IFS

Institute for Fiscal Studies

Sarah Cattan
Daniel A. Kamhöfer
Martin Karlsson
Therese Nilsson

21/06

Working paper

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

UKRI Economic and 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*

February 20, 2021

*Cattan: Institute for Fiscal Studies, London, and IZA, Bonn; Kamhöfer: Heinrich Heine University Düsseldorf and IZA, Bonn; 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, Teresa Molina, 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. Therese Nilsson gratefully acknowledge financial support from the Swedish Research Council (dnr 2019-03553) and the Gyllenstierna Krapperup Foundation. 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.

Abstract

Despite the relatively uncontested importance of promoting school attendance in the policy arena, little evidence exists on the causal effect of school absence on long-run socio-economic outcomes. We address this question by combining historical and administrative records for cohorts of Swedish individuals born in the 1930s. We find that absence significantly reduces contemporaneous academic performance, final educational attainment and labor income throughout the life-cycle. The findings are consistent with a dynamic model of human capital formation, whereby absence causes small immediate learning losses which cumulate to larger human capital losses over time and lead to worse labor market performance.

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

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

1 Introduction

Student absence from school is pervasive around the world. While raising school attendance has long been the focus of policy in developing countries, the issue has also gained prominence in developed countries over the past decade. In the US for example, 19 per cent of fourth-graders were absent from school for three or more days in the last month in 2015. In some schools, absence has reached such alarming level that commentators talk about an "empty-desk epidemic".[1] State and national governments have started taking concrete measures to reduce absenteeism, ranging from better monitoring and public awareness campaigns to monetary fines.

Despite the relatively uncontested importance of promoting school attendance in the policy arena, there is still little causal evidence of the effect of absence on socio-economic outcomes. The few papers that credibly do so establish that absences in elementary and secondary school have detrimental impacts on academic achievement and school graduation (Goodman, 2014, Aucejo and Romano, 2016, Liu et al., 2019). While these papers are important in going beyond the correlational evidence that was so far available, they all focus on educational outcomes and on the US context. Much remains to be known as to whether (i) impacts found on educational outcomes reflect human capital losses that translate into long-term impacts beyond education, such as in the labor market; and (ii) whether these impacts generalize to other contexts. Indeed, given that the overall impact of absences will depend on how teachers and parents help children catch up for learning losses resulting from absences and on how such losses translate into long-term outcomes, impacts could differ across contexts.

This paper fills those two gaps by providing evidence of the impact of student absence on educational achievement, labor market outcomes and mortality in the context of Sweden. We construct a unique panel dataset following a representative sample of cohorts born between 1930 and 1935 in Sweden, which links digitized records of absence and educational performance in elementary school with census and tax

[1] See Chicago Tribune (2012).

register data on socio-economic outcomes throughout the life-cycle. These include final education, employment (at ages 25–30 and 35–40), labor market earnings (at ages 35–40), pension income from past labor market activity (measured at ages 67–72) and mortality. To our knowledge, this is the first dataset ever built that allows for an analysis of the impact of absences over such long time horizon. This is an important innovation in light of examples in the literature showing 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., 2017).

Even with such rich data, identifying the short- and long-term impact of student absences requires a credible strategy to tease out the causal impact of absence from the vast array of unobserved confounding factors. Students who miss school may have poorer health, less engaged parents, and/or less inspiring teachers, which could lead to spurious correlations between absence and achievement. To deal with the endogeneity of absence, we exploit two features of our data. To identify the short-term impact of student absences on educational performance, we exploit within-student, between-grade variation in absence and performance at two time points (grade 1 and grade 4) to control for all individual-level fixed determinants of achievement that could be correlated with absence. This is a similar strategy as that employed by Goodman (2014) and Aucejo and Romano (2016), but one added contribution of our dataset is that we are able to anchor the test scores to adult income and translate the effect on performance in school into its association with adult earnings.

To identify the long-term impact of student absences, we use the fact that our dataset contains many sibling pairs and exploit within-family variation in absence and long-term outcomes to purge the correlation between them from all family-level time-invariant unobserved factors. Unlike Liu et al. (2019) who use selection-on-unobservables bounding methods to partially identify the impacts of absence on high school graduation, this strategy allows us to point identify the long-term effects of absence on socio-economic outcomes. The identifying assumption underlying this strategy is stronger than the one we need to make for the analysis of short-term outcomes, since we are only able to control for family fixed effects (in addition to observable individual characteristics, teacher and school fixed effects), but we present a series of robustness checks to suggest that the biases associated with our estimates are unlikely to be neither statistically nor economically large.

Our analysis yields two main findings. First, we find a negative and very precisely estimated impact of student absence on academic performance in elementary school equivalent to 4.4 per cent of a standard deviation for ten days of absence (the average number of absences in our sample). Interestingly, this impact is very close to estimates presented in the literature for much more recent cohorts of elementary school children growing up in the US (Goodman, 2014, Aucejo and Romano, 2016).[2] Thanks to our anchoring exercise, we are able to show that, when translated in terms of potential impact on long-term income, this effect is of very modest economic size in the context of interest.

Second, our estimates of the long-term effects of absence are most pronounced for our income measures, both income measured at ages 35–40 and pension income, which is a good proxy for lifetime earnings. These effects are of moderate size, with ten days of absence in elementary school estimated to decrease these income measures by between 1 and 2 per cent. The estimates of the impacts of absences on other long-term outcomes are similar in size, but not as precisely estimated with the exception of a significant negative impact on men's likelihood to enroll in secondary education. Overall, the results are consistent with a model of skill accumulation, where, through the self-productivity of skills, early losses in human capital resulting from school absences lead to growing later skill deficiencies, which affect educational achievement, employment and income, but are only large enough to be picked up precisely on later outcomes in the life-cycle.

[2] Liu et al. (2019) also study the impact of student absence on academic performance in the US, but focus on absences in secondary school.

Our paper relates 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., 2019)[3]; 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 resulting from extreme weather events (e.g., Marcotte, 2007, Marcotte and Hemelt, 2008, Marcotte and Hansen, 2010, and Hansen, 2011).

When it comes to individual absence from school, three recent studies, Goodman (2014), Aucejo and Romano (2016), and Liu et al. (2019), analyze the contemporaneous effects of individual school absence on educational achievement 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, Goodman (2014) and Aucejo and Romano (2016) show that school results are negatively affected by absence. Both studies control for institutional heterogeneity using school, teacher and individual fixed effects, as we do in this paper.[4]

In contrast with the other two papers aforementioned, Liu et al. (2019) estimate the impact of school absences during secondary school on educational outcomes. They use between-subject, within-individual variation in absence and test scores to identify the impact of absences in secondary school on contemporaneous achievement

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

[4] To corroborate their results, both studies also implement an instrumental variables (IV) approach using local variation in snowfall (Goodman, 2014) and infectious diseases (Aucejo and Romano, 2016) to instrument school absence. In contrast with those papers, we do not implement an IV approach and discuss the conceptual issues associated with it in the context of estimating the impact of school absences in Section 4.

and estimate bounds around the impact of absence on high school graduation and college enrolment using the method of Oster (2019). A unique feature of their paper is that they know the timing of absence so they can test for a differential impact of absences happening early versus late in the academic year. While our historical records of absence do not allow us to make these distinctions, we complement their paper by exploiting the presence of siblings in the dataset to point identify the impact of elementary school absence on a wide range of adult outcomes, in addition to final educational achievement.

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, pensions, 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 institutional context and the educational system are relevant factors for observed effects (see e.g., Pischke, 2007; Galama et al., 2018; Gathmann et al., 2015), but the impact of individual school absence has not yet been analyzed outside the US. Sweden makes a particularly interesting case in comparison to the US, given that its labor market was characterized by active labor market policies and highly compressed wages (Erixon, 2008) – and embedded in a Social Democratic welfare state providing comprehensive social insurance against most of the health and social risks that a worker faces (Bergh, 2014).

Our results show that these innovations to the literature 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 century. The long-term estimates indicate that the short-term human capital losses measured by impacts on test scores may translate into long-term penalties on the labor market. The latter finding underlies the importance of using outcomes measured at different points of the career, as impacts measured early in the career would miss the impact of absences beyond compulsory schooling.

The remainder of the paper proceeds as follows. Section 2 provides some background on the schooling system in Sweden in the 1930s and 1940s. Section 3 describes the data and presents some descriptive statistics. Section 4 discusses our empirical strategy, while Sections 5 and 6 present our results for short-term and long-term effects, respectively. Section 7 concludes.

2 The Swedish schooling system

In the 1930s and 1940s, the period during which our cohorts grew up, all children in Sweden were required to attend public school, *Folkskola*, starting at the age of seven. The first four grades of *Folkskola* were mandatory for all students. We refer to these grades as elementary school. Students progressing to secondary schools (to receive more than compulsory schooling) generally matriculated after grade 4 of *Folkskola*. Admissions to secondary school depended on performance. Students not pursuing post-compulsory schooling, remained in *Folkskola* for a fifth and sixth grade.[5]

Students attended elementary school full time, six days a week.[6] The educational system exhibited several features of a modern educational system – like absence of tuition fees and joint instruction of boys and girls at all educational levels (Erik-

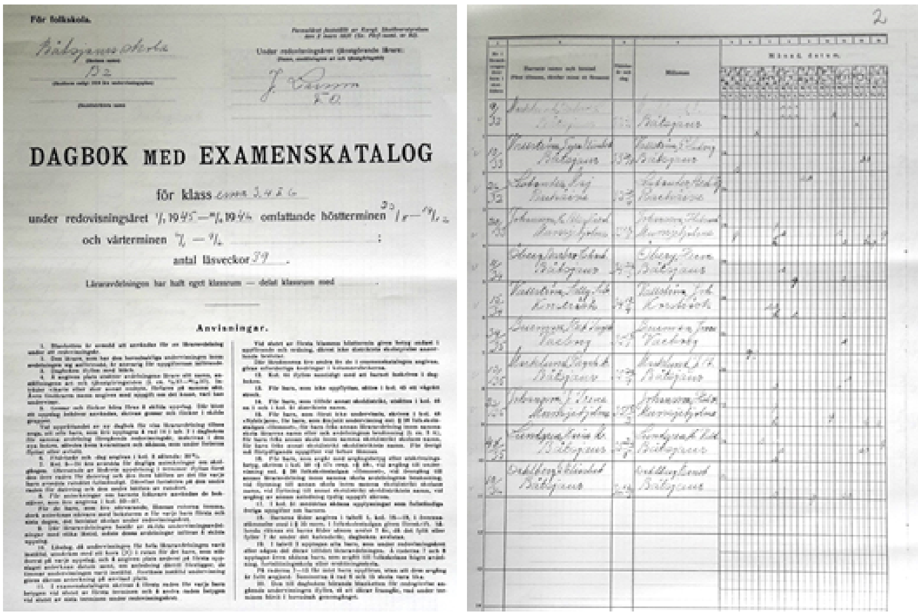
[5]After a compulsory schooling reform in 1936, a seventh grade of *Folkskola* was mandatory for students not enrolled in secondary schools, see footnote 29. Another reform was the expansion of the instructional time per school year from 34.5 to 36.4 and 39 weeks. Although this affected the potential days of absence, we find no pattern that the school year length expansion affects days of absence. We nevertheless control for the school year length in our regression analyses. For an analysis of the compulsory schooling reform and the school year length, see (Fischer et al., 2019).

[6]Instruction ended at noon on Saturdays. Following an exception rule, schools in rural areas had the possibility to offer half-time reading but this option was very limited and only 0.5 per cent of our sample took half-time reading.

son and Jonsson, 1993) – but education was very selective and a clear majority of students only completed compulsory education (Fischer et al., 2019).[7]

The responsibility for providing primary education was decentralized to 2,400 school districts, but the Ministry of Ecclesiastical Affairs provided nationwide standards that applied to all school districts, including the *Folkskola* curricula. Three theoretical subjects were taught in elementary school: math, reading and speaking, and writing. The government established several grading principles (SOU, 1942), which dictated that teachers should reward the quality of knowledge and regularly take notes to ensure that grading reflected performance through the year. Appendix C provides details on the school system and grading principles.

Figure 1: Example of an exam catalog



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

As their main organizational tool teachers kept daily records in an exam catalog (see Figure 1). In the catalog, the teachers recorded students' performance and absences and noted whether absences were due to sickness, natural obstacles, inappropriate clothes and shoes, other valid reasons for absence, or no valid excuses (truancy). At

[7]In rural areas, children in several grades were taught together, when there were not enough students to have separate classes. In the regression analyses, we control for the lowest and highest grade taught to students in the same classroom.

the end of the school year, the teachers summarized the days of absence by type and the final grades by subject. These catalogs provide a unique source of historical data on the achievement and absence of cohorts born almost a hundred years ago. To conduct the analysis presented in this paper, we have digitized these exam catalogs and linked this information to a variety of administrative data sources on long-term outcomes, which we detail in the next section.

3 Data

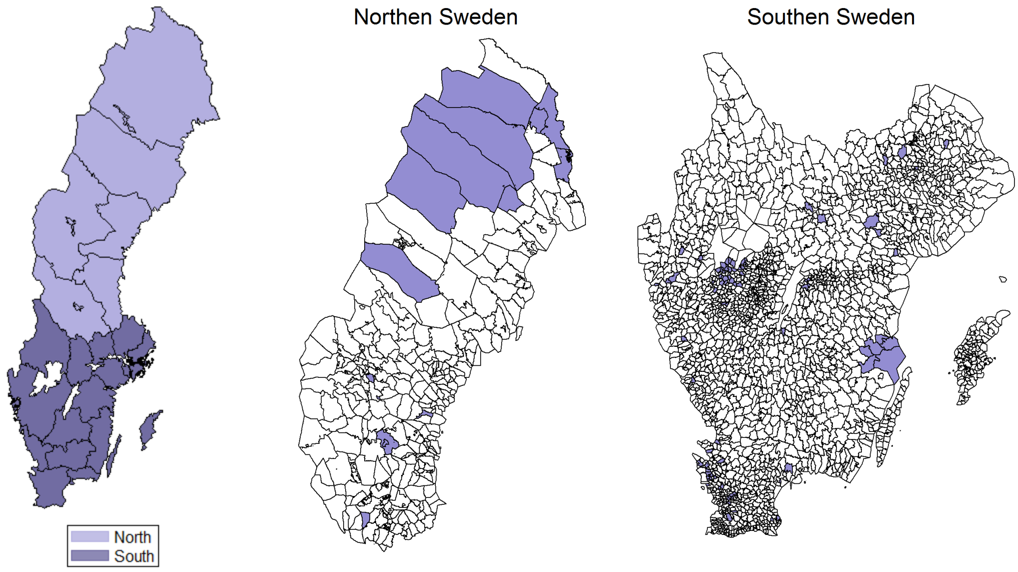
3.1 Data sources and long-term outcomes variables

The data we use in this paper combine several historical and administrative data sources. The base of 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 2,400 Swedish parishes.[8] Figure 2 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' full names, birth date, whether the child was born in a hospital, whether the birth was a twin birth, and whether the child was born out of wedlock. These information in the church records allow to identify siblings. Church records also include information on the parents' occupation at the time of the child's birth, which we use to create an indicator for whether the child's mother was employed at the child's birth and indicators for whether the child's father was an agricultural, a production or a service worker.[9] These occupational indicators are the best proxies available in the data to measure family socio-economic status during the individual's childhood. We will use all these vari-

[8] Appendix D.1 provides details on the construction and representativeness of the base data.

[9] To construct these indicators, we use the first digit of the Historical International Standard Classification of Occupations (HISCO) code for the fathers' occupation. The HICSO code is historical version of today's International Standard Classification of Occupations (ISCO) code, see van Leeuwen et al. (2002).

Figure 2: Spatial 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.

ables as controls in our analysis, and we present summary statistics in Appendix Table B1.[10]

Individual schooling information was collected from historical archives in each parish. Specifically, we collected the exam catalogs in which teachers made systematic notes about types of absence and reported educational performance for each student, in all elementary schools of the 133 parishes in our base dataset. As shown in Figure 1 each student is listed with their first name, surname, date of birth and parents' name. Using this information, we merge the schooling information onto the base dataset. We were able to match schooling information for 18,056 out of the 30,150 children with church records born between 1930 and 1935.[11] We focus on grade 1 and grade 4 and digitize the end-of-school-year summary information of the exam

[10]Appendix D.4 provides a list of definitions for all variables in our data.

