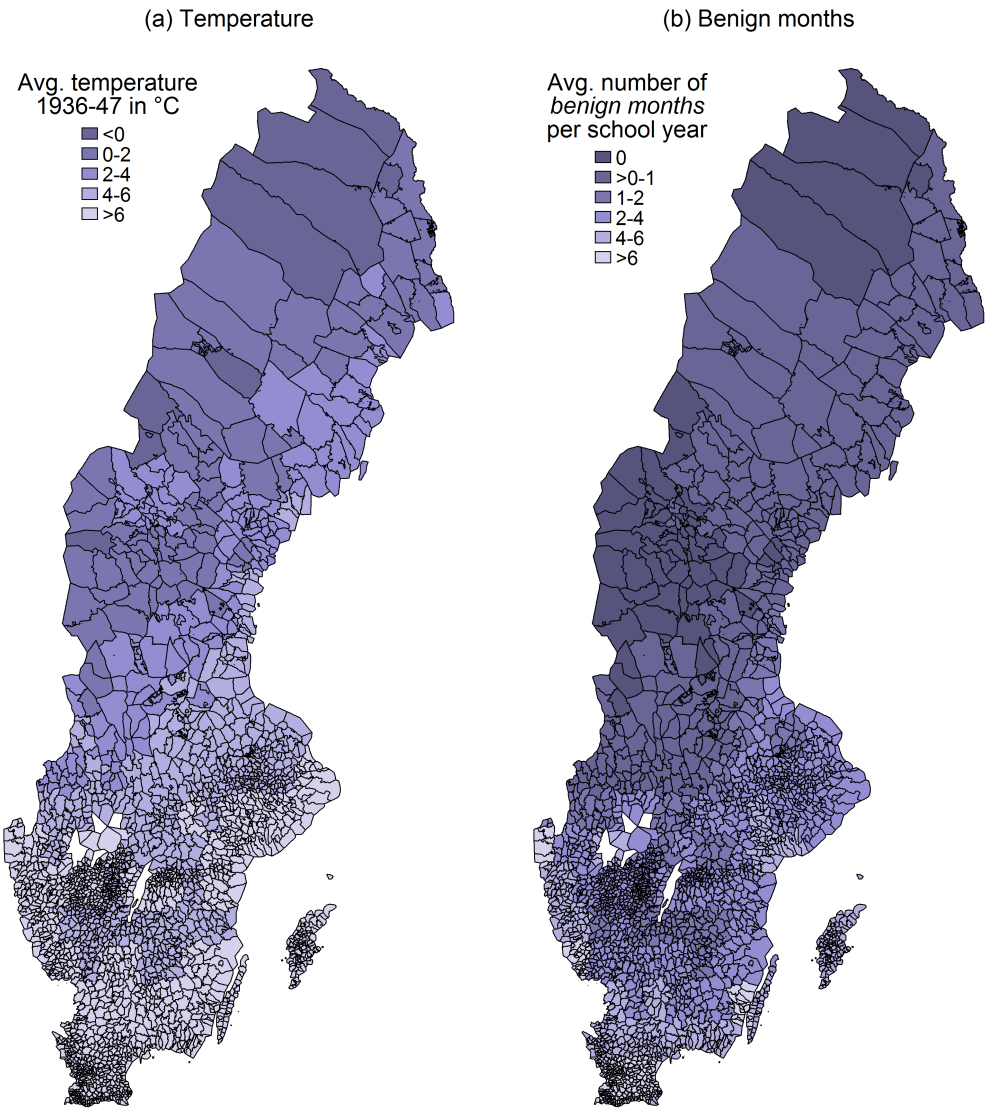
Notes: Own calculations based on exam catalog information.

Figure O3: Nonlinearities in the short-term effect of grouped sickness absence



Notes: To detect nonlinearities in the effect of sickness absence we regress performance on indicator variables giving the number of days of sickness absence in groups of 3. This graph plots the coefficients of the indicator variables. The size of the marker given the relative number of observations for which the group indicator is 1. In total grade 1 and 4 information on 10,682 individuals is used. The spikes around the markers state the 95% confidence interval. The orange line depicts the linear effect of an additional day of absence in the baseline results.