[11]The reason for most of the missing information is that archives of certain schools were accidentally destroyed, which does not relate to individual selection. Another reason why we do not get a perfect match is that individuals moved out of the parish they were born before school age. We significantly reduce matching problems related to migration by tracking migrants and collecting school records from their destination parishes. See Appendix D.2 for details and tests related to matching of school data.

catalogs, that is absence by type (sickness absence and non-sickness absence)[12] and final grades by subject (math, writing, reading and speaking).[13] Grade 1 and grade 4 are pivotal as grade 1 represents the first occasion at which educational performance can be observed and grade 4 represents the last as some students proceed to secondary schooling afterwards. With cohorts born in 1930–1935 the schooling data covers the school years 1936/37 to 1946/47.[14]

We add information on highest educational level completed obtained from the 1970 population census (SCB, 1972). To match individuals we use their full name, date of birth and parish of birth. Given that individuals are aged 35–40 in 1970, this information should reflect their final education. We measure educational attainment with an indicator that takes the value 1 if an individual attains a more advanced level than *Folkskola*, and 0 otherwise.

Turning to labor market outcomes, we gather information about employment at ages 25–30 and 35–40 from the 1960 and 1970 population censuses (SCB, 1972, 1962). We construct our employment outcome variables as indicators that equal 1 if the individual is employed (part-time or full-time), and 0 otherwise. We also create two measures of labor market income. The first one, from the 1970 population census, corresponds to labor income (measured in Swedish krona) at age 35–40, when individuals in our sample were in prime working age. In our main set of results, we construct a measure of income where we have imputed zeros for individuals who were not employed at the time.[15] The second measure of income comes from the tax

[12]The exam catalogs include columns for several reasons for non-health related absence, but teachers often only noted other absence without naming the reason. Therefore we focus only on sickness absence and non-sickness absences.

[13]Given how resource-intensive the task of digitizing historical archives is, we focused on digitizing the yearly summaries of absence only, as opposed to individual absences throughout the year. This means that the data at hand do not allow us to identify the length of absence spells and, as opposed to Liu et al. (2019), we are not able to distinguish the effect of absences earlier versus later in the school year.

[14]The WWII falls in the time under review. Sweden was neutral in the war and there was an oversupply of teachers (Paulsson, 1946). We have not found any historical sources suggesting that the war caused major disruptions in education. Moreover, the probability that we found exam catalogs in the local archives was the same for the war years and non-war years.

[15]In robustness checks in Appendix Table B8, we also run the analysis with income and log income that exclude zeros.

register data. Specifically, we use average annual pensions in 2003–2008, i.e., when our youngest cohort is aged 68–73. For our cohorts full pensions require thirty years of contributions and is based on the fifteen highest income years (Sundén, 2006).[16] Pensions are thus a proxy of lifetime earnings and should be less sensitive to year-to-year fluctuations in labor supply than earnings at a particular point in time. This is desirable feature in evaluating the long-run effects of absence, especially for women as child-related career interruptions, such as parental leave, would be less likely to affect pensions than annual earnings.[17]

Finally, we use church records and the Swedish Death Index of Federation of Swedish Genealogical Societies (Federation of Swedish Genealogical Societies, 2014) to obtain information the exact date of death for all individuals that passed away between 1901 and 2013. Based on this we generate indicators for whether the individual passed away before 1960, 1970 and 2003, respectively, which we use in our robustness checks.

For outcomes measured in adult ages we match individuals in the matched schooling data in the different registers using either combined information on first, middle and last name, sex, date of birth and parish of birth or a unique social security number. We are able to match 18,072 of the individuals in the schooling dataset to the 1960 census and 15,947 individuals to the 1970 census. For pensions we are able to match 13,701 individuals. Missing information on labor market outcomes might be due to individuals migrating from Sweden or passing away. We can directly investigate mortality and we see that 50 per cent of the unmatched individuals died before the 1960 census enumeration, 37 per cent died before the 1970 census enumeration, and 66 per cent died before we have tax records 2003. Appendix D.3 provides details about the matching procedures involved in generating the dataset including adult and later-life outcomes.

[16]Widows were in some cases entitled to a certain share of their spouse's earnings after their death. The widow pension represent the most important deviation from the general rules.

[17]For details about the pension system and related rules applying to our cohorts, see Appendix E.

3.2 Educational achievement measure

We measure educational performance using the grades teachers assigned to students at the end of the school year. Unlike tests that take place on a certain date, the frequent recording of student performance throughout the year reduces the likelihood that educational achievement scores are driven by teaching-to-the-test behavior of teachers and/or idiosyncratic factors on the day of the test. Moreover, because teachers kept records of performance throughout the year, the yearly scores are unlikely to suffer from recall bias.

While exams were not standardized across the country, teachers were provided with clear grading principles laid down by a Royal Commission. The grading scale included seven levels, where 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 scale remained unchanged during the period of interest, the grading principles changed slightly. From the school year 1940/41 onwards, teachers were advised to award the grade BA ("passed with credit") for an average performance. 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.

Table 1: 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 point 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.

We assign to each letter grade a numerical value, ranking from 1 (lowest grade) to 15 (highest grade) in order to take into account that teachers could assign a plus and a minus sign to a student's grade. Table 1 gives the mapping of the potentially ordinal grades into grade points.[18] 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, writing, 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. As we show in Section 3.4 the distribution of grades observed in our sample is remarkably in line with the Royal Commission's grading principles. This gives us confidence that it is a meaningful measure to compare achievement across students.[19]

3.3 Sample selection

Table 2 provides an overview of how our matched schooling data corresponds to number of individuals, siblings and families. As mentioned in Section 3.1 we are able to match schooling information in either grade 1 or 4 for 18,056 out of the 30,150 children in the church record base data born 1930-1935, resulting in almost 29,000 student–grade observations. Using information on parents, we can identify siblings born between 1930–1935 in the matched schooling data. Our sibling panel consists of 8,567 siblings for which we have more than 14,000 student–grade observations

[18] In the appendix (Table B4), we also test the robustness of our results to replacing 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 grading principles into account.

[19] As we discuss in Section 4, our empirical strategy to recover the causal effect of absence on school performance includes a comparison of school grades between siblings. One concern about non-standardized measures of achievement is that, if siblings have different teachers, grades may not be comparable to each other. Teachers were supposed to follow the central grading guidelines and were instructed to make no adjustment for school form, wherefore the grades should reflect an absolute standard and not the relative position of a student in their class. Yet, in order to account for the possibility that siblings had different teachers applying different grading, we include teacher fixed effects in all specifications. Appendix Table B5 shows that estimates based on specifications controlling only for parish and school fixed effects are very similar to those controlling for teacher fixed effects as well, thus providing reassurance that subjective grading is unlikely to be a serious issue in these data.

Table 2: Sample composition

Sample	Number of... student–grade observations	students (individuals)	families	families w/ >2 children
Exam catalog data+church records				
– grades 1 and 4, cohorts 1930–35	28,931	18,056	13,205	960
– siblings panel	14,066	8,567	3,716	960
– siblings+individual panel	8,934	4,467	1,987	445

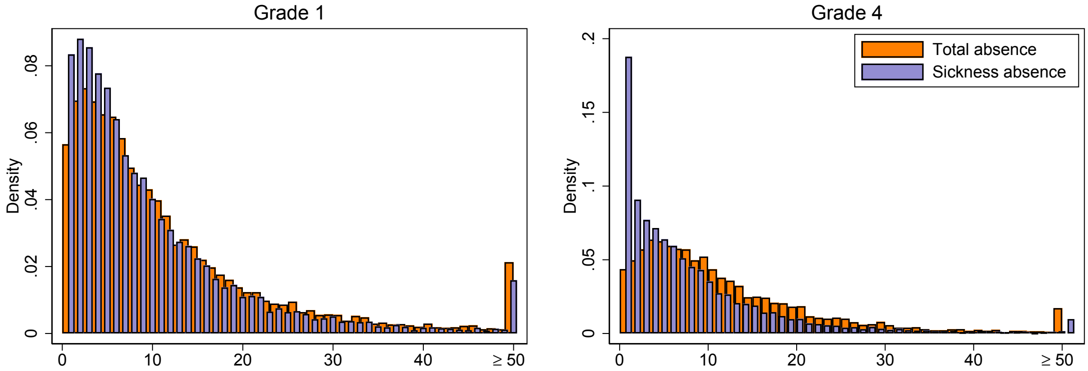
Notes: Own calculations based on exam catalogs and church records. All samples are restricted to observations with non-missing information in the key variables (grade points, absence, and grade). Starting are all 30,150 children born between 1930 and 1935 in our sample of 133 parishes, representative for all approximately 2,500 Swedish parishes at the time. The first row gives the number of individuals, individual–grade, and sibling pair observations we are able to match to exam catalog data based on the name, the date of birth, and the parish of birth. Parental information in the church records allows to identify siblings. The second row restricts the sample to siblings pairs, where we observe each sibling in at least one grade. The final row states the number of pairs of siblings for that we observe both grades for each sibling.

with schooling information from grade 1 and/or grade 4. These siblings stem from 3,716 families, and for 960 families we even have matched information for more than two siblings. As we elaborate further below, this sibling sample will be used to identify the long-term effects of absences using a family fixed effect estimator. To identify short-term effects of absences on contemporaneous achievement, we employ an individual fixed effects approach and we further restrict the sibling sample to individuals for whom we have schooling information both in grade 1 and grade 4. This gives an individual panel with 4,467 individuals, resulting in 8,934 student–grade observations.

3.4 Descriptive statistics

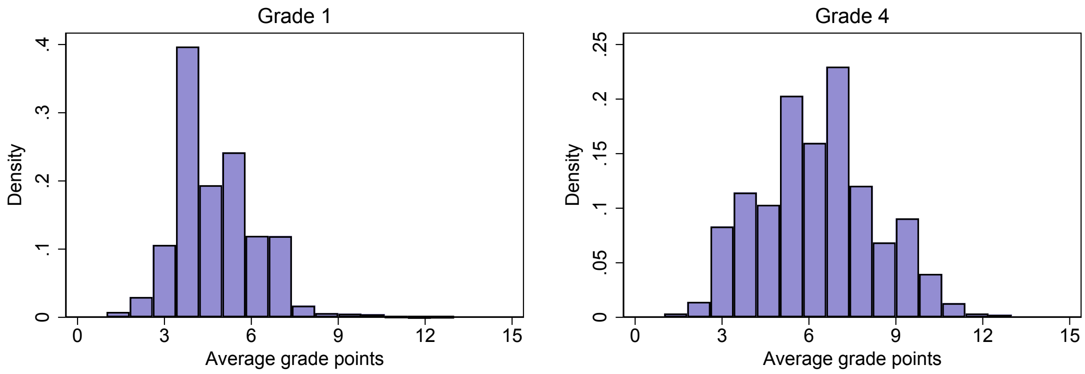
Our main explanatory variable is the number of missed school days in grade 1 and in grade 4, and we can distinguish between absences due to sickness and absences due to other reasons. Figure 3 shows the distribution of individual days of total absence and sickness absence in grade 1 and 4, respectively. In grade 1, 63 per cent of all students miss 10 or less days and about 6 per cent of all students have no absence. The average number of missed days in grade 1 is 11.1 days (median 7 days). In grade 4, students tend to miss slightly more days (mean 11.6 days, median 8 days). 60 per

Figure 3: Distribution of (sickness) absence in grade 1 and grade 4



Notes: Own calculations based on exam catalog information. 14,066 observations.

Figure 4: Distribution of average grade points in grade 1 and grade 4



Notes: Own calculations based on exam catalog information. 14,066 observations.

cent of all students miss 10 or less days and 4 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 of absence in elementary school (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 here.[20]

Figure 3 further illustrates that most absences are sickness absences. The average number of missed days for other reasons than sickness is 1.7 in grade 1 and 3.3 in

[20]In the empirical analysis in Section 5 and Section 6 we winsorize our data at top two per cent of days of sickness absence and top two per cent of days of non-sickness absence.

Table 3: Descriptive statistics on long-term outcomes

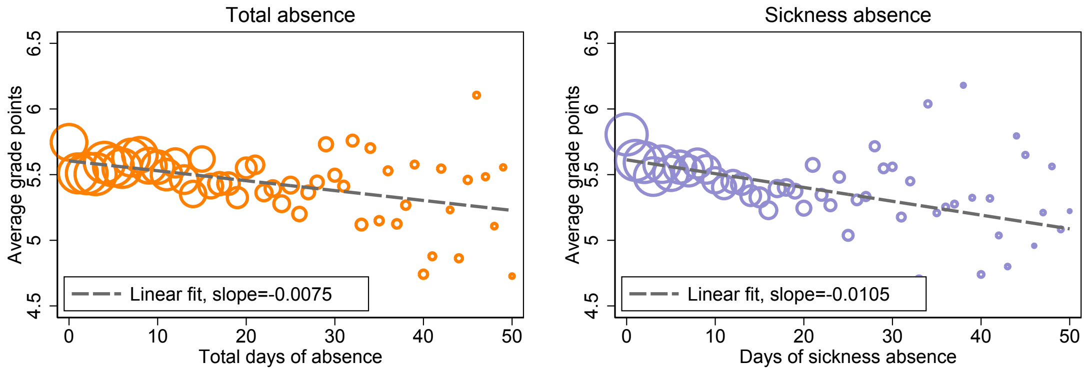
	Age range	Mean All	Mean Female	Mean Male	# obs	% female
Education						
More than *Folkskola* (in per cent)		11.8	12.1	11.4	5,976	49.6
Employed (in per cent)						
in 1960	25–30	65.7	36.6	94.1	7,434	49.4
in 1970	35–40	74.3	55.9	92.6	5,976	49.6
Earnings (in Swedish krona, current values)						
Labor market income 1970	35–40	19,275	10,039	28,475	5,886	49.9
Pensions 2003–2008, if>0	68–78	160,237	138,568	183,468	4,772	51.7

Notes: Own calculations. The numbers in this table refer to the number of individuals in the siblings sample (second row of Table 2) we are able to match to the long-term outcomes. The 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. The education indicator takes the value 1 if the individual has more than compulsory *Folkskola* education, and 0 otherwise. Employment in 1960 and 1970 is taken from the census information in these years. The employment indicators take the value 1 if the individual is employed, and 0 otherwise. Labor market income 1970 is based on census 1970, unemployment enters the labor market income as zero. Pensions information is taken from tax registers and averaged over 2003–2008. Zero pension are dropped (very few cases). Labor market income and pensions are measured in Swedish krona in the year the information refers to.

grade 4. In grade 1 and 4, 61 per cent and 39 per cent of all students, respectively, never miss a day for other reasons than sickness. Appendix Figure A1 plots the within-family and within-individual distributions of days of (any) absence and days of sickness absence.

Turning to school achievement, Figure 4 shows the distribution of the raw average grade points over math, writing, and reading and speaking, by grade. In line with the guidelines of the Royal Commission (SOU, 1942), 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. Appendix Table A2 shows the distributions of grade points by subject and school grade.

Figure 5: Descriptive relationship between (sickness) absence and academic performance



Notes: Own calculations based on exam catalog information for 14,066 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. For legibility we only plot cells up to 50 days of absence. The fitted line is taken from a simple linear regression of performance on total absence and sickness absence, respectively, using all information.

Table 3 presents descriptive statistics for long-term outcomes.[21] The highly selective nature of the education system is reflected in only about 12 per cent of the individuals in our sample having more than compulsory education. Employment measured in 1960 and 1970, when our sample is aged 25–30 and 35–40 respectively, show that the labor market was gender-segmented for our cohorts, with men significantly more likely being employed than women. This is also mirrored by the differences in the 1970 earnings measure corresponding to incomes in prime working age (35–40 years old). As expected pensions are more equal between men and women than 1970 earnings are.

To set the stage for the analysis to come, we conclude this section by documenting the associations between the number of days of absence (across grade 1 and grade 4) and our main outcomes of interest. As expected, the correlation between total days of absence and academic performance is negative, as shown in Figure 5. The linear

[21]The numbers of observations in Table 3 differ from the numbers of matched observations due to the applied sample selection. For all outcomes with exception of completed eduction, the sample selection does not change the mean value significantly. The share of individuals with more than *Folkskola* in the sample not restricted to siblings in 16.5 per cent. Appendix Table B7 uses the unrestricted sample to estimate the effect of absence and absence interacted with a siblings indicator on long-term outcomes. In spite of the different means, the effect of absence on education does not significantly differ for siblings.

fit indicates that the raw correlation is slightly more strongly negative for sickness absence than it is for any absence. Similarly to Figure 5, Figure 6 shows the raw association between the number of days of absence and our long-term outcomes. Plot (a) shows a negative relationship between absence and the probability of having more than *Folkskola*, while plots (b) show a slightly negative relationship between absence and employment both in 1960 and 1970. In plots (c), we graph the distribution of income in 1970 and of pension income for groups of students with less than 5, between 5 and 20 days, and more than 20 days of absence. Kolmogorov–Smirnov tests of equality between the densities confirm that there are significant differences between them.

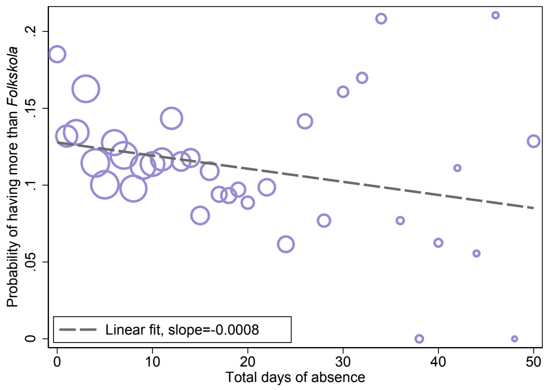
In all cases, the p-value is either below conventional significance thresholds or (for the 5–20 and more than 20 days comparison for 1970 and the less than 5 and 5–20 days comparison for pensions) only slightly above 0.10 (see figure note).[22]

While these correlations point to a potentially negative effect of school absences on short- and long-term outcomes, they are obviously not 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 of such selection, Figure 7 shows a boxplot of the total days of absence in groups defined by different observable characteristics. The indicator for whether an individual was born in 1933 exhibits the highest inter-quartile range in days of absence, followed by the indicators for being born out of wedlock and for the child's father being employed in the service sector. Overall we note quite limited signs of selection based on these observables, though this obviously does not rule out selection on unobservable characteristics, including individual health. In addition, the statistics in Figure 7 indicate that it is not obvious whether we should expect students to select positively or negatively

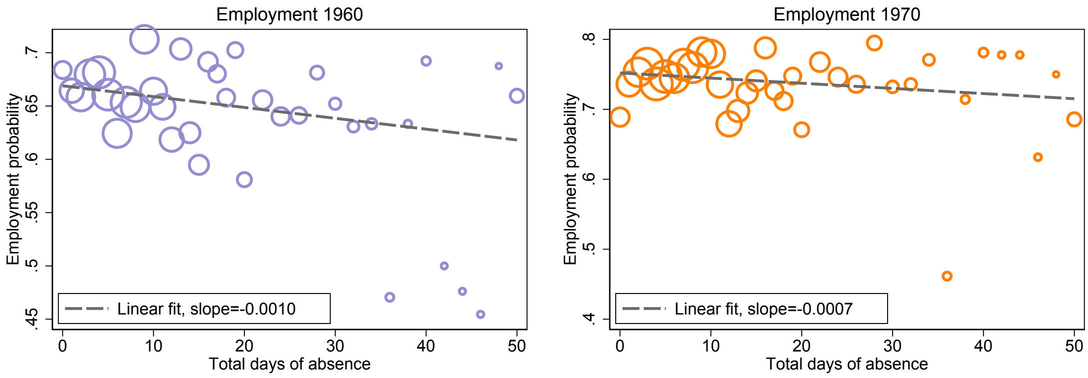
[22]Appendix Figure A3 shows the survival rate of individuals who have missed less than 5 days, 5 to 20 days or more than 20 days, respectively. In contrast to income, the differences between the lines are small and statistically insignificant, suggesting no systematic association between absence and mortality.

Figure 6: Descriptive relationship between total days of absence and long-term outcomes

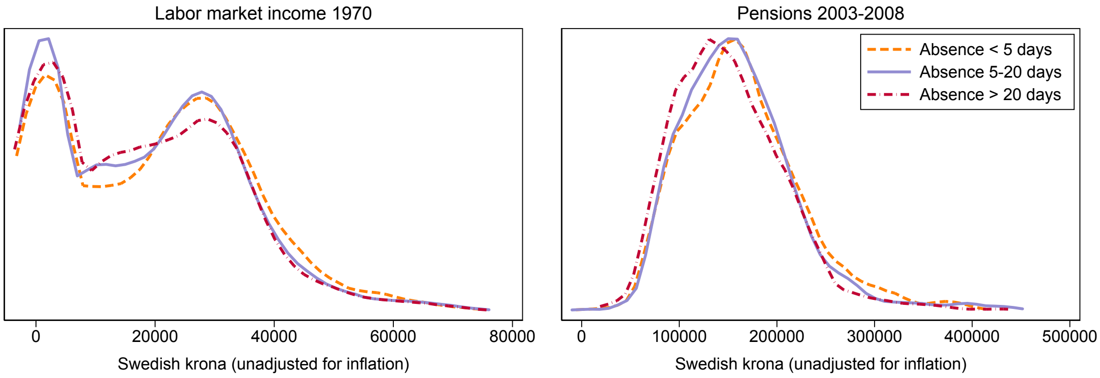
(a) Final education



(b) Employment



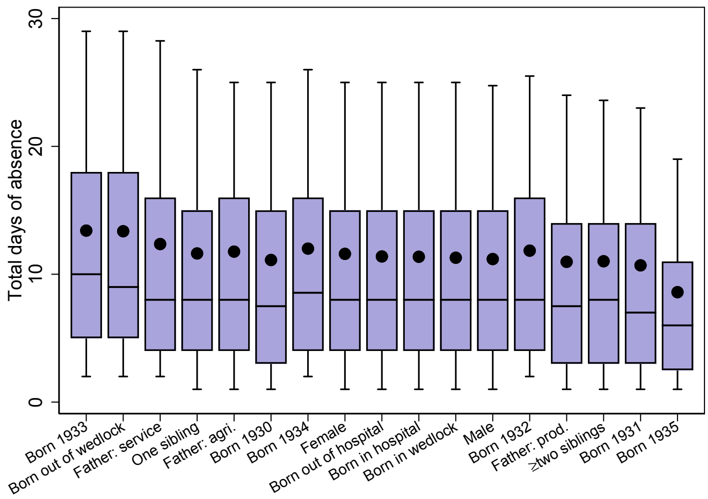
(c) Income



Notes: Own calculation based on exam catalog, census 1960 and 1970 aa well as tax register information. Number of observations as in Table 3. Labor market income is zero for unemployed. In panels (a) and (b), the y-axis gives the expectation value of the outcome conditional on the days of absence value on the x-axis (daily bins below 20 days, two-day bin from 20–50 days). Marker size gives relative number of observations in the days-of-absence cell. The line states the linear fit. In panel (c), we use a Kolmogorov–Smirnov test for the equality of the distributions. Income 1970: corrected p-value are 0.019 for <5 days versus 5–20 days, 0.001 for <5 days versus >20 days, and 0.102 for 5–20 days versus >20 days. Pensions. 0.108 for <5 days versus 5–20 days, 0.000 for <5 days versus >20 days, and 0.001 for 5–20 days versus >20 days.

into absence. Together this emphasizes the need of a sound empirical strategy to deal with potential selection.

Figure 7: Total days of absence by socio-economic characteristics



Notes: Own calculations based on exam catalog information for 14,066 observations. Each characteristic (on the x-axis of the figure) gives a boxplot of the mean values of total days of absence (depicted by the dot), the median (the horizontal line within the box), the 75$^{\text{th}}$ and 25$^{\text{th}}$ percentile (the upper and lower hinge, respectively), and the 90$^{\text{th}}$ and 10$^{\text{th}}$ percentile (the upper and lower adjacent line, respectively) o the y-axis. The characteristics are ordered along the inter-quartile range of days of absence.

4 Empirical strategy

4.1 Identification of short-term effects of school absence

Our analysis first aims to identify the effect of the number of days of absences in an academic year on academic achievement measured at the end of that year. We denote Y_{ifstg} the grade point in grade g of a student i born in family f attending school s taught by teacher t. A natural starting point is to assume that Y_{ifstg} depends linearly on the number of days student i was absent from school in grade g (D_{ig}), a set of individual-specific controls (X_{it}) and a set of school/grade-specific controls (Q_{stg}):[23]

$$Y_{ifstg} = \beta_0 + \tau D_{ig} + X_{it}\beta_1 + Q_{stg}\beta_2 + \varepsilon_{ifstg}, \tag{1}$$

[23]We present the identification strategy in the context of this linear model, but in the empirical application we also test for non-linear effects, see Section 5.2.

where ε_{ifstg} is an error term capturing unobservable determinants of academic achievement. X_{it} includes the student's gender, month–year of birth, age (in months), mother and father's year of birth dummies, father's occupation at the time of the child's birth, indicators for whether the student was born out of wedlock, whether he or she was born in hospital, and whether he or she has a twin, as well as parish fixed effects. In addition, we may include past performance – which turns the model into a value-added model (cf. Todd and Wolpin, 2003) – or past absences. Vector Q_{sg} includes a variable measuring class size in grade g, variables measuring lowest and highest grade taught to students in the same classroom, and length of the school year in weeks.[24]

In order to interpret τ in equation (1) as the incremental effect of one additional day of absence on grade points in the same school year, we need to assume that $\mathbb{E}(\varepsilon_{ifstg}|D_{ig}, X_i, Q_{stg}) = 0$. There are several reasons why this assumption may not hold. To structure our discussion of the threats to identification, we decompose the error term into four unobserved terms, related to four sets of factors, plus an i.i.d. error:

$$\varepsilon_{ifstg} = \rho_s + \eta_t + \alpha_f + \lambda_i + u_{ifstg} \tag{2}$$

where ρ_s is a school-level unobserved factor, η_t is a teacher-level unobserved factor, α_f is a family-level unobserved factor, λ_i is a child-specific unobserved factor, and u_{ifstg} is an idiosyncratic term varying across children and grades.

It is possible to think of reasons why each of these factors could simultaneously affect academic performance and school absence. For example, school- and teacher-level unobserved factors, ρ_s and η_t respectively, could be correlated with school absence if a particularly well-managed school or a particularly good teacher is also more effective at keeping students engaged and making them less willing to miss class.

[24]Class size is incorporated as a spline. 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.

Family-level unobserved factors, α_f, such as parenting style, may also be correlated with both achievement and absences if more engaged parents were more effective at supporting their children's learning or more likely to keep their children's healthy.

Beyond school, teacher and family-specific factors, child-specific unobserved factors, λ_i, may also determine both school performance and school absence. For example, a child with a higher (time-invariant) health endowment may not only be more able to concentrate and learn new material, but also less likely to fall sick and miss school. Similarly, a child with a more conscientious personality may be more likely to do her homework and less likely to play truant. Finally, there could be child- and grade-specific unobservables u_{ifstg} that correlate with both performance and absence. This could be the case, for example, if parental investments vary with the child's age, or if there is a particularly severe weather shock in one year, which hampered the student's learning not only by making him/her miss days of school due to sickness, but also by affecting his/her teacher's quality that year.

The nature of our dataset allows us to handle the first three sources of endogeneity. We can control for teacher- and school-specific time-invariant unobservables by including school and teacher fixed effects. Furthermore, we can exploit the fact that we observe academic performance and absences in two school grades (1 and 4) in order to control for individual-specific time-invariant unobservable factors through the inclusion of individual fixed effects. This strategy is similar to that used by Aucejo and Romano (2016) and, in our case, operationalized by estimating the following equation:

$$Y_{ifstg} = \beta_0 + \tau D_{ig} + Q_{stg}\beta_2 + S_{ig} + T_{itg} + \delta_i + u_{ifstg}, \tag{3}$$

where δ_i is an individual fixed effect. Effectively, this strategy controls for both family-level and individual-level time-invariant factors, α_f and λ_i, in equation (2). In this specification, we also control for school fixed effects S_{isg} and teacher fixed effects

T_{itg} in order to control for school- and teacher-level shocks to academic performance that could also affect student absences.

Equation (3) recovers the causal effect of an additional day of absence on achievement under the assumption that $\mathbb{E}(\varepsilon_{ifstg}|D_{ig}, X_i, Q_{sg}, \rho_s, \eta_t, \alpha_f, \lambda_i) = 0$ and that the effect of a day of absence on achievement in grade 1 is equal to the effect of a day of absence on achievement in grade 4. That is, under these assumptions, the parameter τ will be estimated to be:

$$\tau = \sum_{w=1}^{W} \pi_{w-1} \mathbb{E}\left[Y_i\left(w\right) - Y_i\left(w-1\right)\right], \tag{4}$$

for a causal response function $Y_i\left(w\right)$ specifying the outcome for any realization of the absence variable. π_{w-1} represents the histogram of absence days in the population. The main identifying assumption is that any unobserved grade- and child-specific factor u_{ijsg} affecting student i's achievement in grade g are uncorrelated with the child's absences, conditional on the observables and fixed effects included in the model. This assumption is similar to that maintained in Aucejo and Romano (2016), but stronger than that in Liu et al. (2019) who exploit within-grade, between-subject variation in absences to control for time-varying individual-level unobserved shocks.[25]

To further investigate the validity of our identifying assumption, we perform two robustness exercises (see Section 5.3). First, in light of the fact that most absences are due to sickness, we compare the effects of sickness absences to that of non-sickness absence on contemporaneous academic achievement. The idea underlying this test is that the most likely time-varying individual-level unobserved factor that could drive variation in absence between different school grades are idiosyncratic health shocks. If that were the case, we would expect that the impact of sickness absence on academic achievement reflects the effect of the loss in instructional time as well as that of the health shock, while the impact of a non-sickness absence only

[25] On the other hand, Liu et al. (2019) require an absence of spillover effects across subjects for identification.

would reflect the effect of the loss in instructional time. If we estimate these effects to be undistinguishable from each other (as we do), we should have more confidence that our estimate of the effect of absence is not driven by an idiosyncratic health shock.

Of course, this test does not address the presence of other type of time-varying individual-level factors. We therefore also complement our point estimates of the short-term (and long-term) effect of absences with bounds allowing for extreme forms of correlation between absence and the unobservables that we cannot control for, following the method of Oster (2019).

4.1.1 Anchoring of educational achievement measure

As argued by Bond and Lang (2013), which representation of educational achievement we use can significantly alter the conclusion of the analysis. To circumvent this problem and provide economically meaningful interpretations of our estimated effects of absence on academic achievement, we anchor the grade point scale to pension income, a policy-relevant outcome measured in a meaningful metric (Swedish krona). To do so, we re-run our main models replacing our measure of educational achievement (grade points) as dependent variable with the fitted value of the following auxiliary anchoring regression:[26]

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

where y_{ig}^{anchor} is individual i's average pension income between 2003–2008, 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. The anchoring is performed separately for grade points in grade 1 and for grade points in grade 4.[27]

[26]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.

[27]Educational achievement anchored in earnings potential should not be confused with the effect of school absence on pensions. The anchored effect of absence still gives the short-term effect on educational performance, but scaled in units of Swedish krona instead of the somewhat hard-to-interpret standardized numerical grade points.

4.2 Identification of long-term effects of school absence

Our analysis then turns to estimating the effects of school absences in elementary school on adult and late-life outcomes. Here, we define our treatment variable as the average number of days of absence in grade 1 and 4. Because the average number of absences in all grades of elementary school is fairly stable across all grades however, we can think of the treatment variable as closely measuring the average number of days of absence throughout elementary school.[28]

For long-term outcomes, it is obviously not possible to exploit within-individual variation in absences as in the case of short-term outcomes. Instead, we leverage another strength of our dataset, which is the fact that we can identify sibling pairs in our data. Using this information, we identify the impact of school absence on long-term outcomes from within-family variation and estimate the impact of an additional day of school absence with the following equation:

$$W_{if} = \beta_0 + \tau D_i + X_i\beta_1 + Q_{sg}\beta_2 + S_i + T_i + \delta_f + v_{if}, \tag{5}$$

where W_{if} is an adult outcome of individual i of family f, X_i is the set of individual-specific time-invariant controls and the vector Q_{sg} includes information on class size, the lowest and highest grade taught to students in the same classroom and length of the school year in weeks, averaged across grade 1 and grade 4. We also control for school fixed effects S_i and teacher fixed effects T_i based on information for grade 1, and also add a dummy if the individual changes school or teacher between grade 1 and 4.[29]

[28] In this sample, the average number of absent days are 12.2 (10.4 due to sickness) in grade 1, 10.7 (8.5) in grade 2, 11.9 (9.3) in grade 3, and 11.8 (9.1) in grade 4.