Figure O4: Spatial temperature distribution



Notes: Own illustration. Data on monthly temperatures are taken from Matsuura and Willmott (2012).

Online Appendix: tables

Table O1: Estimation of earnings potential

	(1)	(2)
	Earnings potential in pensions 2002	
	Grade 1	Grade 4
Math points 1	−96266.3***	19367.1***
	(32873.0)	(6668.3)
Math points 2	−21115.4**	−27084.9***
	(9183.3)	(5701.6)
Math points 3	−25267.2***	−24420.8***
	(7611.9)	(5470.9)
Math points 4	−21651.7***	−16571.9***
	(4400.2)	(3985.3)
Math points 5	−13210.3***	−5952.7
	(4608.6)	(4489.5)
Math points 6	−28367.5**	−10389.4**
	(13940.7)	(5242.2)
Math points 8	−15765.1	3756.6
	(10779.7)	(4195.3)
Math points 9	*omitted*	11150.2*
		(6752.2)
Math points 10	−5489.0	21965.2***
	(13303.2)	(4368.4)
Math points 13	26109.9	38576.6***
	(65184.5)	(8914.4)
Reading points 1	21166.6	−6229.5
	(26801.0)	(12175.4)
Reading points 2	−14140.2	12179.3
	(8862.1)	(9248.4)
Reading points 3	−10991.7	6074.4
	(7257.6)	(7185.5)
Reading points 4	−7235.8*	4654.9
	(4331.4)	(3980.4)
Reading points 5	−304.0	4637.7
	(4683.1)	(4363.7)
Reading points 6	−11745.8	9996.8*
	(12893.1)	(5364.9)
Reading points 8	−6410.5	−334.9
	(13150.2)	(4167.6)
Reading points 9	19499.8	−4478.6
	(37223.7)	(7066.7)
Reading points 10	6939.5	5217.8
	(14187.8)	(4251.6)
Reading points 13	−64483.0	10075.0
	(47935.8)	(15234.3)

Continued on next page

Table O1 – *continued*

	(1)	(2)
	Earnings potential in pensions 2002	
	Grade 1	Grade 4
Writing points 1	91143.2***	−7982.3
	(29263.6)	(9019.7)
Writing points 2	15566.0*	620.5
	(9243.1)	(6484.4)
Writing points 3	17213.0**	8285.3
	(7209.3)	(5955.1)
Writing points 4	13766.5***	7793.7*
	(4945.8)	(4049.7)
Writing points 5	8260.5	8874.8**
	(5427.1)	(4209.8)
Writing points 6	6036.5	−1645.7
	(16110.9)	(5563.9)
Writing points 8	13814.6	−2000.9
	(14438.2)	(4343.8)
Writing points 9	74533.5	−12542.6
	(48343.4)	(7671.2)
Writing points 10	−35879.9*	−7030.0
	(21749.1)	(4852.7)
Writing points 13	−14917.4	8006.3
	(102333.7)	(13230.3)
Constant	158638.4***	148308.6***
	(3915.1)	(3075.5)

Notes: Dependent variable: pensions taken from tax registers 2002. Explanatory variables: binary indicators of the points in math, reading and speaking and writing (15-point scale, reference category is 7 points). Grade points of 14 and 15 are not considered as such grades are barely awarded. Standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

Table O2: Baseline results using the full sample

	(1) OLS	(2) Indi. FE
Average grade points in units of SD		
Days of absence	-0.0035^{***}	-0.0047^{***}
	(0.0005)	(0.0007)
Average grade points in units of pension 2002		
Days of absence	-46.9136^{***}	-74.0147^{***}
	(6.9904)	(9.3396)
# observations	28946	28946
# individuals		10682

Notes: See note to the baseline results table. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

Table O3: Short-term effects measuring performance on a 7-point grading scale

	(1) OLS	(2) Sibl. FE	(3) Indi. FE
Average grade points (7-point scale) in units of SD			
Days of absence	-0.0035^{***}	-0.0049^{***}	-0.0045^{***}
	(0.0009)	(0.0011)	(0.0014)