[29] About half of the individuals in our dataset are affected by the 1936 compulsory schooling reform increasing compulsory schooling by one year (from 6 to 7). Although all our specifications include a full set of parish fixed effects in order to control for differences across parishes, differences in the timing of the introduction of this reform might still be unaccounted for by parish fixed effects. In principle, it is possible that siblings fixed effects do also not absorb difference in compulsory schooling if the first cohort affected by the reform lies between the births of the first and second sibling. However, given that our data only cover six cohorts, this only affects very few pairs of siblings and our results do not change noteworthily when controlling for years of compulsory schooling.

Going back to equation (2) and our discussion about various threats to identification, this sibling-fixed effect strategy controls for family-level fixed unobservables that determine educational achievement and absences (α_f). This design controls for any unobserved individual characteristics that have the same additive effect on outcomes of both siblings. To identify the causal effect of absence, we need to assume that any child-specific unobservable affecting adult outcomes is uncorrelated with the child's absence. That is $\mathbb{E}(v_{if}|D_i, X_i, Q_{sg}, \rho_s, \eta_t, \alpha_f) = 0$.

While this family fixed effect strategy allows us to robustly point identify the effect of absences in elementary school on long-term outcomes, it is not without limitations. One threat to the validity of our strategy comes from individual-specific factors that vary across siblings, such as child-specific parenting inputs or health shocks that are correlated with absence and could affect individual later outcomes. We provide suggestive evidence that our estimates are unlikely to be biased by this type of confounding factors by performing a series of robustness checks (see Section 6.2). First, we estimate the short-term effects of school absences on academic achievement using a sibling fixed effect strategy and compare those with our main individual fixed effects. If these estimates were very different from each other, we could be concerned that individual-level unobservable factors (which we are not able to control for via the sibling fixed effect strategy) may be highly correlated with school absences. Conversely, if there are no notable differences between individual and sibling fixed effects estimates in this short-term perspective (as we show in Section 6), it would seem safer to conclude that the estimated effect measures the impact of absence *per se*.

Second, 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 long-run outcomes as non-sickness absence, it seems safe to conclude that the main component of the treatment is not poor health, but rather the loss of instructional time resulting from the absence. Such 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.

Our third robustness check also aims to rule out the possibility that our estimates of the long-term impacts of absence reflect an independent effect of health shocks by showing evidence that absence does not directly affect later health outcomes. To do this, we estimate equation (5) where we use as outcomes indicators for whether the individual has passed away in 1960, 1970 and 2003, the three years in which we measure employment and incomes outcomes. While we realize that mortality is an extreme indicator of health, it is the only available one that we can construct from our data sources.

Lastly, we also estimate bounds around our impacts of absences on adult outcomes. We show that, even when assuming high levels of correlation between the unobservables and school absences, the main conclusions emanating from the analysis remain unchanged.

4.3 On instrumental variable strategies to identify the impact of absences

The main identification strategy followed in this paper is based on fixed effects controlling for individual-level and/or family-level time-invariant potential confounders in the relationship between sickness absence and outcomes. As we will show through our discussion of results, our main results are remarkably robust across the various estimators that we are able to implement (OLS, sibling and/or individual fixed effects). We take this as reassuring evidence that unobserved, time-varying confounders at the individual level may be of limited importance, but, as discussed, acknowledge these strategies do not strictly rule out the possibility that our estimates are biased due to these confounders.

To further evidence the validity of their results, related papers in the literature have proposed using IV strategies, in addition to the individual fixed effect strategies that we also implement. Specifically, to instrument absence, Goodman (2014) uses moderate snowfall and Aucejo and Romano (2016) use variation in flu outbreaks at the county level. It is important to notice that using an IV strategy based on weather or flu outbreaks not only changes the *estimand* – it identifies a local treatment effect instead of a population parameter (Imbens and Angrist, 1994) – but also the *interpretation* of results. Indeed, both of these sources of variation could affect other students in the class, as well as their teacher, and an estimate based on a weather or flu IV will, by construction, no longer estimate the impact of individual school absence, holding everything else constant.

The research in this paper is motivated by the insight that individual sickness absence is different from a school closure, because it potentially increases the risk of a student falling behind the class. In contrast, an estimate based on a weather or flu IV will reduce the difference between the two reasons for absence, since it is based on an exposure common to all students in a location. This not only calls the exclusion restriction into question, but also potentially biases estimates since, by construction, it will give disproportionate weight to events that cause a large number of students to stay home sick.

5 Estimates of the short-term effects of school absence

5.1 Main results

Table 4 reports the estimates for three different specifications of the value-added model presented in Section 4. In the first column, we only include absences as right-hand side variables, whereas specifications (2) and (3) include past performance. Specifications (1) and (3) show that each additional day of absence is associated with a reduction in performance of 0.005 standard deviations (SD) in the same school year. Specifications (2) and (3) show that the outcome in grade 4 is strongly dependent on past performance: a one-SD increase in performance in grade 1 is associated with an improvement by 0.45 SD in grade 4. The results in the table also suggest that the past performance variable also captures the effect of past absences:

the estimated impact in specification (1) of grade-1 absences (-0.0014) is close to the product of the direct effect of absence on performance (-0.0052), and the impact of lagged performance on current performance (0.4545). Besides, when we control for past performance in specification (2), the estimate related to past absence is diminished and gets the 'wrong' sign.

In Table 5 we turn to individual fixed effects estimates according to equation 3.[30] The first column of the table reports the OLS estimates (i.e., a specification only including socio-demographic and classroom-level controls as well as parish fixed effects) and shows that one additional day of absence is significantly associated with a 0.51 per cent of an SD decrease in average performance. The individual fixed effects estimate, reported in column (2), is slightly smaller (0.0045) but also statistically significant at the 1 per cent level. However, the two estimates are not significantly different from each other, nor are they significantly different from the previous estimates based on the value-added model.[31] Appendix Table B5 reports the coefficient estimates associated with all the control variables included in these specifications.[32]

Assuming linearity, the effect of ten days of absence – about the average in our sample – corresponds to 4.5 per cent of a SD in student performance. 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 recent US data, Goodman (2014) finds an effect of 0.8

[30] The specification used in Table 5 assumes homogeneous effect of absence in grade 1 and grade 4. This is an assumption that is necessary to interpret the individual fixed effects estimate as the marginal effect of a day of absence and cannot be tested in the context of the individual fixed effects model. However, it can be tested in the context of the sibling fixed effects model. The lower panel of Appendix Table B3 interacts days of absence with a grade indicator. In the siblings fixed effects specification, absence in grade 1 does not affect performance statistically different than days of absence in general.

[31] As described in Section 3.3 we restrict the individual fixed effects sample to contain individuals with siblings in the dataset. Appendix Table B6 shows the corresponding result when we do not make this restriction. The individual fixed effects point estimates are very similar, suggesting that siblings and singletons react in the same way to an additional day of absence.

[32] We also replace the baseline outcome (performance measured on the 15-point grading scale) to a 7-point grading scale that takes the change in the Royal Commission's guidelines into account, see Appendix Table B4. The findings do not change our interpretation of the results.

Table 4: Estimates of the impact of absence on average grade points in grade 4 (in standard deviations)

	(1)	(2)	(3)
	Dependent variable: grade-4 performance. Lagged grade-1 variables:		
	absence	absence and performance	performance
Absence in grade 1	−0.0014	0.0007	
	(0.0015)	(0.0018)	
Absence in grade 4	−0.0057**		−0.0052***
	(0.0020)		(0.0018)
Grade points in grade 1		0.4553***	0.4545***
		(0.0191)	(0.0188)
# observations	8,934	8,934	8,934
# individuals/families	4,467	4,467	4,467

Notes: All specifications include parish fixed effects and the following set of controls. 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 controls 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. Estimates use the individual fixed effects sample (4,467 individuals). Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

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.

An advantage of our study is that we can anchor student performance, which is measured on a somewhat arbitrary scale, to pension incomes. This way, 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 (Swedish krona) that is economically more meaningful than standardized grade points. The estimates for the anchoring regressions are reported in Appendix Table B2. The results of the short-term effect of absence on school performance anchored to pensions are reported in the second row of Table 5. In our individual fixed effects specification, the impact of ten additional days of absence on school performance

Table 5: Estimates of the short-term effect of school absence on academic performance

	(1) OLS	(2) Individual fixed effects
Average grade points in units of SD (mean: 0, SD 1)		
Days of absence	-0.0051^{***}	-0.0045^{***}
	(0.0007)	(0.0013)
Average grade points in units of pension (mean pension in sample: 159,953 Swedish krona)		
Days of absence	-54.7344^{***}	-16.1192
	(7.4068)	(20.0252)
# observations	14,066	8,934
# individuals/families		4,467

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, see the data description in the text for details. The models control for the following time-variant conditional variables: grade, range of grades instructed in the same classroom, length of the school year in weeks. The OLS specification also controls for the following time-invariant control variables: female, born out of wedlock, twin birth, mother employed at the time of birth, born in hospital. Socio-economics controls 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: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

translates into a non-significant decrease in earnings potential of 16 Swedish krona, which is equivalent to a 0.1 per cent decrease in average pension income, a very small effect. A similar result is found when anchoring educational achievement to income at age 35–40.[33]

[33] As pensions mirror the best fifteen years they are less sensitive to fluctuations in labor supply than earnings and should thus be more representative, especially for women. Together with the fact that the pension system at the time had a progressive component (see Appendix E and Selén and Ståhlberg, 2007, for details), the effect in terms of earnings potential using incomes in middle age is somewhat larger in magnitude than the estimates anchoring to pensions.

5.2 Non-linearities and heterogeneity

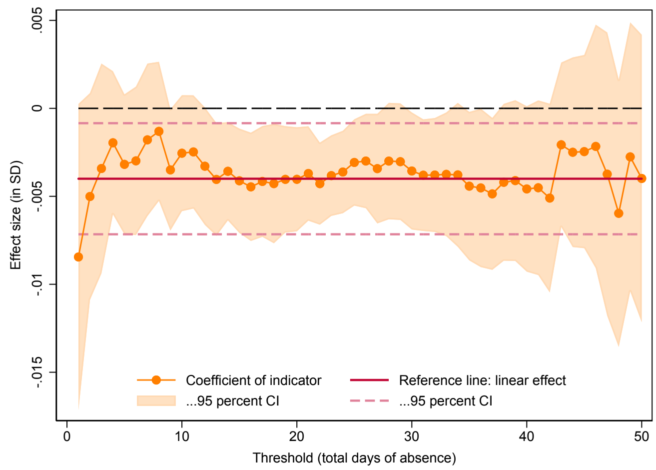
We next explore the extent to which impacts of student absences have non-linear effects on contemporaneous achievement. While a student may be able to easily catch up on a few days of absence, this may be less possible for longer periods out of school. If this were the case, it would result in a non-linear relationship between absence and educational performance. To investigate the presence of non-linearities we employ the individual fixed effects specification (similar to column 2 in Table 5), where we estimate the marginal effect of an additional day of absence on performance. For a threshold or cutoff value c, we estimate the effect of an increase in absence from $c-1$ to c days.[34] Appendix F provides the estimation details. We vary c from 1 to 50.

Figure 8 plots the marginal effects (on the y-axis) along the cut-off value (on the x-axis). If the effect of days of absence is linear, we would expect the marginal effect to be equal to the average effect. The average effect is given by the red reference line in Figure 8. A comparison of the orange markers and red line indicates that the per-day effect of absence of the non-linear estimations using the marginal effect does not substantially differ from the linear effect: the confidence intervals of the marginal effect estimates always include the estimated average effect. This finding is in line with Aucejo and Romano (2016) who do not find evidence of non-linearities either.

Next, we explore the extent to which the effect of absence are heterogeneous across socio-demographic subgroups. If there are heterogeneous effects across groups, these results could inform policy by identifying who may benefit the most from reforms targeting school absence. In Appendix Table B3, we report the results of OLS and individual fixed effects models where we include an interaction between our mea-

[34]Because the number of students that miss a large number of days in one school year is often rather low, it is not meaningful to regress performance on a full set of binary indicators for each number of days in a single regression. Appendix Figure A4 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 8: Non-linearities in the short-term effect of absence for different threshold values



Notes: Own calculations based on exam catalog information for 8,934 observations. This graph assesses linearity in days of absence. The solid red line gives the baseline effect, the dashed red lines the 95 per cent confidence interval. The orange markers give the marginal effect of an additional day of absence, when absence increases from $c-1$ to c days, where c is the cutoff value given on the x-axis. See Appendix F for the exact calculation. To ease the computational burden (each marginal effect is calculated using a separate regression), we apply the individual fixed effects specification, excluding parish, school, and teacher fixed effects. For this reason, the baseline effect is not exactly the same as in Table 5.

sure of absence and a dummy for indicating membership to the particular subgroup. Panel A reports the estimates for men and the interaction shows the difference in impacts between men and women. While the interaction is negative (suggesting a possibly stronger impact of absences for women than for men), it is highly insignificant.

Panel B exploits heterogeneity of impacts across children whose fathers are agricultural workers versus other children. As mentioned above, paternal occupation is our main proxy for socio-economic status (with agricultural workers expected to earn the least), but it could also proxy for prior ability. Aucejo and Romano (2016) find evidence that students in the lowest tercile of the prior attainment distribution are the most adversely affected by an additional absence, which is consistent with the hypothesis that lower ability students or poorer students have a harder time making up missed work. In contrast, in our context, we find no significant difference in impacts between children whose fathers are agricultural workers and those whose fathers are production or service workers. This absence of a social gradient might be due to the comparably low levels of socio-economic inequality observed in Sweden during this period: income and wealth inequalities were at least as large as in other Western countries at the turn of the 20^{th} century; however, the first decades of the 20^{th} century were characterised by rapidly declining earnings inequality (Roine and Waldenström, 2015; Bengtsson and Prado, 2020). In particular the years around the elementary schooling period of these individuals have been characterised as an "equality revolution" (Gärtner and Prado, 2016) due to a sharp reduction in inequality.

5.3 Robustness checks

As discussed in Section 4.1, the main threat to the identification of the causal effect of a day of absence is the presence of time-varying individual-level unobservables correlated with both absences and achievement. Given that absences are mainly driven by sickness absences, a particularly concerning threat is that idiosyncratic health shocks would confound the effect of the loss of instructional time resulting from a day of absence with that of the health shock. To address this concern, we estimate a version of equation 3 where we also include a variable N_{ig} measuring the number of days of absence due to non-sickness absence in grade g. In this model, the coefficient on D_{ig} measures the marginal effect of a day of sickness absence and the coefficient on N_{ig} measures the difference between the marginal effect of a day of non-sickness absence and that of a day of sickness absence. Our hypothesis is that if there is no difference between those two marginal effects, our strategy is unlikely to yield biased estimates of the effect of a day of absence due to idiosyncratic health shocks.

Table 6 reports the coefficient on D_{ig} and N_{ig} across specifications. In neither specification is the coefficient on N_{ig} statistically significant, and in the individual fixed effect model, it becomes extremely small (0.0004), which indicates that time-varying health shocks are unlikely to confound our estimates of τ. This evidence supports an interpretation of the findings where the reduced academic performance

Table 6: Short-term effects – total absence versus sickness absence

	(1) OLS	(2) Individual fixed effects
Average grade points in units of SD		
Total days of absence	-0.0038^{***}	-0.0045^{***}
	(0.0007)	(0.0013)
Days of non-sick. absence	-0.0035	-0.0031
	(0.0029)	(0.0022)
# observations	14,066	8,934
# individuals/families		4,467

Notes: Each column shows the coefficients of total days of absence and days of non-sickness absence when student performance is jointly regressed on both kinds of absences and the control variables according to the baseline specification. The first row gives the coefficient of total absence, whereas the second row gives the coefficient of non-sickness absence. A one-day increase in non-sickness absence, given the total days of (sickness and non-sickness) absence, gives the relative effect of missing a day of school for reasons other than health compared to missing school for health. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

associated with absence is driven by the loss of instructional time, and not by a shock in the student's health.

While a health shock is likely to be the most common type of grade-specific unobservable factor possibly not being controlled for by the individual fixed effect strategy, it is of course possible that there are other factors whereby this test would not be particularly informative. To guard against this possibility, we also provide bounds following the approach by Oster (2019). The goal of this exercise is to bound the effect of absence assuming that the selection of unobservables is as strong as the selection on observables. Given the discussion in Section 3.4 we consider the case where 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. Notably the exercise is only helpful if the observables are informative with respect to selection, therefore we control for a large array of control variables

Table 7: Coefficient bounds for $\delta = 1$ and $\delta = -1$ selection

Dependent variable	(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	-0.0047^{***}	-0.0045^{***}	-0.00986	-0.00194
	(0.0011)	(0.0013)		
	[0.00]	[0.65]		
Long-term outcomes				
More than *Folkskola*	-0.0009^{*}	-0.0005	-0.73850	-0.00014
	(0.0005)	(0.0009)		
	[0.00]	[0.41]		
Employment 1960	-0.0013	-0.0007	0.00306	-0.00177
	(0.0008)	(0.0013)		
	[0.00]	[0.60]		
Employment 1970	-0.0008	-0.0020	-0.04959	-0.00155
	(0.0008)	(0.0014)		
	[0.00]	[0.50]		
Labor market income 1970	-65.4768^{***}	-80.7042^{**}	-107.88081	-73.15848
	(22.3882)	(37.2151)		
	[0.00]	[0.64]		
Pensions 2003–2008	-384.5674^{***}	-396.0887^{**}	-90.55889	-423.49521
	(81.8270)	(169.3640)		
	[0.00]	[0.59]		

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 presented in Tables 4 and 7, respectively. Column 3 and Column 4 report the lower and upper bounds of the effects. Number of observations: short-term performance 14,066; long-term outcomes: more than *Folkskola* 8,567, employment 1960 7,434, employment 1970 8,567, income 1970 5,886, pensions 4,770. Parish-clustered standard errors for the regression coefficients in parentheses. The resulting R^2s are given in brackets.

and the full set of school, teacher and individual fixed effects. We discuss the bounds approach in more detail in Appendix G.

The upper panel of Table 7 shows that the estimated bounds for the effect of absence on academic performance are -0.009 and -0.001. Thus, even under extreme levels of selection on the basis of unobservables, the effect of absence in one grade on achievement measured at the end of this grade is negative and close to our individual fixed effects estimates (and clearly within its confidence interval). Having established that our estimates of the short-term impacts of absence on educational achievement are likely to be robust to the presence of individual-level time-varying unobserved heterogeneity, we now turn to discussing the results of our analysis of the long-term effects of absence.

6 Estimates of the long-term effects of absence

6.1 Main results

Table 8 reports the effects on our set of long-term outcomes of the average number of absences across grades 1 and 4 in columns 1 (OLS) and 3 (siblings fixed effects), respectively, using the sibling fixed effect strategy. To ease the interpretation, we also report next to these point estimates the effect of ten days of absence relative to the mean of the outcome being considered. We calculate this by multiplying the relevant coefficient by five (since the treatment variable measures the average number of days of absences across two grades) and dividing it by the mean of the outcome.

Our estimates point to a consistently negative effect of absences in elementary school on economic outcomes through the life-cycle, with more pronounced and statistically significant impacts on educational attainment for men and on income both for genders (see Appendix Table B7). Interestingly, our estimates are very similar both in terms of magnitude and statistical significance using either the OLS or siblings fixed effect estimators, which aligns with one of the conclusions from the short-term analysis that selection into absences based on unobservable characteristics may not have been particular prominent in the context under study.

Looking at each estimate of Table 8 in turn (and only focusing on the sibling fixed effects estimators for the purpose of this discussion), we find that ten days of absence in elementary school leads to 2.2 per cent reduction in secondary school enrolment relative to baseline, though this impact is not statistically significant at conventional levels across the whole sample. Given important gender differences in educational attainment in the period under study, we explore, in Appendix Table B7, whether impacts are different for men and women. The results show that, indeed, the effect of absence on secondary school enrolment is much larger and statistically significant for men, while it is close to zero for women. The negative effect of elementary school absence on secondary schooling is consistent with secondary school admissions depending on elementary school performance.

Next, we turn to studying impacts of absence on labor market outcomes. Impacts on employment – whether measured earlier in the career (ages 25–30) or later (ages 35–40) – are not statistically significant, but the point estimates suggest that impacts on employment may grow slightly with age, from a 0.5 per cent reduction in employment at ages 25–30 to a 1.3 per cent reduction ten years later. The next two rows in Table 8 present our estimates of the impact of absence in elementary school on our two measures of labor market income. As mentioned earlier, while income in 1970 represents working age income, pension income is closer to a measure of lifetime earnings and would therefore be less likely to be affected by labor supply fluctuations over the life-cycle. Our sibling fixed effects estimates show that a one day increase in the average number of days of absence in elementary school leads to a 80 Swedish krona reduction in labor market income in 1970 and a 396 Swedish krona reduction in pension income, and both of these coefficient estimates are significant at the 5 per cent level. Relative to the means of these variables in the dataset, the impacts are non-negligible and fairly similar in magnitude to each other, between 1 and 2 per cent reduction in income. As mentioned earlier, our measures of income include zero incomes, so these effects could reflect impacts of absence both on wages and on labor supply. While our estimate of the impact of absences on employment are not precisely estimated, the point estimates do suggest that the effects we find on labor market income could be driven by both impacts on employment and wages. And indeed, our estimates of impacts on log(income) and on non-zero income, presented in Appendix B8, suggest a possibly negative, but small impact on wages.[35]

[35]As with our short-term analysis, we explore non-linearities and heterogeneity in the effects of school absences on long-term outcomes. Appendix Figure A5 reports the results of an exercise where we examine linearity for our long-term outcomes. Because of the fewer observations for

Overall, these results point to a long-lasting influence of absences in elementary school on labor market outcomes through the life-cycle. While the magnitudes of the effects are relatively small in absolute terms, they are larger than what could have implied from the short-run effects of absence on school performance and the link between that and long-term outcomes. For comparison, our estimates of the effect of school performance, as measured by average grade points, on secondary schooling enrollment range between 0.08 (siblings fixed effects specification) and 0.11 (OLS, results available upon request). Multiplying this estimate with the estimated effect of absence on school performance of -0.0051 and -0.0045 (cf. Table 5), we would expect an effect of absence on enrollment of -0.0004 and -0.0006 in the OLS and siblings fixed effects specification, respectively. Our OLS estimate of the impact of absence on enrolment is twice as large as this indirect estimate. Moreover, we also find that the effects become slightly larger and more significant as they are measured further and further into the life-cycle. Taken together, these results are consistent with a dynamic model of skill accumulation where small skill deficits resulting from absences in elementary school translate into increasingly lower skill levels over the life-cycle. Finally, these results underline the importance of measuring impacts of absence at various points of, and especially quite late in, the life-cycle.

6.2 Robustness checks

A potential threat to our identification strategy for long-run outcomes using sibling fixed effects is that there are individual-specific unobserved factors of importance. As a first exercise to assert the validity of our strategy, we re-estimate the short-term effect of school absence on educational achievement this time using a sibling fixed effect strategy and compare them to the baseline estimates reported in Table 5 based on the individual fixed effect strategy. As reported in Table 9, the two

the long-term outcomes, we only run the absence indicator up to the threshold of 30 or more days of absence. Overall, there is no strong evidence that absence has non-linear impacts on final educational achievement, adult employment, income, pensions or mortality. Appendix Table B7 reports the results of our heterogeneity analysis. With the exception of a stronger effect of absence on male secondary education, which we commented on earlier, we do not find significant differences in the effect of absence on long-term outcomes between gender or socio-economic groups.

Table 8: Long-term effects of school absence

	(1)	(2)	(3)	(4)
	OLS		Siblings fixed effects	
	coeff.	rel. size	coeff.	rel. size
More than Folkskola *(1=yes)*				
Absence (average, grades 1 and 4)	-0.0007	-0.0297	-0.0005	-0.0215
	(0.0005)		(0.0009)	
Employment 1960 (1=yes)				
Absence (average, grades 1 and 4)	-0.0012^{*}	-0.0088	-0.0007	-0.0053
	(0.0007)		(0.0013)	
Employment 1970 (1=yes)				
Absence (average, grades 1 and 4)	-0.0008	-0.0050	-0.0020	-0.0133
	(0.0007)		(0.0014)	
Labor market income 1970				
Absence (average, grades 1 and 4)	-54.2973^{**}	-0.0141	-80.7042^{**}	-0.0209
	(22.7088)		(37.2151)	
Pensions 2003–2008				
Absence (average, grades 1 and 4)	-330.0940^{***}	-0.0103	-396.0887^{**}	-0.0124
	(71.2631)		(169.3640)	

Notes: Each panel states the coefficient of total days of absence (average over grades 1 and 4). Number of observations: More than *Folkskola* 5,976, employment 1960 7,434, employment 1970 5,976, income 1970 5,886, pensions 4,772. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

estimators yield very similar estimates of the impact on achievement in elementary school, suggesting that individual-level unobservables that affect educational achievement and vary between siblings are unlikely to be correlated with absences. While this test does not fully rule out the possibility that the same could be true of unobservables determining the long-term outcomes we consider, this evidence does reinforce the message emanating from the results that selection into absence based on unobservables is unlikely to be strong in our context.

As a second robustness check we take advantage of the fact that our data allows us to distinguish between absences due to sickness and absences due to other reasons. Similarly to the robustness check we performed for the short-term effects of absences in Section 5.3, we test whether the long-term effects of sickness absence are different from the effect of non-sickness absence by including an additional variety in the model that which measures the average number of non-sickness days across grades 1

Table 9: Short-term effects using siblings fixed effects

	(1) Individual fixed effects	(2) Sibling fixed effects
Average grade points in units of SD (mean: 0, SD 1)		
Days of absence	-0.0045^{***}	-0.0055^{***}
	(0.0013)	(0.0009)
# observations	8,934	14,066
# individuals/families	4,467	3,716

Notes: Each cell states the coefficient of days of absence for a separate regression. 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, 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: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

and 4. Estimates of this specification are reported in Table 10. Across all long-run outcomes, the coefficient on the average days of non-sickness absence is small and statistically insignificant, suggesting that the long-term effects of absence reported in Table 8 are driven by the loss of instructional time rather than by poor health.

We further test the possibility that our long-term effects are confounding the effect of a health shock with that of absence by testing whether school absences have an effect on long-term health, as measured by mortality. To do so, we re-estimate our sibling fixed effect model, this time using indicators for mortality at three points in time (1960, 1970 and 2003) as outcomes. As Table 11 indicates, we cannot reject that the coefficient on absence in any of these regressions is zero, thus further suggesting that the effect of absence we identify through the sibling fixed effect strategy is likely to capture the effect of losing instructional time (and human capital) rather than the effect of a health shock associated with the absence.

Table 10: Long-term effects – total absence versus non-sickness absence

	(1)	(2)	(3)	(4)
	OLS		Siblings fixed effects	
	total absence	non-sick. absence	total absence	non-sick. absence
More than Folkskola *(1=yes)*				
Absence (average, grades 1 and 4)	0.0001	-0.0060^{***}	-0.0006	0.0006
	(0.0006)	(0.0017)	(0.0007)	(0.0022)
Employment 1960 (1=yes)				
Absence (average, grades 1 and 4)	-0.0001	-0.0052^{***}	-0.0005	-0.0036
	(0.0005)	(0.0017)	(0.0010)	(0.0045)
Employment 1970 (1=yes)				
Absence (average, grades 1 and 4)	-0.0005	-0.0005	-0.0015	0.0003
	(0.0006)	(0.0022)	(0.0012)	(0.0038)
Labor market income 1970				
Absence (average, grades 1 and 4)	-37.9619^{**}	-13.7218	-50.0327	7.4814
	(19.0897)	(61.6756)	(32.7365)	(154.4145)
Pensions 2003–2008				
Absence (average, grades 1 and 4)	-217.2837^{***}	-87.9434	-319.0513^{***}	706.7839
	(57.2104)	(268.0873)	(121.5442)	(430.9648)

Notes: Each panel states the coefficient of total days of absence (average over grades 1 and 4). Number of observations: More than *Folkskola* 5,976, employment 1960 7,434, employment 1970 5,976, income 1970 5,886, pensions 4,772. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

Finally, we address the caveat that the sibling fixed effect estimates could be biased if there are unobservable factors that correlate with individual absence and outcomes *and* that vary over time or across siblings by implementing the bounding approach of Oster (2019). As mentioned above, this 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 siblings fixed effects. This removes factors such as constant family resources, parental preferences and genetic endowment that are likely to be 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 particularly useful in the application at hand. The bottom panel of Table 7 presents the bound analysis results for the long-term outcomes. First comparing estimates in columns (1) and (2) gives a sense of the extent to which the raw correlation between absence and the outcome is robust

Table 11: Estimated effects on mortality by the time long-term outcomes are observed

	(1)	(2)	(3)
	Passed away before		
	1960	1970	2003
Absence (average, grades 1 and 4)	0.00001 (0.00030)	0.00036 (0.00049)	0.00056 (0.00078)

Notes: Each panel states the coefficient of total days of absence (average over grades 1 and 4). Number of observations: 8,567. Mean values: 0.0239 (1960), 0.0336 (1970), 0.2197 (2003). Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

to conditioning on a large number of observables. The estimates are remarkably robust and hardly change at all when including the full set of observable controls. This should to be kept in mind when considering the estimated bounds in columns (3) and (4). For all long-term outcome variables, all bounds point towards the same direction as in the baseline model, this reinforces our belief that serious omitted variable bias is unlikely.

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: increasing total absence over a school year by ten days leads to a reduction in grade point average of 4.4 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. For men, this immediate impact on school performance spills over onto secondary school admissions, which were based on elementary school performance. This effect 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, we find consistent evidence that there is a penalty to absence in elementary school: estimates have the expected negative sign for all long-term outcomes, and they are statistically significant for earnings along the entire life-cycle for both genders.

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 the long-term results. 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 that a persistent health shock from elementary school has an independent effect on long-term outcomes. Nevertheless, we exploit the fact that our data includes 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 are most likely capturing the impact of reduced instructional time.

Our findings are obviously specific to a particular context – cohorts born in the 1930s in Sweden. But in the absence of evidence about the long-run effects of student absence for other contexts and more recent cohorts, our study makes a useful contribution to both academics and policy-makers concerned with high rates of absences and with the dramatic increase in student absence due to the COVID-19 pandemic. Our findings suggest that the learning losses associated with school absences in elementary school lead to non-negligible reductions in income through the life-cycle. And although our impacts of absence on educational attainment and employment are not precisely estimated, they are consistent with a dynamic model of skill accumulation where early instructional losses lead to lower skill levels, which accumulate over the life-cycle and eventually create non-negligible income penalties. Our findings hone in on the impact of individual absences, as opposed to school closures. In this light, they may only partly be relevant to predict the long-term effects of the school closures during the early phases of the pandemic. But, as student absence become increasingly driven by individual students self-isolating, our results can provide useful evidence that associated learning losses may have a long-term impact if not appropriately compensated.

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. 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.

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.

Bengtsson, E. and S. Prado (2020). The rise of the middle class: the income gap between salaried employees and workers in Sweden, ca. 1830–1940. *Scandinavian Economic History Review 68*(2), 91–111.

Bergh, A. (2014). *Sweden and the revival of the capitalist welfare state.* Edward Elgar Publishing.

Bhuller, M., M. Mogstad, and K. G. Salvanes (2017). Life-Cycle Earnings, Education Premiums, and Internal Rates of Return. *Journal of Labor Economics 35*(4), 993–1030.

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.

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.

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).

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

Erixon, L. (2008). The Swedish third way: an assessment of the performance and validity of the Rehn–Meidner model. *Cambridge Journal of Economics 32*(3), 367–393.

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

Fischer, M., M. Karlsson, T. Nilsson, and N. Schwarz (2019). The Long-Term Effects of Long Terms – Compulsory Schooling Reforms in Sweden. *Journal of the European Economic Association 18*(6), 2776–2823.

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.

Galama, T. J., A. Lleras-Muney, and H. van Kippersluis (2018, January). The Effect of Education on Health and Mortality: A Review of Experimental and Quasi-Experimental Evidence. NBER Working Papers 24225, National Bureau of Economic Research, Inc.

Gärtner, S. and S. Prado (2016). Unlocking the social trap: Inequality, trust and the Scandinavian welfare state. *Social Science History 40*(1), 33–62.

Gathmann, C., H. Jürges, and S. Reinhold (2015). Compulsory schooling reforms, education and mortality in twentieth century Europe. *Social Science & Medicine 127*, 74–82.