Notes: See note to the baseline results table. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

Table O4: Long-term effect of absence in school – alternative outcome measures

	(1)	(2)	(3)	(4)
	Pooled effect		Separate effects	
	OLS	Sibl. FE	OLS	Sibl. FE
Years of schooling				
Total absence (avg. both grades)	−0.0040	−0.0010		
	(0.0038)	(0.0035)		
Total absence in grade 1			0.0004	−0.0010
			(0.0025)	(0.0027)
Total absence in grade 4			−0.0043**	−0.0000
			(0.0021)	(0.0020)
Gymnasium-*track education (1=yes)*				
Total abs. (avg. both grades)	−0.0001	−0.0001		
	(0.0006)	(0.0005)		
Total absence in grade 1			0.0003	0.0001
			(0.0004)	(0.0004)
Total absence in grade 4			−0.0004	−0.0001
			(0.0003)	(0.0004)
Labor market income 1970 (incl. zero incomes)				
Total abs. (avg. both grades)	−54.4394**	−59.2415		
	(25.7378)	(49.9878)		
Total absence in grade 1			−29.8358**	−25.4044
			(14.1129)	(24.7863)
Total absence in grade 4			−24.6486	−33.2657
			(16.7865)	(31.0227)
Pensions 2002 (incl. zero pensions)				
Total abs. (avg. both grades)	−235.8604*	−229.8668		
	(119.0779)	(234.4330)		
Total absence in grade 1			−63.5785	5.0779
			(123.0447)	(170.3856)
Total absence in grade 4			−167.9506**	−216.4590*
			(75.2518)	(124.7150)
Pensions 2002 (incl. non-labor market income)				
Total abs. (avg. both grades)	−305.5247	−265.5650		
	(184.8329)	(254.3537)		
Total absence in grade 1			−143.9064	−19.9836
			(171.0246)	(184.6732)
Total absence in grade 4			−160.9356	−228.0035
			(110.7283)	(161.8834)

Notes: Number of observations: education measures 3,019 (in 1,373 families), income 1970 3,092 (1,398), income 1970 2,137 (985), pensions 2002 measures 2,468 (1,129). Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

Table O5: Short-term IV results for days of sickness absence

	(1)	(2)	(3)	(4)
	Specification			
	2SLS		TS2SLS	
	grade 1	grade 4	grade 1	grade 4
First-stage effect of weather on absence				
# benign months	-0.4033^{***}	-0.3231^{**}	-0.3744^{***}	-0.3744^{***}
	(0.1238)	(0.1285)	(0.0617)	(0.0617)
Second-stage effect of fitted absence on performance				
Absence	-0.0034	-0.0242	-0.0036	-0.0209
	(0.0199)	(0.0284)	(0.0568)	(0.0566)
# observations first stage	13,884	14,152	28,036	28,036
# observations second stage	13,884	14,152	13,884	14,152
F-statistic instrument	10.61	6.32	36.83	36.83

Notes: Each cell (but the first stage in column 3 and 4) states a separate regression. In the two-sample two-stage least square (TS2SLS) specification the first stage is estimated jointly over both grades. All standard errors in parentheses. The standard errors of the second stage in columns 1 and 2 are Stata's default standard errors clustered on parish level. In columns 3 and 4 the standard errors of the second stage are calculated using the Delta method and clustered on parish level. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

Table O6: Long-term IV using the TS2SLS approach – First-stage results for total days of absence by grade