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.

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

Imbens, G. W. and J. D. Angrist (1994). Identification and estimation of local average treatment effects. *Econometrica 62*(2), 467–475.

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.

Liu, J., M. Lee, and S. Gershenson (2019). The short- and long-run impacts of secondary school absences. (12613).

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.

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. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics 37*(2), 187–204.

Paulsson, E. (1946). *Om folkskoleväsendets tillstånd och utveckling i Sverige under 1920-och 1930-talen (till omkring år 1938).* Länstryckeriaktiebologet.

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.

Roine, J. and D. Waldenström (2015). Long-run trends in the distribution of income and wealth. In *Handbook of income distribution*, Volume 2, pp. 469–592. Elsevier.

SCB (1962). Population and housing census 1960. *National Bureau of Statistics*.

SCB (1972). Population and housing census 1970. *National Bureau of Statistics*.

Selén, J. and A.-C. Ståhlberg (2007). Why Sweden's pension reform was able to be successfully implemented. *European Journal of Political Economy 23*(4), 1175–1184.

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

SOU (1942). Betänkande med utredning och förslag angående betygsättningen i folkskolan, angivet av inom ecklesiastikdepartementet tillkalade sakkunniga. *Statens offentliga utredningar 1942:11*.

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

Todd, P. E. and K. I. Wolpin (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal 113*(485), F3–F33.

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

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

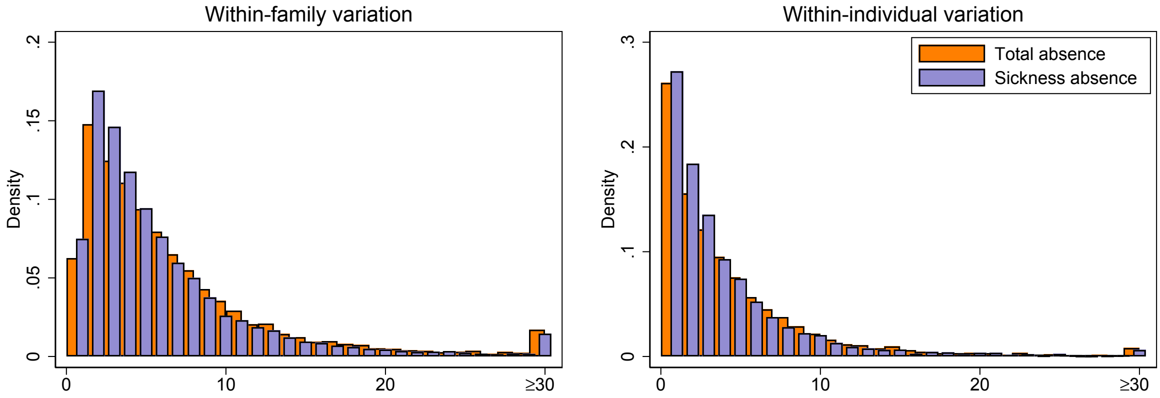
Sarah Cattan, Daniel A. Kamhöfer, Martin Karlsson and Therese Nilsson

February 20, 2021

ONLINE APPENDIX

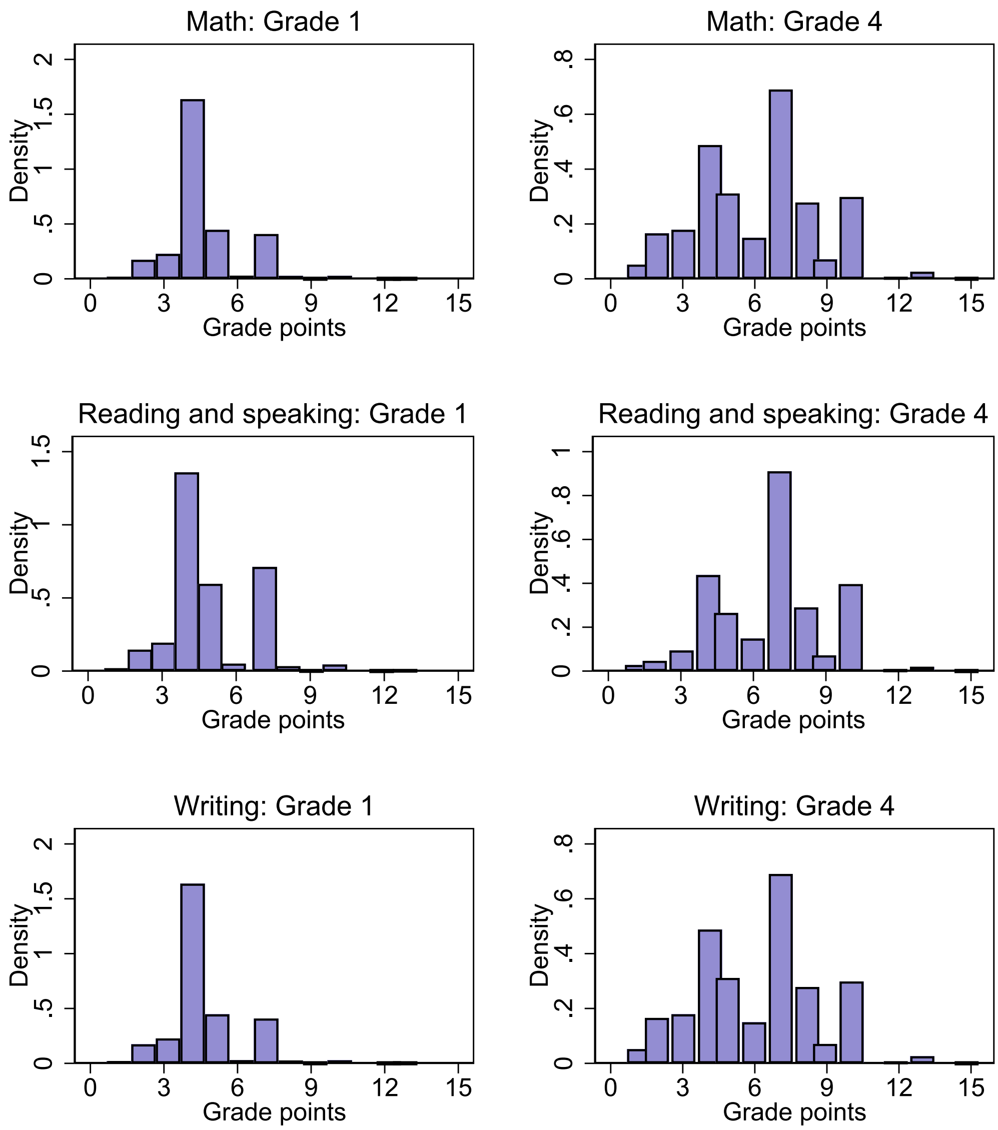
A Appendix: Figures

FIGURE A1. Distribution of the within-family and within-individual variation in (sickness) absence



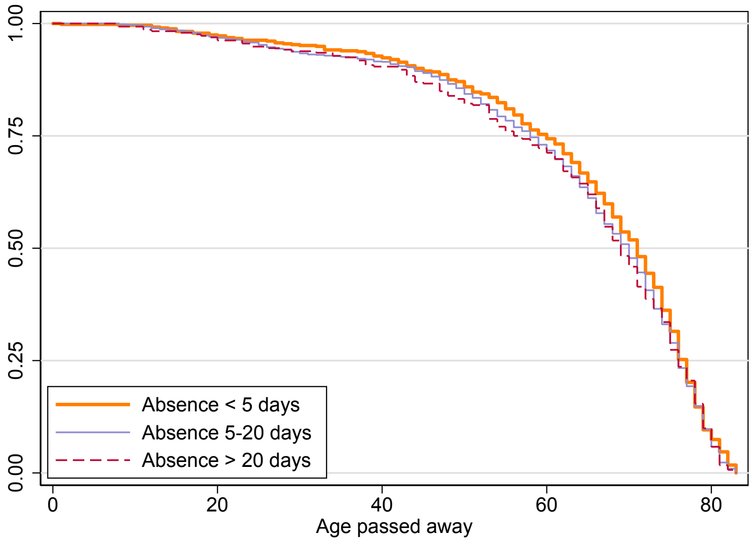
Notes: Own calculations based on exam catalog information. 14,066 observations.

FIGURE A2. Distribution of grades by subject



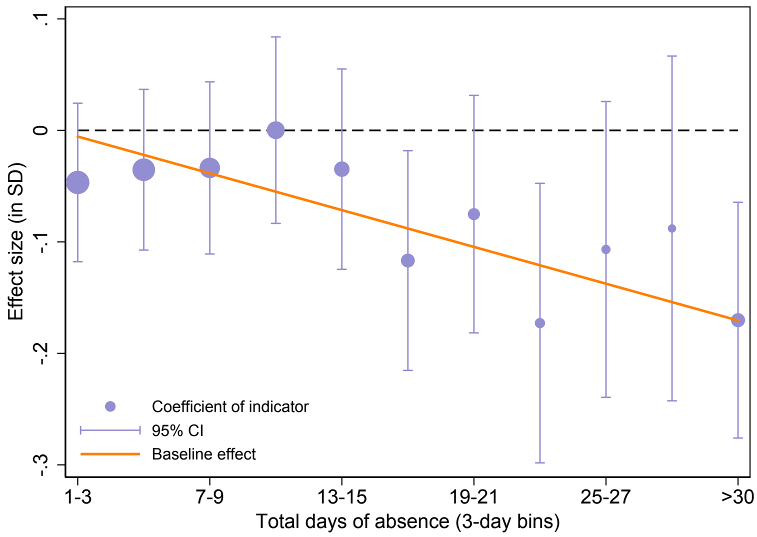
Notes: Own calculations based on exam catalog information.

FIGURE A3. Kaplan–Meier survival function by total days of absence



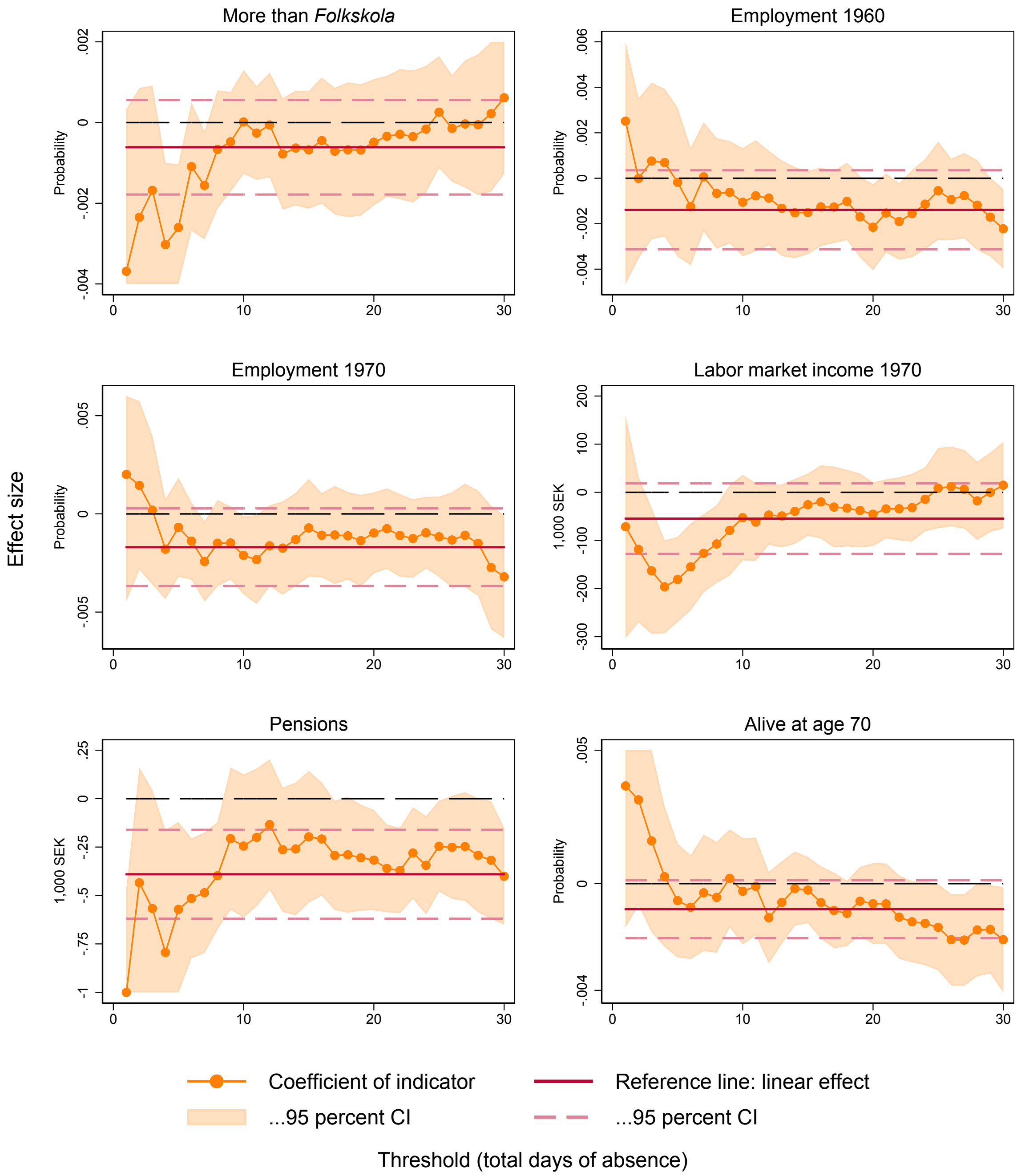
Notes: Own calculations based on exam catalog and Swedish Death Index information, 8,567 observations. A Kolmogorov–Smirnov test for the equality of the distributions indicates that the conditional distributions to not differ significantly.

FIGURE A4. Nonlinearities in the short-term effect of grouped sickness absence



Notes: Own calculations based on exam catalog information. 14,066 observations. To detect non-linearities 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 markers give the relative number of observations for which the group indicator is 1. The spikes around the markers state the 95 per cent confidence interval. The orange line depicts the linear effect of an additional day of absence in the baseline short-term results.

FIGURE A5. Non-linearities in the effect of average days of absence in both grades on long-term outcomes using individual fixed effects



Notes: Own calculations, data sources and number of observations as for the baseline ling-term results. This graph plots the coefficient of a siblings FE regression of the long-term outcome on a binary indicator for average days of total absence over grades 1 and 4 and the same control variables as in the baseline long-term specification. 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. Number of observations: more than *Folkskola* 8,567, employment 1960 7,434, employment 1970 8,567, income 1970 4,154, pensions 4,770, alive at age 70 8,567.

B Appendix: Tables

TABLE B1. Summary statistics for control variables

Time-invariant variables	Mean
Female (in %)	49.5
Number of siblings	1.5
Year of birth (in %)	
1930	18.1
1931	16.5
1932	18.0
1933	17.5
1934	16.5
1935	13.4
(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	39.8
Father: agricultural worker	33.2
Father: service and sales worker	50.9
Father: production workers	7.4
Mother: employed (binary)	2.5
Living at the time of time and birth conditions (in %)	
Born out of wedlock (in %)	4.8
Born in hospital (in %)	8.3
Twin birth (in %)	4.0
Mother's year of birth	1902
Father's year of birth	1898
(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 %)	34.3	30.2
Some classmates in lower grade (in %)	0.0	63.085
Some classmates in higher grade (in %)	65.7	30.6
Class size (number of students)	13.6	15.9
(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. Observations: 14,066. 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, taking 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 B2. Estimation of earnings potential

	(1)	(2)
	Earnings potential in pensions 2002	
	Grade 1	Grade 4
Math points: 3 or less	$-26,452.6^{***}$	$-14,287.7^{***}$
	$(3,842.6)$	$(2,827.6)$
Math points: 4	$-19,508.2^{***}$	$-16,831.8^{***}$
	$(2,658.8)$	$(2,814.0)$
Math points: 5	$-11,637.7^{***}$	$-8,314.0^{***}$
	$(2,844.8)$	$(3,191.0)$
Math points: 6	$-21,180.0^{***}$	$-6,795.9^{*}$
	$(7,369.1)$	$(3,714.3)$
Math points: 8	$3,113.8$	$2,679.6$
	$(6,576.2)$	$(3,005.4)$
Math points: 9 or more	$-2,801.4$	$18,667.0^{***}$
	$(8,052.8)$	$(2,874.1)$
Reading points: 3 or less	$-1,551.2$	$6,843.7^{*}$
	$(3,708.6)$	$(4,028.2)$
Reading points: 4	$-2,427.1$	$7,352.8^{***}$
	$(2,591.0)$	$(2,837.2)$
Reading points: 5	-324.7	$2,339.9$
	$(2,839.3)$	$(3,170.1)$
Reading points: 6	$-11,507.4^{*}$	$8,788.0^{**}$
	$(6,797.9)$	$(3,820.1)$
Reading points: 8	$-13,723.4^{*}$	$2,647.9$
	$(8,121.1)$	$(2,921.7)$
Reading points: 9 or more	$21,555.8^{***}$	$7,555.2^{***}$
	$(7,864.9)$	$(2,838.0)$
Writing points: 3 or less	$12,126.8^{***}$	$7,947.8^{**}$
	$(4,436.2)$	$(3,331.3)$
Writing points: 4	$10,753.2^{***}$	$10,457.5^{***}$
	$(3,359.7)$	$(2,871.8)$
Writing points: 5	$8,668.6^{**}$	$9,694.4^{***}$
	$(3,898.6)$	$(3,020.7)$
Writing points: 6	$11,533.2$	-269.0
	$(10,783.5)$	$(3,975.3)$
Writing points: 8	563.7	$3,466.5$
	$(11,193.1)$	$(3,041.9)$
Writing points: 9 or more	$-37,322.5^{***}$	$2,179.2$
	$(13,198.7)$	$(3,222.5)$

Notes: Dependent variable: pensions taken from tax registers 2003–2008. Explanatory variables: binary indicators of the points in math, reading and speaking and writing (reference category is 7 points). Standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

TABLE B3. Heterogeneity in the short-term effects by subgroup

	(1) OLS	(2) Sibl. FE	(3) Indi. FE
Gender			
Absence	-0.0040^{***}	-0.0052^{***}	-0.0039^{***}
	(0.0011)	(0.0014)	(0.0015)
Absence×female	-0.0022	-0.0006	-0.0014
	(0.0016)	(0.0019)	(0.0023)
Father's occupation			
Absence	-0.0037^{***}	-0.0059^{***}	-0.0046^{***}
	(0.0010)	(0.0011)	(0.0018)
Absence×agri. worker	-0.0035^{**}	0.0010	-0.0001
	(0.0012)	(0.0011)	(0.0027)
Grade			
Absence	-0.0071^{***}	-0.0071^{***}	
	(0.0013)	(0.0013)	
Absence×grade 1	0.0037^{***}	0.0030	
	(0.0014)	(0.0018)	
# observations	14,066	14,066	8,934

Notes: Each panel states the coefficient of total days of absence differentiated for different subgroups. Columns 1 and 2 use the siblings panel and employ OLS and siblings fixed effects estimation, respectively. Column 3 gives individual fixed effects results using the individual–grades panel. In all models, days of absence enter the specification linearly and interacted with a subgroup indicator. Other control variables are as in the baseline short-term results. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

TABLE B4. 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	-0.0050^{***}	-0.0056^{***}	-0.0053^{***}
	(0.0008)	(0.0010)	(0.0016)
# observations	$14,066$	$14,066$	$8,934$

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

TABLE B5. 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	-0.0051^{***}	-0.0058^{***}	-0.0060^{***}	-0.0055^{***}	-0.0045^{***}
	(0.0007)	(0.0009)	(0.0009)	(0.0009)	(0.0013)
Female	0.2949^{***}	0.3084^{***}	0.3126^{***}	0.3374^{***}	
	(0.0239)	(0.0235)	(0.0253)	(0.0301)	
Birth year: 1931	0.0277	0.0908	0.1238	−0.1809	
	(0.0915)	(0.1014)	(0.1122)	(0.1188)	
Birth year: 1932	0.4758	0.4509^{***}	0.8561^{**}	−0.7520	
	(0.4682)	(0.1627)	(0.4123)	(0.5501)	
Birth year: 1933	0.7233^{*}	0.5821^{***}	1.1033^{**}	−0.1541	
	(0.3924)	(0.2084)	(0.4722)	(0.3457)	
Birth year: 1934	−0.0145	0.0029	0.0407	−0.8640	
	(0.0929)	(0.1100)	(0.1241)	(0.5882)	
Birth year: 1935	0.5722	0.4061^{**}	0.8614^{**}	−0.3603	
	(0.4187)	(0.1914)	(0.4088)	(0.3722)	
Wedlock	0.0154^{***}	0.0185^{***}	0.0182^{***}	0.0135^{***}	
	(0.0044)	(0.0039)	(0.0032)	(0.0016)	
Hospital	0.1122^{***}	0.1148^{***}	0.1132^{**}	0.0547	
	(0.0401)	(0.0425)	(0.0461)	(0.0543)	
Twin	-0.1436^{**}	-0.1653^{**}	-0.1673^{**}	-0.1901^{**}	
	(0.0574)	(0.0703)	(0.0798)	(0.0927)	
Grade 4	1.4707^{***}	1.3496^{***}	1.3366^{***}	0.6706^{***}	0.5571^{***}
	(0.1614)	(0.1340)	(0.1463)	(0.1608)	(0.0378)
Occup. father: agriculture	0.0590	0.0290	0.0277	0.0003	
	(0.0621)	(0.0582)	(0.0629)	(0.0677)	
Occup. father: services	0.1910^{***}	0.2113^{***}	0.2155^{***}	−0.1033	
	(0.0498)	(0.0550)	(0.0546)	(0.0869)	
Occup. father: farmer	0.0051	0.0339	0.0303	−0.0458	
	(0.0576)	(0.0505)	(0.0524)	(0.0676)	
Occup. father: unknown	0.0166	0.1420^{***}	0.1566^{***}	0.0861	
	(0.0638)	(0.0480)	(0.0443)	(0.0564)	
Mother employed	-0.1279^{*}	-0.1181^{*}	-0.1269^{*}	−0.0638	
	(0.0715)	(0.0647)	(0.0694)	(0.0614)	
Classmates in lower grade	0.0668	0.0851	0.1456^{***}	0.1548^{***}	0.1957^{***}
	(0.0494)	(0.0540)	(0.0431)	(0.0380)	(0.0659)
Classmates in higher grade	0.0217	−0.0545	−0.0765	−0.0793	−0.0365
	(0.0321)	(0.0337)	(0.0760)	(0.0672)	(0.0557)
Class size 1–5	-0.0304^{*}	−0.0239	−0.0213	−0.0136	0.0306
	(0.0178)	(0.0241)	(0.0268)	(0.0281)	(0.0397)
Class size 6–10	0.0083	0.0067	0.0038	−0.0030	0.0011
	(0.0097)	(0.0118)	(0.0122)	(0.0107)	(0.0238)
Class size 11–15	−0.0096	−0.0049	0.0023	0.0110	0.0194
	(0.0080)	(0.0081)	(0.0084)	(0.0106)	(0.0195)
Class size 16–20	0.0032	0.0149^{*}	0.0000	0.0033	0.0035
	(0.0082)	(0.0079)	(0.0130)	(0.0109)	(0.0186)

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

TABLE B5 – *continued*

	(1) OLS	(2) School FE	(3) Teacher FE	(4) Sibl.	(5) Indi.
Class size 21–25	0.0025	−0.0089	0.0051	−0.0019	0.0138
	(0.0101)	(0.0098)	(0.0148)	(0.0166)	(0.0243)
Class size > 25	−0.0004	0.0109**	0.0122***	0.0131***	0.0131*
	(0.0038)	(0.0043)	(0.0045)	(0.0049)	(0.0070)
# observations	14,066	14,066	14,066	14,066	8,934
# units		955	1,639	3,716	4,467

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

TABLE B6. Short-term estimates using the full sample (not restricted too siblings)

	(1) OLS	(2) Indi. FE
Average grade points in units of SD (mean: 0, SD 1)		
Days of absence	-0.0040^{***}	-0.0052^{***}
	(0.0006)	(0.0007)
# observations	28,931	21,750
# families/individuals		10,875

Notes: Column 1 gives the OLS estimate using all observation with non-missing information in the key variables and no other sample restrictions. Column 2 shows an individual fixed effects estimation, requiring that an individual is observed in grades 1 and 4. Unlike to the baseline short-term results, the individual fixed effects estimate in this table is not restricted to siblings, but also includes singletons and individuals with siblings born before 1930 and after 1935. In terms of control variables, the specification is similar to the baseline short-term model. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

TABLE B7. Heterogeneity in the long-term effects by subgroup

	(1)	(2)	(3)	(4)	(5)
	Dependent variable				
	> *Folkskola*	Empl. 1960	Empl. 1970	Income 1970	Pensions
Gender (siblings FE)					
Absence	−0.0015**	−0.0005	−0.0022	−102.6306*	−359.9496*
	(0.0006)	(0.0013)	(0.0017)	(56.9857)	(216.3130)
Absence× female	0.0016*	−0.0004	0.0003	38.8400	−64.9566
	(0.0009)	(0.0015)	(0.0019)	(58.8859)	(290.5578)
Father's occupation (siblings FE)					
Absence	−0.0006	−0.0010	−0.0012	−68.4765*	−484.4240**
	(0.0008)	(0.0010)	(0.0016)	(38.7539)	(234.0316)
Absence× agri. worker	0.0000	0.0008	−0.0019	−29.2021	195.8226
	(0.0011)	(0.0017)	(0.0022)	(47.2187)	(284.0396)
Sibling in sample (teacher FE)					
Absence	0.0003	0.0001	−0.0007	−20.4367	−135.9070
	(0.0005)	(0.0005)	(0.0005)	(17.3118)	(84.7254)
Absence× sibling	−0.0008	−0.0016**	−0.0005	−32.8386	−112.3591
	(0.0005)	(0.0007)	(0.0008)	(26.1065)	(132.0163)

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. The first two panels give results of a siblings FE specification. Number of observations: more than *Folkskola* 8,567, employment 1960 7,434, employment 1970 8,567, income 1970 4,154, pensions 4,770, alive at age 70 8,567. The third panel uses all individuals with exam catalog and church record data to investigate whether the family size (and our sample selection for the baseline analysis) matters for the results. While cannot employ siblings FE in this specification, we do use teacher FE. The number of observations are: more than *Folkskola* 18,056, employment 1960 16,579, employment 1970 18,056, income 1970 11,921, pensions 12,456, alive at age 70 18,056. About half of all individual have at least one sibling. Dependent variables defined as in the baseline long-term results. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

TABLE B8. Long-term effects of school absence on alternative measures of adult income

	(1)	(2)	(3)	(4)
	OLS		Siblings FE	
	coeff.	rel. size	coeff.	rel. size
Income 1970, exluding zero (in SEK)				
Absence (average, grades 1 and 4)	−0.0023	−0.0012	−0.0033	−0.0017
	(0.0023)		(0.0047)	
Log income 1970 (exluding zero)				
Absence (average, grades 1 and 4)	−37.5179	−0.0079	−22.2282	−0.0047
	(29.4510)		(65.1444)	

Notes: Each panel states the coefficient of total days of absence (average over grades 1 and 4). Number of observations: income 1970 excluding zero 4,225, log income 1970 4,154. Parish-clustered standard errors in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

C Appendix: Education in the 1930s

C.1 Schooling and the schooling system

Compulsory schooling in Sweden was introduced in 1842. In the 1930s and 1940s, the period during which our cohorts grew up, all children were required to attend public school, *Folkskola*, for at least six years, starting at the age of seven. In 1936 the government decided to increase compulsory schooling by one year in all districts over a twelve year period, and in 1937 the school year length was increased stepwise from 34.5 to 39 weeks (Fischer et al., 2019).[1] Students attended elementary school full time, six days a week, with instruction ending at noon on Saturdays.[2]

The responsibility for providing compulsory education was decentralized to 2,400 school districts, but 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*. Instruction was generally done in classes separated by grade. When the number of students was low, schools were 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., 2019). Students who decided to take more than compulsory education followed a tracking system and generally left *Folkskola* after grade 4 or grade 6 to enter lower secondary school (*Realskola*). All other students remained in *Folkskola* until they reached the

[1] About half of the individuals in our dataset are affected by the reform extending compulsory schooling by one year. Although all our specifications include a full set of parish fixed effects in order to control for differences across parishes, differences in the timing of the introduction might still be unaccounted for by parish fixed effects. We therefore additionally control for years of compulsory schooling in the long-term outcomes specifications, but this does not affect our results. Regarding the length of the school year, we control for length of the school year in weeks in our specifications.

[2] 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.

compulsory years of schooling. From 1939 and onwards the admission to *Realskola* was based on grades received in elementary school. After four or five years of lower secondary schooling, students either entered upper secondary school (*Gymnasium*) or finished their education.

C.2 The grading system

Three theoretical subjects were taught in *Folkskola*: math; reading and speaking; and writing. A 1940 Royal Commission established precise guidelines for teachers to evaluate and grade their students' performance (SOU, 1942). 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. While all students had to take math, writing, and reading and speaking, writing was not always graded in the very first school year.

While exams were not standardized across the country, teachers were provided with clear grading guidance. The grading scale included seven levels, where 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 during the period of interest, 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.[3] 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 be expected to have knowledge corresponding to this level and receive this grade. The recommendation was to also have BA as the normal mark in grade 1 and grade 2, but the the recommendation to teachers were to be restrictive in the use of any high or low marks for children in these grades.

[3]One third of students of a cohort should receive a better grade and one third a poorer grade.

As their main organizational tool teachers kept daily records in an exam catalog called *Dagbok med examenskatalog*. In these catalogues, 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 (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 type and the final grades by subject in a separate column for end-of-school-year information.

The WWII falls in the time when we observe our sample individuals in primary school. Sweden was neutral in the war and there was an oversupply of teachers (Paulsson, 1946). 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. About 50,000 Finnish children came to Sweden 1941-1944.[4]

[4] Children spent on average two years in the country (Santavirta, 2012) and were very evenly distributed across Swedish counties (as share of total population)(De Geer, 1986). The Finnish children were taught by the regular teacher in the same class as natives. The government granted money to school districts to provide four extra hours of Swedish language class for migrant children in school age (Fredriksson, 1971)

D Appendix: Data

D.1 The administrative church records data and its representativeness

The data we use in this paper combine several historical and administrative data sources. The base of 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. The base dataset was initially used to evaluate an infant and maternal health program that the Swedish Government introduced between 1931 and 1933 (see Bhalotra et al. (2017) and Bhalotra et al. (2018)).

The base data includes information on all births as recorded in church records.[5] These records contain individual information on name, gender, date of birth and parish of birth. The records also provide information on whether the child was born in a hospital, whether the birth was a twin birth, and whether the child was born out of wedlock. It also includes information on the parents' occupation at the time of the child's birth, which we use to create an indicator for whether the child's mother was employed at the child's birth and indicators for whether the father is an agricultural, a production or a service worker.[6]

The 133 parishes for which individual level data was collected includes two cities and 57 parishes where the infant care trial was introduced in the early 1930's, and an identified suitable control group for these areas based on observable parish characteristics in the 1930 census. The best matches in terms of observable characteristics were identified using the Mahalanobis distance metric and the observable parish characteristics average income; net wealth; employment shares in manufacturing and agriculture; population density; proportion of fertile married women; and a dummy variable for urban locations. The matching was done in random order and without replacement.

Match quality was assessed by conducting tests of differences in means for a set of

[5]Since the eighteenth century, the Swedish clergy created an information system that included all individuals in their parishes. Membership of the Church of Sweden could actively be cancelled if an individual wanted to enter another denomination, but church records nevertheless covered all citizens.

[6]To construct these indicators, we use the first digit of the Historical International Standard Classification of Occupations (HISCO) code for the fathers' occupation. The HICSO code is historical version of today's International Standard Classification of Occupations (ISCO) code, see van Leeuwen et al. (2002).

observable covariates. Table D9 includes information from (Bhalotra et al., 2017). Panel A presents summary statistics for observable characteristics from the 1930 census on which the matching was done. Tests for balance (column 6) indicate that the treated (column 2) and matched-controls (column 5) are balanced on observable characteristics,. Important for our purposes we also show the standardised difference between the treatment group and the rest of Sweden (column 3, labelled "Other") also indicate balance, which suggests that the sample of 133 parishes were representative of Sweden.[7]

TABLE D9. Characteristics of Matched and Control Districts

	All (1)	Treated (2)	Other (3)	Std. Dif. (2) vs. (3)	Matched (5)	Std. Dif. (2) vs. (5)
Matching Characteristics from the 1930 Census.						
Agriculture	0.340	0.324	0.340	-0.040	0.302	0.054
Manufacturing	0.318	0.340	0.318	0.096	0.345	-0.018
Fertile Women	0.121	0.101	0.121	-0.135	0.100	0.060
Income	811	839	810	0.042	847	-0.013
Wealth	2,525	2,703	2,521	0.080	2,655	0.022
Urban	0.334	0.439	0.331	0.158	0.437	0.003
Population	6,271,266	258,418	6,004,052		160,987	

The table contains local characteristics from the 1930 census. We compare statistics for treated (column 2) and matched (column 5) areas with national averages in the column All (the whole of Sweden) and Other (including all non-treated parishes and cities in Sweden).'*Std Dif.*' presents the standardised difference (cf. Imbens and Wooldridge, 2009); a standardised difference of less than 0.25 is generally viewed as acceptable.