	(1)	(2)	(3)	(4)	(5)	(6)
	Sample for					
	> *Folk-skola*	Empl. 1960	Empl. 1970	Income 1970	Pensions 2002	Passed away≤70
Grade 1 only (and only complete information)						
# benign months	−0.3206**	−0.3215**	−0.3431***	−0.2797**	−0.2780*	−0.3431***
	(0.1303)	(0.1342)	(0.1257)	(0.1127)	(0.1424)	(0.1257)
# observations	10,978	12,664	13,821	9,121	9,380	13,821
F-statistic instr.	6.05	5.74	7.46	6.16	3.81	7.46
Grade 4 only (and only complete information)						
# benign months	−0.2922*	−0.2951*	−0.3435**	−0.2904*	−0.3324**	−0.3435**
	(0.1491)	(0.1562)	(0.1581)	(0.1521)	(0.1554)	(0.1581)
# observations	11,253	12,993	14,092	9,338	9,629	14,092
F-statistic instr.	3.84	3.57	4.72	3.65	4.58	4.72
T2SSLS first-stage results (all first-stage information used)						
# benign months	−0.3527***	−0.3527***	−0.3527***	−0.3527***	−0.3527***	−0.3527***
	(0.0749)	(0.0749)	(0.0749)	(0.0749)	(0.0749)	(0.0749)
# observations	27,913	27,913	27,913	27,913	27,913	27,913
F-statistic instr.	22.19	22.19	22.19	22.19	22.19	22.19

Notes: Each cell states a separate regression. In rows 1 and 2 only observations with complete second-stage information were used to estimate the first stage. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

Table O7: Long-term IV strategy treating sickness absences in grades 1 and 4 as two endogenous variables – First-stage results

	(1)	(2)	(3)	(4)	(5)	(6)
	Sample for					
	> *Folk-skola*	Empl. 1960	Empl. 1970	Income 1970	Pensions 2002	Passed away≤70
Endogenous variable: first-grade sickness absence						
# benign months grade 1	-0.2834^{**}	-0.3592^{**}	-0.3706^{***}	-0.2983^{***}	-0.3872^{**}	-0.3706^{***}
	(0.1317)	(0.1451)	(0.1403)	(0.1059)	(0.1769)	(0.1403)
# benign months grade 4	-0.1958	-0.1959	-0.2282	-0.2019	-0.2306	-0.2282
	(0.1833)	(0.1848)	(0.1811)	(0.1745)	(0.2229)	(0.1811)
# observations	10,978	12,664	13,821	9,121	9,380	13,821
F-statistic instr.	1.49	3.08	4.99	2.39	2.87	4.99
Endogenous variable: fourth-grade sickness absence						
# benign months grade 1	0.1471	0.1533	0.1258	0.1503	0.1856	0.1258
	(0.1391)	(0.1338)	(0.1250)	(0.1355)	(0.1484)	(0.1250)
# benign months grade 4	-0.1794	-0.1877^{*}	-0.2074^{*}	-0.1515	-0.1919^{*}	-0.2074^{*}
	(0.1118)	(0.1111)	(0.1090)	(0.1038)	(0.1150)	(0.1090)
# observations	11,253	12,993	14,092	9,338	9,629	14,092
F-statistic instr.	3.63	5.19	5.15	2.86	4.13	5.15

Notes: Each cell states a separate regression. The first-stage F-statistic of the instrument is calculated using the method suggested by Sanderson and Windmeijer (2016). Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

Table O8: Long-term IV strategy treating sickness absences in grades 1 and 4 as two endogenous variables – Second-stage results

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable					
	> *Folk-skola*	Empl. 1960	Empl. 1970	Income 1970	Pensions 2002	Passed away≤70
Sickness absence grade 1	0.0051	-0.0100	-0.0137	185.65	-695.65	0.0097
	(0.0243)	(0.0162)	(0.0143)	(570.65)	(2759.25)	(0.0105)
Sickness absence grade 4	-0.0152	-0.0003	0.0090	73.05	-2104.72	0.0042
	(0.0151)	(0.0084)	(0.0129)	(554.44)	(2826.38)	(0.0126)
# observations	8,417	9,677	10,491	6,976	7,161	10,491

Notes: Each cell states a separate regression. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

Online Appendix: sample selection

The socio-economic characteristics assessed in the church books used to gather the base dataset are available for all 30,150 individuals born in the sampled parishes. If exam catalog information are missing at random, the mean value of those characteristics between individuals we are able to trace down in school should equal the mean of the full sample of all 30,150 individuals. Table O9 shows the results. The first two columns show the means and standard deviations over all individuals, while the second two columns give the means and SD of the subsample of individuals for that we have exam catalog information. The right-most column indicates whether the difference of the means is statistically significant at any of the conventional levels. Only the share of individuals born in certain years differs occasionally at the 5 per cent level. Individuals we are able to trace down are more likely to be born in 1935. However, the absolute difference is quite small and we do not think that this is somehow correlated with the relationship between absence and performance. A likely reason of the difference is that exam catalogs are often missing for entire schools and school years so that the data are missing for a larger number of individuals. All in all, Table O9 does not indicate systematic sample selection.

Still, to investigate this further Table O10 gives the baseline short-term effects separately for individuals who did not move parishes between birth and grade 4 and individuals who have moved. If moving is selective with respect to absence and performance in school, the effects would differ between the samples. This does not seem to be the case. In fact, for the siblings FE model, the point estimates are the same, even thus the association is only statistically different from zero for same-parish matches due to the fewer movers observations.

Table O9: Balancing check for church and school data samples

	(1)	(2)	(3)	(4)	(5)
	Background variable in				
	full sample		exam catalog sample		Difference
Variable	mean	SD	mean	SD	significant
Female	0.49	(0.50)	0.47	(0.50)	
Year of birth: 1930	0.18	(0.38)	0.17	(0.38)	*
Year of birth: 1931	0.17	(0.38)	0.16	(0.37)	**
Year of birth: 1932	0.17	(0.38)	0.16	(0.37)	**
Year of birth: 1933	0.16	(0.36)	0.16	(0.37)	
Year of birth: 1934	0.16	(0.37)	0.15	(0.36)	**
Year of birth: 1935	0.15	(0.36)	0.19	(0.39)	
Father: farmer, fisherman, hunter	0.32	(0.47)	0.26	(0.44)	
Father: agricultural worker	0.27	(0.44)	0.22	(0.42)	
Father: service and sales worker	0.09	(0.29)	0.12	(0.32)	
Father: production worker	0.57	(0.50)	0.60	(0.49)	
Father: occupation unknown	0.23	(0.42)	0.31	(0.46)	
Mother employed	0.04	(0.19)	0.08	(0.27)	
Born out of wedlock	0.08	(0.28)	0.15	(0.36)	
Born in hospital	0.11	(0.32)	0.13	(0.34)	
Twin birth	0.02	(0.15)	0.04	(0.19)	
Observations	30,150		17,771		

Notes: Own calculations on church records. Columns 1 and 2 gives the mean value and the standard deviation (SD), respectively, for the full sample. Columns 3 and 4 state the corresponding values for the sample restricted to individuals for that we are able to find exam catalog information. Column 5 indicates whether the difference in the means is statistically significant based on the p-value of a t-test of equal means. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

Table O10: Short-term effects for same-parish matches and movers

	(1) OLS	(2) Sibl. FE	(3) Indi. FE
Same-parish matches			
Average grade points in units of SD			
Days of absence	-0.0038^{***}	-0.0039^{***}	-0.0031^{**}
	(0.0009)	(0.0012)	(0.0013)
Average grade points in units of income 1970			
Days of absence	-8.5387^{***}	-6.8373^{**}	-6.6752
	(2.9067)	(3.2566)	(4.1866)
Average grade points in units of pension 2002			
Days of absence	-55.7055^{***}	-37.7550^{*}	-44.8117^{**}
	(15.1663)	(21.7071)	(20.6301)
# observations	8173	8173	8173
# individuals/families	4110	1851	4110
Movers			
Average grade points in units of SD			
Days of absence	-0.0025	-0.0039	-0.0119^{**}
	(0.0031)	(0.0035)	(0.0059)
Average grade points in units of income 1970			
Days of absence	-9.4861	-2.1847	5.1823
	(10.4066)	(6.9127)	(6.4421)
Average grade points in units of pension 2002			
Days of absence	-43.2471	-34.0697	-85.7676
	(43.2815)	(62.9909)	(105.1412)
# observations	769	769	769
# individuals/families	408	202	408

Notes: See note to the baseline results table. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.