D.2 Matching of base data and schooling data

Individual schooling information was collected from historical archives in each parish. Specifically, we collected the exam catalogs in which teachers made systematic notes about types of absence and reported grades for each student, for each primary school in the 133 parishes in our base dataset. As shown in Figure 1 each student is listed with their first name, surname, date of birth and parents' name. Using this information, we merge the schooling information onto the base dataset. We focus on information from garde 1 and grade 4 when students are 7 and 10 years old.

[7]The programme documentation of the infant trial indicates that the National Board of Health selected areas where the trail was implemented "randomly" to be representative of the country, and Table D9 confirms that they managed well with this task.

We are able to match schooling information for 17,999 children out of the 30,150 children with church records born between 1930 and 1935. The reasons why we are not able to get a perfect match are that (i) exam catalogues were destroyed or cannot be found in the archives, (ii) there is not sufficient information for identifying an individual, (iii) an individual left the sample parish and moved before school age, and/or (iv) 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. We significantly reduce the matching problem related to migration by tracing migrants and their exam catalogues in a different parish than their birth parish 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 school absence and educational performance given the socio-economic background.

The socio-economic characteristics in the church books 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 the matched school data should equal the mean of the base data. Table D10 shows the results from an exercise where we test this. 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 sub-sample 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. But the absolute difference is quite small and we do not see why this should be 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 D10 does not indicate systematic sample selection.

Still, to investigate sample selection further Table D11 gives the baseline short-term effects on academic performance separately for individuals who did not move parishes between birth and grade 4 (same-parish matches) and individuals who did move (movers) in our sibling-grade panel. If moving is selective with respect to absence and performance in school, the effects should differ between the samples. This does not seem to be the case. In fact, for the siblings FE model, the point estimates are exactly the same, even if the association is only statistically different from zero for non-movers while few observations reduces power in the mover sample.

TABLE D10. 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 D11. 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: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

D.3 Matching of later life outcomes

Educational attainment, employment and income

We merged individuals in the matched school data set to the 1960 and the 1970 census which cover the entire population of Sweden on 1st November in these two years (SCB, 1972, 1962). To match individuals we use information on first name, date of birth and parish of birth. The censuses contains information on employment status, and the 1970 census also have information on final education and income.[8] We observe 11,570 and 10,246 individuals in the matched school data set in 1960 (at ages 25–30) and 1970 (at ages 35–40), respectively. Upon matching to death registers, we see that 37% of the unmatched individuals died before the 1970 census enumeration.

Pensions

The matched school data set is also merged with official tax registers available on an annual basis for 2003-2014 measured in SEK.[9] The tax register includes records for pension income. Individuals were matched to the matched school data set using their unique social security number. For our cohorts full pensions require thirty years of contributions, and the level of the pension is based on the best fifteen years (Sundén, 2006). This means that pensions mirrors life time income. Another advantage of using pension income compared to annual income is that it is insensitive to career interruptions, such as those associated with childbearing, which could influence the income observed in 1970. Upon matching official tax registers to death registers, we see that 66% of the unmatched individuals had died before the year 2003.

[8] The educational attainment measure we use is an indicator taking the value 1 if an individual leaves *Folkskola* after grade 4 and attends lower secondary or if the individual leaves *Folkskola* after the compulsory years of schooling and enrols into secondary education afterwards. The indicator takes the value of 0 for anyone living in districts with 8 or 9 years of compulsory schooling, but not completing further education, as well as anyone dropping out of lower secondary education.

[9] While tax registers are available from 2002, data for 2002 is of worse quality compared to the other years wherefore we use annual information from 2003 onwards. Our pension measure refers to average pensions 2003-2014.

Mortality

The parish records include subsequent mortality for infants covering all deaths for the period up until 1946. A very detailed and strict reporting procedure regarding death causes was applied where local clergymen had to make monthly reports to Statistics Sweden in cases where no doctor had been involved.[10] The church records thus allow us to track mortality during the first 10-15 years of life for our sample.

To identify mortality beyond age 10-15, and to validate information on mortality during childhood, we use the Swedish Death Index (Federation of Swedish Genealogical Societies, 2014) which includes the universe of all deaths occurring between 1901–2013. Individual records were matched based on date of birth, sex, forename, surname, and birth parish. To validate the matching, we use a dataset containing burial records (Swedish Genealogical Society, 2012). As a second source of validation of adult mortality we also use the official tax records as individuals identified as dead before 2003 should not show up in these registers. Most individuals can be uniquely matched based on first name, date of birth, and parish of birth. To get information in surnames, which is especially important for women who generally change their names if getting married, we used the 1970 census. Surnames were used to validate the match, and in the rare cases of duplicates, to identify the correct match.

[10] For details on the reporting of deaths, cf. Karlsson et al. (2014); Statistics Sweden (1915) and Hultkvist (1940).

D.4 Variable definitions

Information from Church Parish Records

Birth year Birth year.

Month of birth Month of birth.

Female Dummy variable taking on the value one for female births.

Twin Dummy variable taking on value one for (mono- and dizygotic) twins.

Wedlock Dummy variable taking on value one for children born to married mothers.

Hospital birth Dummy variable taking on value one for child being born in hospital.

SES Classification of head of household profession according to HISCO 9-point scale (Leeuwen et al., 2002).

Mother employed Dummy variable taking on value one for mother of child being employed at the time of birth.

Mother birth year Mother birth year.

Father birth year Father birth year

Variables from Exam Catalogues:

Days of absence Days spent absent in grade 1 or 4.

Days of sickness absence Days spent absent due to sickness in grade 1 or 4.

Days of non-sickness absence Days of total absence minus days spent absent due to sickness in grade 1 or 4.

Math Mark for "math" in grade 1 or 4.

Writing Mark for "writing" in grade 1 or 4.

Reading and Speaking Mark for "reading and speaking" in grade 1 or 4.

Average grade points Grade point average of subjects "math", "reading and speaking" and "writing" in grade 1 or 4.

Class size Class size in grade 1 or 4.

Grade range Lowest and highest grade taught to students in the same classroom in grade 1 or 4.

Weeks Length of the school year measured in weeks in grade 1 or 4.

School School in grade 1 or 4.

Teacher Teacher in grade 1 or 4.

Parish Parish of residence in grade 1 or 4.

Variables from 1960 Population and Household Census:

Employed Dummy variable taking on value one for someone working parttime (at least 20 hours per week) or fulltime (at least 35 hours per week).

Variables from 1970 Population and Household Census:

More than *Folkskola* Dummy variable taking on value one for someone having more than compulsory Folkskola education.

Employed Dummy variable taking on value one for someone working parttime (at least 20 hours per week) or fulltime (at least 35 hours per week).

Labor market income Taxable labour earnings in SEK.

Variables from Tax register:

Pensions Average annual pensions 2003–2014 in SEK.

Variables from the Swedish Death Index:

Passed away before 1960 Dummy variable taking value one for someone passing away before Census enumeration in 1960.

Passed away before 1970 Dummy variable taking value one for someone passing away before Census enumeration in 1970.

Passed away before 2003 Dummy variable taking value one for someone passing away before 2003 (the first year with pension income information).

E Appendix: The pension system

Sweden implemented its first public old age pension system in 1913 whereby all citizens were entitled to a pension (Lundberg and Åmark, 2001). The 1913 version of the system was modified in two larger reforms in 1959 and in 1998, respectively. For cohorts born 1930–1935 the system implemented in 1959, the so-called ATP system, is the only scheme of relevance.

The ATP system, with the first pensions paid in 1963, was a pay-as-you-go defined-benefit scheme with a stated pension age at 65, with the possibility to receive an early pension from age 60 and to postpone retirement to age 70 (Kridahl, 2017). The pension scheme consisted of two main parts: (i) a flat benefit independent of previous income financed by the national budget and (ii) an earnings-related benefit covering all employees. The earnings-related benefit corresponded to 60 per cent (up to a ceiling) of 15 years of the highest earnings during 30 years of active labor force participation (Selén and Ståhlberg, 2007).[11] The pension right included earnings, but also social security transfers, e.g. unemployment insurance and parental benefits. Individuals with low or no earnings-related benefit received a supplementary benefit.[12]

With pensions based on the 15 years of highest earnings, pensions in the ATP system mirrors earnings at advanced stages of the career. With earnings generally levelling off around age 40-45 a majority of workers did not get any substantial increase in pension benefits from continuing an employment after 65 (Laun and Wallenius, 2015). Still, individuals could postpone retirement until age 70 and in 2010 20 per cent of all men in the ages ages 65–69 and ten per cent in the ages 70–74 were still employed (SOU, 2015).[13]

Married women born before 1945 whose husband passed away before 1990 could receive an additional widow pension. The widow pension represent the most important deviation from the general pension rules. The widow pension corresponded to 40 per cent of the husband's earnings-related benefit and lasted until age 65 (Olofsson, 1993). After age 65

[11] If an individual had worked less than 30 years a deduction in relation to the number of missing years was done.

[12] Individuals who could not perform gainful employment, could claim retirement through the disability insurance scheme (Johansson et al., 2014).

[13] Individuals claiming an early pension from age 60 received a reduced flat pension.

the widow pension was still 40 per cent of the husband's earnings-related benefit if the widow had never been employed, but was reduced if the widow herself had an earnings-related pension. This reduction also to some extent related to the birth year of the widow.[14] The widow pension came to an end in 1990, but individuals with very low pensions could then get guarantee pension, starting by age 65.

[14]For a widow born 1930 this share was 60 per cent. The share declined gradually by two percentage points for each cohort until cohorts born 1935 or later, for whom the share was 50 per cent.

F Appendix: Linearity test

We want to test if the effect of absence $\tau(D)$ is linear in D. For this purpose, estimate with a moving cutoff c: $\mathbb{E}[\tau(D) \mid D \geq c] - \mathbb{E}[\tau(D) \mid D < c], \forall c = 1, \ldots, \infty$. This is estimated using the regression equation[15]

$$Y = \beta_0 + \tau_c 1(D \geq c) + X\beta_1 + Q\beta_2 + S + T + \delta + u. \tag{F1}$$

This is similar to our preferred short-term estimation specification, but allows the coefficient of absence to differ in the days of absence. If we replace this regression function with another one, defined as

$$Y = \beta_0 + \tau_c' 1(D \geq c) \cdot \left(\bar{D}_c - \bar{D}_{-c}\right) + X_i\beta_1 + Q\beta_2 + S + T + \delta_i + u, \tag{F2}$$

where $\bar{D}_c = \mathbb{E}[D \mid D \geq c]$ and $\bar{D}_{-c} = \mathbb{E}[D \mid D < c]$, the estimand becomes

$$\tau_c' = \frac{\mathbb{E}[\tau(D) \mid D \geq c] - \mathbb{E}[\tau(D) \mid D < c]}{\mathbb{E}[D \mid D \geq c] - \mathbb{E}[D \mid D < c]} \tag{F3}$$

Under the assumption of linearity, $\tau(D) = \beta D$, this simplifies to:

$$\tau_c' = \frac{\mathbb{E}[\tau(D) \mid D \geq c] - \mathbb{E}[\tau(D) \mid D < c]}{\mathbb{E}[D \mid D \geq c] - \mathbb{E}[D \mid D < c]} = \frac{\beta\left(\mathbb{E}[D \mid D \geq c] - \mathbb{E}[D \mid D < c]\right)}{\mathbb{E}[D \mid D \geq c] - \mathbb{E}[D \mid D < c]} = \beta \tag{F4}$$

Hence, if the effect of sickness absence is linear in D, we should find that $\tau_c' = \beta \forall c$.

[15] For legibility, we leave subscripts implicit.

G Appendix: Bounds analysis

We employ the bounding approach suggested by Oster (2019), 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 particularly useful in the application at hand.

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:[16]

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

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).[17]

Table 9 in the main text shows the bound estimates for the short- and long-term effects of days of absence. Comparing estimates in columns (1) and (2) gives an idea of

[16] This expression is only an approximation, see Oster (2019) 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.

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

the extent to which the raw correlation between absence and the outcome is robust to conditioning on a large number of observables. The estimates are remarkably robust and hardly change at all when including the full set of observables. This should to be kept in mind when considering the estimated bounds in columns (3) and (4).

References

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

Bhalotra, S., Karlsson, M., and Nilsson, T. (2017). Infant Health and Longevity: Evidence from a Historical Intervention in Sweden. *Journal of the European Economic Association*.

Bhalotra, S., Karlsson, M., Nilsson, T., and Schwarz, N. (2018). Infant health, cognitive performance and earnings: Evidence from inception of the welfare state in sweden.

De Geer, E. (1986). *De finländska krigsbarnen i Sverige 1941-1948: Suomalaiset sotalapset Ruotsissa 1941-1948*. Föreningen Norden, Stockholm.

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

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

Fischer, M., Karlsson, M., Nilsson, T., and Schwarz, N. (2019). The Long-Term Effects of Long Terms – Compulsory Schooling Reforms in Sweden. *Journal of the European Economic Association*, 18(6):2776–2823.

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

Hener, T., Rainer, H., and Siedler, T. (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.

Hultkvist, G. (1940). Några anmärkningar till vår nya dödsorsaksstatistik. *Allmänna Svenska Läkartidningen*, 17:58–60.

Imbens, G. and Wooldridge, J. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1):5–86.

Johansson, P., Laun, L., and Palme, M. (2014). Pathways to retirement and the role of financial incentives in sweden. In *Social Security Programs and Retirement around the World: Disability Insurance Programs and Retirement*, pages 369–410. University of Chicago Press.

Karlsson, M., Nilsson, T., and Pichler, S. (2014). The impact of the 1918 spanish flu epidemic on economic performance in sweden: An investigation into the consequences of an extraordinary mortality shock. *Journal of health economics*, 36:1–19.

Kridahl, L. (2017). *Time for Retirement: Studies on how leisure and family associate with retirement timing in Sweden.* PhD thesis, Department of Sociology, Stockholm University.

Laun, T. and Wallenius, J. (2015). A life cycle model of health and retirement: The case of swedish pension reform. *Journal of Public Economics*, 127:127–136.

Lundberg, U. and Åmark, K. (2001). Social rights and social security: The swedish welfare state, 1900-2000. *Scandinavian journal of history*, 26(3):157–176.

Olofsson, G. (1993). Det svenska pensionssystemet 1913–1993: historia, struktur och konflikter. *Arkiv för studier i arbetarrörelsens historia*, (58-59):29–84.

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Paulsson, E. (1946). *Om folkskoleväsendets tillstånd och utveckling i Sverige under 1920- och 1930-talen (till omkring år 1938).* Länstryckeriaktiebologet.

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.

SCB (1962). Population and housing census 1960. *National Bureau of Statistics.*

SCB (1972). Population and housing census 1970. *National Bureau of Statistics.*

Selén, J. and Ståhlberg, A.-C. (2007). Why Sweden's pension reform was able to be successfully implemented. *European Journal of Political Economy*, 23(4):1175–1184.

SOU (1942). Betänkande med utredning och förslag angående betygsättningen i folkskolan, angivet av inom ecklesiastikdepartementet tillkalade sakkunniga. *Statens offentliga utredningar*, 1942:11.

SOU (2015). Längre liv, längre arbetsliv - förutsättningar och hinder för äldre att arbeta längre: Delrapport av pensionsåldersutredningen(sou 2012:128). *Statens offentliga utredningar.*

Statistics Sweden (1915). *Dödsorsaker år 1911.* Stockholm: Kungl. Statistiska Centralbyrån.

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

Swedish Genealogical Society (2012). *Begravda i Sverige 2*. Stockholm: Swedish Genealogical Society.

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